



# Code of Practice on Disinformation – Report of TikTok for the period 1 January 2024 - 30 June 2024



# **Table of Contents**

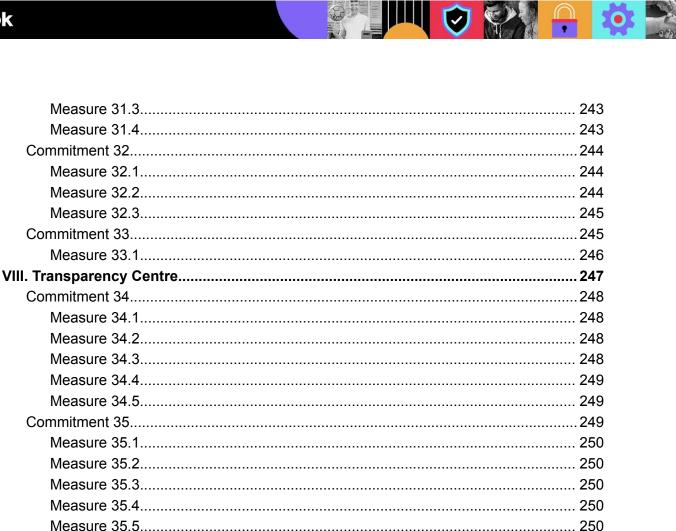
Exe	cutive summary	12
II. S	crutiny of Ad Placements	.17
(	Commitment 1	18
	Measure 1.1	19
	Measure 1.2	27
	Measure 1.3	29
	Measure 1.4	30
	Measure 1.5	30
	Measure 1.6	30
(	Commitment 2	31
	Measure 2.1	32
	Measure 2.2	35
	Measure 2.3	35
	Measure 2.4	38
(	Commitment 3	40
	Measure 3.1	41
	Measure 3.2	41
	Measure 3.3	42
III. P	Political Advertising	.43
(	Commitment 4	44
	Measure 4.1	44
	Measure 4.2	44
(	Commitment 5	46
	Measure 5.1	46
(	Commitment 6	47
	Measure 6.1	47
	Measure 6.2	48
	Measure 6.3	48
	Measure 6.4	48
	Measure 6.5	48
	Measure 7.1	49
	Measure 7.2	51
	Measure 7.3	51
	Measure 7.4	51
(	Commitment 8	52
	Measure 8.1	52
	Measure 8.2	52
(	Commitment 9	53
	Measure 9.1	53
	Measure 9.2	53
(	Commitment 10	54
	Measure 10.1	54



Measure 10.2	54
Commitment 11	
Measure 11.1	
Measure 11.2	
Measure 11.3	
Measure 11.4	
Commitment 12	
Measure 12.1	
Measure 12.2	
Measure 12.3	
Commitment 13	
Measure 13.1	
Measure 13.2	
Measure 13.3	
IV. Integrity of Services	
Commitment 14	
Measure 14.1	
Measure 14.2	
Measure 14.3	
Commitment 15	
Measure 15.1	
Measure 15.2	
Commitment 16	
Measure 16.1	
Measure 16.2	
V. Empowering Users	
Commitment 17	
Measure 17.1	
Measure 17.2	
Measure 17.3	
Commitment 18	
Measure 18.1	
Measure 18.2	
Measure 18.3	
Commitment 19	
Measure 19.1	
Measure 19.2	
Commitment 20	
Measure 20.1	
Measure 20.2	
Commitment 21	
Measure 21.1	
Measure 21.2	
Measure 21.3	



	Commitment 22	. 187
	Measure 22.1	188
	Measure 22.2	188
	Measure 22.3	189
	Measure 22.4	189
	Measure 22.5	189
	Measure 22.6	190
	Measure 22.7	190
	Commitment 23	. 191
	Measure 23.1	192
	Measure 23.2	195
	Commitment 24	. 196
	Measure 24.1	197
	Commitment 25	. 205
	Measure 25.1	206
	Measure 25.2	206
VI.	Empowering the research community	.207
	Commitment 26	. 208
	Measure 26.1	209
	Measure 26.2	210
	Measure 26.3	215
	Commitment 27	. 215
	Measure 27.1	216
	Measure 27.2	216
	Measure 27.3	217
	Measure 27.4	217
	Commitment 28	. 217
	Measure 28.1	218
	Measure 28.2	221
	Measure 28.3	221
	Measure 28.4	221
	Commitment 29	. 222
	Measure 29.1	222
	Measure 29.2	223
	Measure 29.3	223
VII.	Empowering the fact-checking community	225
	Commitment 30	. 226
	Measure 30.1	227
	Measure 30.2	233
	Measure 30.3	234
	Measure 30.4	234
	Commitment 31	. 235
	Measure 31.1	235
	Measure 31.2	236



 IX. Permanent Task-Force.
 252

 Commitment 37.
 253

 Measure 37.1.
 253

 Measure 37.2.
 253

 Measure 37.3.
 253

 Measure 37.4.
 254

 Measure 37.5.
 254

 Measure 37.6.
 254

 X. Monitoring of Code
 255

 Commitment 38
 256

 Measure 38.1
 256

 Commitment 39
 257

 Commitment 40
 258

 Measure 40.1
 259

 Measure 40.2
 259

 Measure 40.3
 259

 Measure 40.4
 259

 Measure 40.5
 259





Measure 40.6	259
Commitment 41	259
Measure 41.1	260
Measure 41.2	260
Measure 41.3	260
Commitment 42	260
Commitment 43	261
Commitment 44	262
War of aggression by Russia on Ukraine	264
Israel - Hamas Conflict	277
2024 European Parliament Elections	290





Commitments	Measures	Service A
II. Scrutiny of Ad Placements		
	Measure 1.1	
	Measure 1.2	
4	Measure 1.3	
1	Measure 1.4	
	Measure 1.5	
	Measure 1.6	
	Measure 2.1	
	Measure 2.2	
2	Measure 2.3	
	Measure 2.4	
	Measure 3.1	
3	Measure 3.2	
	Measure 3.3	
	III. Political advertising	
4	Measure 4.1	
4	Measure 4.2	
5	Measure 5.1	
	Measure 6.1	
	Measure 6.2	
6	Measure 6.3	
	Measure 6.4	
	Measure 6.5	
	Measure 7.1	
7	Measure 7.2	
/	Measure 7.3	
	Measure 7.4	



8	Measure 8.1	
0	Measure 8.2	
0	Measure 9.1	
9	Measure 9.2	
40	Measure 10.1	
10	Measure 10.2	
	Measure 11.1	
11	Measure 11.2	
"	Measure 11.3	
	Measure 11.4	
	Measure 12.1	
12	Measure 12.2	
	Measure 12.3	
	Measure 13.1	
13	Measure 13.2	
	Measure 13.3	
IV. Integrity of services		
	Measure 14.1	
14	Measure 14.2	
	Measure 14.3	
15	Measure 15.1	
15	Measure 15.2	
16	Measure 16.1	
10	Measure 16.2	
V. Empowering users		
	Measure 17.1	
17	Measure 17.2	
	Measure 17.3	
18	Measure 18.1	





	Measure 18.2	
	Measure 18.3	
19	Measure 19.1	
	Measure 19.2	
20	Measure 20.1	
	Measure 20.2	
	Measure 21.1	
21	Measure 21.2	
	Measure 21.3	
	Measure 22.1	
	Measure 22.2	
	Measure 22.3	
22	Measure 22.4	
	Measure 22.5	
	Measure 22.6	
	Measure 22.7	
	Measure 23.1	
23	Measure 23.2	
24	Measure 24.1	
0.5	Measure 25.1	
25	Measure 25.2	
	VI. Empowering the research commu	nity
	Measure 26.1	
26	Measure 26.2	
	Measure 26.3	
	Measure 27.1	
_	Measure 27.2	
27	Measure 27.3	
	Measure 27.4	
	l .	





	Measure 28.1  Measure 28.2	
20		
20	Measure 28.3	
	Measure 28.4	
	Measure 29.1	
29	Measure 29.2	
	Measure 29.3	
	VII. Empowering the fact-checking com	munity
	Measure 30.1	
30	Measure 30.2	
30	Measure 30.3	
	Measure 30.4	
	Measure 31.1	
24	Measure 31.2	
31	Measure 31.3	
	Measure 31.4	
	Measure 32.1	
32	Measure 32.2	
	Measure 32.3	
33	Measure 33.1	
	VIII. Transparency centre	
	Measure 34.1	
	Measure 34.2	
34	Measure 34.3	
	Measure 34.4	
	Measure 34.5	
	Measure 35.1	
35	Measure 35.2	
	Measure 35.3	





	Measure 35.4	
	Measure 35.5	
	Measure 35.6	
	Measure 36.1	
36	Measure 36.2	
	Measure 36.3	
	IX. Permanent Task-Force	
	Measure 37.1	
	Measure 37.2	
37	Measure 37.3	
31	Measure 37.4	
	Measure 37.5	
	Measure 37.6	
	X. Monitoring of the Code	
38	Measure 38.1	
39	-	
	Measure 40.1	
	Measure 40.2	
40	Measure 40.3	
40	Measure 40.4	
	Measure 40.5	
	Measure 40.6	
	Measure 41.1	
41	Measure 41.2	
	Measure 41.3	
42	-	
43	-	
44		





# About TikTok

TikTok's mission is to inspire creativity and bring joy. We know that ensuring the safety of our community is critical to achieving that goal - this includes making sure our users have a trustworthy experience on TikTok. In a global community such as ours it is natural for people to have different opinions, so we seek to operate on a shared set of facts and reality when it comes to topics that impact people's safety.

This is why we treat disinformation with the utmost seriousness, remain committed to preventing its spread, elevate authoritative information, and continue investing in media literacy to help build resilience among our community against misinformation. At TikTok, we place considerable emphasis on proactive content moderation and are proud that the vast majority of violative content is identified and removed proactively before it receives any views or is reported to us. During Q1 2024 of this reporting period, more than 97% of videos violating our misinformation policies were removed proactively globally.

We are also committed to continuing to keep pace with evolving issues that affect our users. As part of the Digital Services Act (DSA) compliance programme, under which the Code of the Practice of Disinformation (the Code) will soon find a new legislative home, we have <u>implemented</u> a range of measures designed to keep our users safe across a number of key areas including disinformation. For more information please refer to our dedicated <u>European Online Safety Hub</u>.

Our work under the Code continues to reflect our strong commitment to combatting disinformation on our platform and to providing transparency to our wider community about the measures we take.

# Our fourth report under the Code - 1 January to 30 June 2024

Since 2020, TikTok has been committed to, and heavily involved in, the Code process. We continue to meaningfully engage in the Code's Taskforce and all of its working groups and subgroups, including those dedicated to Elections and Generative AI.

In line with our previous reports, in this fourth report we continue to provide detailed descriptions of the measures we take to fight disinformation in compliance with our commitments and disclose robust, granular data in support of those measures.

With more than half of the world's population going to the polls, 2024 represents a significant year for elections and civic processes. This report provides additional details on the significant and ambitious resources we are devoting to protect the integrity of our platform during this time. This includes an update on the recent European Parliament elections ('EU Elections') which was the third largest election in the world in terms of votes cast, with elections taking place in 27 EU countries between 6-9 June 2024.

The mitigation measures we implemented during this critical period include: establishing a Mission Control Centre to provide consistent and dedicated coverage of potential elections-related issues; participating in the Code's Rapid Response System to streamline the exchange of information among civil society organisations, fact-checkers and platforms; ensuring fact-checking coverage for at least one official language of every EU Member State; and launching localised media literacy campaigns.

We are proud of our continuing efforts to protect our community against disinformation and to empower them with the tools they need to be resilient in the face of new misinformation trends and with skills to assess the accuracy of the information they are exposed to.





Our Integrity and Authenticity policies aim to promote a trustworthy, authentic experience for our users. Our policies focus on misinformation and deceptive behaviours. We do not allow false or misleading content that may cause significant harm to individuals or society, regardless of intent. These policies are carefully designed to apply to a wide range of content and the constantly changing nature of misinformation trends, often based on what's happening in the world. We also tackle deceptive behaviour by removing accounts that seek to mislead people or engage in platform manipulation.

We continue to review and develop our policies. In the first half of 2024, we've introduced new policies (e.g. hack and leak policy) and refined existing ones, including our Edited Media and Al-generated content (AIGC) policy to support transparent and responsible content creation while keeping misleading and harmful AIGC off TikTok.

# **Enforcing our policies**

At TikTok, over 40,000 safety professionals work every day to keep our community safe, including over 6,000 people dedicated to moderating EU language content. To do this effectively at scale, we continue to strengthen our automated review process and invest in training for our Trust and Safety team. We have provided more transparency to our community on <a href="https://www.moderate">how we moderate</a> and what <a href="moderate">moderation actions we take</a>. This includes more detail about what content we make <a href="moderate">ineligible for the</a> For You Feed. We disclose the underlying metrics in this report.

While we rely on automated moderation when our systems have a high degree of confidence that content is violative, disinformation differs from other content issues and context and fact-checking are critical. While we use machine learning models to help detect potential misinformation, our approach entails having our trained misinformation moderation team assess, confirm, and remove harmful misinformation violations. Our moderators have access to <u>independent fact-checking partners</u> and our database of previously fact-checked claims to help assess the accuracy of content.

In the first half of 2024, we continued to expand our fact-checking programme in the EU and were thus able to provide fact-checking coverage in at least one official language of each of the Member States during the EU Elections.

# **Transparency and Scrutiny of Advertising**

Ads are reviewed and must comply with our <u>ad policies</u> before being allowed on our platform. These <u>policies</u> specifically prohibit misleading, inauthentic and deceptive behaviours. In particular, our <u>misinformation policies</u> in the EEA covers dangerous misinformation, medical misinformation, dangerous conspiracy theories and synthetic and manipulated media. We have been continuously working on improving the implementation and enforcement of these policies, and are constantly considering new focus areas for which we should develop new policies.

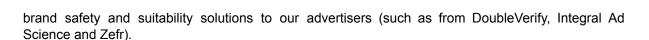
Like all users of our platform, participants in content monetisation programs must adhere to our <a href="Community Guidelines">Community Guidelines</a>, including our <a href="Integrity and Authenticity policies">Integrity and Authenticity policies</a>. Those policies make clear that we do not allow activities that may undermine the integrity of our platform or the authenticity of our users. They also make clear that we remove content or accounts, including those of creators, which contain misleading information that causes significant harm or deceptive behaviours. In certain scenarios, we may remove a creator's access to a creator monetisation feature.

We continue to engage with external stakeholders in order to increase the effectiveness of our scrutiny of ad placements on TikTok. We have expanded the functionality (including choice and ability) of the TikTok Inventory Filter, our in-house pre-campaign safety tool and continue to offer third party

.

<sup>&</sup>lt;sup>1</sup> At TikTok, we take action to moderate or remove content containing harmful misinformation, irrespective of intent (i.e. whether or not the content in question amounts to disinformation), in accordance with our Community Guidelines. For the purposes of alignment with the Code, throughout this report, we use "misinformation" and "disinformation" interchangeably.





# **Prohibiting Political Ads**

TikTok is first and foremost an entertainment platform, and we're proud to be a place that brings people together through creative and entertaining content. While sharing political beliefs and engaging in political conversation is allowed as organic content on TikTok, our policies prohibit our community, including politicians and political party accounts, from placing political ads or posting political branded content. We also prevent governments, politicians and political party accounts from accessing our monetisation features and campaign fundraising.

We welcome the EU Regulation on Transparency and Targeting of Political Advertising which came into force earlier this year and are continuing to focus on enforcement of our political ads prohibition in advance of the majority of the provisions applying next year.

By prohibiting political advertising, campaign fundraising and limiting access to certain monetisation features, we're aiming to strike a balance between enabling people to discuss the issues that are relevant to their lives while also protecting the creative, entertaining platform that our community wants.

# **Ensuring the Integrity of Services**

Our Integrity and Authenticity policies (<u>I&A policies</u>) robustly prohibit deceptive behaviours and we use a range of tactics, techniques and procedures to enforce those policies.

Our updated Edited Media and Al-Generated Content (AIGC) policy (formerly referred to as "synthetic media" policy) addresses the use of content created or modified by AI on our platform. While we welcome the creativity that new AI may unlock, users must proactively disclose when their content is AI-generated or manipulated but shows realistic scenes. Building on our existing AI-generated label for creators to disclose AIGC, we have expanded our efforts in this space by implementing the Coalition for Content Provenance and Authenticity (C2PA) Content Credentials technology, which enables our systems to recognise and automatically label AIGC that originated on other major platforms.

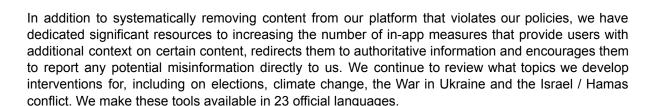
We also continue to fight against covert influence operations (CIO) and we do not allow attempts to sway public opinion while misleading our platform's systems or community about the identity, origin, operating location, popularity, or purpose of the account. In order to provide more regular and detailed updates about the CIOs we disrupt, we have introduced a new dedicated transparency report on CIOs, which is <u>available in TikTok's Transparency Centre</u> and in which we have added new information about operations that we have previously removed and that have attempted to return to our platform with new accounts.

Last but not least, TikTok is proud to have taken an active participation in the set up of the Code's Rapid Response System, which was in place from April to June 2024, in order to streamline the exchange of information among civil society organisations, fact-checkers, and online platforms. The Rapid Response System was a first of its kind time-bound dedicated framework for cooperation among signatories during the EU Elections, which allowed non-platform signatories to flag time sensitive content, accounts or trends that may present threats to the integrity of the electoral process. This initiative has proved that information sharing between platforms and civil society is key to effectively combating disinformation.

# **Empowering Users**

We offer our community easy-to-use in-app and online reporting tools so they can flag any content or account they feel is in violation of our Community Guidelines or contains suspected illegal content.





We are involved in a variety of on- and off-platform media literacy campaigns in close collaboration with our fact-checking partners, particularly in relation to elections. We rolled out localised election integrity campaigns in each EU country in advance of the EU Elections, as well as ahead of local elections which took place in Croatia, Finland, France and Slovakia. In total, the 27 Election Centers we launched ahead of the EU Elections were visited more than 7.5 million times in the four weeks leading up to and during the election.

We believe that collaboration helps strengthen our efforts to mitigate harm and misuse on our platform. To better inform our approach to the EU Election, as part of our existing Election Speaker Series, we invited a growing number of external experts, including from the fact-checking community, to share their insights and market expertise with our internal teams.

We also strengthened our approach and efforts to <u>countering influence attempts</u> by making state-affiliated media accounts that attempt to reach communities outside their home country on current global events and affairs ineligible for recommendation in our users' For You feeds, and by prohibiting state-affiliated media accounts in all markets where our state-controlled media labels are available, including in the EU, from advertising outside of the country with which they are primarily affiliated.

# **Empowering Researchers**

We recognise the important role of researchers in helping to identify disinformation trends and practices.

Since 2020, we have published quarterly <u>Community Guidelines Enforcement Reports</u> to provide ongoing insights into the action we take against content and accounts that violate our Community Guidelines, Terms of Service or Advertising Policies. As part of our continued efforts to make it easy to study the TikTok platform, the report also offers access to aggregated data in a downloadable data file. In 2023 we launched, and continued to iterate, our <u>Research API</u> (providing researchers in Europe with access to public data on content and accounts from our platform) as well as our <u>Commercial Content API</u> (bringing transparency to paid advertising and other content that is commercial in nature on TikTok). In 2024, we continued to advance our efforts by launching the new <u>Virtual Compute Environment</u> (VCE) which offers broader access to user data to qualifying non-academic not-for-profit researchers, while ensuring robust security and privacy protections.

# **Empowering the Fact-Checking Community**

TikTok recognises the important contribution of our fact-checking partners in the fight against disinformation.

Our fact-checking programme incorporates fact-checker input into our broader content moderation efforts. This means that this feedback is relayed to TikTok's moderation teams so that they can ensure it is factored into their moderation work. This approach effectively produces a force multiplier to the underlying fact-checking output, ensuring that the disinformation content or trend is more comprehensively and broadly addressed. Our fact-checking repository, that we have continued to expand, ensures that our teams and systems leverage the full scope of insights our fact-checking partners submit to TikTok.

We have continued to progress the scale-up of our fact-checking programme across Europe where we now have coverage in at least one official language of every EU Member State.

# **Looking forward**















TikTok remains fully committed to our commitments under the Code, which will soon become a DSA Code of Conduct. We look forward to continued and meaningful collaboration with the industry, civil society and EU authorities in our collective efforts to prevent our platforms from being misused through deceptive behaviour and the spread of harmful disinformation.





# II. Scrutiny of Ad Placements Commitments 1 - 3



# **II. Scrutiny of Ad Placements**

# Commitment 1

Relevant signatories participating in ad placements commit to defund the dissemination of disinformation, and improve the policies and systems which determine the eligibility of content to be monetised, the controls for monetisation and ad placement, and the data to report on the accuracy and effectiveness of controls and services around ad placements

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	Yes
If yes, list these implementation measures here [short bullet points].  Guidance to support the identification of policies. Improving identification. Improvement of the enforcement of the policies themselves (not the policy wording).	<ul> <li>Improved the enforcement of four granular harmful misinformation ad policies in the EEA. As mentioned in our H2 2023 report, the policies cover:         <ul> <li>Medical Misinformation</li> <li>Dangerous Misinformation</li> <li>Synthetic and Manipulated Media</li> <li>Dangerous Conspiracy Theories</li> </ul> </li> <li>Expanded the functionality (including choice and ability) in the EEA of our in-house pre-campaign brand safety tool, the TikTok Inventory Filter.</li> <li>Co-chaired the working subgroup on Elections (Crisis Response) as well as continued to regularly participate in the working group on Ad Scrutiny.</li> </ul>
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No



If yes, which further implementation measures do you plan to put in place in the next 6 months?

We are continuously reviewing and improving our tools and processes to fight disinformation and will report on any further development in the next COPD report.

Measure 1.1



### **QRE 1.1.1**

To help keep our platform welcoming and authentic for everyone, we are focused on ensuring it is free from harmful misinformation.

## (I) Our policies and approach

Our Integrity & Authenticity (**I&A**) policies within our CGs are the first line of defence in combating harmful misinformation and deceptive behaviours on our platform. All users are required to comply with our CGs, which set out the circumstances where we will remove, or otherwise limit the availability of, content.

Paid ads are also subject to our <u>ad policies</u> and are reviewed against these policies before being allowed on our platform. Our ad policies specifically prohibit inaccurate, misleading, or false content that may cause significant harm to individuals or society, regardless of intent. They also prohibit other misleading, inauthentic and deceptive behaviours. Ads deemed in violation of these <u>policies</u> will not be permitted on our platform, and accounts deemed in severe or repeated violation may be suspended or banned.

Last year, in order to improve our existing ad policies, we launched four new, more granular policies in the EEA. The policies cover:

- Medical Misinformation
- Dangerous Misinformation
- Synthetic and Manipulated Media
- Dangerous Conspiracy Theories

We have been constantly working on improving the implementation of these policies, and reflecting on whether there are further focused areas for which we should develop new policies.

Our <u>ad policies</u> require advertisers to meet a number of requirements regarding the landing page. For example, the landing page must be functioning and must contain complete and accurate information including about the advertiser. Ads risk not being approved if the product or service advertised on the landing page does not match that included in the ad.

In line with our approach of building a platform that brings people together, not divides them, we have long prohibited political ads and political branded content. Specifically, we do not allow paid ads (nor landing pages) that promote or oppose a candidate, current leader, political party or group, or content that advocates a stance (for or against) on a local, state, or federal issue of public importance in order to influence a political decision or outcome. Similar rules apply in respect of branded content. We also classify certain accounts as Government, Politician, and Political Party Accounts (GPPPA) and we have introduced restrictions on these at an account level. This means accounts belonging to the government, politicians and political parties will automatically have their access to advertising features turned off. We make exceptions for governments in certain circumstances e.g., to promote public health.



We make various brand safety tools available to advertisers to assist in helping to ensure that their ads are not placed adjacent to content they do not consider to fit with their brand values. While any content that is violative of our CGs, including our I&A policies, is removed, the brand safety tools are designed to help advertisers to further protect their brand. For example, a family-oriented brand may not want to appear next to videos containing news-related content. In support of these principles, we voluntarily adopted GARM's Brand Safety Floor and Suitability Framework (the "GARM Framework").

## (II) Verification in the context of ads

We provide verified badges on some accounts including certain advertisers. Verified badges help users make informed choices about the accounts they choose to follow. It's an easy way for notable figures to let users know they're seeing authentic content, and it helps to build trust among high-profile accounts and their followers. For individuals, non-profits, institutions, businesses, or official brand pages, this badge builds an important layer of clarity with the TikTok community. We consider a number of factors before granting a verified badge, such as whether the notable account is authentic, unique, and active.

In the EU, we apply an internal label to accounts belonging to a government, politician, or political party. Once an account has been labelled in this manner, a number of policies will be applied that help prevent misuse of certain features e.g., access to advertising features and solicitation for campaign fundraising are not allowed. We strengthened our approach to countering influence attempts by:

- Making state-affiliated media accounts that attempt to reach communities outside their home country on current global events and affairs ineligible for recommendation, which means their content won't appear in the For You feed.
- Prohibiting state-affiliated media accounts in all markets where our state-controlled media labels are available from advertising outside of the country with which they are primarily affiliated.
- Investing in our detection capabilities of state-affiliated media accounts.
- Working with third party external experts to shape our state-affiliated media policy and assessment of state-controlled media labels.

# SLI 1.1.1 Numbers by actions enforcing policies above (specify if at page and/or domain level)

# Methodology of data measurement:

We have set out the number of ads that have been removed from our platform for violation of our political content policies, as well as our more granular policies on medical misinformation, dangerous misinformation, synthetic and manipulated media and dangerous conspiracy theories.

The majority of ads that violate our four granular misinformation ad policies would have been removed under our existing policies. In cases where an ad is deemed violative for other policies and also for our newer granular misinformation policies, the removal is counted under the existing policy. Therefore, the second column below shows only the number of ads removed



where the sole reason was one of these four additional misinformation policies, and does not include ads already removed under our existing policies or where misinformation policies were not the driving factor for the removal.

The data below suggests that our existing policies (such as political content and other policy areas such as our inaccurate, misleading, or false content policy) already cover the majority of harmful misinformation ads, due to their expansive nature of coverage.

Note that numbers have only been provided for monetised markets and are based on where the ads were displayed.

	Number of ad removals under the political content ad policy	Number of ad removals under the four granular misinformation ad policies
Member States		
Austria	32	0
Belgium	73	2
Bulgaria	N/A	N/A
Croatia	N/A	N/A
Cyprus	N/A	N/A
Czech Republic	44	0
Denmark	14	0
Estonia	N/A	N/A
Finland	18	0
France	255	67
Germany	133	15



29	0
24	0
14	0
63	2
N/A	N/A
N/A	N/A
0	0
N/A	N/A
72	1
173	0
5	0
88	0
N/A	N/A
N/A	N/A
163	0
32	0
N/A	N/A
N/A	N/A
8	0
	24 14 63 N/A N/A 0 N/A 72 173 5 88 N/A N/A 163 32 N/A N/A



Total EU	1,232	87
Total EEA	1,240	87

This additional Service Level Indicator provides an estimated financial value of the actions taken by Signatories to demonetise disinformation sources (under SLI 1.1.1). It is based on media metrics available to Signatories (query/bid<sup>2</sup> or impression<sup>3</sup>) and applying an agreed-upon conversion factor provided by a third party designated by the Task-force of the Code (Ebiquity plc.).

# advertising or content that are disinformation. designated as disinformation

SLI 1.1.2 - Preventing The EU Code of Practice Permanent Task-force appointed Ebiquity Plc to serve in the capacity of 'third-party estimator', the flow of legitimate responsible for providing periodic Cost Per Mille ('CPM') conversion factors for use by signatories to derive financial estimates solely in relation to this SLI. This SLI applies the conversion factor developed by Ebiquity to the impressions of ads that have been investment to sites removed from our platform and aims to provide an estimate of the unrealised value of revenue related to the demonetisation of

> The CPM data is provided on a specific EU market basis where available, and currently covers 26 of the 27 EU markets (excluding Luxembourg). Note that we have only provided numbers for monetised markets, based on where the ads were displayed.

> Ebiguity creates the CPM estimates based on aggregated and anonymized media investment data from their Media Data Vault, its proprietary database infrastructure used to service its clients. Ebiquity has developed three dedicated channel CPMs (Programmatic - Display, Programmatic - Online Video, and Paid Social) as well as a Blended CPM, which can be used for general application and is not identifiable to any specific client, agency, vendor, platform, or ad format.

> We have calculated the approximate financial value by using the "Blended CPM" value provided by Ebiquity. We note that this SLI provides an approximate financial value only, and does not reflect how TikTok monetises ad content in practice. It is therefore not an accurate representation of TikTok's actual revenue.

> TikTok was not involved in the development of the conversion factor and so cannot endorse it or the underlying methodology and data. We remain committed to increasing transparency in this chapter, including by participating in the working group to better understand the underlying methodology and ensure the conversion factor provides, as accurately as possible, an approximation of the financial value of actions taken to demonetise disinformation. Note that most of the conversion factors provided by Ebiquity have changed since the last report.

<sup>&</sup>lt;sup>2</sup> Request placed between a seller and buyer of advertising that can detail amongst other things website, specific content, targeting data inclusive of audience or content.

<sup>&</sup>lt;sup>3</sup> Comprehensive calculation of the number of people who have been reached by a piece of media content by passive exposure (viewing a piece of content) or active engagement (visiting a destination).



Member States	Euro value of ads demonetised			
	Number of impressions for ads removed under the political content ad policy	Number of impressions for ads removed under the more granular misinformation policies	Ebiquity Conversion	Euro Value of Ads Demonetised/1000 (accounting for Cost per thousand)
Austria	248,818	0	3	746.454
Belgium	2,714,652	0	5.5	14930.586
Bulgaria	6,335	N/A	1	6.335
Croatia	N/A	N/A	2	0
Cyprus	N/A	N/A	2.5	0
Czech Republic	527,438	10,614	4	2152.208
Denmark	2,399,140	0	4.5	10796.13
Estonia	18,094	N/A	2.5	45.235
Finland	4,722,613	0	4	18890.452
France	2,689,389	0	4	10772.944



	1	I	Т	
Germany	4,071,969	0	5.5	22395.8295
Greece	3,419,384	0	1.5	5129.076
Hungary	618,982	44,010	1.5	994.488
Ireland	70,614	0	3	211.842
Italy	5,270,848	0	4	21093.848
Latvia	N/A	N/A	3	0
Lithuania	N/A	N/A	2.5	0
Luxembourg	57,301	0	0	0
Malta	N/A	N/A	3	0
Netherlands	1,165,938	0	5.5	6412.659
Poland	3,226,973	1,192	1.5	4842.2475
Portugal	653,336	0	2.5	1633.34
Romania	3,155,352	26,977	1.5	4773.4935



Slovakia	N/A	N/A	3	0
Slovenia	N/A	N/A	2	0
Spain	1,391,622	0	3	4199.799
Sweden	423,725	0	3	1288.626
Iceland	N/A	N/A	3	0
Liechtenstein	N/A	N/A	0	0
Norway	127,222	0	5.5	699.721
Total EU	36,852,523	82,793	N/A	131,315.59
Total EEA	36,979,745	82,793	N/A	132,015.31

Measure 1.2	
QRE 1.2.1	We do not currently offer ad revenue sharing for creators in the EEA.  During the reporting period, we offered creator monetisation opportunities such as the Creator Rewards Program. These programs offer creators meeting certain eligibility criteria the opportunity to monetise their content based on a range of factors, including for example the level of user interaction with the videos they post.

	All creators must comply with TikTok's CGs, including our I&A policies. Where creators fail to comply with our CGs, this may result in loss of access to monetisation and / or loss of account access. Users in all EU member states are notified by an in-app notification in their relevant local language where there has been a restriction of their ability to monetise, restriction of their access to a feature, removal or otherwise restriction of access to their content, or a ban of their account.  Our policies prohibit accounts verified as belonging to a government, politician or political party from accessing monetisation features. They will, for instance, be ineligible for participation in content monetisation programs such as our Creator Rewards Program. Along with our existing ban on political advertising, this means that accounts belonging to politicians, political parties and governments will not be able to give or receive money through TikTok's monetisation features, or spend money promoting their content (although exemptions are made for governments in certain circumstances such as for public health).  In addition, we launched the Creator Code of Conduct in April 2024. These are the standards we expect creators involved in TikTok programs, features, events and campaigns to follow on and off-platform, in addition to our Community Guidelines and Terms of Service. Being a part of these creator programs is an opportunity that comes with additional responsibilities,			
	and <u>lerms of Service</u> . Being a part of these creator programs is an opportunity that comes with additional responsibilities, and this code will also help provide creators with additional reassurance that other participants are meeting these standards too. We are actively improving our enforcement guidance and process around this, including building on proactive signalling of off-platform activity.			
SLI 1.2.1	Methodology of data measurement:			
	Our I&A policies within our CGs are the first line of defence in combating harmful misinformation and deceptive behaviours on our platform. All creators are required to comply with our CGs, which set out the circumstances where we will remove, or otherwise limit the availability of, content. TikTok's Creator Rewards Program does not relate to advertising and we do not currently offer ad revenue sharing to creators in the EEA. Creators who breach the Community Guidelines or Terms of Service are not eligible to receive rewards. We have set out the number of ads that have been removed from our platform for violation of our political content policies as well as our four granular policies on medical misinformation, dangerous misinformation, synthetic and manipulated media and dangerous conspiracy theories in SLI 1.1.1. Further, SLI 1.1.2 aims to provide an estimate of the potential impact on revenue of demonetising disinformation. We are working towards being able to provide more data for this SLI.			
Member States	N/A	N/A	N/A	N/A
Total EU				
Total EEA				

Measure 1.3	
	We partner with a number of industry leaders to provide a number of controls and transparency tools to advertising buyers with regard to the placement of ads:
	<b>Controls</b> : We offer pre-campaign solutions to advertisers so they can put additional safeguards in place before their campaign goes live to mitigate the risk of their advertising being displayed adjacent to certain types of user-generated content. These measures are in addition to the CGs, which provide overarching rules around the types of content that can appear on TikTok and are eligible for the For You feed:
QRE 1.3.1	TikTok Inventory Filter: This is our proprietary system which enables advertisers to choose the profile of content they want their ads to run adjacent to. The Inventory Filter is available in 18 jurisdictions in the EEA and is embedded directly in TikTok Ads Manager, the system through which advertisers purchase ads. We have expanded the functionality of this Inventory Filter in various EEA countries. More details can be found <a href="here">here</a> . The Inventory Filter is informed by the GARM Framework and policies include topics which may be susceptible to disinformation.
	TikTok Brand Safety by Integral Ad Science ("IAS"): Advertisers can select IAS (available in France and Germany) within the platform to ensure their ads run adjacent to content verified by IAS. IAS verifies content against the GARM Framework. Some disinformation content may be filtered by IAS as a result of the existing GARM Brand Safety Floor & Brand Suitability Framework categories, such as Sensitive Social Issues.
	<b>Transparency</b> : We have partnered with third parties to offer post-campaign solutions that enable advertisers to assess the suitability of user content that ran immediately adjacent to their ad in the For You feed, against their chosen brand suitability parameters:
	Zefr: Through our partnership with Zefr, advertisers can obtain campaign insights into brand suitability and safety on the platform. Zefr aligns with the GARM Framework.
	IAS: Advertisers can measure brand safety, viewability and invalid traffic on the platform with the IAS Signal platform (post campaign is available in 20 countries in the EEA). As with IAS's pre-bid solution covered above, this aligns with the GARM Framework.
	DoubleVerify: We are partnering with DoubleVerify to provide advertisers with media quality measurement for ads. DoubleVerify is working actively with us to expand their suite of brand suitability and media quality solutions on the platform.



Measure 1.4	
QRE 1.4.1	When TikTok advertises, we buy advertising space only through ad networks (either directly, through publishers or agencies) which allows for direct measurement of brand safety and suitability, via tagging, using leading brand safety tools across all digital media channels. This allows us to mitigate the risk of TikTok ads appearing next to sources of disinformation and be in control of the environment our content is appearing next to.  We use DoubleVerify to ensure our own ads run on or near suitable content, whilst running and monitoring brand safety and suitability metrics across other placements, always updating the context and content of our blocklists as well as to ensure the TikTok brand is protected in any context.  For instance, we monitor the placement of our ads very closely, especially in the context of politically sensitive events such as the War in Ukraine or the Israel / Hamas conflict, and in the event of our ads appearing adjacent to or on sources of disinformation, we are able to identify and investigate the content in question to assess risks using DoubleVerify dashboards. Once identified, we will then adjust any filters or add the publication to our blocklist (which is regularly reviewed and updated) to prevent recurrence.
Measure 1.5	
QRE 1.5.1	We have achieved the TAG Brand Safety Certified seal and the TAG Certified Against Fraud seal by the Trustworthy Accountability Group ("TAG") in the EEA and globally. This required appropriate verification by external auditors. Details of our TAG seal can be found by searching for "TikTok" on their public register which can be found <a href="https://example.com/here">here</a> .  We have been certified by the Interactive Advertising Bureau ("IAB") for the IAB Ireland Gold Standard 2.1 (listed <a href="here">here</a> ) and IAB Sweden Gold Standard 1.0 (listed <a href="here">here</a> ).
QRE 1.5.2	We have achieved the TAG Brand Safety Certified and TAG Certified Against Fraud seals and the IAB Ireland Gold Standard and IAB Sweden Gold Standard 1.0.
Measure 1.6	



QRE 1.6.1	We offer a variety of brand safety tools for preventing ads from being placed beside specific types of content. We continue to invest in our existing partnerships with leading third party brand safety and suitability providers (including DoubleVerify, Integral Ad Science, and Zefr).  We evaluate, on an ongoing basis, whether there are potential new partnerships, including with researchers, that may be appropriate for our platform. Furthermore, our advertising policies help to ensure that the categories of content which are most likely to require such checks and integration of information do not make it onto the platform in the first place.	
QRE 1.6.2	We only purchase ads through ad networks which make robust and reputable brand safety tools available to us. All of our media investment is therefore protected by such tools.	
QRE 1.6.3	We have partnered with several third parties (IAS, Double Verify and Zefr) to offer post-campaign solutions that enal advertisers to assess the suitability of user content that ran immediately adjacent to their ad in the For You feed.	
QRE 1.6.4	Not applicable as TikTok does not rate sources.	
SLI 1.6.1	N/A	

# Commitment 2 Relevant Signatories participating in advertising commit to prevent the misuse of advertising systems to disseminate Disinformation in the form of advertising messages. In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No] Yes If yes, list these implementation measures here [short bullet points]. • Improved the enforcement of four granular harmful misinformation ad policies in the EEA. As mentioned in our H2 2023 report, the policies cover:

		<ul> <li>Medical Misinformation</li> <li>Dangerous Misinformation</li> <li>Synthetic and Manipulated Media</li> <li>Dangerous Conspiracy Theories</li> </ul> <li>Expanded the functionality (including choice and ability) in the EEA of our in-house pre-campaign brand safety tool, the TikTok Inventory Filter.</li> <li>Co-chaired the working subgroup on Elections (Crisis Response) as well as continued to regularly participate in the working group on Ad Scrutiny.</li>
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]		No
If yes, which further implementation measures do you plan to put in place in the next 6 months?		We are continuously reviewing and improving our tools and processes to fight disinformation and will report on any further development in the next COPD report.
Measure 2.1		
QRE 2.1.1	Paid ads are subject to our strict <u>ad policies</u> , which specifically prohibit misleading, inauthentic and deceptive behaviours. All ads are reviewed against these policies before being allowed on our platform. In order to improve our existing ad policies, we launched four more granular policies in the EEA in 2023 (covering Medical Misinformation, Dangerous Misinformation, Synthetic and Manipulated Media and Dangerous Conspiracy Theories) which advertisers also need to comply with.	

Bulgaria

Croatia

Cyprus

N/A

N/A

N/A



# Methodology of data measurement: We have set out the number of ads that have been removed from our platform for violation of our political content policies, as well as our four granular policies on medical misinformation, dangerous misinformation, synthetic and manipulated media and dangerous conspiracy theories. The majority of ads that violate our newly launched misinformation policies would have been removed under our existing policies. In cases where an ad is deemed violative for other policies and also for these additional misinformation policies, the removal is counted under the older policy. Therefore, the second column below shows only the number of ads removed where the sole reason was one of these four additional misinformation policies, and does not include ads already removed under our SLI 2.1.1 - Numbers by existing policies or where misinformation policies were not the driving factor for the removal. actions enforcing policies above The data below suggests that our existing policies (such as political content and other policy areas such as our inaccurate, misleading, or false content policy) already cover the majority of harmful misinformation ads, due to their expansive nature of coverage. Note that numbers have only been provided for monetised markets and are based on where the ads were displayed. Number of ad removals under the four granular Number of ad removals under the political content ad misinformation ad policies policy Member States Austria 32 0 73 2 Belgium

N/A

N/A

N/A



Czech Republic	44	0
Denmark	14	0
Estonia	N/A	N/A
Finland	18	0
France	255	67
Germany	133	15
Greece	29	0
Hungary	24	0
Ireland	14	0
Italy	63	2
Latvia	N/A	N/A
Lithuania	N/A	N/A
Luxembourg	0	0
Malta	N/A	N/A
Netherlands	72	1
Poland	173	0
Portugal	5	0
Romania	88	0
Slovakia	N/A	N/A



Total EEA	1,240	87
Total EU	1,232	87
Norway	8	0
Liechtenstein	N/A	N/A
Iceland	N/A	N/A
Sweden	32	0
Spain	163	0
Slovenia	N/A	N/A

Measure 2.2	
QRE 2.2.1	In order to identify content and sources that breach our ad policies, all ads go through moderation prior to going "live" on the platform.  After the ad goes live on the platform, users can report any concerns using the "report" button, and the ad will be reviewed again and appropriate action taken if necessary.  TikTok also operates a "recall" process whereby ads already on TikTok will go through an additional stage of review if certain conditions are met, including reaching certain impression thresholds. TikTok also conducts additional reviews on random samples of ads to ensure its processes are functioning as expected.
Measure 2.3	
QRE 2.3.1	In order to identify content and sources that breach our ad policies, all ads go through moderation prior to going "live" on the platform.

	After the ad goes live on the platform, users can report any concerns using the "report" button, and the ad will be reviewed again and appropriate action taken if necessary.  TikTok also operates a "recall" process whereby ads already on TikTok will go through an additional stage of review if certain conditions are met, including reaching certain impression thresholds. TikTok also conducts additional reviews on random samples of ads to ensure its processes are functioning as expected.				
	We are pleased to be able to report on the ads removed for breach of our political content policies, as well as our more granular misinformation ad policies, including the impressions of those ads in this report.				
SLI 2.3.1		Number of ad removals under the four granular misinformation ad policies		for ads removed under	
Member States					
Austria	32	0	248,818	0	
Belgium	73	2	2,714,652	0	
Bulgaria	N/A	N/A	6,335	N/A	
Croatia	N/A	N/A	N/A	N/A	
Cyprus	N/A	N/A	N/A	N/A	
Czech Republic	44	0	527,438	10,614	
Denmark	14	0	2,399,140	0	
Estonia	N/A	N/A	18,094	N/A	
Finland	18	0	4,722,613	0	
France	255	67	2,689,389	0	



Germany	133	15	4,071,969	0
Greece	29	0	3,419,384	0
Hungary	24	0	618,982	44,010
Ireland	14	0	70,614	0
Italy	63	2	5,270,848	0
Latvia	N/A	N/A	N/A	N/A
Lithuania	N/A	N/A	N/A	N/A
Luxembourg	0	0	57,301	0
Malta	N/A	N/A	N/A	N/A
Netherlands	72	1	1,165,938	0
Poland	173	0	3,226,973	1,192
Portugal	5	0	653,336	0
Romania	88	0	3,155,352	26,977
Slovakia	N/A	N/A	N/A	N/A
Slovenia	N/A	N/A	N/A	N/A
Spain	163	0	1,391,622	0
Sweden	32	0	423,725	0
Iceland	N/A	N/A	N/A	N/A
Liechtenstein	N/A	N/A	N/A	N/A



Norway	8	0	127,222	0
Total EU	1,232	87	36,852,523	82,793
Total EEA	1,240	87	36,979,745	82,793

Measure 2.4				
QRE 2.4.1	We are clear with advertisers that their ads must comply with our strict ad policies (see <u>TikTok Business Help Centre</u> ). We explain that all ads are reviewed before being uploaded on our platform - usually within 24 hours. Where an advertiser has violated an ad policy they are informed by way of a notification. This is visible in their TikTok Ads Manager account, or similar or, where an advertiser has booked their ad through a TikTok representative, then the representative will inform the advertiser of any violations. Advertisers are able to make use of functionality to appeal rejections of their ads in certain circumstances. Ads already on TikTok may go through an additional stage of review if they are reported, if certain conditions are met (e.g., reaching certain impression thresholds) or because of random sampling conducted at TikTok's own initiative.			
SLI 2.4.1	We are pleased to be able to share the number of appeals for ads removed under our political content ad policies and of four granular misinformation ad policies as well as the number of respective overturns. The data shows a reduced number appeals for ads removed under the political content policy evidencing our improved moderation and decision making processes.			
	Number of appeals for ads removed under the the four granular misinformation ad policies	removed under political	Number of overturns of appears under the four more granular misinformation policies	appeal under political
Member States				
Austria	0	12	0	4
Belgium	0	19	0	5



Bulgaria	N/A	N/A	N/A	N/A
Croatia	N/A	N/A	N/A	N/A
Cyprus	N/A	N/A	N/A	N/A
Czech Republic	0	5	0	1
Denmark	0	14	0	5
Estonia	N/A	N/A	N/A	N/A
Finland	0	14	0	3
France	1	90	0	33
Germany	0	60	0	17
Greece	0	9	0	2
Hungary	0	11	0	1
Ireland	0	14	0	1
Italy	0	36	0	5
Latvia	N/A	N/A	N/A	N/A
Lithuania	N/A	N/A	N/A	N/A
Luxembourg	0	0	0	0
Malta	N/A	N/A	N/A	N/A
Netherlands	0	22	0	2
Poland	0	21	0	3



Total EU	1	410	0	105
Norway	0	15	0	2
Liechtenstein	N/A	N/A	N/A	N/A
Iceland	N/A	N/A	N/A	N/A
Sweden	0	17	0	3
Spain	0	48	0	18
Slovenia	N/A	N/A	N/A	N/A
Slovakia	N/A	N/A	N/A	N/A
Romania	0	11	0	1
Portugal	0	7	0	3

## **II. Scrutiny of Ad Placements**

#### Commitment 3

Relevant Signatories involved in buying, selling and placing digital advertising commit to exchange best practices and strengthen cooperation with relevant players, expanding to organisations active in the online monetisation value chain, such as online e-payment services, e-commerce platforms and relevant crowd-funding/donation systems, with the aim to increase the effectiveness of scrutiny of ad placements on their own services.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]

Yes

	Ø		<b>O</b>	2017	

		·	
If yes, list these implementation measures here [short bullet points].		<ul> <li>Continued to regularly participate in the working group on Ad Scrutiny and the working subgroup on Elections (Crisis Response).</li> <li>Sponsored the Global Fact 11 Global Fact-Checking Conference hosted by the International Fact-Checking Network in June 2024 and hosted a panel discussion on our approach to countering harmful misinformation.</li> </ul>	
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]		No	
If yes, which further implementation measures do you plan to put in place in the next 6 months?		We are continuously reviewing and improving our tools and processes to fight disinformation and will report on any further development in the next COPD report.	
Measure 3.1	·		
QRE 3.1.1	As set out later on in this report, we cooperate with a number of third parties to facilitate the flow of information that may be relevant for tackling purveyors of harmful misinformation. This information is shared internally to help ensure consistency of approach across our platform.  At the end of June 2024, TikTok sponsored and presented at GlobalFact 11 2024 a fact checking summit hosted by the International Fact-Checking Network ("IFCN"). We look forward to attending again in 2025.  We also continue to be actively involved in the Task-force working group for Chapter 2, specifically the working subgroup on Elections (Crisis Response) which we co-chaired. We work with other signatories to define and outline metrics regarding the monetary reach and impact of harmful misinformation. We are in close collaboration with industry to ensure alignment and clarity on the reporting of these code requirements.		
Measure 3.2	and the same of th		



QRE 3.2.1	TikTok continues to believe in the effectiveness and ongoing value of the GARM Framework and is committed to removing harmful misinformation from monetisation. The misinformation guidelines have been developed in coordination with the European Commission and in consultation with NGO partners, such as Consumers International and Reporters without Borders. We see benefits in common standards and definitions to support consistency of categorising content, adjacency standards, & measurement where we submit transparency data, including that of misinformation within their Aggregated Measurement Report.  We work with industry partners to discuss these relevant topics, in appropriate fora. We work closely with IAB Sweden, IAB Ireland and other organisations such as TAG in the EEA and globally. We are also on the board of the Brand Safety Institute.  We continue to share relevant insights and metrics within our quarterly transparency reports, which aim to inform industry peers and the research community. We continue to engage in the sub groups set up for insights sharing between signatories and the Commission.
Measure 3.3	
QRE 3.3.1	We continue to work closely with GARM, IAB Sweden, IAB Ireland and other organisations such as TAG in the EEA and globally.





# III. Political Advertising Commitments 4 - 13



### III. Political Advertising Commitment 4 Relevant Signatories commit to adopt a common definition of "political and issue advertising". In line with this commitment, did you deploy new Yes implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No] If yes, list these implementation measures here [short We are pleased that Regulation EU 2024/900 on the Transparency and Targeting of Political Advertising set out a common definition of "political advertising". We have been reviewing our bullet points]. policies to ensure that our prohibition of political advertising is at least as broad as the regulation. We co-chair the Elections (Crisis Response) working subgroup and have been working with the other Signatories as part of this. Do you plan to put further implementation measures As we prohibit political advertising we are continuing to focus on enforcement of this policy in light in place in the next 6 months to substantially improve of Regulation EU 2024/900 on the Transparency and Targeting of Political Advertising coming into force and the majority of provisions applying from October 2025. the maturity of the implementation of this commitment? [Yes/No] If yes, which further implementation measures do you N/A plan to put in place in the next 6 months? Measure 4.1 Measure 4.2 **QRE** 4.1.1 TikTok is first and foremost an entertainment platform, and we're proud to be a place that brings people together through (for creative and entertaining content. While sharing political beliefs and engaging in political conversation is allowed as organic measures 4.1 and 4.2) content on TikTok, our policies prohibit our community, including politicians and political party accounts, from placing political ads or posting political branded content.



Specifically, our Politics, Culture and Religion policy prohibits ads and landing pages which:

- reference, promote, or oppose candidates or nominees for public office, political parties, or elected or appointed government officials;
- reference an election, including voter registration, voter turnout, and appeals for votes;
- include advocacy for or against past, current, or proposed referenda, ballot measures, and legislative, judicial, or regulatory outcomes or processes (including those that promote or attack government policies or track records); and
- reference, promote, or sell, merchandise that features prohibited individuals, entities, or content, including campaign slogans, symbols, or logos.

Where accounts are designated as <u>Government</u>, <u>Politician</u>, <u>and Political Party Accounts</u> ("**GPPPA**"), those accounts are banned from placing ads on TikTok, accessing monetisation features and from campaign fundraising. We may allow some cause-based advertising and public services advertising from government agencies, non-profits and other entities if they meet certain conditions and are working with a TikTok sales representative.

We prohibit <u>political content in branded content</u> i.e. content which is posted in exchange for payment or any other incentive by a third party.

We are pleased that political advertising has now been defined by the Regulation EU 2024/900 on the Transparency and Targeting of Political Advertising and we have been reviewing our policies to ensure that our prohibition is at least as broad as the definition set out in the regulation. Our prohibition on political advertising is one part of our <u>election integrity</u> efforts which you can read more about in the EU Elections report.

QRE 4.1.2 (for measures 4.1 and 4.2)

Not applicable at this stage.



#### Commitment 5

Relevant Signatories commit to apply a consistent approach across political and issue advertising on their services and to clearly indicate in their advertising policies the extent to which such advertising is permitted or prohibited on their services.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]		No
If yes, list these implementation measures here [short bullet points].		N/A
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]		We prohibit political advertising and are continuing to focus on enforcement of this policy in light of Regulation EU 2024/900 on the Transparency and Targeting of Political Advertising coming into force and the majority of provisions applying from October 2025.
If yes, which further implementation measures do you plan to put in place in the next 6 months?		N/A
Measure 5.1		
QRE 5.1.1 Not applie	Not applicable as TikTok does not allow political advertising, as outlined in our Politics, Culture and Religion policy. We do not	

party accounts and non-political advertisers expressing political views in advertising.

TikTok takes the enforcement of its political advertising prohibition very seriously. In order to enforce our policies we:

• Use automated and human review to identify advertisements that may violate our political advertising rules;

allow featuring political content in any form of advertising, extending this prohibition to both government, politician, or political

**QRE 6.1.1** 



- Invite uses to report violative content; and
- Remove political advertising content from the platform promptly upon becoming aware of it.

### III. Political Advertising

### Commitment 6

Relevant Signatories commit to make political or issue ads clearly labelled and distinguishable as paid-for content in a way that allows users to understand that the content displayed contains political or issue advertising

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	No
If yes, list these implementation measures here [short bullet points].	N/A
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	We prohibit political advertising and are continuing to focus on enforcement of this policy in light of Regulation EU 2024/900 on the Transparency and Targeting of Political Advertising coming into force and the majority of provisions applying from October 2025.
If yes, which further implementation measures do you plan to put in place in the next 6 months?	N/A
Measure 6.1	

TikTok's CoPD Report January to June 2024 47

Not applicable as TikTok does not allow political advertising.



Measure 6.2				
QRE 6.2.1	TikTok prohibits political ads on the platform. For ads which are permitted, we provide enhanced information accessible through our "About this Ad" functionality, which was refined and enhanced last year in order to ensure compliance with our Digital Services Act Article 26(1) transparency obligations. This feature displays information including who has presented and paid for the ad, the parameters used to determine why the user has received the ad and about how to change those parameters.			
QRE 6.2.2	Not applicable as TikTok does not allow political advertising.			
SLI 6.2.1 – numbers for	N/A			
actions enforcing policies above				
Member States	N/A	N/A	N/A	
Total EU				
Total EEA				

Measure 6.3	
QRE 6.3.1	Not applicable as TikTok does not allow political advertising.
Measure 6.4	
QRE 6.4.1	Not applicable as TikTok does not allow political advertising.
Measure 6.5	
QRE 6.5.1	Not committed. This commitment is not applicable as TikTok is not a messaging app.



#### Commitment 7

Relevant Signatories commit to put proportionate and appropriate identity verification systems in place for sponsors and providers of advertising services acting on behalf of sponsors placing political or issue ads. Relevant signatories will make sure that labelling and user-facing transparency requirements are met before allowing placement of such ads.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	No
If yes, list these implementation measures here [short bullet points].	N/A
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	We prohibit political advertising and are continuing to focus on enforcement of this policy in light of Regulation EU 2024/900 on the Transparency and Targeting of Political Advertising coming into force and the majority of provisions applying from October 2025.
If yes, which further implementation measures do you plan to put in place in the next 6 months?	N/A
Measure 7.1	

Member States

N/A

QRE 7.1.1	Where accounts are designated as Government, Politician, and Political Party Accounts ("GPPPA"), those accounts are banned from placing ads on TikTok (with the exception of certain governmental entities) and from monetisation features. We publish the details of our GPPPA policy on our website, where we set out who we consider to be a GPPPA and the	
	publish the details of our GPPPA policy on our website, where we set out who we consider to be a GPPPA and the restrictions on those types of account.  In the EU, we apply an internal label to accounts belonging to a government, politician, or political party. Once an account has been labelled in this manner, a number of policies will be applied that help prevent misuse of certain features e.g., access to advertising features and solicitation for campaign fundraising are not allowed.	
SLI 7.1.1 – numbers for actions enforcing policies above (comparable metrics as for SLI 6.2.1)		

Total EU		
Total EEA		

N/A



Measure 7.2	
QRE 7.2.1	Not applicable as TikTok does not allow political advertising.  Our Actor Policy aims to protect the integrity and authenticity of our community and prevent actors from evading our tools and processes. If an actor consistently demonstrates behaviour that deceives, misleads or is inauthentic to users and/or to TikTok we apply account level enforcement. This is not exclusive to ads containing political content.  TikTok is dedicated to investigating and disrupting confirmed cases of CIO on the platform. Covert influence operations (CIOs) are organised attempts to manipulate or corrupt public debate while also misleading TikTok systems or users about identity, origin, operating location, popularity, or overall purpose.
QRE 7.2.2	Not applicable as TikTok does not allow political advertising.
Measure 7.3	
QRE 7.3.1	Not applicable as TikTok does not allow political advertising.
QRE 7.3.2	Not applicable as TikTok does not allow political advertising.
Measure 7.4	
QRE 7.4.1	Not applicable as TikTok does not allow political advertising.



## Commitment 8

Relevant Signatories commit to provide transparency information to users about the political or issue ads they see on their service.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	No
If yes, list these implementation measures here [short bullet points].	N/A
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	We prohibit political advertising and are continuing to focus on enforcement of this policy in light of Regulation EU 2024/900 on the Transparency and Targeting of Political Advertising coming into force and the majority of provisions applying from October 2025.
If yes, which further implementation measures do you plan to put in place in the next 6 months?	N/A
Measure 8.1	
Measure 8.2	
QRE 8.2.1 (for measures 8.1 & 8.2)	Not applicable as TikTok does not allow political advertising. More generally, TikTok provides transparency on TikTok through the "About this ad" feature, which allows our users to find information about how an ad was personalised using various targeting information selected by the advertiser.



## Commitment 9

Relevant Signatories commit to provide users with clear, comprehensible, comprehensive information about why they are seeing a political or issue ad.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	No
If yes, list these implementation measures here [short bullet points].	N/A
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	We prohibit political advertising and are continuing to focus on enforcement of this policy in light of Regulation EU 2024/900 on the Transparency and Targeting of Political Advertising coming into force and the majority of provisions applying from October 2025.
If yes, which further implementation measures do you plan to put in place in the next 6 months?	N/A
Measure 9.1	N/A
Measure 9.2	N/A
QRE 9.2.1 (for measures 9.1 & 9.2)	Not applicable as TikTok does not allow political advertising.



#### Commitment 10

Relevant Signatories commit to maintain repositories of political or issue advertising and ensure their currentness, completeness, usability and quality, such that they contain all political and issue advertising served, along with the necessary information to comply with their legal obligations and with transparency commitments under this Code.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	No
If yes, list these implementation measures here [short bullet points].	N/A
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	We prohibit political advertising and are continuing to focus on enforcement of this policy in light of Regulation EU 2024/900 on the Transparency and Targeting of Political Advertising coming into force and the majority of provisions applying from October 2025.
If yes, which further implementation measures do you plan to put in place in the next 6 months?	N/A
Measure 10.1	
Measure 10.2	



#### QRE 10.2.1 (for measures 10.1 & 10.2)

In compliance with our obligations pursuant to the Digital Services Act, TikTok maintains a publicly searchable Ad Library that features ads that TikTok has been paid to display to users, including those that are not currently active or have been paused by the advertisers. This includes information on the total number of recipients reached, with aggregate numbers broken down by Member State for the group or groups of recipients that the ad specifically targeted, including for political ads which have been removed. Each ad entry is available for the duration that it is shown on TikTok and for a year afterwards in compliance with the Digital Services Act.

Article 39(3) of the Digital Services Act requires that such libraries should not include the content of the ad, the identity of the person on whose behalf it was presented, or who paid for it where an ad has been <u>removed for incompatibility</u> with a platform's terms and conditions. As political ads are prohibited on TikTok, in order to comply with its legal obligations TikTok must remove these specific details of any political ads that have been removed from its platform (as such ad breach its terms and conditions). For this reason TikTok's ad library is required to display different information in respect of political ads in comparison to platforms that do allow them.

## III. Political Advertising

#### Commitment 11

Relevant Signatories commit to provide application programming interfaces (APIs) or other interfaces enabling users and researchers to perform customised searches within their ad repositories of political or issue advertising and to include a set of minimum functionalities as well as a set of minimum search criteria for the application of APIs or other interfaces."

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]

No



If yes, list these implementation measures here [short bullet points].	N/A
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	N/A
Measure 11.1	
Measure 11.2	
Measure 11.3	
Measure 11.4	
QRE 11.1.1 (for measures 11.1-11.4)	Whilst we do not consider this provision applies to TikTok, given that TikTok bans political ads, we have completed the roll out of the Commercial Content API in the EEA, to comply with our obligations under the Digital Services Act. This had previously been tested with a limited number of researchers. We have built commercial content related APIs that include ads, ad and advertiser metadata, and targeting information. These APIs will allow customised advertiser name or keyword based searches on ads and other commercial content data that is stored in the Commercial Content Library.
QRE 11.4.1	We value feedback on where we can improve any of our APIs. We will continue to receive feedback from researchers on their use of the Research and Commercial Content APIs and to make updates to better support independent research and enhance transparency about TikTok content.



## Commitment 12

Relevant Signatories commit to increase oversight of political and issue advertising and constructively assist, as appropriate, in the creation, implementation and improvement of political or issue advertising policies and practices.

and improvement of political or issue advertising policies and practices.		
In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	No	
If yes, list these implementation measures here [short bullet points].	TikTok did not subscribe to this commitment. Commitment 12 and associated measures are applicable to Civil Society only.	
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No	
If yes, which further implementation measures do you plan to put in place in the next 6 months?	TikTok did not subscribe to this commitment. Commitment 12 and associated measures are applicable to Civil Society only.	
Measure 12.1		
Measure 12.2		
Measure 12.3		
QRE 12.1.1 (for measures 12.1-12.3)	Commitment 12 and measures 12.1, 12.2 and 12.3 are applicable to Civil Society only.	



## Commitment 13

Relevant Signatories agree to engage in ongoing monitoring and research to understand and respond to risks related to Disinformation in political or issue advertising.

advertising.		
In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	No	
If yes, list these implementation measures here [short bullet points].	N/A	
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	We prohibit political advertising and are continuing to focus on enforcement of this policy in light of Regulation EU 2024/900 on the Transparency and Targeting of Political Advertising coming into force and the majority of provisions applying from October 2025.	
If yes, which further implementation measures do you plan to put in place in the next 6 months?	N/A	
Measure 13.1		
Measure 13.2	TikTok does not allow political advertising and this continues in blackout periods.	
Measure 13.3		



QRE 13.1.1 (for measures 13.1-13.3)	Whilst we do not allow political advertising, we remain engaged with discussions being held through the Task-force and other fora to ensure our policies and processes remain current and emerging threats are addressed in our policies and enforcement.





## IV. Integrity of Services Commitments 14 - 16



#### **IV. Integrity of Services**

#### Commitment 14

In order to limit impermissible manipulative behaviours and practices across their services, Relevant Signatories commit to put in place or further bolster policies to address both misinformation and disinformation across their services, and to agree on a cross-service understanding of manipulative behaviours, actors and practices not permitted on their services. Such behaviours and practices, which should periodically be reviewed in light with the latest evidence on the conducts and TTPs employed by malicious actors, such as the AMITT Disinformation Tactics, Techniques and Procedures Framework, include:

The following TTPs pertain to the creation of assets for the purpose of a disinformation campaign, and to ways to make these assets seem credible:

- 1. Creation of inauthentic accounts or botnets (which may include automated, partially automated, or non-automated accounts)
- 2. Use of fake / inauthentic reactions (e.g. likes, up votes, comments)
- 3. Use of fake followers or subscribers
- 4. Creation of inauthentic pages, groups, chat groups, fora, or domains
- 5. Account hijacking or impersonation

The following TTPs pertain to the dissemination of content created in the context of a disinformation campaign, which may or may not include some forms of targeting or attempting to silence opposing views. Relevant TTPs include:

- 6. Deliberately targeting vulnerable recipients (e.g. via personalised advertising, location spoofing or obfuscation)
- 7. Deploy deceptive manipulated media (e.g. "deep fakes", "cheap fakes"...)
- 8. Use "hack and leak" operation (which may or may not include doctored content)
- 9. Inauthentic coordination of content creation or amplification, including attempts to deceive/manipulate platforms algorithms (e.g. keyword stuffing or inauthentic posting/reposting designed to mislead people about popularity of content, including by influencers)
- 10. Use of deceptive practices to deceive/manipulate platform algorithms, such as to create, amplify or hijack hashtags, data voids, filter bubbles, or echo chambers
- 11. Non-transparent compensated messages or promotions by influencers
- 12. Coordinated mass reporting of non-violative opposing content or accounts

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]

Yes



If yes, list these implementation measures here [short bullet points].

- Updating our existing 'Synthetic and Manipulated Media' policy and renamed it the
   'Edited Media and Al-Generated Content (AIGC)' policy, which became effective in
   May 2024. Through these updates, we have adopted more commonly used and easily
   understood language when referring to AIGC, and clarified our existing prohibitions on
   AIGC showing fake authoritative sources or crisis events, or falsely showing public
   figures in certain contexts including being bullied, making an endorsement, or being
   endorsed.
- Building on our new <u>Al-generated label</u> for creators to disclose content that is completely Al-generated or significantly edited by Al, we have expanded our efforts in the AIGC space by:
  - Implementing the Coalition for Content Provenance and Authenticity (C2PA)
     <u>Content Credentials.</u> which enables our systems to instantly recognize and automatically label AIGC.
  - Supporting the coalition's working groups as a C2PA General Member.
  - Joining the Content Authenticity Initiative (CAI) to drive wider adoption of the technical standard.
  - Publishing a new Transparency Center article <u>Supporting responsible</u>, <u>transparent Al-generated content</u>.
- In accordance with our commitments as a launch partner for PAI Responsible Practices on Synthetic Media, we worked on a case study outlining how the Practices informed our policy making on synthetic media. In addition, representatives from our team attended a series of workshops on content provenance and watermarking organised by PAI, which was attended by representatives of peer platforms and media outlets.
- Joined industry partners as a party to the "Tech Accord to Combat Deceptive Use of Al
  in 2024 Elections' which is a joint commitment to combat the deceptive use of Al in
  elections.
- Continued to participate in the working groups on integrity of services and Generative Al.
- We have continued to enhance our ability to detect covert influence operations. To provide more regular and detailed updates about the covert influence operations we disrupt, we have introduced a new dedicated Transparency Report on covert influence operations, which is available in TikTok's transparency centre. In this report, we have also added new information about operations that we have previously removed and that have attempted to return to our platform with new accounts
- Creating a new <u>Harmful Misinformation Guide</u> in our Safety Center.





Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight disinformation and will report on any further development in the next COPD report.
Measure 14.1	



#### **QRE 14.1.1**

As well as our I&A policies in our CGs which safeguard against harmful misinformation (see QRE 18.2.1), our I&A policies also expressly prohibit deceptive behaviours. Our policies on deceptive behaviours relate to the TTPs as follows:

TTPs which pertain to the creation of assets for the purpose of a disinformation campaign, and to ways to make these assets seem credible:

Creation of inauthentic accounts or botnets (which may include automated, partially automated, or non-automated accounts)

Our I&A policies which address **Spam and Deceptive Account Behaviours** expressly prohibit account behaviours that may spam or mislead our community. You can set up multiple accounts on TikTok to create different channels for authentic creative expression, but not for deceptive purposes.

We do not allow spam including:

- accounts that are operated in bulk, through unauthorised automation, or in order to distribute high-volume commercial content; and
- operating networks of accounts that represent similar entities or post similar content to lead others to specific locations (on or off-platform), such as other accounts, websites, and businesses.

We also do not allow impersonation including:

- Accounts that pose as another real person or entity, such as using someone's name, biographical details, content, or image without disclosing it
- Presenting as a person or entity that does not exist (a fake persona) with a demonstrated intent to mislead others on the platform

If we determine someone has engaged in any of these deceptive account behaviours, we will ban the account, and may ban any new accounts that are created.

Use of fake / inauthentic reactions (e.g. likes, up votes, comments) and use of fake followers or subscribers

Our I&A policies which address **fake engagement** do not allow the trade of services that attempt to artificially increase engagement or deceive TikTok's recommendation system. We do not allow our users to:



- facilitate the trade of services that artificially increase engagement, such as selling followers or likes; or
- provide instructions on how to artificially increase engagement on TikTok.

If we become aware of accounts or content with inauthentically inflated metrics, we will remove the associated fake followers or likes. Content that tricks or manipulates others as a way to increase engagement metrics, such as "like-for-like" promises and false incentives for engaging with content is ineligible for our For You feed.

#### Creation of inauthentic pages, groups, chat groups, fora, or domains

TikTok does not have pages, groups, chat groups, fora or domains. This TTP is not relevant to our platform.

#### Account hijacking or Impersonation

Again, our policies prohibit **impersonation** which refers to accounts that pose as another real person or entity or present as a person or entity that does not exist (a fake persona) with a demonstrated intent to mislead others on the platform. Our users are not allowed to use someone else's name, biographical details, or profile picture in a misleading manner. In order to protect freedom of expression, we do allow accounts that are clearly parody, commentary, or fan-based, such as where the account name indicates that it is a fan, commentary, or parody account and not affiliated with the subject of the account.

We also have a number of policies that address account hijacking. Our privacy and security policies under our CGs expressly prohibit users from providing access to their account credentials to others or enable others to conduct activities against our CGs. We do not allow access to any part of TikTok through unauthorised methods; attempts to obtain sensitive, confidential, commercial, or personal information; or any abuse of the security, integrity, or reliability of our platform. We also provide practical <u>guidance</u> to users if they have concerns that their account may have been hacked.

TTPs which pertain to the dissemination of content created in the context of a disinformation campaign, which may or may not include some forms of targeting or attempting to silence opposing views:



Deliberately targeting vulnerable recipients (e.g. via personalised advertising, location spoofing or obfuscation), inauthentic coordination of content creation or amplification, including attempts to deceive/manipulate platforms algorithms (e.g. keyword stuffing or inauthentic posting/reposting designed to mislead people about popularity of content, including by influencers), use of deceptive practices to deceive/manipulate platform algorithms, such as to create, amplify or hijack hashtags, data voids, filter bubbles, or echo chambers and coordinated mass reporting of non-violative opposing content or accounts.

We fight against **CIOs** as our policies prohibit attempts to sway public opinion while also misleading our systems or users about the identity, origin, approximate location, popularity or overall purpose.

Where our teams have a high degree of confidence that an account is engaged in inauthentic coordination of content creation or amplification, uses deceptive practices to deceive/manipulate platform algorithm or coordinated mass reporting of non-violative opposing content/accounts accounts and is engaged in or is connected to networks we took down in the past as part of a CIO, it is removed from our Platform in accordance with our CIO policy.

When we investigate and remove these operations, we focus on behaviour and assessing linkages between accounts and techniques to determine if actors are engaging in a coordinated effort to mislead TikTok's systems or our community. In each case, we believe that the people behind these activities coordinate with one another to misrepresent who they are and what they are doing. We know that CIOs will continue to evolve in response to our detection and networks may attempt to reestablish a presence on our platform. We continue to iteratively research and evaluate complex deceptive behaviours on our platform and develop appropriate product and policy solutions as appropriate in the long term. We publish all of the CIO networks we identify and remove voluntarily in a dedicated report within our transparency centre here.

Use "hack and leak" operation (which may or may not include doctored content)

We have a number of policies that address hack and leak related threats (some examples are below):



- Our CIO policy addresses use of leaked documents to sway public opinion as part of a wider operation
- Our Edited Media and Al-Generated Content (AIGC)captures materials that has been digitally altered without an appropriate disclosure
- Our harmful misinformation policies combats conspiracy theories related to unfolding events and dangerous misinformation
- Our Trade of Regulated Goods and Services policy prohibits trading of hacked goods.

In addition to the above, we recently launched a new hack and leak policy, which aims to further reduce the harms inflicted by the unauthorised disclosure of hacked materials on the individuals, communities and organisations that may be implicated or exposed by such disclosures.

#### Deceptive manipulated media (e.g. "deep fakes", "cheap fakes"...)

We have updated our existing 'Synthetic and Manipulated Media' policy and renamed it the 'Edited Media and Al-Generated Content (AIGC)' policy, which became effective in May 2024. Through these updates, we have adopted more commonly used and easily understood language when referring to AIGC, and clarified our existing prohibitions on AIGC showing fake authoritative sources or crisis events, or falsely showing public figures in certain contexts including being bullied, making an endorsement, or being endorsed.

For the purposes of our updated policy, AIGC refers to content created or modified by artificial intelligence (AI) technology or machine-learning processes, which may include images of real people, and may show highly realistic-appearing scenes, or use a particular artistic style, such as a painting, cartoons, or anime. 'Significantly edited content' is content that shows people doing or saying something they did not do or say, or altering their appearance in a way that makes them difficult to recognise or identify.

In accordance with our policy, we prohibit AIGC that features:

- Realistic-appearing people under the age of 18
- The likeness of adult private figures, if we become aware it was used without their permission
- Misleading AIGC or edited media that falsely shows:



- Content made to seem as if it comes from an authoritative source, such as a reputable news organisation
- o A crisis event, such as a conflict or natural disaster
- o A public figure who is:
  - being degraded or harassed, or engaging in criminal or antisocial behaviour
  - taking a position on a political issue, commercial product, or a matter of public importance (such as an elections)
  - being politically endorsed or condemned by an individual or group

As AI evolves, we continue to invest in combating harmful AIGC by evolving our proactive detection models, consulting with experts, and partnering with peers on shared solutions.

#### Non-transparent compensated messages or promotions by influencers

Our <u>Terms of Service</u> and <u>Branded Content Policy</u> require users posting about a **brand or product in return for any payment or other incentive** to disclose their content by enabling the branded content toggle which we make available for users. We also provide functionality to enable users to report suspected undisclosed branded content which reminds the user who posted the suspected undisclosed branded content of our requirements and prompts them to turn the branded content toggle on if required. We made this requirement even clearer to users in our **Commercial Disclosures and Paid Promotion** policy in our March 2023 CG refresh, by increasing the information around our policing of this policy and providing specific examples.



## QRE 14.1.2

At TikTok, we place considerable emphasis on proactive content moderation and use a combination of technology and safety professionals to detect and remove harmful misinformation (see QRE 18.1.1) and deceptive behaviours on our Platform, *before* it is reported to us by users or third parties.

For instance, we take proactive measures to prevent inauthentic or spam accounts from being created. Thus, we have created and use detection models and rule engines that:

- prevent inauthentic accounts from being created based on malicious patterns; and
- remove registered accounts based on certain signals (ie, uncommon behaviour on the platform).

We also manually monitor user reports of inauthentic accounts in order to detect larger clusters or similar inauthentic behaviours.

However, given the complex nature of the TTPs, human moderation is critical to success in this area and TikTok's moderation teams therefore play a key role assessing and addressing identified violations. We provide our moderation teams with detailed guidance on how to apply the I&A policies in our CGs, including providing case banks of harmful misinformation claims to support their moderation work, and allow them to route new or evolving content to our fact-checking partners for assessment.

In addition, where content reaches certain levels of popularity in terms of the number of video views, it will be flagged for further review. Such review is undertaken given the extent of the content's dissemination and the increase in potential harm if the content is found to be in breach of our CGs including our I&A policies.

We have also set up specifically-trained teams that are focused on investigating and detecting CIO on our Platform. When we investigate and remove these operations, we focus on behaviour and assessing linkages between accounts and techniques to determine if actors are engaging in a coordinated effort to mislead TikTok's systems or our community. In each case, we believe that the people behind these activities coordinate with one another to misrepresent who they are and what they are doing.

#### Measure 14.2



#### **QRE 14.2.1**

The implementation of our policies is ensured by different means, including specifically-designed tools (such as toggles to disclose branded content - see QRE 14.1.1) or human investigations to detect deceptive behaviours (for CIO activities - see QRE 14.1.2).

The implementation of these policies is also ensured through enforcement measures applied in all Member States.

CIO investigations are resource intensive and require in-depth analysis to ensure high confidence in proposed actions. Where our teams have the necessary high degree of confidence that an account amounts to an **impersonation**, or is engaged in or is connected to networks we took down in the past as part of a CIO, it is removed from our Platform.

We are building and testing data infrastructure that can provide information required at a high level of fidelity.

Similarly, where our teams have a high degree of confidence that a specific content violates one of our TTPs-related policies (See QRE 14.1.1), such content is removed from TikTok.

Lastly, we may reduce the discoverability of some content, including by making videos ineligible for recommendation in the For You feed section of our platform. This is, for example, the case for content that tricks or manipulates users in order to inauthentically increase followers, likes, or views.



## SLI 14.2.1 - SLI 14.2.4 TTP OR ACTION1 TTP No. 1: Creation of inauthentic accounts or botnets (which may include automated, partially automated, or non-automated accounts) Methodology of data measurement We have based the number of: (i) fake accounts removed; and (ii) followers of the fake accounts (identified at the time of removal of the fake account), on the country the fake account was last active in. We have updated our methodology to report the ratio of monthly average of fake accounts over of monthly active users, based on the latest publication of monthly active users, in order to better reflect TTPs related content in relation to overall content on the service. **SLI 14.2.1** SLI 14.2.2 **SLI 14.2.3** SLI 14.2.4 TTPs related Nr of actions taken by Interaction/ engagement before action content in type relation to overall content on the service



List actions per member states (see example table above)	Number of fake accounts removed	Number of followers of fake accounts identified at the time of removal		Ratio of monthly average of Fake accounts over monthly active users	
Member States					
Austria	219,852	370176			
Belgium	552,800	510,618			
Bulgaria	278,481	64,902			
Croatia	108,797	556,404			
Cyprus	288,900	136,415			
Czech Republi c	1,641,721	186,222			
Denmark	119,933	79,984			
Estonia	73,500	158,025			
Finland	132,853	141,705			
France	11,806,409	11,119,872			
Germany	13,105,844	8,456,293			



Greece	92,874	181,444			
Hungary	168,391	171,164			
Ireland	409,089	284,361			
Italy	5,294,570	1,911,744			
Latvia	60,251	43,977			
Lithuania	103,818	273,285			
Luxembourg	32,621	51,932			
Malta	40,314	1,314			
Netherlands	690,743	485,871			
Poland	1,502,923	236,962			
Portugal	459,766	213,810			
Romania	261,749	4,006,803			
Slovakia	75,660	55,519			
Slovenia	44,784	44,099			
Spain	4,185,885	1,428,005			
Sweden	174,798	202,323			
Iceland	42,100	40,358			
Liechtenstein	610	165			
Norway	306,586	236,467			



Total EU	41,927,326	31,373,229		0.047	
Total EEA	42,276,622	31,650,219			

TTP OR ACTION 2	TTP no. 2: Use of fake / inauthentic reactions (e.g. likes, up votes, comments)  Methodology of data measurement:  We based the number of fake likes that we removed on the country of registration of the user. We also based the number of fake likes prevented on the country of registration of the user.					
	SLI 14.2.4					
	Nr of actions taken by type	Interaction/ engagement before action				
List actions per member states (see example table above)		Number of fake likes prevented				
Austria	242,505	60,664,030				
Belgium	303,774	93,103,042				



		•	
Bulgaria	50,717	63,316,256	
Croatia	18,188	27,875,338	
Cyprus	69,822	10,207,634	
Czech Rep	65,264	53,378,137	
Denmark	85,811	21,929,011	
Estonia	21,016	21,177,388	
Finland	38,718	42,117,056	
France	1,248,502	760,851,271	
Germany	2,406,770	6,504,594,882	
Greece	91,824	71,653,337	
Hungary	52,337	45,188,210	



Ireland	69,132	54,124,972	
Italy	1,204,640	391,649,620	
Latvia	31,675	18,158,891	
Lithuania	63,516	18,028,451	
Luxembour g	38,446	6,899,029	
Malta	23,588	2,788,524	
Netherlands	407,076	499,588,856	
Poland	299,375	214,033,877	
Portugal	74,177	64,390,320	
Romania	270,655	115,675,193	
Slovakia	19,814	23,982,947	



Total EEA	8,241,036	9,731,416,675	
Total EU	8,036,544	9,655,221,876	
Norway	186,032	69,023,707	
Liechtenstei n	275	95,778	
Iceland	18,185	7,075,314	
Sweden	302,560	74,800,209	
Spain	525,378	382,425,290	
Slovenia	11,264	12,620,105	

# TTP OR ACTION 3

#### TTP No. 3: Use of fake followers or subscribers

#### Methodology of data measurement:

We based the number of fake followers that we removed on the country of registration of the user. We also based the number of fake followers prevented on the country of registration of the user.



	SLI 14.2.1	SLI 14.2.2	SLI 14.2.3	SLI 14.2.4
	Nr of actions taken by type	Interaction/ engagement before action		
List actions per member states (see example table above)	Number of fake followers removed	Number of fake follows prevented		
Member States				
Austria	525,800	16,547,046		
Belgium	566,959	25,528,915		
Bulgaria	132,462	8,820,853		
Croatia	38,971	4,994,758		
Cyprus	83,778	5,066,024		
Czech Repub lic	126,808	16,250,328		



Denmark	191,945	28,098,854	
Estonia	21,774	3,495,448	
Finland	65,151	10,510,190	
France	3,573,580	436,627,274	
Germany	6,748,336	251,514,408	
Greece	233,152	13,622,711	
Hungary	197,509	9,347,289	
Ireland	125,524	11,377,127	
Italy	1,293,172	284,517,343	
Latvia	33,838	4,423,898	
Lithuania	57,238	6,845,974	
Luxembourg	155,227	3,696,004	



Malta	11,530	505,550	
Netherlands	858,448	31,767,217	
Poland	367,590	26,211,680	
Portugal	134,977	58,649,607	
Romania	513,990	27,783,780	
Slovakia	48,610	4,847,157	
Slovenia	9,786	1,732,488	
Spain	936,640	213,734,494	
Sweden	502,942	43,082,288	
Iceland	44,366	14,409,563	
Liechtenstein	0	543,803	
Norway	312,421	51,386,389	



Total EU	17,555,737	1,549,598,705	
Total EEA	17,912,524	1,615,938,460	

# TTP OR ACTION 4

#### TTP No. 4: Creation of inauthentic pages, groups, chat groups, fora, or domains

TikTok does not have pages, groups, chat groups, fora or domains. This TTP is not relevant to our platform.

## TTP OR ACTION 5

#### TTP No. 5: Account hijacking or impersonation

#### Methodology of data measurement:

The number of accounts removed under our impersonation policy is based on the approximate location of the users. We have updated our methodology to report the ratio of monthly average impersonation accounts banned over of monthly active users, based on the latest publication of monthly active users, in order to better reflect TTPs related content in relation to overall content on the service.

	SLI 14.2.1	SLI 14.2.2	SLI 14.2.3	SLI 14.2.4		
	Nr of actions taken by type			TTPs related content in relation to overall content on the service		
Member States	Number of account banned under impersonation policy			Impersonation accounts over monthly active users		



List actions per member states (see example table above)			
Austria	218		
Belgium	343		
Bulgaria	95		
Croatia	21		
Cyprus	40		
Czech Rep ublic	83		
Denmark	92		
Estonia	18		
Finland	108		
France	3,662		
Germany	2,267		
Greece	153		
Hungary	141		
Ireland	174		
Italy	691		



Latvia	26		
Lithuania	30		
Luxembour g	27		
Malta	2		
Netherlands	521		
Poland	930		
Portugal	109		
Romania	1,113		
Slovakia	48		
Slovenia	20		
Spain	1,006		
Sweden	268		
Iceland	14		
Liechtenstei n	2		
Norway	210		
Total EU	12,206		0.000014
Total EEA	12,432		



## TTP OR ACTION 6

TTP No. 6. Deliberately targeting vulnerable recipients (e.g. via personalised advertising, location spoofing or obfuscation)

#### Methodology of data measurement:

The number of new CIO network discoveries found to be targeting EU markets relates to our public disclosures for the period January 1st 2024 to June 30th 2024. We have categorised disrupted CIO networks by the country we assess that the network originated in. We have included any network which we assess to have targeted one or more European markets. We publish all of the CIO networks we identify and remove within our transparency reports <a href="https://example.com/here/beta/fig/46/2024/">https://example.com/here/beta/fig/46/2024/</a>.

CIO networks identified and removed are detailed below, including the assessed geographic location of network operation and the assessed target audience of the network, which we assess via technical and behavioural evidence from proprietary and open sources. The number of followers of CIO networks has been based on the number of accounts that followed any account within a network as of the date of that network's removal.

	SLI 14.2.	1	SLI 14.2.2	SLI 14.2.2		SLI 14.2.3	SLI 14.2.4
Member States (Network origin)		Nr of actions taken by type	Interaction/ engagement before action	Views/ impressions after action	Interaction/ engagement after action	Trends on targeted audiences	

January - June 2024



Germany	46 accoun ts in the networ k	46 removed accounts	12,837  accounts followed an account within the network as at the date of removal	0	0	We assess that this network operated from Germany and targeted a German audience. The individuals behind this network created opaque entities in order to artificially amplify narratives in German language favorable to the AfD (Alternative für Deutschland) political party, promoting Christianity and criticizing Islam, attempting to manipulate Germany's domestic politics. We assess
			orremoval			political party, promoting Christianity and criticizing Islam, attempting to manipulate Germany's
						former Muslims in order to insert divisive rhetoric about Islam into spaces on TikTok where Islam-related hashtags were being searched.



Sweden	35 accoun ts in the networ k	35 removed accounts	accounts followed an account within the network as at the date of removal	0	0	We assess that this network operated from Sweden and targeted a Swedish audience. The individuals behind this network created inauthentic accounts in order to artificially amplify narratives in Swedish favorable to the Sweden Democrats political party and its agenda. The network was found to create misleading personas, including accounts impersonating government departments.
--------	---	---------------------	---	---	---	---

TTP OR ACTION 7	TTP No. 7: Deploy deceptive manipulated media (e.g. "deep fakes", "cheap fakes")  We have based the following numbers on the country in which the video was posted: videos removed because of violations Synthetic and Manipulated Media policyThe number of views of videos removed because of violation of each of these policies is based on the approximate location of the user.				
Member States	Number of videos removed because of violation of synthetic media policy	Number of views of videos removed because of violation of synthetic media policy	Number of unique videos labelled with AIGC tag of "Creator labeled as AI-generated"	Number of unique videos labelled with AIGC tag of "AI-generated"	
Austria	111	13,595	75,157	41,506	



Belgium	182	44,894	136,844	82,145
Bulgaria	30	122,302	45,064	71,357
Croatia	22	63,556	18,268	20,248
Cyprus	15	925	32,976	2,534
Czech Rep ublic	41	5,950	42,630	112,919
Denmark	96	4,716,805	31,822	31,221
Estonia	11	856	14,268	13,543
Finland	30	1,726	39,854	53,795
France	954	12,016,565	1,031,014	475,265
Germany	1,553	6,907,249	925,044	541,169



Greece	61	7,089	134,590	8,027
Hungary	11	412,593	48,779	97,514
Ireland	111	252,728	24,501	31,507
Italy	460	191,672	545,417	406,189
Latvia	32	1,159	28,703	19,636
Lithuania	33	6,527	25,896	21,305
Luxembour g	23	556	29,892	4,353
Malta	14	2,285	8,844	2,663
Netherlands	178	2,834,922	141,406	26,503
Poland	121	3,838,027	166,465	429,221



Total EU	4,682	61,281,948	4,519,165	3,267,961
Norway	144	652,489	41,254	40,124
Liechtenstei n	5	1,542	184	61
Iceland	42	1,095	2,369	4,462
Sweden	68	1,823,170	102,276	93,994
Spain	270	2,886,615	522,037	507,059
Slovenia	14	290	7,216	6,685
Slovakia	38	23,403,840	17,972	37,021
Romania	106	1,713,376	230,754	46,014
Portugal	97	12,676	91,476	84,568



Total EEA	4,873	61,937,074	4,562,972	3,312,608

#### TTP OR ACTION 8 Member

TTP No. 8: Use "hack and leak" operation (which may or may not include doctored content)

We have recently launched our new hack and leak policy, and so we are therefore unable to share meaningful metrics in this report.

## TTP OR ACTION 9

TTP No. 9: Inauthentic coordination of content creation or amplification, including attempts to deceive/manipulate platforms algorithms (e.g. keyword stuffing or inauthentic posting/reposting designed to mislead people about popularity of content, including by influencers)

We have provided data on the CIO networks that we have disrupted in the reporting period under TTP No. 6.

#### TTP OR ACTION 10

TTP No. 10: Use of deceptive practices to deceive/manipulate platform algorithms, such as to create, amplify or hijack hashtags, data voids, filter bubbles, or echo chambers

We have provided data on the CIO networks that we have disrupted in the reporting period under TTP No. 6.



TTP OR ACTION 11	TTP No. 11. Non-transparent compensated messages or promotions by influencers  Methodology of data measurement:  We are unable to provide this metric due to insufficient data available for the reporting period.			
	SLI 14.2.1	SLI 14.2.2	SLI 14.2.3	SLI 14.2.4
	Nr of actions taken by type			
Member States				
List actions per member states (see example table above)	Number of times the branded content toggle has been used to disclose the existence of a commercial relationship			



TTP OR ACTION 12	TTP No. 12: Coordinated mass reporting of non-violative opposing content or accounts  We have provided data on the CIO networks that we have disrupted in the reporting period under TTP No. 6.			
	SLI 14.2.1	SLI 14.2.2	SLI 14.2.3	SLI 14.2.4
Member States				
List actions per member states (see example table above)				

Measure 14.3	
QRE 14.3.1	We engaged with the Integrity of Services working group to set up the first list of TTPs.



### **IV. Integrity of Services**

#### Commitment 15

Relevant Signatories that develop or operate AI systems and that disseminate AI-generated and manipulated content through their services (e.g. deep fakes) commit to take into consideration the transparency obligations and the list of manipulative practices prohibited under the proposal for Artificial Intelligence Act.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]

Yes





If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight disinformation and will report on any further development in the next COPD report.
Measure 15.1	



#### **QRE 15.1.1**

We have updated our existing 'Synthetic and Manipulated Media' policy and renamed it the 'Edited Media and Al-Generated Content (AIGC)' policy, which became effective in May 2024. Through these updates, we have adopted more commonly used and easily understood language when referring to AIGC, and clarified our existing prohibitions on AIGC showing fake authoritative sources or crisis events, or falsely showing public figures in certain contexts including being bullied, making an endorsement, or being endorsed. As AI evolves, we continue to invest in combating harmful AIGC by evolving our proactive detection models, consulting with experts, and partnering with peers on shared solutions.

While we welcome the creativity that new AI may unlock, in line with our updated policy, users must proactively disclose when their content is AI-generated or manipulated but shows realistic scenes (i.e. fake people, places or events that look like they are real). We launched an AI toggle in September 2023, which allows users to self-disclose AI-generated content when posting. When this has been turned on, a tag "Creator labelled as AI-generated" is displayed to users. Alternatively, this can be done through the use of a sticker or caption, such as 'synthetic', 'fake', 'not real', or 'altered'.

#### We do not allow:

- Realistic-appearing people under the age of 18
- The likeness of adult private figures, if we become aware it was used without their permission
- Misleading AIGC or edited media that falsely shows:
  - Content made to seem as if it comes from an authoritative source, such as a reputable news organisation
  - o A crisis event, such as a conflict or natural disaster
  - A public figure who is:
    - being degraded or harassed, or engaging in criminal or antisocial behaviour
    - taking a position on a political issue, commercial product, or a matter of public importance (such as an elections)
    - being politically endorsed or condemned by an individual or group

We have developed media literacy campaigns with guidance from expert organisations like Mediawise and Witness, that teach our community how to spot and label Al-generated content.



Measure 15.2	
QRE 15.2.1	<ul> <li>We have a number of measures to ensure the AI systems we develop uphold the principles of fairness and comply with applicable laws. To that end:</li> <li>We have in place internal guidelines and training to help ensure that the training and deployment of our AI systems comply with applicable data protection laws, as well as principles of fairness.</li> <li>We have instituted a compliance review process for new AI systems that meet certain thresholds, and are working to prioritise review of previously developed algorithms.</li> <li>We are proud to be a launch partner of the Partnership on AI's Responsible Practices for Synthetic Media. In accordance with our commitments as a launch partner, we worked on a case study outlining how the Practices informed our policy making on synthetic media. In addition, representatives from our team attended a series of workshops on content provenance and watermarking organised by PAI, which was attended by representatives of peer platforms and media outlets.</li> </ul>

### **IV. Integrity of Services**

#### Commitment 16

Relevant Signatories commit to operate channels of exchange between their relevant teams in order to proactively share information about cross-platform influence operations, foreign interference in information space and relevant incidents that emerge on their respective services, with the aim of preventing dissemination and resurgence on other services, in full compliance with privacy legislation and with due consideration for security and human rights risks.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]

Yes





If yes, list these implementation measures here [short bullet points].	Actively engaged with the Crisis Response working group, sharing insights and learnings about relevant areas including CIOs.
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight disinformation and will report on any further development in the next COPD report.
Measure 16.1	

QRE 16.1.1	Central to our strategy for identifying and removing CIO on our platforms is working we stakeholders from civil society to user reports. This approach facilitates us - and other disrupting the network's operations in their early stages. In addition to continuously er our in-house capabilities, we proactively engage in comprehensive reviews of our peepublicly disclosed findings and swiftly implement necessary actions in alignment with policies.			
	To provide more regular and detailed updates about the CIO we disrupt, we have introduced a new dedicated Transparency Report on covert influence operations, which is available in TikTok's transparency centre. In this report, we have also added new information about operations that we have previously removed and that have attempted to return to our platform with new accounts. The insights and metrics in this report aim to inform industry peers and the research community.			
	We share relevant insights and metrics within our quarterly transparency reports, which aim to inform industry peers and the research community.			
	We continue to engage in the sub groups set up for insights sharing between signatories and the Commission.			
	As we have detailed in other chapters to this report, we have robust monetisation integrity policies in place and have established joint operating procedures between specialist CIO investigations teams and monetisation integrity teams to work on joint investigations of CIOs involving monetised products.			
SLI 16.1.1	N/A			
Numbers of actions as a result of information sharing				
Data				
Measure 16.2				
QRE 16.2.1	We publish all of the CIO networks we identify and remove within our transparency reports <a href="here">here</a> . As new deceptive behaviours emerge, we'll continue to evolve our response, strengthen enforcement capabilities, and publish our findings.			





V. Empowering Users Commitments 17 - 25



### V. Empowering Users

#### Commitment 17

In light of the European Commission's initiatives in the area of media literacy, including the new Digital Education Action Plan, Relevant Signatories commit to continue and strengthen their efforts in the area of media literacy and critical thinking, also with the aim to include vulnerable groups.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No] Yes



If yes, list these implementation measures here [short bullet points].

- Rolled out two new media literacy campaigns in Denmark and the Netherlands in collaboration with our trusted fact-checking partners on general media literacy and critical thinking skills.
- Delivered localised media literacy campaigns in all 27 EU Member States ahead of the European Parliamentary Elections in June 2024.
- Promoted election integrity in the run up to country-level EU elections in Croatia, Finland, France, and Slovakia, the Irish Referendum, and the EU Parliamentary Elections by:
  - o Inviting regional experts to share their insights with our internal teams as part of our Election Speaker Series.
  - Launching additional in-app election campaigns, using our intervention tools to connect users to authoritative information sources, including bespoke media literacy videos.
- Expanding fact-checking coverage to all EU Member States in advance of the EU Parliamentary Elections. Continued our
  in-app interventions, including video tags, search interventions and in-app information centres, available in 23 official EU
  languages and Norwegian and Icelandic for EEA users, around the elections (including country-level EU elections and the
  EU Parliament Elections), the Israel-Hamas Conflict, Climate Change, Covid-19 and Covid-19 Vaccine, Holocaust Denial,
  Mpox and the War in Ukraine.
- Strengthened our approach to countering influence attempts by:
  - Making state-affiliated media accounts that attempt to reach communities outside their home country on current global events and affairs ineligible for recommendation, which means their content won't appear in the For You feed.
  - Prohibiting state-affiliated media accounts in all markets where our state-controlled media labels are available, including in the EU, from advertising outside of the country with which they are primarily affiliated.
  - Investing in our detection capabilities of state-affiliated media accounts.
  - Working with third party external experts to shape our state-affiliated media policy and assessment of state-controlled media labels.
- Building on our new <u>Al-generated label</u> for creators to disclose content that is completely Al-generated or significantly edited by Al, we have expanded our efforts in the AIGC space by:
  - Implementing the Coalition for Content Provenance and Authenticity (C2PA) <u>Content Credentials</u>, which enables our systems to instantly recognize and automatically label AIGC.
  - o Supporting the coalition's working groups as a C2PA General Member.
  - o Joining the Content Authenticity Initiative (CAI) to drive wider adoption of the technical standard.
  - o Publishing a new Transparency Center article <u>Supporting responsible, transparent Al-generated content.</u>
- Partnered with Amadeu Antonio Stiftung in Germany on the Demo:create project, an educational project that supports some young TikTok users to effectively deal with hate speech, disinformation and misinformation on platforms.
- Progressed our partnership with <u>Verified for Climate</u> by:

|--|--|

	<ul> <li>Bringing together a team of Verified Champions, including trusted messengers and experts from Brazil, the United Arab Emirates, and Spain, with select TikTok creators, to tackle climate misinformation while driving climate action within the TikTok community.</li> <li>Working with the United Nations and social impact agency Purpose to launch the first phase of the Verified for Climate program in Brazil, the UAE and Spain.</li> <li>Brought greater transparency about our systems and our integrity and authenticity efforts to our community by:         <ul> <li>Updating our Introduction to the TikTok recommendation system page and adding a new section on System Guides in the Transparency Center.</li> <li>Publishing new articles on Supporting responsible, transparent Al-generated content and Protecting the integrity of elections in our Transparency Center.</li> <li>Creating a new Harmful Misinformation Guide in our Safety Center.</li> <li>Introducing a new monthly transparency report focused on our work to eliminate covert influence operations.</li> </ul> </li> <li>Continued to invest in training and development for our human moderation teams.</li> <li>TikTok continues to co-chair the working group on Elections.</li> </ul>
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report
Measure 17.1	



#### **QRE 17.1.1**

In addition to systematically removing content from our platform that violates our I&A policies, we continue to dedicate significant resources to: expanding our in-app measures that show users additional context on certain content; redirecting them to authoritative information; and making these tools available in 23 EU official languages (plus, for EEA users, Norwegian & Icelandic).

We work with external experts to combat harmful misinformation. For example, we worked with the World Health Organisation (WHO) on medical information, and our global fact-checking partners, taking into account their feedback, as well as user feedback, to continually identify new topics and consider which tools may be best suited for raising awareness around that topic.

Since 2020, on topical issues such as elections (including country-level EU elections and the EU Parliament Elections), the Israel-Hamas Conflict, <u>Climate Change</u>, <u>Covid-19</u>, <u>Covid-19 Vaccine</u>, <u>Holocaust Denial</u>, MPox and the War in Ukraine, we have deployed a combination of in-app **user intervention tools**.

#### Video notice tags.

A video notice tag is an information bar at the bottom of a video which is automatically applied to a specific word or hashtag (or set of hashtags). The information bar is clickable and invites users to "Learn more about [the topic]". Users will be directed to an in-app guide, or reliable third party resource, as appropriate.



















Holocaust Facts Video Notice Tag

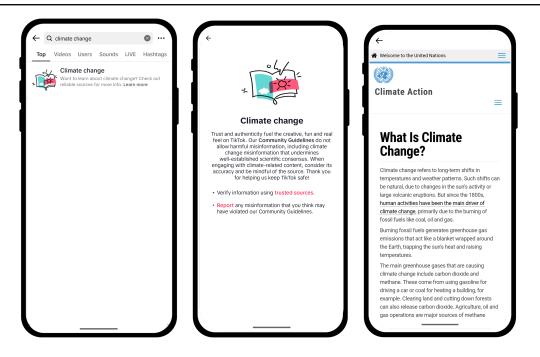
Mpox Video Notice Tag

Covid-19 Video Notice Tag

**Search intervention.** If users search for terms associated with a topic, they will be presented with a banner encouraging them to verify the facts and providing a link to a trusted source of information. Search interventions are not deployed for search terms that violate our Community Guidelines, which are actioned according to our policies.

• For example, in April 2023, we launched a <u>climate change search intervention tool</u> in 23 official EU languages (plus Norwegian and Icelandic for EEA users), which redirects users looking for climate change-related content to authoritative information and encourages them to report any potential misinformation they see. We worked with the UN to source the authoritative information and redirect users to their <u>resources</u>.





**Public service announcement (PSA).** If users search for a hashtag on the topic, they will be served with a public service announcement reminding them about our Community Guidelines and presenting them with links to a trusted source of information.

**Unverified content label.** In addition to the above mentioned tools, to encourage users to consider the reliability of content related to an emergency or unfolding event, which has been assessed by our fact-checking partners but cannot be verified as accurate i.e., 'unverified content', we apply warning labels and we prompt people to <u>reconsider sharing</u> such content. Details of these warning labels are included in our <u>Community Guidelines</u>.

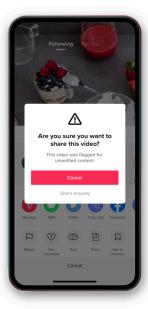
Where users continue to post despite the warning:

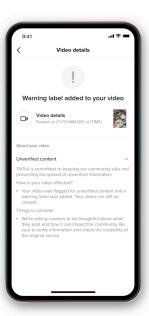
• To limit the spread of potentially misleading information, the video will become ineligible for recommendation in the For You feed.



• The video's creator is also notified that their video was flagged as unsubstantiated content and is provided additional information about why the warning label has been added to their content. Again, this is to raise the creator's awareness about the credibility of the content that they have shared.







State-controlled media label. Our state-affiliated media policy is to label accounts run by entities whose editorial output or decision-making process is subject to control or influence by a government. We apply a prominent label to all content and accounts from state-controlled media. The user is also shown a screen pop-up providing information about what the label means, inviting them to "learn more", and redirecting them to an <u>in-app page</u>. The measure brings transparency to our community, raises users' awareness, and encourages users to consider the reliability of the source. We continue to work with experts to inform our approach and explore how we can continue to expand its use.



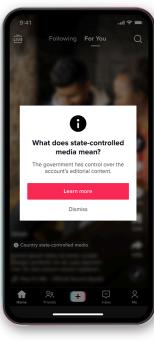


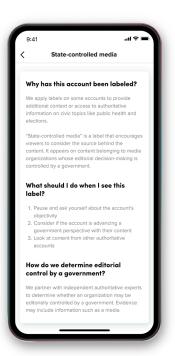












In the EU, Iceland and Liechtenstein, we have also taken steps to restrict access to content from:

- RT- Russia Today UK
- RT Russia Today Germany
- RT Russia Today France
- RT- Russia Today Spanish
- Sputnik
- Rossiya RTR / RTR Planeta
- Rossiya 24 / Russia 24
- TV Centre International
- NTV/NTV Mir
- Rossiya 1
- REN TV
- Pervyi Kanal / Channel 1

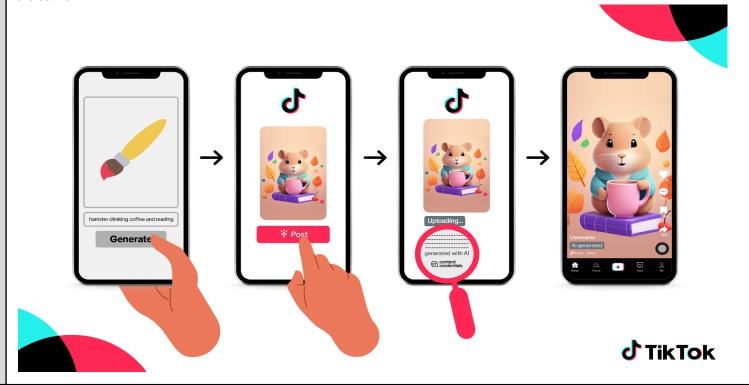


- RT Arabic
- Sputnik ArabicRT Balkan
- Oriental Review
- TsargradNew Eastern Outlook
- Katehon
- Voice of EuropeRIA Novosti

- Izvestija Rossiiskaja Gazeta



Al-generated content labels. As more creators take advantage of Artificial Intelligence (AI) to enhance their creativity, we want to support transparent and responsible content creation practices. Last year, TikTok launched a AI-generated label for creators to disclose content that is completely AI-generated or significantly edited by AI. The launch of this new tool to help creators label their AI-generated content was accompanied by a <u>creator education campaign</u>, a <u>Help Center page</u>, and a <u>Newsroom Post</u>. In May 2024, we started using the Coalition for Content Provenance and Authenticity (C2PA) <u>Content Credentials</u>, which enables our systems to instantly recognize and automatically label AIGC. In the interests of transparency, we also renamed TikTok AI effects to explicitly include "AI" in their name and corresponding effects label, and updated our guidelines for Effect House creators to do the same.



	_		
	Dedicated online and in-app information resources. The above mentioned tools provide links to users to accurate and up-to-date information from trusted sources. Depending on the topic, or the relevant EU country, users may be directed to an external authoritative source (e.g., a national government website or an independent national electoral commission), an in-app information centre (e.g., War in Ukraine), or a dedicated page in the TikTok Safety Center or Transparency Center.  We use our Safety Center to inform our community about our approach to safety, privacy, and security on our platform. Relevant to combating harmful misinformation, we have dedicated information on:  • Harmful Misinformation. Online challenges, Covid-19. Election integrity, and how to safely share content about tragic events on TikTok.  • Our safety partners page provides details of some of our work with global experts, non-governmental organisations, and industry associations to help build a safe platform for our community.  Users can learn more about our transparency efforts in our dedicated Transparency Center, available in a number of EU languages, which houses our transparency reports, including the new standalone Covert Influence Operations report and the reports we have published under this Code, as well as information on our commitments to maintaining platform integrity e.g., Protecting the integrity of elections, Combating misinformation, Countering influence operation and Supporting responsible, transparent Al-generated content.  We also use Newsroom posts to keep our community informed about our most recent updates and efforts across News, Product, Community, Safety and Product. Users can select their country, including EU, for preferred language where available, and regionally relevant posts. For example, upon publication of our thirdCode report in September 2023, we provided users with an overview of our efforts in Code of Practice on Disinformation: Investing in trusted information advance Al transparency and literacy.		
SLI 17.1.1 - actions	Methodology of data measurement:		
enforcing policies above	The number of impressions, clicks and click through rates of video notice tags, search interventions and public service announcements are based on the approximate location of the users that engaged with the tools. The number of impressions of the Safety Center pages is based on the IP location of the users.		
	Total count of the tool's impressions	Interactions/ engagement with the tool	Other relevant metrics
	Number of impressions of the State-Controlled Media label	Number of clicks of the State-Controlled Media label	Numbers of clicks through rate of the State-Controlled Media label



Member States			
Austria	1,645,007	3,332	0.20%
Belgium	1,927,568	3,750	0.19%
Bulgaria	6,982,915	11,367	0.16%
Croatia	773,492	2,430	0.31%
Cyprus	404,065	610	0.15%
Czech Republic	2,373,881	4,584	0.19%
Denmark	830,303	2,378	0.29%
Estonia	1,238,763	2,559	0.21%
Finland	1,292,029	5,930	0.46%
France	12,436,920	14,867	0.12%
Germany	22,226,816	32,100	0.14%
Greece	1,275,558	3,050	0.24%
Hungary	18,227,371	16,225	0.09%
Ireland	1,304,302	3,067	0.24%
Italy	5,626,259	8,290	0.15%
Latvia	3,541,528	5,394	0.15%
Lithuania	3,080,893	5,619	0.18%
Luxembourg	189,305	429	0.23%



Malta	236,286	368	0.16%
Netherlands	5,611,793	9,409	0.17%
Poland	132,693,648	57,980	0.04%
Portugal	628,500	2,056	0.33%
Romania	15,287,460	26,059	0.17%
Slovakia	1,098,439	2,016	0.18%
Slovenia	351,094	814	0.23%
Spain	15,652,469	31,313	0.20%
Sweden	2,961,533	6,042	0.20%
Iceland	151,998	371	0.24%
Liechtenstein	7,022	15	0.21%
Norway	2,205,435	5,949	0.27%
Total EU	259,898,197	262,038	0.10%
Total EEA	262,262,652	268,373	0.10%

	Number of impressions of Video Notice Tag covered by Intervention (Covid-19)	Number of clicks of Video Notice Tag covered by Intervention (Covid-19)	Click Through Rate of Video Notice Tag covered by Intervention (Covid-19)
Member States			
Austria	17,732,612	36,316	0.2%



		1	
Belgium	16,773,182	40,083	0.2%
Bulgaria	6,402,572	16,287	0.3%
Croatia	3,193,127	7,218	0.2%
Cyprus	3,859,749	8,001	0.2%
Czech Republic	19,479,655	49,992	0.3%
Denmark	3,847,538	10,936	0.3%
Estonia	3,082,015	7,289	0.2%
Finland	13,150,594	28,333	0.2%
France	14,972,025	35,065	0.2%
Germany	152,163,264	292,503	0.2%
Greece	30,663,609	56,947	0.2%
Hungary	20,733,464	49,679	0.2%
Ireland	32,706,696	64,734	0.2%
Italy	256,653,418	568,469	0.2%
Latvia	4,965,970	10,895	0.2%
Lithuania	5,350,668	12,830	0.2%
Luxembourg	1,314,259	3,636	0.3%
Malta	1,818,869	3,890	0.2%
Netherlands	44,307,807	103,498	0.2%



<del></del>		T
66,257,873	124,761	0.2%
22,664,882	62,947	0.3%
38,700,839	76,231	0.2%
4,162,698	10,331	0.2%
3,516,378	8,606	0.2%
179,418,433	351,578	0.2%
27,786,590	82,813	0.3%
1,538,053	6,017	0.4%
55,507	134	0.2%
5,870,911	16,653	0.3%
995,678,786	2,123,868	0.2%
1,003,143,257	2,146,672	0.2%
Number of impressions of Video Notice Tag covered by Intervention (Covid-19 Vaccine)	Number of clicks of Video Notice Tag covered by Intervention (Covid-19 Vaccine)	Click Through Rate of Video Notice Tag covered by Intervention (Covid-19 Vaccine)
5,833,126	11,368	0.19%
10,259,080	23,797	0.23%
3,176,465	6,936	0.22%
2 784 723	6 189	0.22%
	22,664,882  38,700,839  4,162,698  3,516,378  179,418,433  27,786,590  1,538,053  55,507  5,870,911  995,678,786  1,003,143,257  Number of impressions of Video Notice Tag covered by Intervention (Covid-19 Vaccine)  5,833,126  10,259,080  3,176,465	22,664,882       62,947         38,700,839       76,231         4,162,698       10,331         3,516,378       8,606         179,418,433       351,578         27,786,590       82,813         1,538,053       6,017         55,507       134         5,870,911       16,653         995,678,786       2,123,868         1,003,143,257       2,146,672         Number of impressions of Video Notice Tag covered by Intervention (Covid-19 Vaccine)       Number of clicks of Video Notice Tag covered by Intervention (Covid-19 Vaccine)         5,833,126       11,368         10,259,080       23,797



Cyprus	983,361	2,677	0.27%
Czech Republic	3,504,033	9,119	0.26%
Denmark	3,481,566	9,433	0.27%
Estonia	673,662	1,558	0.23%
Finland	3,704,093	8,707	0.24%
France	72,893,482	193,736	0.27%
Germany	56,702,999	103,203	0.18%
Greece	9,270,201	23,490	0.25%
Hungary	6,380,982	11,887	0.19%
Ireland	9,454,634	18,981	0.20%
Italy	144,272,225	325,022	0.23%
Latvia	1,347,056	2,541	0.19%
Lithuania	1,315,874	3,060	0.23%
Luxembourg	631,773	1,735	0.27%
Malta	555,112	1,026	0.18%
Netherlands	19,065,607	37,440	0.20%
Poland	21,647,453	44,919	0.21%
Portugal	7,941,047	21,184	0.27%
Romania	12,657,767	24,065	0.19%



Slovakia	2,163,028	5,880	0.27%
Slovenia	1,415,490	3,063	0.22%
Spain	101,978,544	245,966	0.24%
Sweden	11,086,578	28,572	0.26%
Iceland	393,110	1,153	0.29%
Liechtenstein	19,744	49	0.25%
Norway	3,880,416	10,082	0.26%
Total EU	515,179,961	1,175,554	0.23%
Total EEA	519,473,231	1,186,838	0.23%

	Number of impressions of Video Notice Tag covered by Intervention (Holocaust Misinformation/Denial)	Number of clicks of Video Notice Tag covered by Intervention (Holocaust Misinformation/Denial)	Click Through Rate of Video Notice Tag covered by Intervention (Holocaust Misinformation/Denial)
Member States			
Austria	4,972,962	14,106	0.3%
Belgium	3,640,990	9,650	0.3%
Bulgaria	656,715	2,767	0.4%
Croatia	798,277	3,334	0.4%
Cyprus	318,984	1,390	0.4%
Czech Republic	2,794,893	11,764	0.4%



Denmark	2,079,604	7,020	0.3%
Estonia	343,079	1,282	0.4%
Finland	2,567,136	9,484	0.4%
France	2,229,370	7,006	0.3%
Germany	41,457,275	104,663	0.3%
Greece	2,328,232	11,410	0.5%
Hungary	5,299,700	22,806	0.4%
Ireland	3,887,963	11,174	0.3%
Italy	2,072,974	7,104	0.3%
Latvia	344,002	1,320	0.4%
Lithuania	666,175	2,197	0.3%
Luxembourg	237,505	815	0.3%
Malta	226,411	634	0.3%
Netherlands	12,310,136	30,886	0.3%
Poland	34,591,785	92,730	0.3%
Portugal	2,395,084	8,853	0.4%
Romania	3,736,857	16,156	0.4%
Slovakia	758,539	2,737	0.4%
Slovenia	647,206	2,190	0.3%



7,671,835	35,282	0.5%
5,537,929	22,825	0.4%
263,387	933	0.4%
11,385	44	0.4%
3,279,637	13,035	0.4%
144,571,618	441,585	0.3%
148,126,027	455,597	0.3%
Number of impressions of Video Notice Tag covered by Intervention (MPox)	Number of clicks of Video Notice Tag covered by Intervention (MPox)	Click Through Rate of Video Notice Tag covered by Intervention (MPox)
61,179	98	0.2%
66,087	152	0.2%
55,073	122	0.2%
12,307	59	0.5%
28,473	52	0.2%
24,893	74	0.3%
30,810	121	0.4%
5,072	18	0.4%
60,388	232	0.4%
	5,537,929 263,387 11,385 3,279,637 144,571,618 148,126,027 Number of impressions of Video Notice Tag covered by Intervention (MPox) 61,179 66,087 55,073 12,307 28,473 24,893 30,810 5,072	5,537,929 22,825 263,387 933 11,385 44 3,279,637 13,035 144,571,618 441,585 148,126,027 455,597  Number of impressions of Video Notice Tag covered by Intervention (MPox)  61,179 98 66,087 152 55,073 122 12,307 59 28,473 52 24,893 74 30,810 121 5,072 18



France	274,012	698	0.3%
Germany	909,477	1,926	0.2%
Greece	70,484	266	0.4%
Hungary	37,670	136	0.4%
Ireland	42,646	155	0.4%
Italy	851,993	1,612	0.2%
Latvia	8,734	25	0.3%
Lithuania	9,456	29	0.3%
Luxembourg	5,405	10	0.2%
Malta	10,139	15	0.1%
Netherlands	135,942	351	0.3%
Poland	80,346	272	0.3%
Portugal	38,581	135	0.3%
Romania	113,647	254	0.2%
Slovakia	5,317	5	0.1%
Slovenia	5,868	7	0.1%
Spain	584,189	1,634	0.3%
Sweden	49,904	152	0.3%
Iceland	1,993	10	0.5%



Liechtenstein	66	0	0.0%
Norway	39,657	117	0.3%
Total EU	3,578,092	8,610	0.2%
Total EEA	3,619,808	8,737	0.2%
	Number of impressions of topic covered by video Intervention (Regional Elections)	Number of clicks by video Intervention (Election)	Number of CTR by video Intervention (Election)
Member States			
Austria	170,796,738	293,093	0.2%
Belgium	412,019,441	876,023	0.2%
Bulgaria	96,519,175	164,946	0.2%
Croatia	108,177,236	189,525	0.2%
Cyprus	38,802,802	73,732	0.2%
Czech Republic	93,531,544	168,309	0.2%
Denmark	86,547,328	220,775	0.3%
Estonia	26,061,091	53,709	0.2%
Finland	105,690,526	186,120	0.2%
France	2,374,699,199	5,045,840	0.2%
Germany	1,595,813,167	2,492,447	0.2%
Greece	284,109,560	520,495	0.2%



Total EU	9,813,084,413	18,172,397	0.2%
Norway	2,160,698	2,815	0.1%
Liechtenstein	40,440	90	0.2%
Iceland	256,584	325	0.1%
Sweden	62,105,698	152,494	0.2%
Spain	1,045,520,890	2,000,518	0.2%
Slovenia	46,162,018	82,875	0.2%
Slovakia	66,595,057	125,006	0.2%
Romania	391,304,336	565,643	0.1%
Portugal	271,822,287	550,664	0.2%
Poland	569,218,169	824,553	0.1%
Netherlands	395,352,996	782,111	0.2%
Malta	14,809,128	26,687	0.2%
Luxembourg	20,565,321	49,718	0.2%
Lithuania	63,396,037	115,573	0.2%
Latvia	37,436,285	68,586	0.2%
Italy	1,085,916,157	1,901,592	0.2%
Ireland	139,784,027	307,349	0.2%
Hungary	210,328,200	334,014	0.2%



Total EEA	9,815,542,135	18,175,627	0.2%
	Number of impressions of Search interventions (Covid-19)	Number of clicks of Search interventions (Covid-19)	Click Through Rate of Search interventions (Covid-19)
Member States			
Austria	80,335	68	0.1%
Belgium	84,785	55	0.1%
Bulgaria	30,741	29	0.1%
Croatia	22,420	26	0.1%
Cyprus	10,634	16	0.2%
Czech Republic	69,722	34	0.0%
Denmark	28,535	18	0.1%
Estonia	8,386	11	0.1%
Finland	33,698	28	0.1%
France	724,764	237	0.0%
Germany	767,644	512	0.1%
Greece	85,371	56	0.1%
Hungary	97,710	80	0.1%
Ireland	153,842	61	0.0%
Italy	1,077,291	496	0.0%



Latvia	12,694	9	0.1%
Lithuania	17,648	22	0.1%
Luxembourg	7,302	5	0.1%
Malta	6,523	3	0.0%
Netherlands	120,838	80	0.1%
Poland	505,048	136	0.0%
Portugal	388,797	151	0.0%
Romania	201,540	126	0.1%
Slovakia	37,617	20	0.1%
Slovenia	11,588	8	0.1%
Spain	853,432	522	0.1%
Sweden	102,636	57	0.1%
Iceland	4,786	3	0.1%
Liechtenstein	198	0	0.0%
Norway	49,917	31	0.1%
Total EU	5,541,541	2,866	0.1%
Total EEA	5,596,442	2,900	0.1%



	Number of impressions of Search interventions (Covid-19 Vaccine)	Number of clicks of Search interventions (Covid-19 Vaccine)	Click Through Rate of Search interventions (Covid-19 Vaccine)
Member States			
Austria	38	0	0.0%
Belgium	121	0	0.0%
Bulgaria	8	0	0.0%
Croatia	24	0	0.0%
Cyprus	0	0	0.0%
Czech Republic	36	0	0.0%
Denmark	18	0	0.0%
Estonia	4	0	0.0%
Finland	0	0	0.0%
France	602	1	0.2%
Germany	391	0	0.0%
Greece	12	0	0.0%
Hungary	7	0	0.0%
Ireland	278	0	0.0%
Italy	650	0	0.0%
Latvia	0	0	0.0%



Member States			
	Number of impressions of Search interventions (Holocaust Misinformation/Denial)	Number of clicks of Search interventions (Holocaust Misinformation/Denial)	Click Through Rate of Search interventions (Holocaust Misinformation/Denial)
Total EEA	6,757	2	0.0%
Total EU	6,437	2	0.0%
Norway	40	0	0.0%
Liechtenstein	280	0	0.0%
Iceland	0	0	0.0%
Sweden	56	0	0.0%
Spain	846	0	0.0%
Slovenia	0	0	0.0%
Slovakia	6	0	0.0%
Romania	72	0	0.0%
Portugal	2,908	0	0.0%
Poland	68	1	1.5%
Netherlands	246	0	0.0%
Malta	2	0	0.0%
Luxembourg	40	0	0.0%
Lithuania	4	0	0.0%



304,369	6,457	2%
348,812	5,401	2%
124,110	1,678	1%
174,702	2,551	1%
59,549	468	1%
203,870	2,673	1%
156,388	1,691	1%
39,449	564	1%
286,983	2,200	1%
1,654,401	32,951	2%
2,355,435	68,531	3%
1,210,580	4,439	0%
280,233	6,078	2%
294,619	2,197	1%
2,030,292	19,849	1%
46,159	782	2%
85,437	1,040	1%
19,997	395	2%
13,217	165	1%
	348,812 124,110 174,702 59,549 203,870 156,388 39,449 286,983 1,654,401 2,355,435 1,210,580 280,233 294,619 2,030,292 46,159 85,437 19,997	348,812       5,401         124,110       1,678         174,702       2,551         59,549       468         203,870       2,673         156,388       1,691         39,449       564         286,983       2,200         1,654,401       32,951         2,355,435       68,531         1,210,580       4,439         280,233       6,078         294,619       2,197         2,030,292       19,849         46,159       782         85,437       1,040         19,997       395



933,944	7,179	1%
1,463,351	1,098	0%
232,087	2,319	1%
377,938	5,423	1%
102,871	1,467	1%
62,677	1,139	2%
1,218,219	35,026	3%
434,173	4,692	1%
11,390	211	2%
617	10	2%
223,982	2,907	1%
14,513,862	218,453	2%
14,749,851	221,581	2%
Number of impressions of Search interventions (MPox)	Number of clicks of Search interventions (MPox)	Click Through Rate of Search interventions (MPox)
1,062	0	0%
1,608	0	0%
936	1	0%
564	2	0%
	1,463,351 232,087 377,938 102,871 62,677 1,218,219 434,173 11,390 617 223,982 14,513,862 14,749,851 Number of impressions of Search interventions (MPox)  1,062 1,608	1,463,351       1,098         232,087       2,319         377,938       5,423         102,871       1,467         62,677       1,139         1,218,219       35,026         434,173       4,692         11,390       211         617       10         223,982       2,907         14,513,862       218,453         14,749,851       221,581         Number of impressions of Search interventions (MPox)       Number of clicks of Search interventions (MPox)         1,062       0         1,608       0         936       1



Cyprus	130	1	1%
Czech Republic	998	0	0%
Denmark	1,040	0	0%
Estonia	162	0	0%
Finland	1,420	4	0%
France	3,674	0	0%
Germany	9,032	4	0%
Greece	1,866	3	0%
Hungary	762	3	0%
Ireland	1,864	1	0%
Italy	4,912	7	0%
Latvia	238	1	0%
Lithuania	386	0	0%
Luxembourg	116	0	0%
Malta	60	0	0%
Netherlands	3,816	2	0%
Poland	3,498	4	0%
Portugal	1,428	0	0%
Romania	3,770	6	0%



Slovakia	476	1	0%
Slovenia	146	0	0%
Spain	6,566	0	0%
Sweden	3,228	9	0%
Iceland	40	0	0%
Liechtenstein	6	0	0%
Norway	1,548	1	0%
Total EU	53,758	49	0%
Total EEA	55,352	50	0%

	Number of impressions of Search interventions (Climate change)	Number of clicks of Search interventions (Climate change)	Click Through Rate of Search interventions (Climate change)
Member States			
Austria	190,243	70	0.0%
Belgium	185,641	64	0.0%
Bulgaria	75,215	46	0.1%
Croatia	83,932	27	0.0%
Cyprus	18,270	9	0.0%
Czech Republic	113,079	68	0.1%
Denmark	180,842	38	0.0%



Estonia	27,830	15	0.1%
Finland	197,657	73	0.0%
France	725,332	214	0.0%
Germany	1,551,842	481	0.0%
Greece	171,828	114	0.1%
Hungary	172,584	107	0.1%
Ireland	279,240	39	0.0%
Italy	823,065	343	0.0%
Latvia	34,498	25	0.1%
Lithuania	65,701	45	0.1%
Luxembourg	12,527	4	0.0%
Malta	9,172	5	0.1%
Netherlands	454,563	110	0.0%
Poland	756,713	379	0.1%
Portugal	265,699	45	0.0%
Romania	256,836	120	0.0%
Slovakia	58,904	36	0.1%
Slovenia	33,536	20	0.1%
Spain	1,106,982	231	0.0%



Total EEA	8,550,043	2,898	0.0%
Total EU	8,364,476	2,838	0.0%
Norway	180,076	54	0.0%
Liechtenstein	502	0	0.0%
Iceland	4,989	6	0.1%
Sweden	512,745	110	0.0%

	Number of impressions of Search interventions (Regional Elections)	Number of clicks of Search interventions (Regional Elections)	Number of CTR of Search interventions (Regional Elections)
Member States			
Austria	1,391,597	2,065	0.1%
Belgium	7,384,171	3,882	0.1%
Bulgaria	503,487	432	0.1%
Croatia	821,537	798	0.1%
Cyprus	184,467	163	0.1%
Czech Republic	743,697	465	0.1%
Denmark	796,198	481	0.1%
Estonia	205,668	206	0.1%
Finland	958,746	532	0.1%
France	43,521,161	14,164	0.0%



Germany	15,780,895	24,091	0.2%
Greece	2,902,749	2,117	0.1%
Hungary	869,113	693	0.1%
Ireland	1,658,260	626	0.0%
Italy	12,914,047	11,212	0.1%
Latvia	290,005	358	0.1%
Lithuania	513,440	484	0.1%
Luxembourg	234,104	153	0.1%
Malta	113,452	72	0.1%
Netherlands	2,443,760	1,146	0.0%
Poland	2,939,553	1,761	0.1%
Portugal	10,617,934	3,144	0.0%
Romania	6,849,724	3,831	0.1%
Slovakia	417,230	285	0.1%
Slovenia	263,037	203	0.1%
Spain	8,821,184	4,181	0.0%
Sweden	1,828,769	1,230	0.1%
Iceland	1,872	0	0.0%
Liechtenstein	332	0	0.0%



Norway	13,649	12	0.1%
Total EU	125,967,985	78,775	0.1%
Total EEA	125,983,838	78,787	0.1%

	Number of impressions of Public service announcements (Covid-19)	Number of impressions of Public service announcements (Covid 19 Vaccine)	Number of impressions of Public service announcements (Holocaust Misinformation/Denial)
Member States			
Austria	25	2	19
Belgium	28	4	69
Bulgaria	23	4	9
Croatia	11	3	12
Cyprus	3	1	2
Czech Republic	69	16	40
Denmark	22	4	18
Estonia	4	2	2
Finland	30	6	1,789
France	171	44	502
Germany	311	23	227
Greece	75	9	19



Hungany	23	3	13
Hungary		3	13
Ireland	33	6	21
Italy	189	48	97
Latvia	9	1	7
Lithuania	16	9	9
Luxembourg	0	1	1
Malta	1	1	0
Netherlands	73	11	41
Poland	151	17	74
Portugal	18	2	8
Romania	39	7	14
Slovakia	29	3	10
Slovenia	5	0	3
Spain	156	30	33
Sweden	41	29	4,853
Iceland	1	1	7
Liechtenstein	0	0	0
Norway	21	6	2,611
Total EU	1,555	286	7,892



Total EEA	1,577	293	10,510
	Number of impressions of Public service announcements (MPox)	Number of impressions of the safety center page on Covid-19	Number of impressions of the safety center page on election integrity
Member States			
Austria	0	1,597	252
Belgium	0	1,238	349
Bulgaria	0	868	193
Croatia	0	206	65
Cyprus	0	210	51
Czech Republic	0	485	92
Denmark	0	276	67
Estonia	0	69	23
Finland	2	1,000	219
France	2	12,733	3,188
Germany	22	16,023	3,395
Greece	0	804	215
Hungary	1	1,057	184
Ireland	0	810	402
Italy	3	6,523	1,318
Latvia	1	111	29



Lithuania	0	206	29
Luxembourg	1	143	34
Malta	0	76	13
Netherlands	2	1,959	644
Poland	0	2,012	499
Portugal	1	980	393
Romania	0	1,882	618
Slovakia	0	302	61
Slovenia	0	132	25
Spain	1	4,063	1,156
Sweden	1	925	316
Iceland	0	33	7
Liechtenstein	0	1	0
Norway	2	628	128
Total EU	37	56,690	13,830
Total EEA	39	57,352	13,965

Measure 17.2	
INCASULE 17.2	



## **QRE 17.2.1**

In order to raise awareness among our users about specific topics and empower them, we run a variety of on and off-platform media literacy campaigns. Our approach may differ depending on the topic. We localise certain campaigns (e.g., for elections) meaning we collaborate with national partners to develop an approach that best resonates with the local audience. For other campaigns such as the War in Ukraine, our emphasis is on scalability and connecting users to accurate and trusted resources.

Below are examples of the campaigns we have most recently run in-app which have leveraged a number of the intervention tools we have outlined in our response to QRE 17.1.1 (e.g. search interventions and video notice tags).

(I) Promoting election integrity. As well as the election integrity pages on TikTok's <u>Safety</u> <u>Center</u> and <u>Transparency Center</u>, which bring awareness and visibility to how we tackle election misinformation and covert influence operations on our platform, we launched media literacy campaigns in advance of several major European elections.

**2024 Finnish Presidential Election:** From 1 January 2024, we launched an in-app <u>Election Centre</u> to provide users up-to-date information about the 2024 Finnish presidential elections. The centre contained a section about fighting fake news, which linked to media literacy association <u>Mediataitokoulu</u> and included videos created in partnership with fact-checking organisation <u>Logically Facts</u>.





**2024 Irish Referendum:** From 2 February 2024, we launched an in-app <u>Referendum Centre</u> to provide users with up-to-date information about the 2024 Irish Referendum. The centre contained a section about fighting fake news, which was linked to media literacy association <u>medialiteracyireland.ie</u> and included videos created in partnership with fact-checking organisations <u>The Journal</u> and <u>Logically Facts</u>.















**2024 Slovak Presidential Election:** From 3 March 2024, we launched an in-app <u>Election Centre</u> to provide users up-to-date information about the 2024 Slovak presidential elections. The centre contained a section about fighting-fake news, including videos created in partnership with local media literacy organisation <u>DigiQ</u>.





**2024 EU Parliamentary Election:** From 28 March 2024, we launched an individual in-app Election Centre for each EU member state available in 24 official EU languages. Working with electoral commissions, the European Parliament, and civil society organisations including fact-checkers, these Centres connected people with reliable voting information, including when, where, how to vote and, ultimately, the election results themselves. Between March and June 2024, several countries in the EU, including Croatia, Lithuania, Bulgaria and Belgium, called national elections. In these specific instances we worked with national electoral commissions to include relevant voting information in our in-app EU Election Centre. All 27 Election Centres contained a media literacy section that included videos from the European Parliament, our trusted fact-checking partners and local media literacy bodies: Agence France-Presse (AFP), Deutsche Presse-Agentur (dpa), Demagog.pl, Demagog.cz, Facta, Faktograf, Logically Facts, Newtral, Poligrafo, Delfi.lt, The Journal, Nieuwscheckers, Funky Citizens, DigiQ, Ostro.













**2024 French Parliamentary Election:** From 17 June 2024, we launched an in-app <u>Election Centre</u> to provide users up-to-date information about the 2024 French Legislative elections. The centre contained a section about fighting fake news, including videos created in partnership with fact-checking organisations <u>Agence France-Presse</u> (AFP).





(II) Election Speaker Series. To further promote election integrity, and inform our approach to country-level and 2024 EU Parliamentary Election, we invited suitably qualified local and regional external experts to share their insights and market expertise with our internal teams. Our recent Election Speaker Series heard presentations from the following organisations:

• German Press Agency (dpa): Austria, Germany & The Netherlands

• Agence France-Presse (AFP): France, Belgium, Greece & Cyprus

Facta: ItalyNewtral: SpainPoligrafo: Portugal

• Faktograf: Croatia & Slovenia

Demagog.pl: Poland

Funky Citizens: Romania & Bulgaria
 Demagog.cz: Czechia & Slovakia

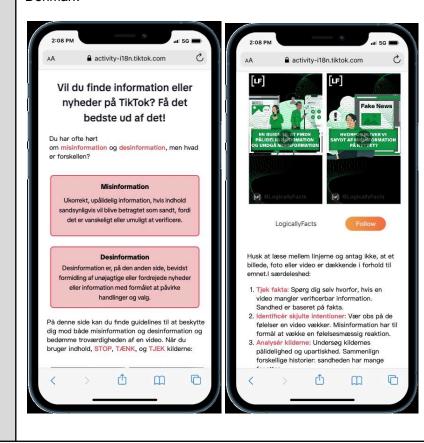
LeadStories: HungaryThe Journal: Ireland



- Delfi.lt: Latvia, Lithuania & Estonia
- Logically Facts: Sweden, Denmark & Finland

(III) Media literacy (General). We rolled out two general media literacy and critical thinking skills campaigns with Logically Facts in Denmark and Nieuwschecker in the Netherlands during this reporting period.

## Denmark

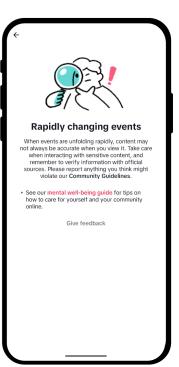


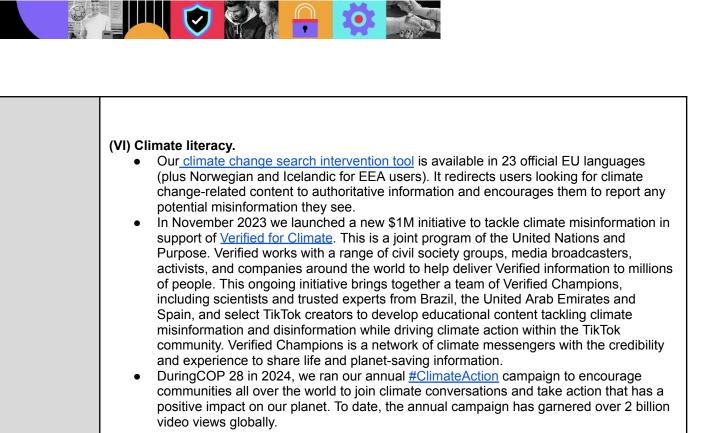


(IV) Media literacy (War in Ukraine). We continue to serve 14 localised media literacy campaigns specific to the war in Ukraine in Austria, Bulgaria, Czech Republic, Croatia, Estonia, Germany, Hungary, Latvia, Lithuania, Poland, Romania, Slovakia, Slovenia, Ukraine. Through these media literacy campaigns, users searching for keywords relating to the war in Ukraine on TikTok are directed to tips prepared in partnership with local media literacy bodies and our trusted fact-checking partners, to help them identify misinformation and prevent its spread on the platform.

**(V) Israel-Hamas conflict.** To help raise awareness and to protect our users, we have launched search interventions which are triggered when users search for neutral terms related to this topic (e.g., Israel, Palestine). These search interventions remind users to pause and check their sources and also directs them to well-being resources.







Over the last year, we've seen a cumulative total of 12 billion views on content pertaining to environmental sustainability. <u>Cleanup</u> content, <u>sustainable living</u> content and the EarthDay hashtag, which was used in over 25,000 videos on Earth Day itself, provided educational detailed <u>explanations</u> about how and why we should protect our planet.

### SLI 17.2.1 - actions enforcing policies above

We are pleased to report metrics on the two general media literacy and critical thinking skills campaigns in Denmark and the Netherlands as well as on the following campaigns that were launched at the end of December 2023 and have continued to run during the reporting period:

- Six [Austria, Bulgaria, Czech Republic, Croatia, Germany and Slovenia] that were specific to the war in Ukraine, including identifying misinformation and manipulated media reports
- Five [Finland, Ireland, Italy, Spain and Sweden] were focused on general media literacy and critical thinking skills.





Member States	Total number of impressions of the H5 Page (Views generated till June 30th, 2024)	Number of impressions of the search intervention	Number of clicks on the search intervention	Click through rate of the search intervention
Denmark (in partnership with Logically Facts)	27,025	118,603	134	0.11%
The Netherlands (in partnership with Nieuwscheckers)	19,607	579,980	656	0.11%
Ireland (in partnership with The journal.ie)	1,504	1,378,004	2,023	0.15%
Finland (in partnership with Logically Facts)	17,314	836,546	1,758	0.21%
Sweden (in partnership with Logically Facts)	32,223	1,599,654	2,671	0.17%
Spain (in partnership with Maldita)	71,952	12,625,290	9,391	0.07%
Italy (in partnership with Facta)	187,930	937,322	1,134	0.12%
Austria and Germany (in partnership with Correctiv)	15,262	16,883,657	25,586	0.15%
Bulgaria	340	299,303	565	0.19%
Croatia	183	356,262	330	0.09%
Czech Republic	613	1,138,561	960	0.08%
Slovenia	103	125,118	174	0.14%

TikTok's CoPD Report 147 January to June 2024











Measure 17.3



#### **QRE 17.3.1**

As documented in the TikTok Safety Center <u>Safety Partners</u> page, we work with an array of industry experts, non-governmental organisations, and industry associations around the world in our commitment to building a safe platform for our community. They include media literacy bodies, to develop campaigns that educate users and redirect them to authoritative resources, and fact-checking partners. Specific examples of partnerships within the campaigns and projects set out in QRE 17.2.1 are:

- (I) Promoting election integrity. We partner with various media organisations and fact-checkers to promote election integrity on TikTok. For more detail about the input our fact-checking partners provide please refer to QRE 30.1.3.
  - EU Parliamentary Election: In advance of the EU Parliamentary Election in June 2024, we launched an individual in-app Election Centre for each EU member state available in 24 official EU languages. All 27 Elections Centres contained a media literacy section that included videos from the European Parliament, our trusted fact-checking partners and local media literacy bodies: Agence France-Presse (AFP), Deutsche Presse-Agentur (dpa), Demagog.pl, Demagog.cz, Facta, Faktograf, Logically Facts, Newtral, Poligrafo, Delfi.lt, The Journal, Nieuwscheckers, Funky Citizens, DigiQ, Ostro.
  - Election speaker series. To further promote election integrity, and inform our approach to country-level EU elections, we invited suitably qualified local and regional external experts to share their insights and market expertise with our internal teams. Our recent Election Speaker Series heard presentations from the following organisations:
    - o German Press Agency (dpa): Austria, Germany & The Netherlands
    - Agence France-Presse (AFP): France, Belgium, Greece & Cyprus
    - Facta: ItalyNewtral: SpainPoligrafo: Portugal
    - > Faktograf: Croatia & Slovenia
    - Demagog.pl: Poland
    - Funky Citizens: Romania & BulgariaDemagog.cz: Czechia & Slovakia
    - LeadStories: HungaryThe Journal: Ireland
    - Delfi.lt: Latvia, Lithuania & Estonia



Logically Facts: Sweden, Denmark & Finland
(II) War in Ukraine.  We continue to run our media literacy campaigns about the war in Ukraine, developed in partnership with our media literacy partners Correctiv in Austria and Germany, Fakenews.pl in Germany and Lead Stories in Bulgaria, Czech Republic, Croatia, Estonia, Hungary, Latvia, Lithuania, Romania, Slovakia, Slovenia, and Ukraine.
(III) Covid-19. In relation to Covid-19 misinformation, we partnered with: (i) the WHO Tech Taskforce; and (ii) our European fact-checking organisations including AFP, Facta, Logically, Lead Stories, Newtral, Science Feedback, Teyit, DPA, Reuters, and Demagog. Our partners provide us with information about Covid-19 trends including new variants, vaccine side effects on adults and children, and information about new treatments (such as antivirals). This information helps us improve our in-app intervention tools.

# V. Empowering Users

### Commitment 18

Relevant Signatories commit to minimise the risks of viral propagation of Disinformation by adopting safe design practices as they develop their systems, policies, and features.

In line with this commitment, did you deploy new
implementation measures (e.g. changes to your
terms of service, new tools, new policies, etc)?
[Yes/No]

Yes

If yes, list these implementation measures here [short bullet points].

 Ensured fact-checking coverage by a combination of permanent and temporary partnerships covering at least one official language of each of the European Member States for the EU Parliamentary Elections in June 2024, and expanded our misinformation moderation teams' coverage accordingly.

	<ul> <li>Onboarded a new fact-checking partner: Faktograf, the first Croatian media specialised in fact-checking.</li> <li>Continued to improve the accuracy of, and overall coverage provided by, our machine learning detection models.</li> <li>Participated in the data access pilot with EDMO trialling the process for sharing data with vetted researchers as designated under the DSA.</li> <li>Actively participating in the EDMO working group for the creation of the Independent Intermediary Body (IIB) to support research on digital platforms.</li> <li>Invested in training and development for our Trust and Safety team, including regular internal sessions dedicated to knowledge sharing and discussion about relevant issues and trends and attending external events to share their expertise and support continued professional learning.         <ul> <li>For example, 12 members of our Trust &amp; Safety team (including leaders of our fact-checking program) attended the Global Fact 11   Global Fact-Checking Conference hosted by the International Fact-Checking Network in June 2024 and participated in an on-the-record mainstage presentation answering questions about our misinformation strategy and partnerships with professional fact-checkers.</li> <li>Continued to participate in, and co-chair, the working group on Elections.</li> </ul> </li> </ul>
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.
Measure 18.1	
QRE 18.1.1	TikTok takes a multi-faceted approach to tackling the spread of harmful misinformation, regardless of intent. This includes our policies, products, practices and external partnerships with fact-checkers, media literacy bodies, and researchers.

- (I) Removal of violating content or accounts. To reduce potential harm, we aim to remove content or accounts that violate our CGs including our I&A policies before they are viewed or shared by other people. We detect and take action on this content by using a combination of automation and human moderation.
  - a) Automated Review We place considerable emphasis on proactive detection to remove violative content. Content that is uploaded to the platform is typically first reviewed by our automated moderation technology, which looks at a variety of signals across content, including keywords, images, captions, and audio, to identify violating content. We work with various external experts, like our <u>fact-checking partners</u>, to inform our keyword lists. If our automated moderation technology identifies content that is a potential violation, it will either be automatically removed from the platform or flagged for further review by our human moderation teams. In line with our safeguards to help ensure accurate decisions are made, automated removal is applied when violations are the most clear-cut. We also carry out targeted sweeps of certain types of violative content including harmful misinformation, where we have identified specific risks or where our fact-checking partners or other experts have alerted us to specific risks.
  - b) Human Moderation While some misinformation can be enforced through technology alone—for example, repetitions of previously debunked content—misinformation evolves quickly and is highly nuanced. That's why we have misinformation moderators with enhanced training and access to tools like our global repository of previously fact-checked claims from the IFCN-accredited fact-checking partners, who help assess the accuracy of content. We also have teams on the ground who partner with experts to prioritise local context and nuance. We may also issue guidance to our moderation teams to help them more easily spot and take swift action on violating content. Human moderation will also occur if a video gains popularity or has been reported. Community members can report violations in-app and on our website. Our fact-checking partners and other stakeholders can also report potential violating content to us directly.



- (II) Safety in our recommendations. In addition to removing content that clearly violates our CGs, we have a number of safeguards in place to ensure the For You feed (as the primary access point for discovering original and entertaining content on the platform) has safety built-in.
  - a. For content that does not violate our CGs but may negatively impact the authenticity of the platform, we reduce its prominence on the For You feed and / or label it. The types of misinformation we may make For You feed ineligible are now made even clearer to users here; general conspiracy theories, unverified information related to an emergency or unfolding event and potential high-harm misinformation that is undergoing a fact-check. We also label accounts and content of state-affiliated media entities to empower users to consider the sources of information. Our moderators take additional precautions to review videos as they rise in popularity to reduce the likelihood of content that may not be appropriate entering our recommended system.
  - b. Providing access to authoritative information is an important part of our overall strategy to counter misinformation. There are a number of ways in which we do this, including launching information centres with informative resources from authoritative third-parties in response to global or local events, adding public service announcements on hashtag or search pages, or labelling content related to a certain topic to prompt our community to seek out authoritative information.
- (III) Safety by Design. Within our Trust and Safety Product and Policy teams, we have subject matter experts dedicated to integrity and authenticity. When we develop a new feature or policy, these teams work closely with external partners to ensure we are building safety into TikTok by design and reflecting industry best practice. For example:
  - We collaborated with <u>Irrational Labs</u> to develop and implement <u>specialised prompts</u> to help users consider before sharing unverified content (as outlined in **QRE 21.3.1**),
  - Yad Vashem created an enrichment program on the Holocaust for our Trust and Safety team. The five week program aimed to give our team a deeper understanding about the Holocaust, its lessons and misinformation related to antisemitism and hatred.
  - We worked with local/regional experts through our Election Speaker Series to ensure their insights and expertise informs our internal teams ahead of the EU Parliamentary Elections of June 2024 (as outlined in **QRE 17.3.1**).

QRE 18.1.2	The For You feed is the interface users first see when they open TikTok. It is central to the TikTok experience and where most of our users spend their time exploring the platform. The system recommends content by ranking videos based on a combination of factors including:  • user selections (i.e. interest categories indicated in the registration process or selecting "not interested" on content);  • user interactions (i.e. videos users like/share, watch in full or skip accounts users follow, accounts that follow users, comments users post and content they create);  • video information (i.e. captions/sounds and hashtags used, number of video views, and the country in which the video was published.); and  • device and account settings (i.e. language preferences, country settings, time zone and day, and device types).  All these factors are weighted based on their value to a user and the recommendation system then ranks them to determine the likelihood of a user's interest in a particular category of content. The weighting of a factor can change dynamically. We share more information about our recommender system in our Transparency Center and below in our response to QRE 19.1.1.
QRE 18.1.3	We take action to prevent and mitigate the spread of inaccurate, misleading, or false misinformation that may cause significant harm to individuals or the public at large. We do this by removing content and accounts that violate our rules, investing in media literacy and connecting our community to authoritative information, and partnering with external experts. Our I&A policies make clear that we do not allow activities that may undermine the integrity of our platform or the authenticity of our users. We remove content or accounts that involve misleading information that causes significant harm or, in certain circumstances, reduce the prominence of content. The types of misinformation we may make ineligible For You feed are set out in our Community Guidelines.
	<ul> <li>Misinformation</li> <li>Conspiracy theories that are unfounded and claim that certain events or situations are carried out by covert or powerful groups, such as "the government" or a "secret society".</li> <li>Moderate harm health misinformation, such as an unproven recommendation for how to treat a minor illness.</li> </ul>



- Repurposed media, such as showing a crowd at a music concert and suggesting it is a political protest.
- Misrepresenting authoritative sources, such as selectively referencing certain scientific data to support a conclusion that is counter to the findings of the study.
- Unverified claims related to an emergency or unfolding event.
- Potential high-harm misinformation while it is undergoing a fact-checking review.

#### Civic and Election Integrity

- Unverified claims about an election, such as a premature claim that all ballots have been counted or tallied.
- Statements that significantly misrepresent authoritative civic information, such as a false claim about the text of a parliamentary bill.

#### Fake Engagement

 Content that tricks or manipulates others as a way to increase gifts, or engagement metrics, such as "like-for-like" promises or other false incentives for engaging with content.

To enforce our CGs at scale, we use a combination of automated review and human moderation. While some misinformation can be enforced through technology alone—for example, repetitions of previously debunked content—misinformation evolves quickly and is highly nuanced. Assessing harmful misinformation requires additional context and assessment by our misinformation moderators who have enhanced training, expertise and tools to identify such content, including our global repository of previously fact-checked claims from the IFCN-accredited fact-checking partners and direct access to our fact-checking partners where appropriate.

Our network of independent fact-checking partners do not moderate content directly on TikTok, but assess whether a claim is true, false, or unsubstantiated so that our moderators can take action based on our Community Guidelines. We incorporate fact-checker input into our broader content moderation efforts through:

Proactive insight reports that flag new and evolving claims they're seeing across the internet.
 This helps us detect harmful misinformation and anticipate misinformation trends on our platform.

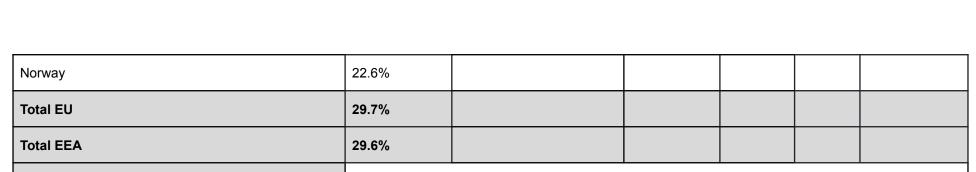
	<ul> <li>A repository of previously fact-checked claims to help misinformation moderators make swift and accurate decisions.</li> </ul>					
	Working with our network of independent fact-checking organisations enables TikTok to identify and take action on misinformation and connect our community to authoritative information around important events. This is an important part of our overall strategy to counter misinformation. There a a number of ways in which we do this, including launching information centers with resources from authoritative third-parties in response to global or local events, adding public service announcement (PSAs) on hashtags or search pages, or labelling content related to a certain topic to prompt our community to seek out authoritative information.  We are also committed to civic and election integrity and mitigating the spread of false or misleading content about an electoral or civic process. We work with national electoral commissions, media literacy bodies and civil society organisations to ensure we are providing our community with accura up-to-date information about an election through our in-app election information centers, election					
	guides, search ii	nterventions, content labels a	iu nasiliay PSA	5.		
SLI 18.1.1 - actions proving effectiveness of measures and policies	Methodology of data measurement:  The share cancel rate (%) following the unverified content label share warning pop-up indicates the percentage of users who do not share a video after seeing the label pop up. This metric is based on the approximate location of the users that engaged with these tools.					
List actions per member states and languages (see example table above)	Share cancel rate (%) following the unverified content label share warning pop-up (users who do not share the					



	video after seeing the pop up)			
Member States				
Austria	28.4%			
Belgium	25.5%			
Bulgaria	26.6%			
Croatia	27.5%			
Cyprus	25.0%			
Czech Republic	24.3%			
Denmark	21.1%			
Estonia	20.8%			
Finland	24.2%			
France	31.8%			
Germany	30.9%			
Greece	27.0%			
Hungary	35.2%			



Ireland	20.8%			
Italy	32.4%			
Latvia	33.3%			
Lithuania	25.7%			
Luxembourg	35.2%			
Malta	30.2%			
Netherlands	20.3%			
Poland	32.0%			
Portugal	27.2%			
Romania	27.7%			
Slovakia	29.0%			
Slovenia	20.8%			
Spain	32.7%			
Sweden	24.5%			
Iceland	26.3%			
Liechtenstein	13.3%			



#### Measure 18.2

#### **QRE 18.2.1**

We take action against misinformation that causes significant harm to individuals, our community, or the larger public regardless of intent. We do this by removing content and accounts that violate our rules, by investing in media literacy and connecting our community to authoritative information, and by partnering with experts.

Our Terms of Service and I&A policies under our CGs are the first line of defence in combating harmful misinformation and (as outlined in more detail in QRE 14.1.1) deceptive behaviours on our platform. These rules make clear to our users what content we remove or make ineligible for the For You feed when they pose a risk of harm to our users and our community.

Specifically, our policies do not allow:

#### Misinformation

- Misinformation that poses a risk to public safety or may induce panic about a crisis event or emergency, including using historical footage of a previous attack as if it were current, or incorrectly claiming a basic necessity (such as food or water) is no longer available in a particular location. Health misinformation, such as misleading statements about vaccines, inaccurate medical advice that discourages people from getting appropriate medical care for a life-threatening disease, or other misinformation which may cause negative health effects on an individual's life
- Climate change misinformation that undermines well-established scientific consensus, such as denying the existence of climate change or the factors that contribute to it.
- o Conspiracy theories that name and attack individual people.
- 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.

## Civic and Election Integrity



- Election misinformation, including:
  - How, when, and where to vote or register to vote;
  - Eligibility requirements of voters to participate in an election, and the qualifications for candidates to run for office:
  - Laws, processes, and procedures that govern the organisation and implementation of elections and other civic processes, such as referendums, ballot propositions, or censuses;
  - Final results or outcome of an election.

#### Edited Media and Al-Generated Content (AIGC)

- Realistic-appearing people under the age of 18.
- The likeness of adult private figures, if we become aware it was used without their permission.
- Misleading AIGC or edited media that falsely shows:
  - Content made to seem as if it comes from an authoritative source, such as a reputable news organisation;
  - A crisis event, such as a conflict or natural disaster.
- A public figure who is:
  - being degraded or harassed, or engaging in criminal or antisocial behaviour;
  - taking a position on a political issue, commercial product, or a matter of public importance (such as an election);
  - being politically endorsed or condemned by an individual or group.

### Fake Engagement

- Facilitating the trade or marketing of services that artificially increase engagement, such as selling followers or likes.
- o Providing instructions on how to artificially increase engagement on TikTok.

We have made even clearer to our users <u>here</u> that the following content is ineligible for the For You feed:

#### Misinformation

- Conspiracy theories that are unfounded and claim that certain events or situations are carried out by covert or powerful groups, such as "the government" or a "secret society"
- Moderate harm health misinformation, such as an unproven recommendation for how to treat a minor illness



- Repurposed media, such as showing a crowd at a music concert and suggesting it is a political protest
- Misrepresenting authoritative sources, such as selectively referencing certain scientific data to support a conclusion that is counter to the findings of the study
- Unverified claims related to an emergency or unfolding event
- Potential high-harm misinformation while it is undergoing a fact-checking review

#### Civic and Election Integrity

- Unverified claims about an election, such as a premature claim that all ballots have been counted or tallied
- Statements that significantly misrepresent authoritative civic information, such as a false claim about the text of a parliamentary bill

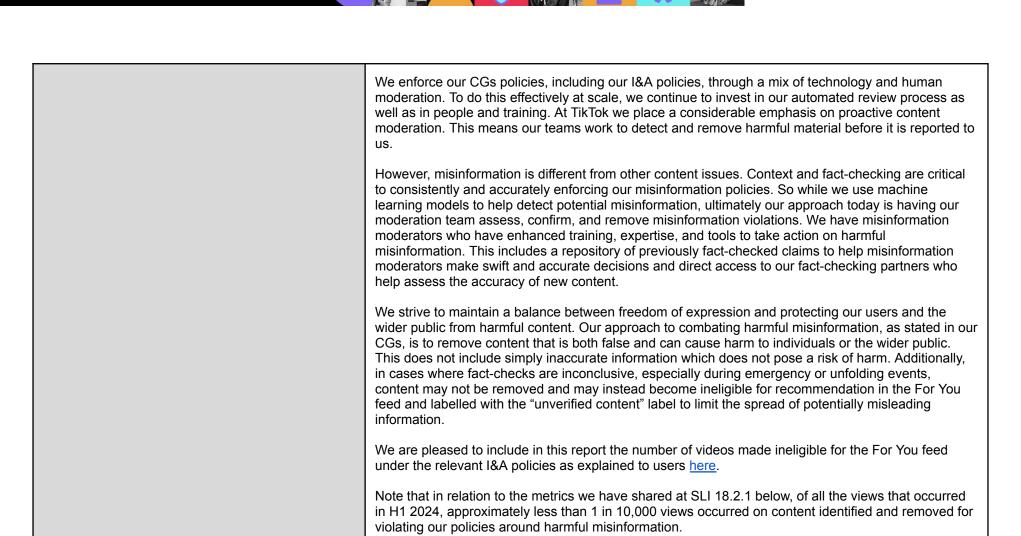
#### Fake Engagement

 Content that tricks or manipulates others as a way to increase gifts, or engagement metrics, such as "like-for-like" promises or other false incentives for engaging with content

As outlined in the QRE 14, we also remove accounts that seek to mislead people or use TikTok to deceptively sway public opinion. These activities range from inauthentic or fake account creation, to more sophisticated efforts to undermine public trust.

We have policy experts within our Trust and Safety team dedicated to the topic of integrity and authenticity. They continually keep these policies under review and collaborate with external partners and experts to understand whether updates or new policies are required and ensure they are informed by a diversity of perspectives, expertise, and lived experiences. In particular, our Safety Advisory Council for Europe, which brings together independent leaders from academia and civil society, represent a diverse array of backgrounds and perspectives, and are made up of experts in free expression, misinformation and other safety topics. They work collaboratively with us to inform and strengthen our policies, product features, and safety processes.

**Enforcing our policies.** We remove content – including video, audio, livestream, images, comments, links, or other text – that violates our I&A policies. Individuals are notified of our decisions and can appeal them if they believe no violation has occurred. We also make clear in our CGs that we will temporarily or permanently ban accounts and/or users that are involved in serious or repeated violations, including violations of our I&A policies.



# SLI 18.2.1 - actions taken in response to policy violations

### Methodology of data measurement:

We have based the following numbers on the country in which the video was posted: videos removed because of violations of our Misinformation, Civic and Election Integrity and Synthetic and Manipulated Media policies.



	The number of views of videos removed because of violation of each of these policies is based on the approximate location of the user.  We also updated the methodology on the number of videos made inelegible for the For You Feed under our Misinformation policy.					
	Total no of violations					
List actions per member states and languages (see example table above)	Number of videos removed because of violation of misinformati on policy	Number of views of videos removed because of violation of misinformation policy	Number of videos made ineligible for the For You feed under the Misinformation policy.			
Member States						
Austria	2,577	5,281,606	2,140			
Belgium	5,957	9,135,478	12,051			
Bulgaria	1,211	4,130,686	4,249			
Croatia	1,050	3,479,647	650			
Cyprus	329	1,492,081	242			
Czech Republic	7,715	5,965,348	5,989			
Denmark	1,227	7,251,467	4,453			



Estonia	131	393,644	186	
Finland	51,960	3,711,975	2,836	
France	54,817	136,142,611	474,524	
Germany	47,499	192,724,825	39,801	
Greece	2,294	11,428,865	2,874	
Hungary	1,357	6,885,084	2,530	
Ireland	1,872	10,018,444	3,928	
Italy	31,343	50,536,321	76,878	
Latvia	248	102,409	343	
Lithuania	273	248,280	329	
Luxembourg	155	25,393	165	
Malta	72	887	78	
Netherlands	4,725	12,386,535	8,560	
Poland	9,346	74,347,417	11,108	
Portugal	1,790	9,065,634	2,830	
Romania	8,087	35,808,475	23,449	



Slovakia	581	2,659,806	468	
Slovenia	197	839,246	118	
Spain	21,430	51,965,214	56,649	
Sweden	4,409	15,751,209	8,081	
Iceland	74	56,096	70	
Liechtenstein	6	206,612	3	
Norway	1,678	9,611,385	2,830	
Total EU	262,652	651,778,587	745,509	
Total EEA	264,410	661,652,680	748,412	
List actions per member states and languages (see example table above)	Number of videos removed because of violation of Civic and Election Integrity policy	Number of views of videos removed because of violation of Civic and Election Integrity policy	Number of videos removed because of violation of Synthetic and Manipulated Media	Number of views of videos removed because of violation of Synthetic and Manipulated Media
Member States				
Austria	395	124,006	111	13,595
Belgium	562	19,378	182	44,894
Bulgaria	170	636,745	30	122,302
Croatia	49	332,549	22	63,556



Cyprus	25	544	15	925
Czech Republic	14,040	198,058	41	5,950
Denmark	251	88,553	96	4,716,805
Estonia	47	25,619	11	856
Finland	31,939	28,790	30	1,726
France	2,273	8,487,560	954	12,016,565
Germany	5,080	1,746,091	1,553	6,907,249
Greece	144	208,358	61	7,089
Hungary	118	20,744	11	412,593
Ireland	300	243,851	111	252,728
Italy	1,701	67,541	460	191,672
Latvia	192	70,702	32	1,159
Lithuania	133	11,025	33	6,527
Luxembourg	140	7,136	23	556
Malta	16	8,427	14	2,285
Netherlands	737	69,616	178	2,834,922
Poland	431	747,214	121	3,838,027
Portugal	312	248,901	97	12,676
Romania	788	4,563,567	106	1,713,376



		I		1		
Slovakia	60	171,706	38	23,403,840		
Slovenia	7	319	14	290		
Spain	1,314	188,132	270	2,886,615		
Sweden	439	66,412	68	1,823,170		
Iceland	83	816	42	1,095		
Liechtenstein	15	2,015 5 1,542				
Norway	277	1,098,224	144	652,489		
Total EU	61,663 18,381,544 4,682 61,281,948					
Total EEA	62,038	19,482,599	4,873	61,937,074		
Measure 18.3						
QRE 18.3.1	We regularly consult with third party experts and researchers in relation to the development of policies and features which are designed to reduce the spread of disinformation. For example, we engaged with experts globally on our Election Misinformation policies, which help inform updates of our I&A policies.  We are proud of our close work with behavioural psychologists, Irrational Labs, which led to the development of the following warning and labelling features (more detail at QRE 21.3.1):  • specialised prompts for unverified content, which alerts viewers to unverified content identified during an emergency or unfolding event and • our state-controlled media label, which brings transparency to our community in relation to state affiliated media entities and raises awareness among users to encourage users to consider the reliability of the source.  We are proud to be a signatory to the Partnership on Al's (PAI) Responsible Practices for Synthetic Media. We contributed to developing this code of industry best practices for Al transparency and responsible innovation, balancing creative expression with the risks of emerging Al technology.					

TikTok's CoPD Report 167 January to June 2024



# V. Empowering Users

# Commitment 19

Relevant Signatories using recommender systems commit to make them transparent to the recipients regarding the main criteria and parameters used for prioritising or deprioritising information, and provide options to users about recommender systems, and make available information on those options.

QRE 19.1.1	The For You feed is the interface users first see when they open TikTok. It's central to the TikTok experience and where most of our users spend their time exploring the platform.
Measure 19.1	
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, list these implementation measures here [short bullet points].	Improved user transparency about our recommender systems. See our newly published     Transparency Center <u>System Guides</u> article and updated Help Center article <u>here</u> .
In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	Yes



We make clear to users in our Terms of Service and CGs (and also provide more context in our Help Center <a href="article">article</a> and Transparency Center <a href="page">page</a>) that each account holder's For You feed is based on a personalised recommendation system. The For You feed is curated to each user. Safety is built into our recommendations. As well as removing harmful misinformation content that violates our CGs, we take steps to avoid recommending certain categories of content that may not be appropriate for a broad audience including general conspiracy theories and unverified information related to an emergency or unfolding event. We may also make some of this content harder to find in search.

**Main parameters.** The system recommends content by ranking content based on a combination of factors including:

- user interactions (e.g. content users like, share, comment on, and watch in full or skip, as well as accounts of followers that users follow back):
- Content information (e.g. sounds, hashtags, number of views, and the country the content was published); and
- User information (e.g. device settings, language preferences, location, time zone and day, and device types).

The main parameters help us make predictions on the content users are likely to be interested in. Different factors can play a larger or smaller role in what's recommended, and the importance – or weighting – of a factor can change over time. For many users, the time spent watching a specific video is generally weighted more heavily than other factors. These predictions are also influenced by the interactions of other people on TikTok who appear to have similar interests. For example, if a user likes videos 1, 2, and 3 and a second user likes videos 1, 2, 3, 4 and 5, the recommendation system may predict that the first user will also like videos 4 and 5.

Building on our existing recommender system transparency we introduced the "Why this video" feature, which allows users to see with any particular video that appears in their For You feed factors that influenced why it appeared in their feed. This feature provides added transparency in relation to how our ranking system works and empowers our users to better understand why a particular video has been recommended to them. The feature essentially explains to users how past interactions on the platform have impacted the video they have been recommended. For further information, see our newsroom post.

	User preferences. Together with the safeguards we build into our platform by design, we also empower our users to customise their experience to their preferences and comfort.  These include a number of features to help shape the content they see. For example, in the For You feed:  • Users can click on any video and select "not interested" to indicate that they do not want to see similar content.				
	<ul> <li>Users are able to automatically filter out specific words or hashtags from the content recommended to them(see <a href="https://here">here</a>).</li> <li>Users are able to <a href="refresh their For You feed">refresh their For You feed</a> if they no longer feel like recommendations are relevant to them or are too similar. When the For You feed is refreshed, users view a number of new videos which include popular videos (e.g., they have a high view count or a high like rate). Their interaction with these new videos will inform future recommendations.</li> <li>As part of our obligations under the DSA (Article 38), we introduced non-personalized feeds on our platform, which provide our European users with an alternative to recommender systems. They are able to turn off personalisation so that feeds show non-personalised content. For example, the For You feed, will instead show popular videos in their regions and internationally. See <a href="here.">here.</a></li> </ul>				
Measure 19.2					
SLI 19.2.1 – user settings	Methodology of data measurement:  The number of users who have filtered hashtags or a keyword to set preferences for For You feed, the number of times users clicked "not interested" in relation to the For You feed, and the number of times users clicked on the For You Feed Refresh are all based on the approximate location of the users that engaged with these tools.  The number for videos tagged with AIGC label includes both automatic and creator-generated labeling.  No of times users actively  No of times users actively				



	engaged with these settings	engaged with these settings		
List actions per member states and languages (see example table above)	Number of users that filtered hashtags or words	Number of users that clicked on "not interested"	Number of times users clicked on the For You Feed Refresh	Number of Videos tagged with AIGC label
Member States				
Austria	43,651	950,524	62,732	116,663
Belgium	59,066	1,409,021	92,612	218,989
Bulgaria	29,992	668,940	41,662	116,421
Croatia	21,707	386,552	24,473	38,516
Cyprus	6,480	151,219	15,097	35,510
Czech Republic	37,897	842,904	31,540	155,549
Denmark	31,414	579,925	34,737	63,043
Estonia	10,694	150,574	14,846	27,811
Finland	34,645	619,981	51,962	93,649
France	274,576	8,324,620	452,755	1,506,279
Germany	424,285	8,488,738	694,068	1,466,213
Greece	53,123	1,237,023	82,262	142,617
Hungary	43,113	1,065,910	34,667	146,293
Ireland	49,115	844,056	61,642	56,008



Total EEA	2,211,200	51,376,306	3,207,443	7,875,580
Total EU	2,162,999	50,589,799	3,156,106	7,787,126
Norway	43,820	725,835	46,875	81,378
Liechtenstein	89	2,015	141	245
Iceland	4,292	58,657	4,321	6,831
Sweden	72,760	1,377,269	100,283	196,270
Spain	247,804	6,556,506	371,258	1,029,096
Slovenia	9,003	163,467	17,857	13,901
Slovakia	16,582	338,009	15,081	54,993
Romania	79,852	2,613,471	190,583	276,768
Portugal	53,593	1,084,136	64,187	176,044
Poland	178,345	3,419,332	164,245	595,686
Netherlands	107,804	2,144,265	168,248	167,909
Malta	2,973	68,514	5,603	11,507
Luxembourg	3,607	68,489	6,099	34,245
Lithuania	19,311	309,017	28,438	47,201
Latvia	13,675	260,065	22,931	48,339
Italy	237,932	6,467,272	306,238	951,606



# V. Empowering Users

# Commitment 20

Relevant Signatories commit to empower users with tools to assess the provenance and edit history or authenticity or accuracy of digital content.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	No
If yes, list these implementation measures here [short bullet points].	TikTok did not subscribe to this commitment. TikTok considers that it would be imprudent to commit to this measure at a time when the underlying technology remains unproven and the standards to be complied with are not yet finalised.
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight disinformation and will report on any further development in the next COPD report.
Measure 20.1	
QRE 20.1.1	<b>Not committed</b> . TikTok considers that it would be imprudent to commit to this measure at a time when the underlying technology remains unproven and the standards to be complied with are not yet finalised. Once the relevant technology is proven and the standards are agreed and identifiable, TikTok will assess whether such an approach would be beneficial, taking into account existing measures in place.
Measure 20.2	



QRE 20.2.1	<b>Not committed</b> . Similarly to Measure 20.1, TikTok considers that it would be imprudent to commit to this measure at this time. TikTok is, however, open to reassessing in the future whether such an approach would be beneficial.
	approach would be beneficial.

# V. Empowering Users

### Commitment 21

Relevant Signatories commit to strengthen their efforts to better equip users to identify Disinformation. In particular, in order to enable users to navigate services in an informed way, Relevant Signatories commit to facilitate, across all Member States languages in which their services are provided, user access to tools for assessing the factual accuracy of sources through fact-checks from fact-checking organisations that have flagged potential Disinformation, as well as warning labels from other authoritative sources.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	Yes
If yes, list these implementation measures here [short bullet points].	<ul> <li>Ensured fact-checking coverage by a combination of permanent and temporary partnerships covering at least one official language of each of the European Member States for the EU Parliamentary Election in June 2024, and expanded our misinformation moderation teams' coverage accordingly.</li> <li>Onboarded a new fact-checking partner:Faktograf, the first Croatian media specialised in fact-checking.</li> <li>Building on our new Al-generated label for creators to disclose content that is completely Al-generated or significantly edited by Al, we have expanded our efforts in the AIGC space by:         <ul> <li>Implementing the Coalition for Content Provenance and Authenticity (C2PA)</li> <li>Content Credentials, which enables our systems to instantly recognize and automatically label AIGC that originated on other major platforms.</li> <li>Supporting the coalition's working groups as a C2PA General Member.</li> </ul> </li> </ul>

	<ul> <li>Joining the Content Authenticity Initiative (CAI) to drive wider adoption of the technical standard.</li> <li>Publishing a new Transparency Center article Supporting responsible, transparent Al-generated content.</li> <li>Progressed our partnership with Verified for Climate by:         <ul> <li>Bringing together a team of Verified Champions, including trusted messengers and experts from Brazil, the United Arab Emirates, and Spain, with select TikTok creators, to tackle climate misinformation while driving climate action within the TikTok community.</li> <li>Working with the United Nations and social impact agency Purpose to launch the first phase of the Verified for Climate program in Brazil, the UAE and Spain.</li> </ul> </li> </ul>
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.
Measure 21.1	
QRE 21.1.1	We currently have 12 IFCN accredited fact-checking partners in Europe who provide fact-checking coverage in 23 official EEA languages, including at least one official language of each EU Member States.  Our response to QRE 30.1.2 sets out the specific organisations we partner with.  We ensure that our users benefit from the context and insights provided by the fact checking organisations we partner with in the following ways:

- Enforcement of misinformation policies. Our fact-checking partners play a critical role in helping us enforce our misinformation policies, which aim to promote a trustworthy and authentic experience for our users. We consider context and fact-checking to be key to consistently and accurately enforcing these policies, so, while we use machine learning models to help detect potential misinformation, we have our misinformation moderators assess, confirm, and take action on harmful misinformation. As part of this process, our moderators can access a repository of previously fact-checked claims and they are able to provide content to our expert fact checking partners for further evaluation. Where fact-checking partners advise that content is false, our moderators take measures to assess and remove it from our platform. Our response to QRE 31.1.1 provides further insight into the way in which fact-checking partners are involved in this process.
- Unverified content labelling. As mentioned above, we partner with fact checkers to assess the accuracy of content. Sometimes, our fact-checking partners determine that content cannot be confirmed or checks are inconclusive (especially during unfolding events). Where our fact-checking partners provide us with a rating that demonstrates the claim cannot yet be verified, we may use our unverified content label to inform viewers via a banner that a video contains unverified content, in an effort to raise user awareness about content credibility. In these circumstances, the content creator is also notified that their video was flagged as unsubstantiated content and the video will become ineligible for recommendation in the For You feed.
- In-app tools related to specific topics:
  - Covid-19. We partnered with a number of fact checkers throughout the EU in order to prevent the spread of harmful misinformation related to Covid-19 on our platform, including AFP, Facta, Logically, Lead Stories, Newtral, Science Feedback, Tevit and DPA.
  - Election integrity. We have launched campaigns in advance of several major elections aimed at educating the public about the voting process which encourage users to fact-check information with our fact-checking partners. For example, the election integrity campaign we rolled out in advance of the Finnish Presidential election in January 2024 included a search intervention and setting up an in-app Election Centre. The centre contained a section about fighting fake news, which linked to media literacy association Mediataitokoulu and included videos created in partnership with fact-checking organisation Logically Facts.

	Climate Change. We launched a search intervention which redirects use seeking out climate change-related content to authoritative information. We worked with the UN to provide the authoritative information (see our newsrood post here).  • User awareness of our fact-checking partnerships and labels. We have published blog posts (available in more than 25 languages) and created pages on our Safety Center and Iabels and to support the work of our fact-checking partners.  Methodology of data measurement:  The share of removals under our harmful misinformation policy, share of proactive removals, share of removals before any views and share of the removals within 24h are relative to the total removals of each policy.  The share cancel rate (%) following the unverified content label share warning pop-up indicates the percentage of users who do not share a video after seeing the label pop up. This metric is based on the approximate location of the users that engaged with these tools.  Reach of labels/ fact-checkers and other authoritative sources  Other pertinent metric  Other pertinent metric  Other pertinent metric				
SLI 21.1.1 - actions taken under measure 21.1					
List actions per member states and languages (see example table above)	Share cancel rate (%) following the unverified content label share warning pop-up (users who do not share the video after seeing the pop up)	Share of removals under misinformation policy	Share of proactive removals under misinformatio n policy	Share of video removals before any views under misinformation policy	Share of video removals within 24h by misinformation policy



Member States					
Austria	28.4%	23.1%	94.8%	76.5%	74.3%
Belgium	25.5%	33.5%	98.3%	76.6%	65.5%
Bulgaria	26.6%	23.3%	95.0%	80.5%	85.9%
Croatia	27.5%	58.4%	95.0%	69.3%	82.2%
Cyprus	25.0%	31.9%	95.7%	62.9%	84.5%
Czech Republic	24.3%	6.5%	97.0%	93.5%	95.9%
Denmark	21.1%	9.5%	93.2%	72.9%	80.1%
Estonia	20.8%	11.1%	93.9%	72.5%	80.2%
Finland	24.2%	18.1%	99.8%	99.5%	99.7%
France	31.8%	44.8%	98.4%	78.8%	79.5%
Germany	30.9%	29.7%	96.4%	69.1%	72.5%
Greece	27.0%	29.5%	87.1%	64.0%	80.8%
Hungary	35.2%	32.7%	87.8%	59.4%	76.5%
Ireland	20.8%	30.6%	92.7%	67.3%	66.2%
Italy	32.4%	47.8%	97.5%	74.6%	76.8%
Latvia	33.3%	2.3%	95.6%	68.5%	76.2%
Lithuania	25.7%	11.8%	94.5%	78.0%	81.3%
Luxembourg	35.2%	10.1%	92.9%	65.2%	77.4%



Austria		3.5%	Integrity policy 86.1%	Integrity policy 72.7%	99.5%
Member States		Share of video removals under Civic and Election Integrity policy	Share of proactive video removals under Civic and Election	Share of video removals before any views under Civic and Election	Share of video removals within 24h under Civic and Election Integrity policy
Total EEA	29.6%	26.6%	97.2%	80.0%	83.6%
Total EU	29.7%	26.6%	97.2%	80.0%	83.7%
Norway	22.6%	25.3%	93.9%	71.6%	79.1%
Liechtenstein	13.3%	1.2%	83.3%	66.7%	0.0%
Iceland	26.3%	11.7%	75.7%	66.2%	73.0%
Sweden	24.5%	26.5%	96.2%	75.6%	86.4%
Spain	32.7%	43.4%	98.1%	78.9%	89.0%
Slovenia	20.8%	30.3%	91.4%	78.2%	87.3%
Slovakia	29.0%	23.9%	71.9%	54.6%	70.1%
Romania	27.7%	35.7%	90.9%	74.6%	89.1%
Portugal	27.2%	30.1%	94.4%	73.6%	85.6%
Poland	32.0%	37.1%	91.3%	66.2%	88.2%
Netherlands	20.3%	18.0%	97.1%	86.3%	89.9%
Malta	30.2%	12.1%	98.6%	75.0%	75.0%



Bulgaria	3.3%	91.2%	84.1%	97.6%
Croatia	2.7%	75.5%	38.8%	93.9%
Cyprus	2.4%	96.0%	96.0%	100.0%
Czech Republic	11.8%	99.8%	99.4%	100.0%
Denmark	1.9%	77.7%	56.2%	99.6%
Estonia	4.0%	85.1%	72.3%	85.1%
Finland	11.1%	99.8%	99.8%	100.0%
France	1.9%	96.8%	62.9%	69.9%
Germany	3.2%	98.5%	95.8%	98.0%
Greece	1.9%	97.2%	89.6%	97.2%
Hungary	2.8%	94.1%	82.2%	99.2%
Ireland	4.9%	81.7%	58.7%	96.3%
Italy	2.6%	95.4%	84.4%	96.7%
Latvia	1.8%	65.6%	37.0%	99.0%
Lithuania	5.7%	78.2%	66.9%	100.0%
Luxembourg	9.2%	86.4%	70.0%	100.0%
Malta	2.7%	68.8%	56.3%	87.5%
Netherlands	2.8%	97.2%	93.8%	99.5%
Poland	1.7%	82.8%	79.4%	98.4%



Portugal	5.2%	87.5%	72.1%	99.4%
Romania	3.5%	93.4%	80.7%	91.9%
Slovakia	2.5%	90.0%	83.3%	95.0%
Slovenia	1.1%	100.0%	28.6%	85.7%
Spain	2.7%	97.6%	91.4%	97.3%
Sweden	2.6%	80.9%	55.8%	98.9%
Iceland	13.1%	97.6%	97.6%	100.0%
Liechtenstein	3.0%	46.7%	46.7%	100.0%
Norway	4.2%	74.7%	45.1%	97.8%
Total EU	6.2%	98.4%	95.2%	98.3%
Total EEA	6.2%	98.3%	95.0%	98.3%
Member States	% video removals under Synthetic Media policy	% proactive video removals under Synthetic Media policy	% video removals before any views under Synthetic Media policy	% video removals within 24h under Synthetic Media policy
Austria	1.0%	94.6%	78.4%	73.9%
Belgium	1.0%	96.2%	56.0%	53.3%
Bulgaria	0.6%	96.7%	30.0%	30.0%
Cractic	 4.00/	00.40/	22.7%	40.9%
Croatia	1.2%	86.4%	22.1%	40.9%



Czech Republic	0.0%	92.7%	48.8%	48.8%
Denmark	0.7%	96.9%	76.0%	82.3%
Estonia	0.9%	90.9%	27.3%	18.2%
Finland	0.0%	100.0%	60.0%	43.3%
France	0.8%	94.0%	31.9%	33.4%
Germany	1.0%	93.9%	60.1%	57.7%
Greece	0.8%	98.4%	44.3%	32.8%
Hungary	0.3%	81.8%	72.7%	45.5%
Ireland	1.8%	89.2%	51.4%	45.0%
Italy	0.7%	97.0%	24.8%	13.9%
Latvia	0.3%	96.9%	56.3%	65.6%
Lithuania	1.4%	72.7%	33.3%	24.2%
Luxembourg	1.5%	100.0%	69.6%	82.6%
Malta	2.3%	78.6%	14.3%	50.0%
Netherlands	0.7%	92.7%	46.1%	41.0%
Poland	0.5%	90.1%	66.9%	44.6%
Portugal	1.6%	92.8%	59.8%	30.9%
Romania	0.5%	93.4%	38.7%	25.5%
Slovakia	1.6%	84.2%	34.2%	26.3%



Slovenia	2.2%	100.0%	64.3%	71.4%
Spain	0.5%	94.4%	20.7%	9.3%
Sweden	0.4%	97.1%	23.5%	23.5%
Iceland	6.6%	97.6%	66.7%	81.0%
Liechtenstein	1.0%	40.0%	0.0%	100.0%
Norway	2.2%	95.1%	69.4%	72.2%
Total EU	0.5%	94.0%	46.3%	42.0%
Total EEA	0.5%	94.0%	47.2%	43.3%

	Methodology of data measurement:			
	The number of videos tagged with the unverified content label is based on the country in which the video was posted.			
SLI 21.1.2 - actions taken under measure 21.1	The share cancel rate (%) following the unverified content label share warning pop-up indicates the percentage of users who do not share a video after seeing the label pop up. This metric is based on the approximate location of the users that engaged with these tools.			
		Number of labels applied to content, such as on the basis of such articles		Meaningful metrics such as the impact of 21.1. measures on user interactions with, or user re-shares of, content fact-checked as false or misleading
List actions per member states and languages (see example table above)		Number of videos tagged with the		Share cancel rate (%) following the unverified



	unverified content label	content label share warning pop-up (users who do not share the video after seeing the pop up)
Member States		
Austria	3,459	28.4%
Belgium	5,793	25.5%
Bulgaria	1,323	26.6%
Croatia	1,024	27.5%
Cyprus	1,033	25.0%
Czech Republic	1,116	24.3%
Denmark	5,464	21.1%
Estonia	665	20.8%
Finland	2,194	24.2%
France	9,960	31.8%
Germany	10,571	30.9%
Greece	1,874	27.0%
Hungary	1,305	35.2%
Ireland	2,783	20.8%
Italy	5,382	32.4%
Latvia	777	33.3%



Lithuania	 817	 25.7%
Luxembourg	1,101	35.2%
Malta	619	30.2%
Netherlands	5,570	20.3%
Poland	2,364	32.0%
Portugal	1,560	27.2%
Romania	2,732	27.7%
Slovakia	775	29.0%
Slovenia	873	20.8%
Spain	5,487	32.7%
Sweden	5,526	24.5%
Iceland	507	26.3%
Liechtenstein	105	13.3%
Norway	 3,128	 22.6%
Total EU	82,147	29.7%
Total EEA	85,887	29.6%

Measure 21.2	

QRE 21.2.1	We undertook consultations with researchers from the Massachusetts Institute of Technology (US) and the University of Regina (Canada) in order to understand the effect of showing users accurate information after they interacted with a misinformation claim.  We continue exploring ways in which we can leverage this research to consider how we engage with users who have interacted with harmful misinformation on our platform and direct them to authoritative information.
Measure 21.3	
QRE 21.3.1	As set out within our response to QRE 17.1.1, we apply our unverified content, state-controlled media labels, and Al-generated labels to certain content in order to empower our community by providing them with an additional layer of context. We ensure these labels are developed and deployed in line with scientific evidence by partnering with fact-checkers and working with external experts, including scientists, in the following ways:  • Unverified content label. In 2021, we partnered with behavioural scientists, Irrational Labs, on the design and testing of the specialised prompts which encourage users to consider content which has been labelled as unverified, before sharing it, as detailed in QRE 17.1.1. On testing the prompts, Irrational Labs found that viewers decreased the rate at which they shared videos by 24%, while likes on such unsubstantiated content also decreased by 7%. Their full report can be found here.  As mentioned above, we partner with a number of IFCN accredited fact-checkers in Europe, who assist with assessing the accuracy of certain content on our platform. Where our fact-checking partners determine that a video is not able to be confirmed or their fact-checks are inconclusive (which is sometimes the case, particularly during unfolding events or emergencies), we may apply our unverified content label to the video.  • State-controlled media label. Since January 2023,, we have been applying state-controlled media labels to accounts or content where there is evidence of clear editorial control and decision-making by members of the state. To inform our state-affiliated media policy, including the updates set out in this report, and our approach to making such designations, we consult with media experts, political scientists, academics, and representatives from international organisations and civil society across North and South America, Africa, Europe, the Middle East, Asia, and Australia. We will



continue to work with these experts to inform our global approach and expansion of the policy.
We worked closely with Irrational Labs on the development of the state-affiliated media policy and the ways in which we could present the label to our users. We tested various copy options across English, Spanish, and Arabic via quantitative surveys and qualitative panels, and found that "[country] state-controlled media" was the option most preferred by users while being the most accurate representation of the relevant media entities' relationship to their respective governments.
<ul> <li>Al-generated content label. In advance of launching our new Al-generated labels for creators to disclose content that is completely Al-generated or significantly edited by Al, we consulted with our Safety Advisory Councils as well as industry experts including MIT's Dr. David G. Rand, who is studying how viewers perceive different types of Al labels. Dr. Rand's research helped guide the design of our Al-generated labels.</li> </ul>

# V. Empowering Users

## Commitment 22

Relevant Signatories commit to provide users with tools to help them make more informed decisions when they encounter online information that may be false or misleading, and to facilitate user access to tools and information to assess the trustworthiness of information sources, such as indicators of trustworthiness for informed online navigation, particularly relating to societal issues or debates of general interest.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]

No

		-		
A. T.			2.00	

If yes, list these implementation measures here [short bullet points].	TikTok did not subscribe to this commitment.
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.
Measure 22.1	
QRE 22.1.1	Not committed. TikTok considers that the implementation of the other commitments under the Code (including, but not limited to, the other commitments relating to empowering users and fact-checking) provide a comprehensive approach to tackling disinformation, including facilitating users in making more informed decisions when they encounter online information that may be false or misleading. This position is also reflected in the Commission's guidance. As such, TikTok does not consider that committing to this measure will materially add to the other measures being adopted under the Code.
	TikTok does, however, keep an open mind. As such, it is prepared to explore such solutions with the relevant providers in due course and to reassess whether such an approach would be beneficial, taking into account existing measures in place.
SLI 22.1.1 - actions enforcing policies above	N/A
	N/A

Measure 22.2	

	Щ	Ø		<b>O</b>	09)	

QRE 22.2.1	Not committed. TikTok does not consider this to be a practical or implementable proposal. In any event, this measure is unnecessary as the combination of other commitments underpinning the Code (including, but not limited to, those relating to fact-checking) represent a comprehensive approach to achieving the goal of providing users with tools to make more informed decisions when they encounter online information that may be false or misleading.
Measure 22.3	
QRE 22.3.1	Not committed. Not relevant, linked to the above Measures.
Measure 22.4	
QRE 22.4.1	<b>Not committed.</b> Measures 22.4, 22.5 and 22.6 are not applicable as TikTok is not a provider of trustworthiness indicators.
SLI 22.4.1 - actions enforcing policies above	N/A
	N/A
Data	<b>Not committed.</b> TikTok does not consider this to be a practical or implementable proposal. In any event, this measure is unnecessary as the combination of other commitments underpinning the Code (including, but not limited to, those relating to fact-checking) represent a comprehensive approach to achieving the goal of providing users with tools to make more informed decisions when they encounter online information that may be false or misleading.
Measure 22.5	
QRE 22.5.1	Not committed. Not relevant, linked to the above Measures.
SLI 22.5.1 - actions enforcing policies above	



<b>Not committed.</b> Measures 22.4, 22.5 and 22.6 are not applicable as TikTok is not a provider of trustworthiness indicators.
trustwortniness indicators.

SLI 22.5.2 - actions enforcing policies above	N/A		
	N/A		
Data			
Measure 22.6			
QRE 22.6.1	<b>Not committed.</b> Measures 22.4, 22.5 and 22.6 are not applicable as TikTok is not a provider of trustworthiness indicators.		
SLI 22.6.1 - actions enforcing policies above	N/A		
	N/A		
Data			
Measure 22.7			
QRE 22.7.1	As per our response to QRE 17.1.1, we have numerous tools (video notice tags, search interventions, public service announcements, in-app information centres and Safety Center pages) that lead users to authoritative sources available in all EU member states and in 23 official EU languages (plus, for EEA users, Norwegian and Icelandic). We also run localised campaigns on specific topics which deploy different engagement techniques depending on the subject matter and / or member state involved, e.g., in-person workshops, radio and newspaper campaigns.		



SLI 22.7.1 - actions enforcing policies above	N/A					

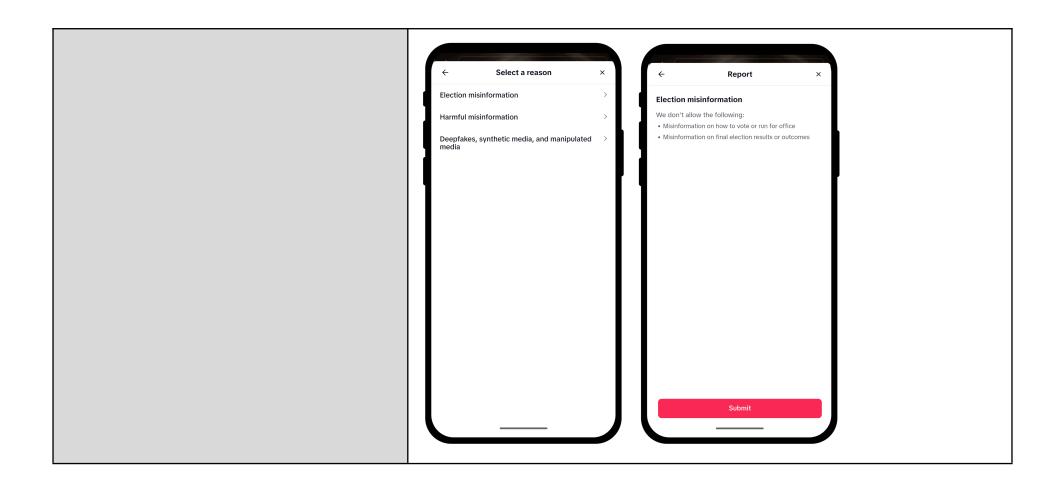
# V. Empowering Users Commitment 23 Relevant Signatories commit to provide users with the functionality to flag harmful false and/or misleading information that violates Signatories policies or terms of service. In line with this commitment, did you deploy new Yes implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No] If yes, list these implementation measures here [short In line with our DSA requirements, we continued to provide a dedicated reporting channel. and appeals process for users who disagree with the outcome, for our community in the bullet points]. European Union to 'Report Illegal Content,' enabling users to alert us to content they believe breaches the law. Do you plan to put further implementation measures in No place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No] If yes, which further implementation measures do you We are continuously reviewing and improving our tools and processes to fight misinformation and plan to put in place in the next 6 months? disinformation and will report on any further development in the next COPD report.



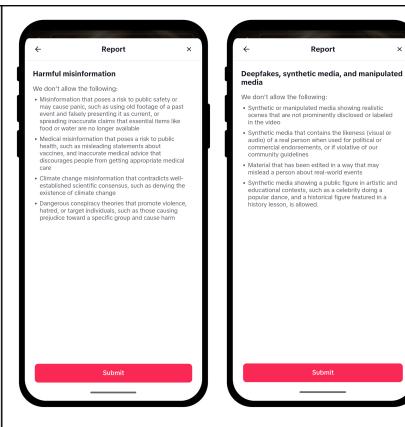
Measure 23.1	
QRE 23.1.1	We provide users with simple, intuitive ways to report/flag content in-app for any breach of our Terms of Service or CGs including for harmful misinformation in each EU Member State and in an official language of the European Union.  By 'long-pressing' (e.g., clicking for 3 seconds) on the video content and selecting the "Report" option.  By selecting the "Share" button available on the right-hand side of the video content and then selecting the "Report" option.  The user is then shown categories of reporting reasons from which to select (which align with the harms our CGs seek to address). We recently updated this feature to make the "Misinformation" categories more intuitive and allow users to report with increased granularity. We have also made changes to implement an additional option to enable users to report illegal content in line with our requirements under DSA.











Users do not need to be logged into an account on the platform to report content, and can also report video content via the TikTok website (by clicking on the "Report" button which is prominently displayed in the upper right hand corner of each video when hovering over a video) or by means of our "Report Inappropriate content" webform which is available in our <u>Support Centre</u>.

We are aware that harmful misinformation is not limited to video content and so users can also report a comment, a suggested search, a hashtag, a sound or an account, again specifically for harmful misinformation.



ΝЛ	00011		വവ	•
IVI	easu	пе	7.0	

#### **QRE 23.2.1**

#### Reporting system

To ensure the integrity of our reporting system, we deploy a combination of automated review and human moderation.

Videos uploaded to TikTok are initially reviewed by our automated moderation technology, which aims to identify content that violates our Community Guidelines. If a potential violation of our CGs is found, the automated review system will either pass it on to our moderation teams for further review or, if there is a high degree of confidence that the content violates our CGs, remove it automatically. Automated removal is only applied when violations are clear-cut, such as where the content contains nudity or pertains to youth safety. We are constantly working to improve the precision of our automated moderation technology so we can more effectively remove violative content at scale, while also reducing the number of incorrect removals.

To support the fair and consistent review of potentially violative content, where violations are less clear-cut, content will be passed to our human moderation teams for further review. Human moderators can take additional context and nuance into account, which cannot always be picked up by technology, and in the context of harmful misinformation, for example, our moderators have access to a repository of previously fact-checked claims to help make swift and accurate decisions and direct access to our fact-checking partners who help assess the accuracy of new content.

We have sought to make our CGs as clear and comprehensive as possible and have put in place robust Quality Assurance processes (including steps such as review of moderation cases, flows, appeals and undertaking Root Cause Analyses).

As part of our requirements under the DSA, we have introduced an <u>additional reporting channel</u> for our community in the European Union to 'Report Illegal Content,' which enables users to alert us to content they believe breaches the law. TikTok will review the content against our Community Guidelines and where a violation is detected, the content may be removed globally. If it is not removed, our illegal content moderation team will further review the content to assess whether it is unlawful in the relevant jurisdiction - this assessment is undertaken by human review. If it is, access to that content will be restricted in that country. Those who report suspected illegal content will be notified of our decision, including if we consider that the content is not illegal. Users who disagree can <u>appeal</u> those decisions using the appeals process.





We also note that whilst user reports are important, at TikTok we place considerable emphasis on proactive detection to remove violative content. We are proud that the vast majority of removed content is identified proactively before it is reported to us.

#### Appeals system.

We are transparent with users in relation to appeals. We set out the options that may be available both to the user who reported the content and the creator of the affected content, where they disagree with the decision we have taken.

The integrity of our appeals systems is reinforced by the involvement of our trained human moderators, who can take context and nuance into consideration when deciding whether content is illegal or violates our CGs.

Our moderators review all appeals raised in relation to removed videos, removed comments, and banned accounts and assess them against our policies. To ensure consistency within this process and its overall integrity, we have sought to make our policies as clear and comprehensive as possible and have put in place robust Quality Assurance processes (including steps such as auditing appeals and undertaking Root Cause Analyses).

If users who have submitted an appeal are still not satisfied with our decision, they can share feedback with us via the <a href="webform">webform</a> on TikTok.com. We continuously take user feedback into consideration to identify areas of improvement, including within the appeals process. Users may also have other legal rights in relation to decisions we make, as set out further <a href="here">here</a>.

## V. Empowering Users

#### Commitment 24

Relevant Signatories commit to inform users whose content or accounts has been subject to enforcement actions (content/accounts labelled, demoted or otherwise enforced on) taken on the basis of violation of policies relevant to this section (as outlined in Measure 18.2), and provide them with the possibility to



appeal against the enforcement action at issue and to hand undue delay where the complaint is deemed to be founded.	lle complaints in a timely, diligent, transparent, and objective manner and to reverse the action without
In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	No
If yes, list these implementation measures here [short bullet points].	<ul> <li>Continued to serve user notifications following action on a user's account or content, which includes a clear explanation about the action taken and a simple way to appeal the decision taken.</li> <li>Continued to provide additional user transparency around our appeals processes (here)</li> </ul>
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.
Measure 24.1	
QRE 24.1.1	Users in all EU member states are notified by an in-app notification in their relevant local language where the following action is taken:  • removal or otherwise restriction of access to their content;  • A ban of the account;  • restriction of their access to a feature (such as LIVE); or  • restriction of their ability to monetise.

	Such notifications are provided in near real time after action has been taken (i.e. generally within several seconds or up to a few minutes at most).					
	Where we have taken any of these decisions, an in-app inbox notification sets out the violation deemed to have taken place, along with an option for users to "disagree" and submit an appeal. Users can submit appeals within 180 days of being notified of the decision they want to appeal. Further information, including about how to appeal a report is set out <a href="https://example.com/here/new/methods/">here</a> .					
	All such appeals raised will be queued for review by our specialised human moderators so as to ensure that context is adequately taken into account in reaching a determination. Users can monitor the status and view the results of their appeal within their in-app inbox.					
	don't agree with the res	sult of their appeal. They	to share feedback with us can do so by using the in- nuously taking user feedb e appeals process.	app function which		
SLI 24.1.1 - enforcement actions	Methodology of data measurement:  The number of appeals/overturns is based on the country in which the video being appealed/overturned was posted. These numbers are only related to our Misinformation, Civic and Election Integrity and Synthetic and Manipulated Media.					
	Number of actions appealed Metrics on results of appealed Number of actions appealed					
List actions per member states and languages (see example table above)	Number of Appeals of videos removed for violation of misinformation policy  Number of Appeals of videos removed for videos removed for violation of misinformation policy  Number of Appeals success rate of videos removed for violation of misinformation policy					
Member States						
Austria	463	196	42.3%			



Belgium	1,493	958	64.2%	
Bulgaria	179	64	35.8%	
Croatia	287	166	57.8%	
Cyprus	104	62	59.6%	
Czech Republic	204	121	59.3%	
Denmark	135	93	68.9%	
Estonia	44	37	84.1%	
Finland	122	77	63.1%	
France	11,706	7,447	63.6%	
Germany	14,557	5,528	38.0%	
Greece	630	347	55.1%	
Hungary	351	257	73.2%	
Ireland	672	481	71.6%	
Italy	2,463	1,902	77.2%	
Latvia	25	15	60.0%	
Lithuania	49	33	67.3%	



Luxembourg	23	13	56.5%	
Malta	7	6	85.7%	
Netherlands	869	651	74.9%	
Poland	3,138	1,410	44.9%	
Portugal	447	247	55.3%	
Romania	1,632	1,063	65.1%	
Slovakia	123	74	60.2%	
Slovenia	55	41	74.5%	
Spain	5,484	3,221	58.7%	
Sweden	1,138	501	44.0%	
Iceland	15	9	60.0%	
Liechtenstein	0	0	0.0%	
Norway	239	177	74.1%	
Total EU	46,400	25,011	53.9%	
Total EEA	46,654	25,197	54.0%	
List actions per member states and languages (see example table above)	Number of appeals of videos removed for	Number of overturns of appeals for violation	Appeal success rate of videos removed for	



	violation of Civic and Election Integrity policy	of Civic and Election Integrity policy	violation of Civic and Election Integrity policy	
Member States				
Austria	10	10	100.0%	
Belgium	42	35	83.3%	
Bulgaria	11	5	45.5%	
Croatia	5	4	80.0%	
Cyprus	0	0	0.0%	
Czech Republic	15	13	86.7%	
Denmark	13	10	76.9%	
Estonia	0	0	0.0%	
Finland	7	3	42.9%	
France	98	82	83.7%	
Germany	310	233	75.2%	
Greece	14	12	85.7%	
Hungary	7	6	85.7%	



Ireland	22	21	95.5%	
Italy	121	103	85.1%	
Latvia	0	0	0.0%	
Lithuania	2	2	100.0%	
Luxembourg	0	0	0.0%	
Malta	0	0	0.0%	
Netherlands	62	53	85.5%	
Poland	57	40	70.2%	
Portugal	31	27	87.1%	
Romania	53	44	83.0%	
Slovakia	3	2	66.7%	
Slovenia	0	0	0.0%	
Spain	77	65	84.4%	
Sweden	36	27	75.0%	
Iceland	0	0	0.0%	
Liechtenstein	0	0	0.0%	



Norway	24	18	75.0%	
Total EU	996	797	80.0%	
Total EEA	1,020	815	79.9%	
List actions per member states and languages (see example table above)	Number of appeals of videos removed for violation of Synthetic and Manipulated Media	Number of overturns of appeals for violation of Synthetic and Manipulated Media	Appeal success rate of videos removed for violation of Synthetic and Manipulated Media	
Member States				
Austria	0	0	0.0%	
Belgium	8	8	100.0%	
Bulgaria	2	2	100.0%	
Croatia	0	0	0.0%	
Cyprus	0	0	0.0%	
Czech Republic	0	0	0.0%	
Denmark	4	4	100.0%	
Estonia	0	0	0.0%	
Finland	0	0	0.0%	
France	29	24	82.8%	



Germany	41	35	85.4%	
Greece	5	2	40.0%	
Hungary	0	0	0.0%	
Ireland	7	3	42.9%	
Italy	20	19	95.0%	
Latvia	0	0	0.0%	
Lithuania	0	0	0.0%	
Luxembourg	0	0	0.0%	
Malta	0	0	0.0%	
Netherlands	9	8	88.9%	
Poland	2	0	0.0%	
Portugal	7	7	100.0%	
Romania	5	5	100.0%	
Slovakia	6	6	100.0%	
Slovenia	3	2	66.7%	
Spain	12	11	91.7%	



Sweden	11	11	100.0%	
Iceland	0	0	0.0%	
Liechtenstein	0	0	0.0%	
Norway	3	2	66.7%	
Total EU	171	147	86.0%	
Total EEA	174	149	85.6%	

# V. Empowering Users

#### Commitment 25

In order to help users of private messaging services to identify possible disinformation disseminated through such services, Relevant Signatories that provide messaging applications commit to continue to build and implement features or initiatives that empower users to think critically about information they receive and help them to determine whether it is accurate, without any weakening of encryption and with due regard to the protection of privacy.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	No
If yes, list these implementation measures here [short bullet points].	This commitment is not applicable as TikTok is not a messaging app.
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the	No



maturity of the implementation of this commitment? [Yes/No]	
If yes, which further implementation measures do you plan to put in place in the next 6 months?	This commitment is not applicable as TikTok is not a messaging app.
Measure 25.1	
QRE 25.1.1	Not committed. This commitment is not applicable as TikTok is not a messaging app.
SLI 25.1.1	N/A
	N/A
Data	
Measure 25.2	
QRE 25.2.1	Not committed. This commitment is not applicable as TikTok is not a messaging app.
SLI 25.2.1 - use of select tools	N/A
	N/A
Data	













# VI. Empowering the research community Commitments 26 - 29



## VI. Empowering the research community

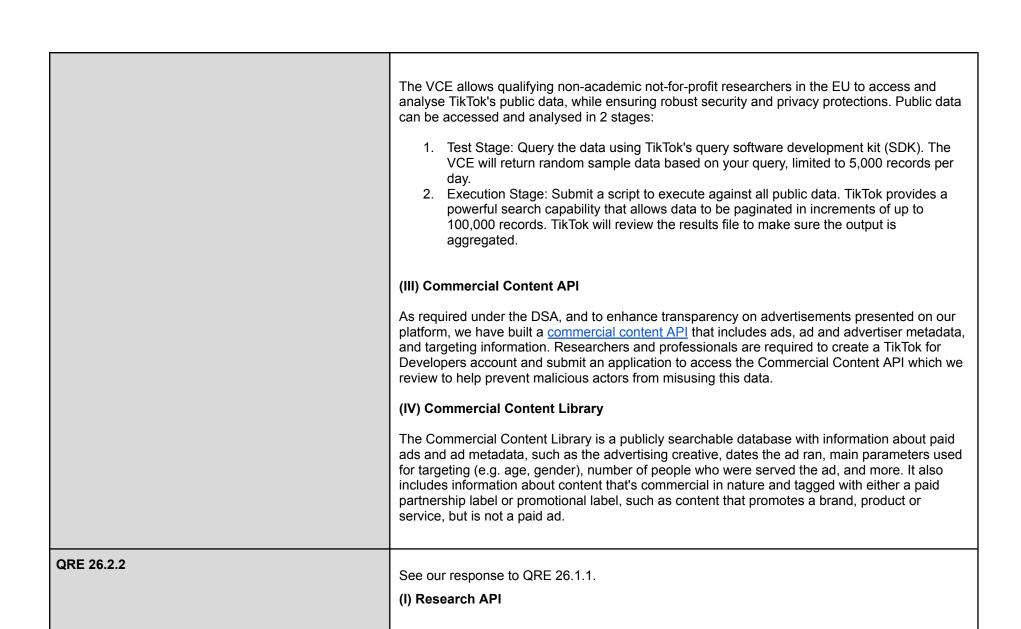
### Commitment 26

Relevant Signatories commit to provide access, wherever safe and practicable, to continuous, real-time or near real-time, searchable stable access to non-personal data and anonymised, aggregated, or manifestly-made public data for research purposes on Disinformation through automated means such as APIs or other open and accessible technical solutions allowing the analysis of said data.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	Yes
If yes, list these implementation measures here [short bullet points].	<ul> <li>Launched the new <u>Virtual Compute Environment (VCE)</u>. The VCE is a robust technical solution that offers broader access to user data to qualifying non-academic not-for-profit researchers, while ensuring robust security and privacy protections.</li> <li>Continued to make the Research API, which supports independent research through access to public data about TikTok content and accounts, available to eligible researchers in Europe.</li> <li>Continued to make the Commercial Content API available in Europe to bring transparency to paid advertising, advertisers and other commercial content on TikTok.</li> <li>Continued to offer our Commercial Content Library, a publicly searchable EU ads database with information about paid ads and ad metadata, such as the advertising creative, dates the ad ran, main parameters used for targeting (e.g. age, gender), number of people who were served the ad, and more.</li> <li>Participated in the <u>data access pilot with the European Digital Media Observatory (EDMO)</u>, which trialled the process for sharing data with vetted researchers designated under the DSA.</li> </ul>
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No

If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.
Measure 26.1	
QRE 26.1.1	<ul> <li>We have a dedicated <u>TikTok Transparency Centre</u> available in a number of EU languages which hosts our:</li> <li><u>COPD Transparency Reports</u>, as part of our commitments to the Code, we publish a transparency report every six months to provide granular data for EU/EEA countries about our efforts to combat online misinformation.</li> <li>Our <u>TikTok Community Guidelines Enforcement Reports</u>, providing proactive quarterly insights into the volume and nature of content and accounts removed from our platform for violating our Community Guidelines, Terms of Service or Advertising Policies since 2019.</li> <li><u>DSA Transparency Reports</u>, building on our proactive approach to transparency in our quarterly TikTok Community Guidelines Enforcement Reports and our obligations under the Digital Services Act ("DSA"), we publish a transparency report every six months to provide granular data for EU countries about our content moderation activities.</li> <li>As part of our commitment to regulatory transparency and accountability, we launched the <u>European Online Safety Hub</u>, which serves as a 'one-stop-shop' for our community to learn more about how we're complying with the DSA. The Hub is currently available in 22 EU languages and at least one official language of each of the EU Member States. Our dedicated <u>TikTok for Developers</u> website hosts our Research Tools and Commercial Content APIs.(detailed below).</li> </ul>
QRE 26.1.2	In our January 2024 Code report, covering the entire H2 2023 period, TikTok shared more than 3200 data points across 30 EU/EEA countries - an increase of more than 500 data points as compared to our previous report. In this latest report, we are pleased to confirm that we are continuing these efforts as we share 3300 data points.  We also provide ongoing insights into the action we take against content and accounts that violate our CGs, Terms of Service or Advertising Policies in our quarterly TikTok Community

	Guideline Enforcement Reports. The report includes a variety of data visualisations, which are designed with transparency and accessibility in mind, including for people with colour vision deficiency. We work hard to supplement the comprehensive data in the report and provide new insights. For example, we recently shared data on our response time to community-reported content, added information on TikTok LIVE enforcement, and provided more detail on how often a video has been viewed before it is removed. As part of our continued efforts to make it easy to study the TikTok platform, the report also offers access to aggregated data in a downloadable data file, which we recently expanded to include additional removal data by policy category for the 50 markets with the highest volumes of removed content.  We provide access to researchers to data that is publicly available on our platform through our Research Tools and through our Commercial Content API for commercial content (detailed below).
SLI 26.1.1 - uptake of the tools and processes	
described in Measure 26.1	
Data	
Measure 26.2	
QRE 26.2.1	(I) Research API
	To make it easier to independently research our platform and bring transparency to TikTok content, we built a Research API that provides researchers with access to public data on accounts and content, including comments, captions, subtitles, number of comments, shares, likes, followers and following lists, and favourites that a video receives on our platform. More information is available <a href="https://example.com/here">here</a> . We released an initial version of the Research API to academic researchers in the US and it has now been expanded to the EEA, UK and Switzerland. We have carefully considered feedback from researchers who have used the API and we have already





Through our Research API, qualifying academic researchers from non-profit academic institutions in the US, EEA, UK or Switzerland can apply to study public data about TikTok content and accounts. This public data includes comments, captions, subtitles, number of comments, shares, likes, followers and following lists, and favourites that a video receives on our platform. More information is available here.

#### (II) Virtual Compute Environment (VCE)

Through our VCE, qualifying non-academic not-for-profit researchers in the EU can query and analyse TikTok's public data. To protect the security and privacy of our users the VCE is designed to ensure that TikTok data is processed within confined parameters. TikTok only reviews the results to ensure that there is no identifiable individual information extracted out of the platform. All aggregated results will be shared as a downloadable link to the approved primary researcher's email.

#### (III) Commercial Content API

Through our Commercial Content API, qualifying researchers and professionals, who can be located in any country,can request public data about commercial content including ads, ad and advertiser metadata, and targeting information. To date, the Commercial Content API only includes data from EU countries.

#### (IV) Commercial Content Library

TikTok's <u>Commercial Content Library</u> is a repository of ads and other types of commercial content posted to users in the European Economic Area (EEA), Switzerland, and the UK only, but can be accessed by members of the public located in any country. Each ad and ad details will be available in the library for one year after the advertisement was last viewed by any user. Through theCommercial Content Library, the public can access information about paid ads and ad metadata, such as the advertising creative, dates the ad ran, main parameters used for targeting (e.g. age, gender), number of people who were served the ad, and more. It also includes information about content that is commercial in nature and tagged with either a paid partnership label or promotional label, such as content that promotes a brand, product or service, but is not a paid ad.



QRE 26.2.3	We make detailed information available to applicants about our Research Tools (Research API and VCE) and Commercial Content API, through our dedicated TikTok for Developers website, including on what data is made available and how to apply for access.  Once an application has been approved for access to our Research Tools, we provide step-by-step instructions for researchers on how to access research data, how to comply with the security steps, and how to run queries on the data.  Similarly with the Commercial Content API, we provide participants with detailed information on how to query ad data and fetch public advertiser data.					
SLI 26.2.1 - meaningful metrics on the uptake, swiftness, and acceptance level of the tools and processes in Measure 26.2	Research Tools, Commercial Content API, and the Commercial Content Library During this reporting period, we received 229 applications to access TikTok's Research Tools (Research API and VCE) from researchers in the EU and EEA.  We received 53 applications to access the TikTok Commercial Content API.					
	Number of Applications received for Researchers API  Number of Number of Applications received for Researchers API  Number of Applications received for Researchers API  Number of Applications received for TikTok Commercial Content API  Number of Applications received for TikTok Commercial Content API  Number of Applications received for TikTok Commercial Content API  Number of Applications received for TikTok Commercial Content API					
Austria	10	4	4	2	1	1
Belgium	3	3	0	0	0	0
Bulgaria	1	1	0	0	0	0
Croatia	4	1	2	0	0	0
Cyprus	0	0	0	0	0	0
Czech Republic	3	1	1	4	4	0



Denmark	1	0	0	0	0	0
Estonia	0	0	0	0	0	0
Finland	4	3	0	1	1	0
France	33	7	20	9	6	2
Germany	48	23	17	8	6	2
Greece	2	1	1	0	0	0
Hungary	1	0	0	0	0	0
Ireland	7	3	0	2	1	0
Italy	26	15	9	2	2	0
Latvia	0	0	0	0	0	0
Lithuania	0	0	0	0	0	0
Luxembourg	2	1	1	0	0	0
Malta	0	0	0	0	0	0
Netherlands	24	7	8	6	5	1
Poland	5	3	2	5	5	0
Portugal	4	1	3	2	2	0
Romania	5	2	3	3	2	1
Slovakia	1	1	0	0	0	0



Slovenia	0	0	0	0	0	0
Spain	32	11	19	8	7	0
Sweden	8	4	2	1	1	0
Iceland	0	0	0	0	0	0
Lichtenstein	0	0	0	0	0	0
Norway	5	2	3	0	0	0
EU Level	224	92	92	53	43	7
EEA Level	229	94	95	53	43	7
Measure 26.3						
QRE 26.3.1	We welcome feedback from researchers on our APIs and have a <u>dedicated support form</u> where researchers can provide feedback about their experience. On foot of recent feedback, we added a number of new query categories, e.g., followers and following lists, and number of video likes. Prior to expanding the Research API to Europe, we <u>acted on feedback</u> from US based researchers by streamlining the application process and enabling greater collaboration through Lab Access.					

## VI. Empowering the research community

## Commitment 27

Relevant Signatories commit to provide vetted researchers with access to data necessary to undertake research on Disinformation by developing, funding, and cooperating with an independent, third-party body that can vet researchers and research proposals.

	Ø		<b>Q</b>	241	

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	Yes	
If yes, list these implementation measures here [short bullet points].	<ul> <li>Completed the <u>data access pilot with EDMO</u>, which trialled the process for sharing data with vetted researchers designated under the DSA.</li> <li>Actively participating in the EDMO working group for the creation of the Independent Intermediary Body (IIB) to support research on digital platforms.</li> <li>Refine our standard operating procedure (SOP) for vetted researcher access to ensure compliance with the provisions of the Delegated Act on Data Access for Research.</li> </ul>	
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No	
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.	
Measure 27.1		
QRE 27.1.1	We have engaged with EDMO and are actively participating in the working group that has been set up in order to implement the Independent Intermediary Body (IIB) that is referred to above.  TikTok was one of the platforms to complete EDMO's data access pilot, trialling the process for sharing data with vetted researchers designated under the DSA. The pilot is aimed at informing the future processes around granting access to data under Article 40(4) of the DSA and we look forward to the publication of the Delegated Act on Data Access for Research.	
Measure 27.2		



QRE 27.2.1	We continue to participate in the working group which has been set up to implement the Independent Intermediary Body (IIB).
Measure 27.3	
QRE 27.3.1	We are participating in the working group which has been set up to implement the Independent Intermediary Body (IIB). TikTok completed the data access pilot with EDMO trialling the process for sharing data with vetted researchers designated under the DSA.
SLI 27.3.1 - research projects vetted by the independent third-party body	N/A
	N/A
Data	
Measure 27.4	
QRE 27.4.1	We completed the data access pilot with EDMO trialling the process for sharing data with vetted researchers designated under the DSA.

VI. Empowering the research community	
Commitment 28	
Relevant Signatories commit to support good faith research into Disinformation that involves their services.	
In line with this commitment, did you deploy new	Yes
implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	

If yes, list these implementation measures here [short bullet points].	<ul> <li>Expanded our Research Tools with the new VCE, a robust technical solution that offers broader access to user data to qualifying non-academic not-for-profit researchers, while ensuring robust security and privacy protections.</li> <li>Continued to make the Research API, which supports independent research through access to public data about TikTok content and accounts, available to researchers in Europe.</li> <li>Continued to bring transparency to paid advertising, advertisers and other commercial content in the EU on TikTok by making the Commercial Content API available to qualifying researchers and professionals, who can be located in any country.</li> <li>Continued to offer open access to the Commercial Content Library, a publicly searchable ads database with information about paid ads and ad metadata, such as the advertising creative, dates the ad ran, main parameters used for targeting (e.g. age, gender), number of people who were served the ad, and more.</li> <li>Completed the data access pilot with EDMO trialling the process for sharing data with vetted researchers as designated under the DSA.</li> </ul>
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.
Measure 28.1	
QRE 28.1.1	TikTok is committed to facilitating research and engaging with the research community.  As set out above, TikTok is committed to facilitating research through our Research Tools,Commercial Content APIs and Commercial Content Library, full details of which are available on our TikTok for Developers website and Commercial Content Library website. We have many teams and individuals across product, policy, data science, outreach and legal



committing time and focus to facilitating research. At TikTok, we believe that accountability and transparency are essential to fostering trust with our community. We're committed to leading the way when it comes to being transparent in how we operate, moderate and recommend content, empower users, and secure our platform. That's why we opened our global Transparency and Accountability Centers (TACs) for invited guests to see first-hand our work to protect the safety and security of the TikTok platform. Guests can also see how teams at TikTok go about the critically important work of securing our community's safety, data, and privacy. Our TACs are located in Dublin, Los Angeles, Singapore, and Washington, DC. In February 2024, we welcomed TikTok's European Safety Advisory Council (ESAC) to the Dublin TAC.

We work closely with our nine regional <u>Safety Advisory Councils</u>, including our European Safety Advisory Council, US Content Advisory Council and new global Youth Advisory Council, which bring together a diverse array of independent experts from academia and civil society and youth perspectives. Advisory Council members provide subject matter expertise and advice on issues relating to user safety, content policy, and emerging issues that affect TikTok and our community, including recently in the development of our <u>Al-generated label</u> for creators to disclose content that is completely Al-generated or significantly edited by Al. These councils are an important way to bring outside perspectives into our company and onto our platform.

In addition to these efforts, there are a plethora of ways through which we engage with the research community in the course of our work.

Our **Outreach & Partnerships Management Team** is dedicated to establishing partnerships and regularly engaging with civil society stakeholders and external experts, including the academic and research community, to ensure their perspectives inform our policy creation, feature development, risk mitigation, and safety strategies. For example, we engaged with numerous academics in Europe as part of the development and implementation of our state-affiliated media policy. We also worked closely with industry experts on our Al-generated content label, including MIT's Dr. David G. Rand, who is studying how viewers perceive different types of Al labels. Dr. Rand's research helped guide the design of our Al-generated labels.

More recently, we engaged with global experts on our Election Misinformation policies, which helped inform recent updates to our I&A policies. This team also plays an important role in our efforts to counter misinformation by identifying, onboarding and managing new partners to our fact-checking programme, for example, In the lead-up to certain elections, we invite suitably qualified external local/regional experts, as part of our Election Speaker Series, to share their market expertise with our internal teams to provide us with insights to better understand areas



that could potentially amount to election manipulation and to inform our approach to the upcoming election. For example, our Election Speaker Series in advance of the EU Parliamentary Elections in June 2024 heard from the following organisations:

German Press Agency (dpa): Austria, Germany & the Netherlands
 Agence France-Presse (AFP): France, Belgium, Greece & Cyprus

Facta: ItalyNewtral: SpainPoligrafo: Portugal

• Faktograf: Croatia & Slovenia

• Demagog.pl: Poland

Funky Citizens: Romania & BulgariaDemagog.cz: Czechia & Slovakia

LeadStories: HungaryThe Journal: Ireland

Delfi.lt: Latvia, Lithuania & Estonia

• Logically Facts: Sweden, Denmark & Finland

TikTok teams and personnel also regularly participate in **research-focused events**. At the end of June 2024, TikTok sent a 12 strong delegation to <u>GlobalFact 11</u> in Sarajevo, Bosnia and Herzegovina. TikTok was one of three top-tier sponsors of GlobalFact 11, the <u>International Fact-Checking Network</u>'s largest gathering for professional fact-checkers. In addition to sponsorship, TikTok participated in an on-the-record mainstage presentation answering questions about our misinformation strategy and partnerships with professional fact-checkers. We also had meetings with many existing and potentially new partners as well as the EFCSN.During the reporting period, TikTok also presented externally on our Research Tools and transparency approach including a <u>virtual webinar hosted by EDMO</u> in April, the <u>Palestine Digital Activism Forum</u>, hosted by 7amleh, member of European Digital Rights (EDRi), in June, the <u>Perspektywy Women in Tech Summit</u> in Warsaw in June, and the <u>Terrorism and Social Media Conference</u> held in Swansea in June. These engagements reached approximately 135 research stakeholders.

As well as opportunities to share context about our approach, research interests, and opportunities to collaborate, these events enable us to learn from the important work being done

Measure 28.3

**QRE 28.3.1** 

Measure 28.4

pinned videos and reposted videos.  • Public content data, such as comments, captions, subtitles, and number of comme shares and likes that a video receives.  Through the VCE, qualifying non-academic not-for-profit researchers in the EU can access analyse TikTok's public data in a secure environment that is subject to strict security contro  Our commercial content related APIs includes ads, ad and advertiser metadata, and target information. These APIs will allow the public and researchers to perform customised - advertiser metadata.		
We have a dedicated TikTok for Developers website which hosts our Research Tools and Commercial Content APIs. With the Research API, researchers can access:  Public account data, such as user profiles, followers and following lists, liked video pinned videos and reposted videos. Public content data, such as comments, captions, subtitles, and number of comme shares and likes that a video receives.  Through the VCE, qualifying non-academic not-for-profit researchers in the EU can access analyse TikTok's public data in a secure environment that is subject to strict security control.  Our commercial content related APIs includes ads, ad and advertiser metadata, and target information. These APIs will allow the public and researchers to perform customised - advertiser.		
Commercial Content APIs. With the Research API, researchers can access:  Public account data, such as user profiles, followers and following lists, liked video pinned videos and reposted videos.  Public content data, such as comments, captions, subtitles, and number of comme shares and likes that a video receives.  Through the VCE, qualifying non-academic not-for-profit researchers in the EU can access analyse TikTok's public data in a secure environment that is subject to strict security contro  Our commercial content related APIs includes ads, ad and advertiser metadata, and target information. These APIs will allow the public and researchers to perform customised - advertiser metadata.	Measure 28.2	
the Commercial Content Library repository. The <u>Library</u> is a searchable database with information about paid ads and ad metadata, such as the advertising creative, dates the acceptance of the commercial Content Library repository. The <u>Library</u> is a searchable database with	QRE 28.2.1	<ul> <li>Commercial Content APIs.</li> <li>With the Research API, researchers can access:</li> <li>Public account data, such as user profiles, followers and following lists, liked videos, pinned videos and reposted videos.</li> <li>Public content data, such as comments, captions, subtitles, and number of comments, shares and likes that a video receives.</li> <li>Through the VCE, qualifying non-academic not-for-profit researchers in the EU can access and analyse TikTok's public data in a secure environment that is subject to strict security controls.</li> <li>Our commercial content related APIs includes ads, ad and advertiser metadata, and targeting information. These APIs will allow the public and researchers to perform customised - advertiser name or keyword based - searches on ads and other commercial content data that is stored in the Commercial Content Library repository. The Library is a searchable database with information about paid ads and ad metadata, such as the advertising creative, dates the ad ran, main parameters used for targeting (e.g. age, gender), number of people who were served the</li> </ul>

TikTok's CoPD Report January to June 2024 221

our platform.

We are eager to engage and cooperate with EDMO as part of this annual consultation and to receive feedback from the research community on how we facilitate their research projects on



QRE 28.4.1	We are committed to regularly engage and collaborate with EDMO including our participation in the pilot trialling the process for sharing data with vetted researchers designated under the DSA and on the creation of the IIB.

# VI. Empowering the research community

### Commitment 29

Relevant Signatories commit to conduct research based on transparent methodology and ethical standards, as well as to share datasets, research findings and methodologies with relevant audiences.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	No
If yes, list these implementation measures here [short bullet points].	This Commitment relates to Research Organisations.
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	This Commitment relates to Research Organisations.
Measure 29.1	
QRE 29.1.1	Not committed.



QRE 29.1.2	Not committed.
QRE 29.1.3	Not committed.
SLI 29.1.1 - reach of stakeholders or citizens	N/A
informed about the outcome of research projects	N/A
Data	
Measure 29.2	
QRE 29.2.1	Not committed.
QRE 29.2.2	Not committed.
QRE 29.2.3	N/A
	N/A
SLI 29.2.1	Not committed.
Data	
Measure 29.3	
QRE 29.3.1	Not committed.





	N/A
SLI 29.3.1 - reach of stakeholders or citizens informed about the outcome of research projects	N/A
Data	





# VII. Empowering the fact-checking community Commitments 30 - 33



### VII. Empowering the fact-checking community

### Commitment 30

Relevant Signatories commit to establish a framework for transparent, structured, open, financially sustainable, and non-discriminatory cooperation between them and the EU fact-checking community regarding resources and support made available to fact-checkers

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	Yes
If yes, list these implementation measures here [short bullet points].	<ul> <li>Ensured fact-checking coverage in at least one official language of each of the EU Member States ahead of the EU Parliamentary Election in June 2024.</li> <li>Onboarded a new fact-checking partner: Faktograf, the first Croatian media specialised in fact-checking.</li> <li>Sponsored the Global Fact 11   Global Fact-Checking Conference in Sarajevo, Bosnia and Herzegovina in June 2024. In addition to top tier sponsorship of the International Fact-Checking Network's largest gathering for professional fact-checkers, TikTok participated in an on-the-record mainstage presentation answering questions about our misinformation strategy and partnerships with professional fact-checkers. We also had meetings with many existing and potentially new partners as well as the EFCSN.</li> <li>Continued to expand our fact-checking repository to ensure our teams and systems leverage the full scope of insights our fact-checking partners submitted to TikTok.</li> <li>Continued to explore ways to improve data sharing in connection with our pilot scheme to share enforcement data with our fact-checking partners on the claims they have provided feedback on.</li> <li>Continued to conduct feedback sessions with our partners to further enhance the efficiency of the fact-checking program. Kept participating in the working group within the Code framework on the creation of an external fact-checking repository.</li> </ul>
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No

If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.
Measure 30.1	and distribution and will report on any further development in the flext COPD report.
QRE 30.1.1	Within Europe, we work with 12 fact-checking partners who provide fact-checking coverage in 23 EEA languages, including at least one official language of every EU Member State. Our partners have teams of fact-checkers who review and verify reported content. Our moderators then use that independent feedback to take action and where appropriate, remove or make ineligible for recommendation false or misleading content or label unverified content.  Our agreements with our partners are standardised, meaning the agreements are based on our template master services agreements and consistent of common standards and conditions. We reviewed and updated our template standard agreements as part of our annual contract renewal process.  The terms of the agreements describe:  • The service the fact-checking partner will provide, namely that their team of fact checkers review, assess and rate video content uploaded to their fact-checking queue.  • The expected results e.g., the fact-checkers advise on whether the content may be or contain misinformation and rate it using our classification categories.  • An option to agree that our fact-checker partners provide regular written reports about disinformation trends identified.  • An option to receive pro-actively flagging of potential harmful misinformation from our partners.  • The languages in which they will provide fact-checking services.  • The ability to request temporary coverage regarding additional languages or support on ad hoc additional projects.
QRE 30.1.2	Within Europe, our IFCN-accredited fact-checking partners are:



- 1. Agence France-Presse (AFP)
- 2. dpa Deutsche Presse-Agentur
- 3. Demagog
- 4. Facta
- 5. Faktograf
- 6. Lead Stories
- 7. Logically Facts
- 8. Newtral
- 9. Poligrafo
- 10. Reuters
- 11. Science Feedback
- 12. Teyit

We can, and have, put in place temporary agreements with these fact-checking partners to provide additional EU language coverage during high risk events like elections or an unfolding crisis. For example, we temporarily expanded our fact-checking coverage to Maltese for the EU Parliamentary Election of June 2024.

Outside of our fact-checking program, we also collaborate with a variety of fact-checking partners to develop media literacy campaigns and work on other discrete projects. We have collaborated with Agence France-Presse (AFP), dpa Deutsche Presse-Agentur (DPA), Demagog.pl, Demagog.cz, Correctiv, Facta, Faktograf, FakeNews.pl., Logically Facts, Newtral, Maldita, Poligrafo, Delfi.lt, The Journal, Nieuwscheckers, Funky Citizens, DigiQ and Ostro on localised media literacy campaigns. For example:

- **2024 Finnish Presidential Election.** We worked with fact-checking organisation <u>Logically Facts</u> to create videos to be displayed in our in-app Finnish Election Centre, which also included information from media literacy association <u>Mediataitokoulu</u>.
- 2024 Irish Referendum. We worked with fact-checking organisations The Journal and Logically Facts to create videos to be displayed in our in-app Irish Referendum Centre, which also included information from media literacy association Media Literacy Ireland.
- 2024 Slovak Presidential Election. We worked with media literacy organisation DigiQ to create videos about fighting fake news on our in-app Slovak Election Centre.
- 2024 French Parliamentary Election. We worked with fact-checking organisation Agence France-Presse (AFP) to create videos about fighting fake news in our in-app French Election Centre.

	<ul> <li>2024 EU Parliamentary Election: We worked with our trusted fact-checking partners and local media literacy bodies to create a media literacy section in our in-app EU Election Centre (for every Member State). These partners and organisations included Agence France-Presse (AFP), Deutsche Presse-Agentur (dpa), Demagog.pl, Demagog.cz, Facta, Faktograf, Logically Facts, Newtral, Poligrafo, Delfi.lt, The Journal, Nieuwscheckers, Funky Citizens, DigiQ and Ostro.</li> <li>We rolled out two general media literacy and critical thinking skills campaigns with Logically Facts in Denmark and Nieuwschecker in the Netherlands during this reporting period.</li> <li>Globally, we have 19 IFCN-accredited fact-checking partners. We are continuously working to expand our fact-checking network and we keep users updated here.</li> </ul>
QRE 30.1.3	We have fact-checking coverage in 23 official EEA languages: Bulgarian, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Italian, Latvian, Lithuanian, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish, Swedish, which are spoken languages of 29 EEA countries.  We have fact-checking coverage in a number of other European languages or languages which affect European users, including Russian and Ukrainian and we can request additional support in Azeri, Armenian, Turkish, and Belarusian.  In terms of global fact-checking initiatives, we currently cover more than 50 languages and assess content in more than 100 countries, thereby improving the overall integrity of the service and benefiting European users.  In order to effectively scale the feedback provided by our fact-checkers globally, we have implemented the measures listed below.  • Fact-checking repository. We have built a repository of previously fact-checked claims to help misinformation moderators make swift and accurate decisions.  • Trends reports. Our fact-checking partners can provide us with regular reports identifying general misinformation trends observed on our platform and across the industry generally, including new/changing industry or market trends, events or topics that generated particular misinformation or disinformation.  • Proactive detection by our fact-checking partners. Our fact-checking partners are authorised to proactively identify content that may constitute harmful misinformation on



our platform and suggest prominent misinformation that is circulating online that may benefit from verification.

- Fact-checking guidelines. We create guidelines and trending topic reminders for our
  moderators on the basis of previous fact-checking assessments. This ensures our
  moderation teams leverage the insights from our fact-checking partners and helps our
  moderators make swift and accurate decisions on flagged content regardless of the
  language in which the original claim was made.
- Election Speaker Series. To further promote election integrity, and inform our approach to country-level EU elections, we invited suitably qualified local and regional external experts to share their insights and market expertise with our internal teams. Our recent Election Speaker Series heard presentations from the following organisations:
  - o German Press Agency (dpa): Austria, Germany & The Netherlands
  - o Agence France-Presse (AFP): France, Belgium, Greece & Cyprus
  - Facta: ItalyNewtral: Spain
  - o Poligrafo: Portugal
  - o Faktograf: Croatia & Slovenia
  - o Demagog.pl: *Poland*
  - Funky Citizens: Romania & BulgariaDemagog.cz: Czechia & Slovakia
  - LeadStories: HungaryThe Journal: Ireland
  - o Delfi.lt: Latvia, Lithuania & Estonia
  - Logically Facts: Sweden, Denmark & Finland

Members of moderation teams receive specialised training on misinformation and have direct access to these tools and measures, which enables them to more accurately take action on violating content across Europe and globally.

We are continuing to invest in building, and improving, models which may allow for the output of these measures to be used to update the machine learning models we use in proactive detection, learning, over time, to search for similar content which can be proactively recalled into our moderation system for review. We use a variety of automated tools, including:

• Computer Vision models, which help to detect objects so it can be determined whether the content likely contains material which violates our policies.

	<ul> <li>Keyword lists and models are used to review text and audio content to detect material in violation of our policies. We work with various external experts, including our fact-checking partners, to inform our keyword lists.</li> <li>Where we have previously detected content that violates our policies, we use de-duplication and hashing technologies that enable us to recognise copies or near copies of such content to prevent further re-distribution of violative content on our platform.</li> <li>We launched the ability to read Content Credentials that attach metadata to content, which we can use to automatically label Al-generated content that originated on other major platforms.</li> <li>Continuing to leverage the fact-checking output in this way enables us to further increase the positive impact of our fact checking programme.</li> </ul>
SLI 30.1.1 - Member States and languages covered by agreements with the fact-checking organisations	
Austria	Fact-checking coverage implemented
Belgium	Fact-checking coverage implemented
Bulgaria	Fact-checking coverage implemented
Croatia	Fact-checking coverage implemented
Cyprus	Fact-checking coverage implemented
Czech Republic	Fact-checking coverage implemented
Denmark	Fact-checking coverage implemented
Estonia	Fact-checking coverage implemented



Finland	Fact-checking coverage implemented
France	Fact-checking coverage implemented
Germany	Fact-checking coverage implemented
Greece	Fact-checking coverage implemented
Hungary	Fact-checking coverage implemented
Ireland	Fact-checking coverage implemented
Italy	Fact-checking coverage implemented
Latvia	Fact-checking coverage implemented
Lithuania	Fact-checking coverage implemented
Luxembourg	Fact-checking coverage implemented
Malta	Temporary Fact-checking coverage during the 2024 EU Parliamentary Elections. We are working hard to continue to expand our fact-checking programme. Meanwhile, our fact-checking repository and other initiatives benefit all European users and ensure the overall integrity of our platform.
Netherlands	Fact-checking coverage implemented
Poland	Fact-checking coverage implemented
Portugal	Fact-checking coverage implemented
Romania	Fact-checking coverage implemented



Slovakia	Fact-checking coverage implemented
Slovenia	Fact-checking coverage implemented
Spain	Fact-checking coverage implemented
Sweden	Fact-checking coverage implemented
Iceland	We are working hard to continue to expand our fact-checking programme. Meanwhile, our fact-checking repository and other initiatives benefit all European users and ensure the overall integrity of our platform.
Liechtenstein	Fact-checking coverage implemented
Norway	Fact-checking coverage implemented
Total EU	22 languages
Total EEA	23 languages

Measure 30.2	
QRE 30.2.1	Our agreements with our fact-checking partners are standardised, meaning the agreements are based on our template master services agreements and consistent of common standards and conditions. These agreements, as with all of our agreements, must meet the ethical and professional standards we set internally including containing anti-bribery and corruption provisions.  Our partners are compensated in a fair, transparent way based on the work done by them using standardised rates. Our fact-checking partners then invoice us on a monthly basis based on work done.

				<b>O</b>	29, 2
--	--	--	--	----------	-------

	All of our fact-checking partners are independent organisations, which are certified through the non-partisan IFCN. Our agreements with them explicitly state that the fact-checkers are non-exclusive, independent contractors of TikTok who retain editorial independence in relation to the fact-checking, and that the services shall be performed in a professional manner and in line with the highest standards in the industry. Our processes are also set up to ensure our fact-checking partners independence. Our partners access flagged content through an exclusive dashboard for their use and provide their assessment of the accuracy of the content by providing a rating. Fact-checkers will do so independently from us, and their review may include calling sources, consulting public data or authenticating videos and images.  To facilitate transparency and openness with our fact-checking partners, we regularly meet them and provide data regarding their feedback and also conduct surveys with them.
QRE 30.2.2	We meet regularly with our fact-checking partners and have an ongoing dialogue with them about how our partnership is working and evolving. We survey our fact-checking partners to encourage feedback about what we are doing well and how we could improve.
QRE 30.2.3	This provision is not relevant to TikTok, only to fact-checking organisations.
Measure 30.3	
QRE 30.3.1	Given our fact-checking partners are all IFCN-accredited, our fact-checking partners already engage in some informal cross-border collaboration through that network.  In addition, we continue to collaborate with our partners to understand how we may be able to facilitate further collaboration through individual feedback sessions with partners.
Measure 30.4	
QRE 30.4.1	We are in regular dialogue with EDMO and the EFCSN on these and other issues. We continue to be open to discussing and exploring what further progress can be made on these points.



# VII. Empowering the fact-checking community

### Commitment 31

Relevant Signatories commit to integrate, showcase, or otherwise consistently use fact-checkers' work in their platforms' services, processes, and contents; with full coverage of all Member States and languages.

Measure 31.1	
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, list these implementation measures here [short bullet points].	<ul> <li>Onboarded a new fact-checking partner: Faktrograf, to cover Croatia</li> <li>Continued to expand our fact-checking repository to ensure our teams and systems leverage the full scope of insights our fact-checking partners submitted to TikTok (regardless of the original language of the relevant content).</li> <li>Participated in the created working group within the Code framework on the creation of an external fact-checking repository.</li> </ul>
In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	Yes



Measure 31.2	
QRE 31.1.1	We see harmful misinformation as different from other content issues. Context and fact-checking are critical to consistently and accurately enforcing our harmful misinformation policies, which is why we work with 12 fact-checking partners in Europe, covering 23 EEA languages.
	While we use machine learning models to help detect potential misinformation, our approach is to have members of our content moderation team, who receive specialised training on misinformation, assess, confirm, and take action on harmful misinformation. This includes direct access to our fact-checking partners who help assess the accuracy of content. Our fact-checking partners are involved in our moderation process in three ways:
	(i) a moderator sends a video to fact-checkers for review and their assessment of the accuracy of the content by providing a rating. Fact-checkers will do so independently from us, and their review may include calling sources, consulting public data, authenticating videos and images, and more.
	While content is being fact-checked or when content can't be substantiated through fact-checking, we may reduce the content's distribution so that fewer people see it. Fact-checkers ultimately do not take action on the content directly. The moderator will instead take into account the fact-checkers' feedback on the accuracy of the content when deciding whether the content violates our CGs and what action to take.
	(ii) contributing to our global database of previously fact-checked claims to help our misinformation moderators make decisions.
	(iii) a proactive detection programme with our fact-checkers who flag new and evolving claims they're seeing on our platform. This enables our moderators to quickly assess these claims and remove violations.
	In addition, we use fact-checking feedback to provide additional context to users about certain content. As mentioned, when our fact checking partners conclude that the fact-check is inconclusive or content is not able to be confirmed, (which is especially common during unfolding events or crises), we inform viewers <u>via a banner</u> when we identify a video with unverified content in an effort to raise users' awareness about the credibility of the content and to reduce sharing. The video may also become ineligible for recommendation into anyone's For You feed to limit the spread of potentially misleading information.



SLI 31.1.1 - use of fact-checks	Methodology of data measurement:  The number of fact checked videos is based on the number of videos have been reviewed by one of our fact-checking partners in the relevant territory.			ve been reviewed by
	Nr of fact-checked articles published			
List actions per member states and languages (see example table above)	Number of fact checked videos (tasks)			
Member States				
Austria	63			
Belgium	408			
Bulgaria	212			
Croatia	378			
Cyprus	17			
Czech Republic	185			
Denmark	147			
Estonia	19			
Finland	114			
France	2,137			
Germany	1,142			



Greece	215
Hungary	243
Ireland	85
Italy	324
Latvia	14
Lithuania	29
Luxembourg	2
Malta	1
Netherlands	158
Poland	689
Portugal	266
Romania	837
Slovakia	177
Slovenia	10
Spain	370
Sweden	487
Iceland	1
Liechtenstein	0
Norway	806



Total EU	8,729		
Total EEA	9,536		

SLI 31.1.2 - impact of actions taken	Methodology of data measurement:  The number of videos removed as a result of a fact-checking assessment and the number of videos removed because of policy guidelines, known misinformation trends and our knowledge based repository is based on the country in which the video was posted.  These metrics correspond to the numbers of removals under the misinformation policy since all of its enforcement are based on the policy guidelines, known misinformation trends and knowledge based repository.  N/A			
List actions per member states and languages (see example table above)	Number of videos removed as a result of a fact checking assessment	Number of videos removed because of policy guidelines, known misinformation trends and knowledge based repository		
Member States				
Austria	15	2,577		
Belgium	71	5,957		
Bulgaria	18	1,211		
Croatia	74	1,050		
Cyprus	2	329		
Czech Republic	28	7,715		



Denmark	8	1,227
Estonia	0	131
Finland	9	51,960
France	589	54,817
Germany	188	47,499
Greece	31	2,294
Hungary	34	1,357
Ireland	14	1,872
Italy	53	31,343
Latvia	3	248
Lithuania	2	273
Luxembourg	0	155
Malta	0	72
Netherlands	14	4,725
Poland	171	9,346
Portugal	39	1,790
Romania	103	8,087
Slovakia	31	581
Slovenia	0	197



Spain	69	21,430	
Sweden	95	4,409	
Iceland	0	74	
Liechtenstein	0	6	
Norway	72	1,678	
Total EU	1,661	262,652	
Total EEA	1,733	264,410	

SLI 31.1.3 – Quantitative information used for contextualisation for the SLIs 31.1.1 / 31.1.2	Methodology of data measurement:  The metric we have provided demonstrates the % of videos which have been removed as a result of the fact checking assessment, in comparison to the total number of videos removed because of violation of our harmful misinformation policy.	
List actions per member states and languages (see example table above)	Number of videos removed as a result of a fact checking assessment / number of removals under harmful misinformation policy	
Austria	0.6%	
Belgium	1.2%	
Bulgaria	1.5%	



Croatia	7.0%
Cyprus	0.6%
Czech Republic	0.4%
Denmark	0.7%
Estonia	0.0%
Finland	0.0%
France	1.1%
Germany	0.4%
Greece	1.4%
Hungary	2.5%
Ireland	0.7%
Italy	0.2%
Latvia	1.2%
Lithuania	0.7%
Luxembourg	0.0%
Malta	0.0%
Netherlands	0.3%
Poland	1.8%
Portugal	2.2%



Romania	1.3%
Slovakia	5.3%
Slovenia	0.0%
Spain	0.3%
Sweden	2.2%
Iceland	0.0%
Liechtenstein	0.0%
Norway	4.3%
Total EU	0.6%
Total EEA	0.7%

Measure 31.3	
	We are participating in the sub-group created for this purpose. We actively worked with all signatories to define clear deliverables and timelines for the creation of an external fact-checking repository, as contemplated in this measure.
Measure 31.4	
	We commit to be an active participant in the discussion about technological solutions to facilitate the efficient use of the common repository across platforms and languages.



### VII. Empowering the fact-checking community

### Commitment 32

Relevant Signatories commit to provide fact-checkers with prompt, and whenever possible automated, access to information that is pertinent to help them to maximise the quality and impact of fact-checking, as defined in a framework to be designed in coordination with EDMO and an elected body representative of the independent European fact-checking organisations.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	Yes
If yes, list these implementation measures here [short bullet points].	<ul> <li>Continued to explore ways to improve data sharing in connection with our pilot scheme to share enforcement data with our fact-checking partners on the claims they have provided feedback on.</li> </ul>
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight disinformation and will report on any further development in the next COPD report.
Measure 32.1	
Measure 32.2	
QRE 32.1.1	Our fact-checking partners access content which has been flagged for review through a dashboard made available for their exclusive use. The dashboard shows our fact-checkers

	<b>6</b>	29, 30

	certain quantitative information about the services they provide, including the number of videos queued for assessment at any one time, as well as the time the review has taken. Fact-checkers can also use the dashboard to see the rating they applied to videos they have previously assessed.  Going forward, we plan to continue to explore ways to further increase the quality of our methods of data sharing with fact-checking partners.	
SLI 32.1.1 - use of the interfaces and other tools	Methodology of data measurement:  N/A. As mentioned in our response to QRE 32.1.1, the dashboard we currently share with our partners only contains high level quantitative information about the services they provide, including the number of videos queued for assessment at any one time, as well as the time the review has taken. We are continuing to work with our fact checking partners to understand what further data it would be helpful for us to share with them.	
Data		
Measure 32.3		
QRE 32.3.1	We continue to participate in the taskforce made up of the relevant signatories' representatives that is being set up for this purpose. Meanwhile we are also engaging with EDMO pro-actively on this commitment.	

# VII. Empowering the fact-checking community

### Commitment 33

Relevant Signatories (i.e. fact-checking organisations) commit to operate on the basis of strict ethical and transparency rules, and to protect their independence.



In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	No	
If yes, list these implementation measures here [short bullet points].	This Commitment and Measure relates to fact checking organisations.	
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No	
If yes, which further implementation measures do you plan to put in place in the next 6 months?	This Commitment and Measure relates to fact checking organisations.	
Measure 33.1		
QRE 33.1.1	Not committed. This Commitment and Measure relates to Fact-Checking organisations.	
SLI 33.1.1 - number of European fact-checkers that are IFCN-certified	N/A	













# VIII. Transparency Centre Commitments 34 - 36



# **VIII. Transparency Centre**

### Commitment 34

To ensure transparency and accountability around the implementation of this Code, Relevant Signatories commit to set up and maintain a publicly available common Transparency Centre website

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	Yes
If yes, list these implementation measures here [short bullet points].	We have been an active participant in the working group that has successfully launched the common Transparency Centre this year. We also took up the position of co-chair of the Transparency working group since September 2023. From January 2024, we have supported the transition of the maintenance and development of the website from the former third-party vendor, to the signatory of the Code, Vost.eu.
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.
Measure 34.1	
Measure 34.2	
Measure 34.3	



Measure 34.4	
Measure 34.5	

# VIII. Transparency Centre

### Commitment 35

Signatories commit to ensure that the Transparency Centre contains all the relevant information related to the implementation of the Code's Commitments and Measures and that this information is presented in an easy-to-understand manner, per service, and is easily searchable.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	Yes
If yes, list these implementation measures here [short bullet points].	<ul> <li>Through our participation in the working group that has successfully launched the common Transparency Centre this year, we have ensured that the Centre will allow the general public to access general information about the Code as well as the underlying reports (and for the Centre to be navigated both by commitment and signatory).</li> <li>We have supported the working group in collecting feedback on the user experience of the website. We have engaged with the vendor to increase the robustness of the website and make improvements to the interface.</li> </ul>
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No



If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.
Measure 35.1	
Measure 35.2	
Measure 35.3	
Measure 35.4	
Measure 35.5	
Measure 35.6	

# Commitment 36 Signatories commit to updating the relevant information contained in the Transparency Centre in a timely and complete manner. In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No] If yes, list these implementation measures here [short bullet points]. We have supported the working group in collecting feedback on the user experience of the website.



Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.
Measure 36.1	
Measure 36.2	
Measure 36.3	
QRE 36.1.1 (for the Commitments 34-36)	We have continued to regularly upload our report, according to the approved deadlines. We supported the creation of a standalone page on the European Parliamentary Elections: <a href="https://disinfocode.eu/eu-elections-2024/">https://disinfocode.eu/eu-elections-2024/</a> .
QRE 36.1.2 (for the Commitments 34-36)	We are committed to maintaining the website and supporting the development of relevant SLIs.
SLI 36.1.1 - (for Measures 34 and 36) meaningful quantitative information on the usage of the Transparency Centre, such as the average monthly visits of the webpage.	We worked with the vendor to develop relevant metrics for this SLI.
Data	Between 1 January 2024 and 30 June 2024, the common Transparency Center has been visited by 62,000 unique visitors. The Signatories' reports were downloaded 76,000 times by 1,800 unique visitors. More specifically, TikTok's previous COPD report was downloaded 422 times by 129 unique visitors.













# IX. Permanent Task-Force Commitment 37



### IX. Permanent Task-Force

### Commitment 37

Signatories commit to participate in the permanent Task-force. The Task-force includes the Signatories of the Code and representatives from EDMO and ERGA. It is chaired by the European Commission, and includes representatives of the European External Action Service (EEAS). The Task-force can also invite relevant experts as observers to support its work. Decisions of the Task-force are made by consensus.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	Yes
If yes, list these implementation measures here [short bullet points].	We have meaningfully engaged in the Task-force and all of its working groups and subgroups.
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.
Measure 37.1	
Measure 37.2	
Measure 37.3	



Measure 37.4  Measure 37.5	
Measure 37.6	
QRE 37.6.1	We have meaningfully engaged in the Task-force and all of its working groups and subgroups:  Monitoring and Reporting Integrity of Services Crisis Response Transparency Centre Outreach and Integration Ad Scrutiny Empowerment of Fact-Checkers Elections Generative Al  We will continue to engage in the Task-force and all of its working groups and subgroups.





X. Monitoring of Code Commitment 38 - 44



### Commitment 38

The Signatories commit to dedicate adequate financial and human resources and put in place appropriate internal processes to ensure the implementation of their commitments under the Code.

	·
In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	Yes.
If yes, list these implementation measures here [short bullet points].	Globally more than 40,000 Trust and Safety professionals are dedicated to keeping TikTok safe.
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.
Measure 38.1	





QRE 38.1.1	TikTok has assigned the highest priority level to the Code, which means that we have, and will continue to have, appropriate resources in place to meet our commitments and compliance.
	Given the breadth of the Code and the commitments therein, our work spans multiple teams, including Trust and Safety, Legal, Monetisation Integrity, Product and Public Policy. Teams across the globe are deployed to ensure that we meet our commitments and compliance with the notable involvement of our Trust and Safety Leadership.
	As of January 2024, TikTok has more than 6,000 people dedicated to the moderation of content in the European Union, as we recognise the importance of local knowledge and expertise as we work to ensure online safety for our users. We take a similar approach to our third party partnerships.

### Commitment 39

Signatories commit to provide to the European Commission, within 1 month after the end of the implementation period (6 months after this Code's signature) the baseline reports as set out in the Preamble.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	No
If yes, list these implementation measures here [short bullet points].	We have shared our baseline report with the Commission in accordance with the agreed timeframes.
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No



If yes, which further implementation measures do you plan to put in place in the next 6 months?	N/A

### Commitment 40

Signatories commit to provide regular reporting on Service Level Indicators (SLIs) and Qualitative Reporting Elements (QREs). The reports and data provided should allow for a thorough assessment of the extent of the implementation of the Code's Commitments and Measures by each Signatory, service and at Member State level.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	Yes
If yes, list these implementation measures here [short bullet points].	We have reported on the SLIs and QREs relevant to the Commitments we signed-up to within this report.
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.

		-	
	•	4	27,

Measure 40.1			
Measure 40.2			
Measure 40.3			
Measure 40.4			
Measure 40.5			
Measure 40.6			
X. Monitoring of Code			
Commitment 41			
Signatories commit to work within the Task-force towards developing Structural Indicators, and publish a first set of them within 9 months from the signature of this Code; and to publish an initial measurement alongside their first full report. To achieve this goal, Signatories commit to support their implementation, including the testing and adapting of the initial set of Structural Indicators agreed in this Code. This, in order to assess the effectiveness of the Code in reducing the spread of online disinformation for each of the relevant Signatories, and for the entire online ecosystem in the EU and at Member State level. Signatories will collaborate with relevant actors in that regard, including ERGA and EDMO.			

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]



If yes, list these implementation measures here [short bullet points].	<ul> <li>We have been an active participant in the working group dedicated to developing Structural Indicators.</li> <li>We supported the publication of the second analysis of Structural Indicators, expanding it to covering 4 markets and increasing the sample size. The Structural Indicators will be published at the same time as the COPD reports.</li> </ul>
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	N/A
Measure 41.1	
Measure 41.2	
Measure 41.3	

### Commitment 42

Relevant Signatories commit to provide, in special situations like elections or crisis, upon request of the European Commission, proportionate and appropriate information and data, including ad-hoc specific reports and specific chapters within the regular monitoring, in accordance with the rapid response system established by the Taskforce.



In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	Yes
If yes, list these implementation measures here [short bullet points].	We have been an active participant in the Crisis Response working group, which resulted in the implementation of the Rapid Response System in the context of the 2024 Elections for the European Parliament, and we publish Crisis Reports specific to the War in Ukraine, the Israel/Hamas conflict and the 2024 Elections for the European Parliament along with this report.
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.

### Commitment 43

Signatories commit to produce reports and provide data following the harmonised reporting templates and refined methodology for reporting and data disclosure, as agreed in the Task-force.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]

Yes



If yes, list these implementation measures here [short bullet points].	<ul> <li>Participated in the monitoring and reporting working group.</li> <li>Provided timely feedback on the harmonised reporting template.</li> </ul>
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.

### Commitment 44

Relevant Signatories that are providers of Very Large Online Platforms commit, seeking alignment with the DSA, to be audited at their own expense, for their compliance with the commitments undertaken pursuant to this Code. Audits should be performed by organisations, independent from, and without conflict of interest with, the provider of the Very Large Online Platform concerned. Such organisations shall have proven expertise in the area of disinformation, appropriate technical competence and capabilities and have proven objectivity and professional ethics, based in particular on adherence to auditing standards and guidelines.

In line with this commitment, did you deploy new implementation measures (e.g. changes to your terms of service, new tools, new policies, etc)? [Yes/No]	Yes
If yes, list these implementation measures here [short bullet points].	TikTok is committed to meeting its obligations under the DSA, including the independent audit obligations. We have taken steps to ensure readiness for our audit obligations and have



	appointed an external independent auditor to assess our compliance with these obligations.
Do you plan to put further implementation measures in place in the next 6 months to substantially improve the maturity of the implementation of this commitment? [Yes/No]	No
If yes, which further implementation measures do you plan to put in place in the next 6 months?	We are continuously reviewing and improving our tools and processes to fight misinformation and disinformation and will report on any further development in the next COPD report.





### Reporting on the service's response during a crisis

### War of aggression by Russia on Ukraine

### Threats observed or anticipated at time of reporting: [suggested character limit 2000 characters].

The war of aggression by Russia on Ukraine (hereinafter, "**War in Ukraine**") is devastating, and, as a platform, has challenged us to confront a complex and rapidly changing environment. At TikTok, the safety of our people and community is of paramount importance and we work continuously to safeguard our platform.

We have set out below some of the main threats we have observed on our platform in relation to the spread of harmful misinformation and covert influence operations (**CIO**) related to the war and note that we are committed to taking action to prevent such content from being shared in this context.

#### (I) Spread of harmful misinformation

We continue to observe and take action where appropriate under our policies. In terms of trends observed, we have seen false or unconfirmed claims about specific attacks and events, the development or use of weapons, the involvement of specific countries in the conflict and statements about specific military activities, such as the direction of troop movement. We also have seen instances of footage repurposed in a misleading way, including from video games or unrelated footage from past events presented as current.

As detailed below, we take a multi-faceted approach to understanding and removing misleading stories and, when it comes to addressing harmful misinformation, we apply the <a href="Integrity & Authenticity policies">Integrity & Authenticity policies</a> (I&A policies) in our <a href="CGs">CGs</a> and we will take action on such content from our platform. We support our moderation teams with detailed policy guidance and direction when moderating on crisis related misinformation using our misinformation policies, including providing case banks of harmful misinformation claims to support their moderation work.

### (II) CIOs

We fight against CIO as our I&A policies prohibit attempts to sway public opinion while also misleading our systems or users about the identity, origin, approximate location, popularity or overall purpose. We have specifically-trained teams which are on high alert to investigate and detect CIOs on our platform.

We have confirmed that during the period from January to June 2024, we took action to remove a total of 3 networks (consisting of 93 accounts in total) that were found to be involved in coordinated attempts to influence public opinion about the War in Ukraine while also misleading individuals, our community, or our systems. We publish all of the CIO networks we identify and remove within our new dedicated CIO transparency report <a href="https://example.com/here/">here</a>.

We know that CIO will continue to evolve in response to our detection and networks may attempt to reestablish a presence on our platform, which is why we continually seek to strengthen our policies and enforcement actions in order to protect our community against new types of harmful misinformation and inauthentic behaviours.



### Mitigations in place at time of reporting: [suggested character limit: 2000 characters].

We are continually working hard to ensure that TikTok is a source of reliable and safe information and recognise the heightened risk and impact of misleading information during a time of crisis.

### (I) Investment in our fact-checking programme

We employ a layered approach to detecting harmful misinformation which is in violation of our CGs. In order to ensure coverage which is comprehensive and responsive to the war, we have invested in more Russian and Ukrainian native language speakers for content moderation, as we recognise the importance of local cultural and linguistic context within the process.

Working closely with our fact-checking partners has been a crucial part of our approach to enforcing harmful misinformation on our platform. Our fact-checking programme includes coverage of Russian, Ukrainian and Belarusian. As part of the scale-up of our fact checking programme, we also onboarded Reuters, who are dedicated to fact-checking content in Russian and Ukrainian.

Collaborating with certain of our fact-checking partners to receive advance warning of emerging misinformation narratives has facilitated proactive responses against high-harm trends and has ensured that our moderation teams have up-to-date guidance.

### (II) Disruption of CIOs

Disrupting CIO networks has also been high priority work for us in the context of the crisis and we published a list of the networks we disrupted in the relevant period within our most recently published transparency report here.

Between January and June 2024, we took action to remove a total of 3 networks (consisting of 93 accounts in total) that were found to be involved in coordinated attempts to influence public opinion about the War in Ukraine while also misleading individuals, our community, or our systems. We now publish all of the CIO networks we identify and remove within our dedicated CIO transparency report <a href="here">here</a>.

### (III) Restricting access to content for state affiliated media

Since the early stages of the war, we have restricted access to content from a number of Russian state affiliated media entities in the EU, Iceland and Liechtenstein. We also expedited the rollout of our state affiliated media policy to bring viewers context to evaluate the content they consume on our platform and labels have since applied to content posted by the accounts of such entities in Russia, Ukraine and Belarus.

Since January 2023, we have scaled up our detection and labelling of state-controlled media accounts following the roll-out of our state-controlled media label policy globally.

### (IV) Mitigating the risk of monetisation of harmful misinformation



Whilst we have long prohibited political advertising on our platform, we have taken measures to further mitigate the risk of monetisation off the back of the War in Ukraine by prohibiting Russian-based advertisers from outbound targeting to EU markets, and suspended TikTok in the Donetsk and Luhansk regions. In addition, the ability to add new video content or Livestream videos to the platform in Russia also continues to be suspended.

### (V) Launching localised media literacy campaigns

Lastly, recognising the importance of proactive measures which are aimed at improving our users' digital literacy and increasing the prominence of authoritative information, during the last six months, we launched six localised media literacy campaigns addressing disinformation related to the War in Ukraine in Austria, Bulgaria, Czech Republic, Croatia, Germany and Slovenia, in close collaboration with our fact-checking partners. Users searching for keywords relating to the War in Ukraine are directed to tips, prepared in partnership with our fact checking partners, to help users identify misinformation and prevent the spread of it on the platform.

[Note: Signatories are requested to provide information relevant to their particular response to the threats and challenges they observed on their service(s). They ensure that the information below provides an accurate and complete report of their relevant actions. As operational responses to crisis/election situations can vary from service to service, an absence of information should not be considered a priori a shortfall in the way a particular service has responded. Impact metrics are accurate to the best of signatories' abilities to measure them].

### **Policies and Terms and Conditions**

Outline any changes to your policies

Outline any changes to your policies		
Policy	Changes (such as newly introduced policies, edits, adaptation in scope or implementation)	Rationale
Integrity and Authenticity Policies  As a nee	We continue to rely on our existing, robust I&A policies, which are an effective basis for tackling content related to the war.	In the context of the War in Ukraine, we have been able to rely on our I&A policies, which are our first line of defence in combating harmful misinformation and deceptive behaviours on our platform.  These guidelines make clear to our users what content we remove or make ineligible for the For You feed when they pose a risk of harm to our users or the wider public. We have also supported our moderation teams with detailed policy guidance and direction when moderating on war-related harmful misinformation using existing policies.
	As such, we have not needed to introduce new policies, or adapt	We have specialist teams within our Trust and Safety department dedicated to the policy issue of I&A, including within the areas of product and policy. Our experienced subject matter experts on I&A continually keep these policies under review and collaborate with external partners and experts when





In a crisis, we keep under review our policies and to ensure moderation teams have supplementary quidance.

understanding whether updates are required.

When situations such as the War in Ukraine arise, these teams work to ensure that appropriate guidance is developed so that the I&A policies are applied in an effective manner in respect of content relating to the relevant crisis (in this case, the war). This includes issuing detailed policy guidance and direction, including providing case banks on harmful misinformation claims to support moderation teams.

### **Scrutiny of Ads Placements**

Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.

# Preventing misuse of our monetisation features

(Commitment 1, Measure 1.1)

### Description of intervention

Since the beginning of the War in Ukraine, we have suspended livestreaming and new content in Russia, and also taken a strict position not to enable monetisation in Russia. Since March 2022, we prohibit Russia-based advertisers from outbound targeting ads to EU markets. We have also suspended TikTok in the Donetsk and Luhansk regions. We also take steps to comply with all other applicable sanctions, such as geoblocking content of identified sanctioned persons from the EU.

### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

The steps we have taken to restrict certain monetisation features aims to ensure that they are not misused.

### Description of intervention

We use a combination of automated and human moderation in order to identify content that breaches our ad policies.

#### Ad content moderation

(Commitment 2, Measure 2.2)

Our Monetisation Integrity department has moderation teams in multiple locations that speak Russian and Ukrainian as we recognise the importance of local cultural and linguistic context within the process.

We enforce our strict ad policies and have expert teams focused on investigating and responding to any attempts to circumvent our policies.



### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

Our efforts on ad moderation practices help to ensure that ads that breach our policies are rejected or removed, both in the context of the War in Ukraine and more broadly on our platform.

### **Political Advertising**

Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.

### Description of intervention

In an effort to maintain both the integrity of the platform and safety of the people who use it, accounts belonging to <u>politicians or political parties</u> are not able to advertise or make money on TikTok.

# Prohibition on Political Advertising

(Commitment 4)

We do not allow paid political promotion, political advertising, or fundraising by politicians and political parties (for themselves or others). Our political ads policy includes both traditional paid ads and creators receiving compensation to support or oppose a candidate for office. This has been a <u>long-standing policy</u> at TikTok since 2019. We don't allow paid or branded content that promotes or opposes a candidate, current leader, political party or group, or issue at the federal, state, or local level – including election-related ads, advocacy ads, or issue ads.

We are pleased that Regulation EU 2024/900 on the Transparency and Targeting of Political Advertising which came into force in the reporting period has now set out a common definition of "political advertising". We have been reviewing our policies to ensure that our prohibition of political advertising is at least as broad as the regulation.

### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

Our monetisation integrity content teams observed a sharp rise in attempts to post ads related to political/war content during the initial period of the conflict commencement. Our risk control methods ensured there was no corresponding increase in violating activity going live on the platform. Since the initial period of the conflict, these numbers have significantly dropped. We have since seen no additional spike.

### **Integrity of Services**



Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.

# Identifying and removing CIO networks

(Commitment 14, Measure 14.1)

### Description of intervention

We fight against CIO as our I&A policies prohibit attempts to sway public opinion while also misleading our systems or users about the identity, origin, approximate location, popularity or overall purpose. Our dedicated investigation teams and automated systems have been on high alert to identify threats related to CIO networks and have removed networks targeting discourse about the War in Ukraine, in accordance with our I&A policies which prohibit deceptive behaviours.

We know that CIO will continue to evolve in response to our detection and networks may attempt to reestablish a presence on our platform, which is why we continually seek to strengthen our policies and enforcement actions in order to protect our community against new types of harmful misinformation and inauthentic behaviours.

### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

Between January and June 2024, we took action to remove the following 3 networks (consisting of 93 accounts in total) that were found to be involved in coordinated attempts to influence public opinion about the Russia-Ukraine war while also misleading individuals, our community, or our systems:

Network Origin: Ukraine

**Description:** We assess that this network operated from Ukraine and targeted a Ukrainian audience. The individuals behind this network created inauthentic accounts and posted content in Ukrainian at scale in order to artificially amplify pro-Ukrainian narratives and generate off-platform traffic, thereby attempting to manipulate discourse about the ongoing war between Russia and Ukraine. The network was observed to be using 'clickbait' tactics, using grabbing sticker text without showing the full details of the video in order to create engagement and attention.

Accounts removed: 52 Followers: 2,624,189

Network Origin: Russia

**Description:** We assess that this network operated from Russia and targeted a US audience. The individuals behind this network created inauthentic accounts in order to artificially amplify narratives that attempt to increase division around immigration in the US and Europe, weaken support for the Ukrainian war effort, and criticize the current US presidential administration. Most accounts within the network posed as fictitious journalists or news brands.



Accounts removed: 9 Followers: 19,261

Network Origin: Ukraine

**Description:** We assess that this network operated from Ukraine and targeted a Ukrainian audience. The individuals behind this network created inauthentic accounts in order to artificially amplify narratives in Ukrainian and Russian languages critical of the current Ukrainian government, in an attempt to manipulate discourse around the end of the presidential term. The accounts within the network were found to amplify the same videos repeatedly by using different overlay text in each video.

Accounts removed: 32 Followers: 39,877

We published this information within our most recently published transparency report here.

### Tackling synthetic and manipulated media

(Commitments 14 and 15, Measures 14.1, 15.1 and 15.2)

### Description of intervention

We do not allow synthetic media (Al-generated audio and visual content) content on the platform that may mislead our users or cause harm.

We have updated our existing 'Synthetic and Manipulated Media' policy and renamed it the 'Edited Media and Al-Generated Content (AIGC)' policy, which became effective in May 2024. Through these updates, we have adopted more commonly used and easily understood language when referring to AIGC, and clarified our existing prohibitions on AIGC showing fake authoritative sources or crisis events, or falsely showing public figures in certain contexts including being bullied, making an endorsement, or being endorsed.

For the purposes of our updated policy, AIGC refers to content created or modified by artificial intelligence (AI) technology or machine-learning processes, which may include images of real people, and may show highly realistic-appearing scenes, or use a particular artistic style, such as a painting, cartoons, or anime. 'Significantly edited content' is content that shows people doing or saying something they did not do or say, or altering their appearance in a way that makes them difficult to recognise or identify.

In accordance with our policy, we prohibit AIGC that features:

- Realistic-appearing people under the age of 18
- The likeness of adult private figures, if we become aware it was used without their permission
- Misleading AIGC or edited media that falsely shows:
  - o Content made to seem as if it comes from an authoritative source, such as a reputable news organisation
  - o A crisis event, such as a conflict or natural disaster
  - o A public figure who is:
    - being degraded or harassed, or engaging in criminal or antisocial behaviour



- taking a position on a political issue, commercial product, or a matter of public importance (such as an elections)
- being politically endorsed or condemned by an individual or group

As AI evolves, we continue to invest in combating harmful AIGC by evolving our proactive detection models, consulting with experts, and partnering with peers on shared solutions.

Prohibited practices are set out in our I&A policies here.

### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

Our efforts support transparent and responsible content creation practices, both in the context of the War in Ukraine and more broadly on our platform.

# Removing harmful misinformation from our platform

(Commitment 14, Measure 14.1)

### Description of intervention

We take action to remove accounts or content which contain inaccurate, misleading, or false content that may cause significant harm to individuals or society, regardless of intent. In conflict environments, such information may include content that is repurposed from past conflicts, content which makes false and harmful claims about specific events, or incites panic. In certain circumstances, we may also reduce the prominence of such content.

We employ a layered approach to misinformation detection, leveraging multiple overlapping strategies to ensure comprehensive and responsive coverage. We place considerable emphasis on proactive content moderation strategies in order to remove harmful misinformation which is in violation of our policies before it is reported to us by users or third parties.

We place significant emphasis on proactive content moderation at TikTok, and are proud that we remove the vast majority of violative videos before they are reported to us by users or other third parties.

### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

In the context of the crisis, we are proud to have proactively removed thousands of videos containing harmful misinformation related to the War in Ukraine. We have been able to do this through a combination of automated review, - human level content moderation, carrying out targeted sweeps of certain types of content (e.g. hashtags/sensitive keyword lists) as well as working closely with our fact-checking partners and responding to emerging trends they identify.

We have invested heavily in Russian and Ukrainian native language speakers for content moderation, as we recognise the importance of local cultural and linguistic context in the misinformation moderation process.

	Ø		0	091	

#### Relevant metrics:

- Number of videos removed because of violation of misinformation policy with a proxy related to the War in Ukraine - 7,726
- Number of videos not recommended because of violation of misinformation policy with a proxy (only focusing on RU/UA) - 8,604
- Number of proactive removals of videos removed because of violation of misinformation policy with a proxy related to the War in Ukraine - 6.954

### **Empowering Users**

Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.

# Not proactively promoting news-type content to our users

(Commitment 18, Measure 18.1)

### Description of intervention

TikTok is primarily an entertainment platform. It may be the case that a user's For You feed shows more news-type content (such as news content relating to the War in Ukraine) over time if that is content that a user is actively looking for and engaging with, but we do not actively promote news content to users on TikTok.

### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

Our approach to news-type content supports our efforts to reduce the spread of harmful disinformation to users.

# Applying our state affiliated media label

(Commitment 17, Measure 17.1)

### Description of intervention

We have restricted access to certain state-affiliated media entities and strengthened our state-affiliated media policy in order to bring viewers context to evaluate content shared by such Russian, Ukrainian and Belarusian entities.

In the EU, Iceland and Liechtenstein, we have taken steps to restrict access to content from media outlets and accounts subject to sanctions.

We have also updated our state-affiliated media policy in order to strengthen our approach to countering influence attempts. These updates included:



- Making state-affiliated media accounts that attempt to reach communities outside their home country on current global events and affairs ineligible for recommendation, which means their content won't appear in the For You feed;
- Prohibiting state-affiliated media accounts in all markets where our state-controlled media labels are available, including in the EU, from advertising outside of the country with which they are primarily affiliated;
- Investing in our detection capabilities of state-affiliated media accounts.

We have also worked with third party external experts to shape our state-affiliated media policy and assessment of state-controlled media labels.

Where our state-affiliated media label is applied to content posted by the accounts of such entities in Russia, Ukraine and Belarus, users across the EEA are automatically shown a full screen pop-up containing information about what the label means and inviting the user to click on "learn more" and be redirected to an <u>in-app page</u>, which explains why the content has been labelled as state-controlled media.

### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

Since January 2023, we have scaled up our detection and labelling of state-controlled media accounts following the roll-out of our state-controlled media label policy globally.

#### Relevant metrics:

- Number of videos tagged with the state affiliated media label for Russia, Belarus and Ukraine 114,218
- Number of impressions of the state affiliated media label for Russia, Belarus and Ukraine 113,527,739

## Creating localised media literacy campaigns

(Commitment 17, Measures 17.2 and 17.3)

### Description of intervention

We recognise the importance of proactive measures which are aimed at improving our users' digital literacy and increasing the prominence of authoritative information.

We launched localised media literacy campaigns relating to the crisis in order to raise awareness amongst our users. We promoted the campaign through a combination of our in-app intervention tools in order to ensure that authoritative information is promoted to our users.

Users searching for keywords relating to the War in Ukraine are directed to tips, prepared in partnership with our fact checking partners, to help users identify misinformation and prevent the spread of it on the platform.



### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

Working with our fact-checking partners, the previous year we had developed and rolled-out six localised media literacy campaigns addressing disinformation related to the War in Ukraine in Austria, Bulgaria, Czech Republic, Croatia, Germany and Slovenia.

Relevant metrics for the six media literacy campaigns (EEA total numbers):

- Total Number of impressions of the search intervention 18,802,901
- Total Number of clicks on the search intervention 27,615
- Click through rate of the search intervention 0.15%

### **Empowering the Research Community**

Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.

### Measures taken to support research into crisis related misinformation and disinformation

(Commitment 27, Measure 26.1 and 26.2)

### Description of intervention

Through our Research API, academic researchers from non-profit universities in the US and Europe can apply to study public data about TikTok content and accounts. This public data includes comments, captions, subtitles, and number of comments, shares, likes, and favourites that a video receives, comments from our platform. More information is available here.

### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

During the period of this COPD report, we have approved 2 applications through the Research API, with an express focus on the War in Ukraine.

### **Empowering the Fact-Checking Community**

Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.

## Applying our unverified content label and

Description of intervention



# making content ineligible for recommendation

(Commitment 31, Measure 31.2)

Where our misinformation moderators or fact-checking partners determine that content is not able to be verified at the given time (which is common during an unfolding event), we apply our unverified content label to the content to encourage users to consider the reliability or source of the content. The application of the label will also result in the content becoming ineligible for recommendation in order to limit the spread of potentially misleading information.

### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

Verifying certain information during dynamic and fast moving events such as a war can be challenging and our moderators and fact-checkers cannot always conclusively determine whether content is indeed harmful misinformation, in violation of our CGs.

Therefore, in order to minimise risk, where our fact-checkers or our trained moderators do not have enough information to verify content which may potentially be misleading, we apply our unverified content label to inform users the content has been reviewed but cannot be conclusively validated. The goal is to raise users' awareness about the credibility of the content and to reduce sharing (see screenshots <a href="here">here</a>). Our unverified content label is available to users in 23 EU official languages (plus, for EEA users, Norwegian and Icelandic).

Where the banner is applied, the content will also become <u>ineligible for recommendation</u> into anyone's For You feed to limit the spread of information relating to unfolding events where details are still developing and which may potentially be misleading.

# Ensuring fact-checking coverage

(Commitment 30, Measure 30.1)

### Description of intervention

Our fact checking efforts cover Russian, Ukrainian, Belarusian and all major European languages (including 23 official EEA languages as well as a number of other languages which affect European users).

### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

Context and fact-checking are critical to consistently and accurately enforcing our harmful misinformation policies, which is why we have ensured that, in the context of the crisis, our fact-checking programme covers Russian, Ukrainian and Belarusian.

More generally, we work with 12-fact-checking partners in Europe, covering the spoken language of 23 EEA countries (as well as a number of other languages which affect European users). We have onboarded a fact-checking partner, Reuters, who are dedicated to fact-checking content in Russian and Ukrainian. To further support our fact-checking efforts in Ukraine specifically, we have also been leveraging additional Ukrainian-speaking reporters who are connected with some of our existing fact checking partners.

#### Relevant metrics:

• Number of fact checked videos with a proxy related to the War in Ukraine - 782



	<ul> <li>Number of videos removed as a result of a fact checking assessment with words related to the War in Ukraine - 155</li> <li>Number of videos not recommend in the For Your Feed as a result of a fact checking assessment with words related to the War in Ukraine - 152</li> </ul>
Collaborating with our	Description of intervention
fact-checking partners in relation to emerging trends	Since the beginning of the crisis, representatives from TikTok's misinformation policy and detection teams have met regularly with our external fact-checking partners, to receive advance warning of emerging misinformation narratives, facilitating proactive responses against high-harm trends.
(Commitment 31, Measure 31.1)	Indication of impact (at beginning of action: expected impact) including relevant metrics when available
<i>31.11</i>	Relevant metrics:
	<ul> <li>Number of videos removed because of violation of misinformation policy with a proxy related to the War in Ukraine - 7,726</li> </ul>





### Reporting on the service's response during a crisis

### Israel - Hamas Conflict

### Threats observed or anticipated at time of reporting: [suggested character limit 2000 characters].

TikTok understands both the significance and sensitivity of the Israel-Hamas conflict (referred to as the "Conflict" throughout this section). Indeed, the Conflict, and commentary about it, illustrates the challenges inherent in successfully engaging in content moderation of violative content at scale while at the same time ensuring that the fundamental rights and freedoms of European citizens are respected and protected. We remain focused on supporting free expression, upholding our commitment to human rights, and maintaining the safety of our community and integrity of our platform during the Conflict. We are also committed to transparency throughout this time and have kept our community informed of our immediate and ongoing response through the following Newsroom post: Our continued actions to protect the TikTok community during the Israel-Hamas war.

Since the start of the Conflict on 7 October 2023, we have been working diligently to remove content that violates our policies. We have set out below some of the main threats both observed and considered in relation to the Conflict and the actions we have taken to address these.

### (I) Spread of harmful misinformation

We believe that trust forms the foundation of our community, and we strive to keep TikTok a safe and authentic space where genuine interactions and content can thrive. TikTok takes a multi-faceted approach to tackling the spread of harmful misinformation, regardless of intent. This includes our: <a href="Integrity & Authenticity policies">Integrity & Authenticity policies</a> (I&A policies) in our <a href="Community Guidelines">Community Guidelines</a>; products; practices; and external partnerships with fact-checkers, media literacy bodies, and researchers. We support our moderation teams with detailed misinformation policy guidance, enhanced training, and access to tools like our global database of previously fact-checked claims from our IFCN-accredited fact-checking partners, who help assess the accuracy of content.

Since 7 October 2023, there has been a rise in misinformation and conspiracy theories relating to the Conflict. We have also seen spikes in deceptive account behaviours and continue to take swift action against fake engagement and accounts, for example, by removing 320 million fake accounts in the six months since the start of the Conflict.

### (II) Covert Influence Operations (CIO)

TikTok's integrity and authenticity policies do not allow deceptive behaviour that may cause harm to our community or society at large. This includes coordinated attempts to influence or sway public opinion while also misleading individuals, our community, or our systems about an account's identity, approximate location, relationships, popularity, or purpose.



We have specifically-trained teams on high alert to investigate CIO, and disrupting CIO networks has been high priority work for us in the context of the Conflict. We now provide regular updates on the CIO networks we detect and remove from our platform, including those we identify relating to the Conflict, in our dedicated CIO transparency report. Between January - June 2024, we took action to remove 2 networks, (consisting of 132 accounts in total) that were found to be related to the Conflict.

We know that CIO will continue to evolve in response to our detection and networks may attempt to reestablish a presence on our platform, which is why we continually seek to strengthen our policies and enforcement actions in order to protect our community against new types of harmful misinformation and inauthentic behaviours.

### Mitigations in place at time of reporting: [suggested character limit: 2000 characters].

We are continually working hard to ensure that TikTok is a source of reliable and safe information and recognise the heightened risk and impact of misleading information during a time of crisis. As part of our crisis management process, we launched a command centre that brings together key members of our 40,000-strong global team of safety professionals, representing a range of expertise and regional perspectives, so that we remain agile in how we take action to respond to this fast-evolving crisis. Since the beginning of the Conflict, we are:

### (I) Upholding TikTok's Community Guidelines

Continuing to enforce our <u>policies</u> against <u>violence</u>, <u>hate</u>, and <u>harmful misinformation</u> by taking action to remove violative content and accounts. For example, we remove content that promotes Hamas, or otherwise supports the attacks or mocks victims affected by the violence. If content is posted depicting a person who has been taken hostage, we will do everything we can to protect their dignity and remove content that breaks our rules. We do not tolerate attempts to incite violence or spread hateful ideologies. We have a zero-tolerance policy for content praising violent and hateful organisations and individuals, and those organisations and individuals aren't allowed on our platform. We also block hashtags that promote violence or otherwise break our rules.

Evolving our proactive automated detection systems in real-time as we identify new threats; this enables us to automatically detect and remove graphic and violent content so that neither our moderators nor our community members are exposed to it.

### (II) Leveraging our Fact-Checking Program

We employ a layered approach to detecting harmful misinformation which is in violation of our Community Guidelines and our global fact-checking program is a critical part of this. The core objective of the fact-checking program is to leverage the expertise of external fact-checking organisations to help assess the accuracy of harmful and difficult to verify claims.

To limit the spread of potentially misleading information, we apply <u>warning labels</u> and prompt users to reconsider sharing content related to unfolding or emergency events, which have been assessed by our fact-checkers but cannot be verified as accurate i.e., 'unverified content'. Mindful about how evolving



events may impact the assessment of sensitive Conflict related claims day-to-day, we have implemented a process that allows our fact-checking partners to update us quickly if claims previously assessed as 'unverified' become verified with additional context and/or at a later stage.

### (III) Scaling up our content moderation capabilities

Since 7 October 2023, we have deployed additional Arabic and Hebrew speaking moderators in order to augment existing content moderation teams in reviewing content and assisting with Conflict-related translations. As we continue to focus on moderator care, we have also deployed additional well-being resources for our human moderation teams during this time.

### (IV) Disruption of CIOs

Disrupting CIO networks has also been high priority work for us in tackling deceptive behaviour that may cause harm to our community or society at large. As noted above, between January and June 2024, we took action to remove 2 networks (consisting of 132 accounts in total) that were found to be related to the Conflict. We now publish all of the CIO networks we identify and remove, including those relating to the conflict, within our dedicated CIO transparency report, here.

### (V) Mitigating the risk of monetisation of harmful misinformation

Making temporary adjustments to policies that govern TikTok features in an effort to proactively prevent them from being used for hateful or violent behaviour in the region. For example, we've added additional restrictions on LIVE eligibility as a temporary measure given the heightened safety risk in the context of the current hostage situation. Our existing political ads policy and safety and civility policies help to mitigate the monetisation of harmful misinformation.

### (VI) Deploying search interventions to raise awareness of potential misinformation

To help raise awareness and to protect our users, we have launched search interventions which are triggered when users search for non-violating terms related to the Conflict (e.g., Israel, Palestine). These search interventions remind users to pause and check their sources and also direct them to well-being resources.

### (VII) Adding opt-in screens over content that could be shocking or graphic

We recognise that some content that may otherwise break our rules can be in the public interest, and we allow this content to remain on the platform for documentary, educational, and counterspeech purposes. Opt-in screens help prevent people from unexpectedly viewing shocking or graphic content as we continue to make <u>public interest exceptions</u> for some content.

In addition, we are committed to engagement with experts across the industry and civil society, such as <u>Tech Against Terrorism</u> and our <u>Advisory Councils</u>, and cooperation with law enforcement agencies globally in line with our <u>Law Enforcement Guidelines</u>, to further safeguard and secure our platform during these difficult times.



[Note: Signatories are requested to provide information relevant to their particular response to the threats and challenges they observed on their service(s). They ensure that the information below provides an accurate and complete report of their relevant actions. As operational responses to crisis/election situations can vary from service to service, an absence of information should not be considered a priori a shortfall in the way a particular service has responded. Impact metrics are accurate to the best of signatories' abilities to measure them].

#### **Policies and Terms and Conditions**

### Outline any changes to your policies

		<del>-</del>
Policy	Changes (such as newly introduced policies, edits, adaptation in scope or implementation)	Rationale
	We continue to rely on our existing, robust I&A policies, which are an effective basis for tackling content related to the Conflict. As such, we have not needed to introduce any new misinformation policies, for the purposes of addressing the crisis.  In a crisis, we keep under review our policies and to ensure moderation teams have supplementary guidance.	In the context of the Conflict, we have been able to rely on our I&A policies, which are our first line of defence in combating harmful misinformation and deceptive behaviours on our platform.  These guidelines make clear to our users what content we remove or make ineligible for the For You feed when they pose a risk of harm to our users or the wider public. We have also supported our moderation teams with detailed policy guidance and direction when moderating on Conflict-related harmful misinformation using existing policies.  We have specialist teams within our Trust and Safety department dedicated to the policy issue of I&A, including within the areas of product and policy. Our experienced subject matter experts on I&A continually keep these policies under review and collaborate with external partners and experts when understanding whether updates are required.  When situations such as the Conflict arise, these teams work to ensure that appropriate guidance is developed so that the I&A policies are applied in an effective manner in respect of content relating to the relevant crisis (in this case, the Conflict). This includes issuing detailed policy guidance and direction, including providing case banks on harmful misinformation claims to support moderation teams.
TikTok Feature Policies	In addition to being able to rely on our I&A policies, we have made	Temporary adjustments have been introduced in an effort to proactively prevent certain features from being used for hateful or violent behaviour in the region.



temporary adjustments to existing policies which govern certain TikTok features. For example, we have added additional restrictions on LIVE eligibility as a temporary measure given the heightened safety risk in the context of the current hostage situation.

### **Scrutiny of Ads Placements**

Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.

### Description of intervention

Our existing political ads policy and safety and civility policies help to mitigate the monetisation of harmful misinformation.

# Preventing misuse of our monetisation features

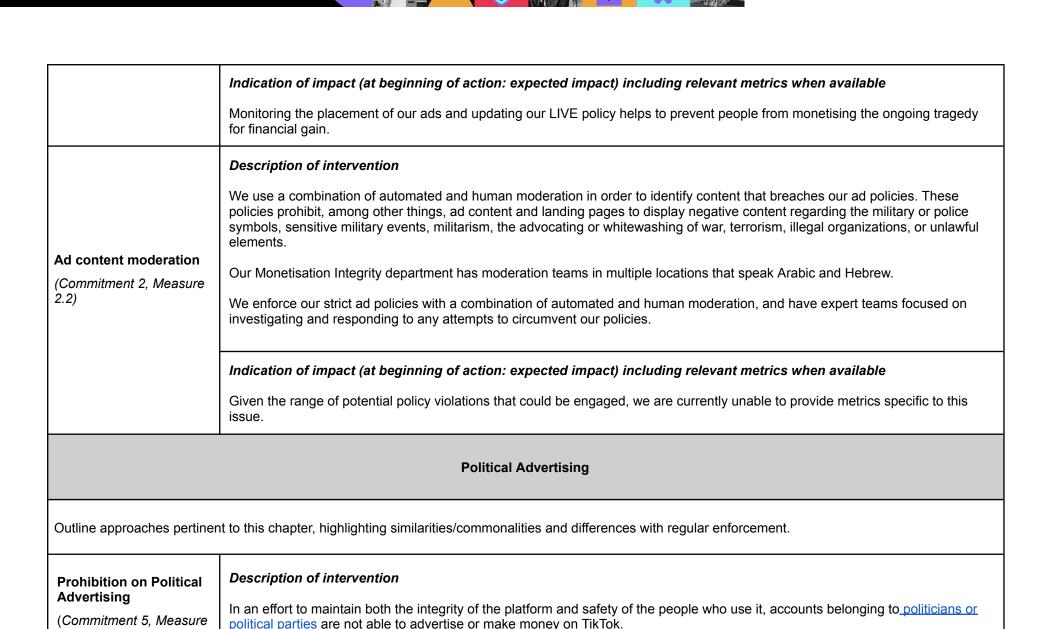
(Commitment 1, Measure 1.1 and 1.4)

We monitor the placement of our ads closely (especially in the context of events such as the Israel-Hamas conflict) and in the event our ads are found to run adjacent to or on sources of disinformation, we will investigate to determine whether the content in question has been correctly identified and verify this with Integral Ad Science (IAS) or GARM as appropriate. We will then adjust any filters or add the publication in question to a blocklist to prevent a recurrence.

NGOs that are vetted by Monetization Integrity and working with our sales team, are permitted to run ad campaigns for fundraising during the Conflict. This process operates with strict oversight to mitigate the risk of false fundraising campaigns. Also, in response to the Conflict, we have updated our LIVE feature guidelines to better prevent people from using monetisation features to exploit the ongoing tragedy for personal gain.



5.1)





We do not allow paid political promotion, political advertising, or fundraising by politicians and political parties (for themselves or others). Our political ads policy includes both traditional paid ads and creators receiving compensation to support or oppose a candidate for office. This has been a <u>long-standing policy</u> at TikTok since 2019. We also do not allow allow <u>ads and landing pages</u> which:

- reference, promote, or oppose candidates or nominees for public office, political parties, or elected or appointed government officials;
- reference an election, including voter registration, voter turnout, and appeals for votes;
- include advocacy for or against past, current, or proposed referenda, ballot measures, and legislative, judicial, or regulatory outcomes or processes (including those that promote or attack government policies or track records); and
- reference, promote, or sell, merchandise that features prohibited individuals, entities, or content, including campaign slogans, symbols, or logos.

We are pleased that Regulation EU 2024/900 on the Transparency and Targeting of Political Advertising which came into force in the reporting period has now set out a common definition of "political advertising". We have been reviewing our policies to ensure that our prohibition of political advertising is at least as broad as the regulation.

### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

By prohibiting political advertising, we help ensure our community can have a creative and authentic TikTok experience, and we reduce the risk of our platform being used to advertise and amplify narratives that may be divisive or false.

### **Integrity of Services**

Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.

# Identifying and removing CIO networks

(Commitment 14, Measure 14.1)

### Description of intervention

We have assigned dedicated resourcing within our specialist teams to proactively monitor for CIO in connection with the Conflict.

We fight against CIO as our I&A policies prohibit attempts to sway public opinion while also misleading our systems or users about the identity, origin, approximate location, popularity or overall purpose. We have specifically-trained and dedicated teams



which are on high alert to investigate and detect CIO networks on our platform and have removed networks targeting discourse about the Conflict, in accordance with our I&A policies which prohibit deceptive behaviours.

We know that CIO will continue to evolve in response to our detection and networks may attempt to reestablish a presence on our platform, which is why we continually seek to strengthen our policies and enforcement actions in order to protect our community against new types of harmful misinformation and inauthentic behaviours.

### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

Between January - June 2024, we took action to remove the following 2 networks (consisting of 132 accounts in total) that were found to be related to the Conflict:

Network Origin: Iran

**Description:** We assess this network operated from Iran and targeted audiences in the US and UK. Prior to October 2023, the individuals behind this network created inauthentic identities and used inauthentic means to gain user engagement on narratives surrounding UK domestic policy discourse. After October 2023, the network operator used the same inauthentic accounts to target the war between Hamas and Israel and artificially amplify pro-Iranian narratives and narratives critical of the US and Israel. The accounts in the network initially posted content associated with travel and tourism in order to build an audience, before switching to political topics.

Accounts in network: 65
Followers of network: 116,612

Network Origin: Iran

**Description:** We assess that this network operated from Iran and targeted audiences in Gulf Countries, particularly Saudi Arabia. The individuals behind this network created inauthentic personas to share narratives criticising social liberalisation. After October 2023, the network switched their messaging to focus on the Israel-Hamas war. The network shared primarily Arabic-language content that was copied from authentic social media users, including audio voiceovers.

Accounts in network: 67 Followers of network: 322,297

We now publish all of the CIO networks we identify and remove, including those relating to the Conflict, within our dedicated CIO transparency report here.

# Tackling synthetic and manipulated media

### Description of intervention



## (Commitment 15, Measures 15.1 and 15.2)

Our Edited Media and Al-Generated Content (AIGC) policy <u>makes it clear</u> that we do not want our users to be misled about crisis events. For the purposes of our policy, AIGC refers to content created or modified by AI technology or machine-learning processes. It includes images of real people, and may show highly realistic-appearing scenes.

We do not allow misleading AIGC or edited media that falsely shows:

- Content made to seem as if it comes from an authoritative source, such as a reputable news organisation,
- A crisis event, such as a conflict or natural disaster,
- A public figure who is:
  - o being degraded or harassed, or engaging in criminal or anti-social behavior
  - taking a position on a political issue, commercial product, or a matter of public importance (such as an elections)

In addition, all AIGC or edited media, including that which depicts public figures, such as politicians, must be clearly labelled as AI generated, and can not be used for endorsements.

In September 2023, we launched a new <u>Al-generated content label</u> for users to easily inform their community when they post AIGC. The label can be applied to any content that has been completely generated or significantly edited by AI which makes it easier to comply with the obligation to disclose AIGC that shows realistic scenes. Creators can do this through this label or through other types of disclosures, like a sticker, watermark or caption.

TikTok is also proud to have joined, earlier this year, the Content Authenticity Initiative (CAI) and the Coalition for Content Provenance and Authenticity (C2PA), and became the first video sharing platform to put Content Credentials into practice. We launched the ability to read Content Credentials that attach metadata to content, which we can use to instantly recognize and label AIGC. This helped us to expand auto-labelling to AIGC created on some other platforms.

### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

Our efforts support transparent and responsible content creation practices, both in the context of the Conflict and more broadly on our platform.

# Removing harmful misinformation from our platform

(Commitment 14, Measure 14.1)

### Description of intervention

We take action to remove accounts or content which contain inaccurate, misleading, or false information that may cause significant harm to individuals or society, regardless of intent. In conflict environments, such information may include content that is repurposed from past conflicts, content which makes false and harmful claims about specific events, or incites panic. In certain circumstances, we may also reduce the prominence of such content.



We employ a layered approach to misinformation detection, leveraging multiple overlapping strategies to ensure comprehensive and responsive coverage. We place considerable emphasis on proactive content moderation strategies in order to remove harmful misinformation which is in violation of our policies before it is reported to us by users or third parties.

### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

In the context of the crisis, we are proud to have proactively enforced over 30,000 videos containing harmful misinformation related to the Conflict. We have been able to do this through a combination of automation and human moderation, carrying out targeted sweeps of certain types of content (e.g. hashtags/sensitive keyword lists) as well as working closely with our fact-checking partners and responding to emerging trends they identify.

We have scaled up our Arabic and Hebrew speaking content moderation as we recognise the importance of language and cultural context in the misinformation moderation process.

#### Relevant metrics:

- Number of videos removed because of violation of misinformation policy with a proxy (IL/Hamas) 29,161
- Number of videos not recommended because of violation of misinformation policy with a proxy (IL/Hamas) -35.447
- Number of proactive removals of videos removed because of violation of misinformation policy with a proxy IL/Hamas 26,617

### **Empowering Users**

Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.

### Deploying search interventions to raise awareness of potential misinformation

(Commitment 21, Measure 21.1)

### Description of intervention

To minimise the discoverability of misinformation and help to protect our users, we have launched search interventions which are triggered when users search for neutral terms related to the Conflict (e.g., Israel, Palestine).

### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

These search interventions remind users to pause and check their sources and also direct them to well-being resources.





# Not proactively

promoting news-type content to our users

(Commitment 18, Measure 18.1)

### Description of intervention

TikTok is primarily an entertainment platform. It may be the case that a user's For You feed shows more news-type content (such as news content relating to the Conflict) over time if that is content that a user is actively looking for and engaging with, but we do not actively promote news content to users on TikTok.

Further, TikTok's recommender system delivers a personalised feed of content based on a user's engagement behaviour with video content types. Every user's feed will look different, according to their own preferences, which are demonstrated through 'likes', watch duration and other related metrics. For more information on how our recommendation system works, see <a href="here">here</a>.

### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

Our approach to news-type content supports our efforts to reduce the spread of harmful disinformation to users.

### **Empowering the Research Community**

Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.

### Measures taken to support research into Conflict- related misinformation and disinformation

(Commitment 26, Measure 26.1 and 26.2)

### Description of intervention

Through our Research API, academic researchers from non-profit universities in the US and Europe can apply to study public data about TikTok content and accounts. This public data includes comments, captions, subtitles, and number of comments, shares, likes, and favourites that a video receives, comments from our platform. More information is available <a href="here">here</a>.

Indication of impact (at beginning of action: expected impact) including relevant metrics when available

Between January and June 2024 4 Research API applications related to the Conflict have been approved.

### **Empowering the Fact-Checking Community**

Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.



### Description of intervention

Applying our unverified content label to make content ineligible for recommendation

(Commitment 31, Measure 31.2)

Where our misinformation moderators or fact-checking partners determine that content is not able to be verified at the given time (which is common during an emergency or unfolding event), we apply our unverified content label to the content to encourage users to consider the reliability or source of the content. The application of the label will also result in the content becoming ineligible for recommendation in order to limit the spread of potentially misleading information

### Indication of impact (at beginning of action: expected impact) including relevant metrics when available

Verifying certain information during dynamic and fast moving events such as the Conflict can be challenging and our moderators and fact-checkers cannot always conclusively determine whether content is indeed harmful misinformation.

Therefore, in order to minimise risk, where our fact-checkers or our moderators do not have enough information to verify content which may potentially be misleading, we apply our unverified content label to inform users the content has been reviewed but cannot be conclusively validated. The goal is to raise users' awareness about the credibility of the content and to reduce sharing (see screenshots here). Our unverified content label is available to users in 23 EU official languages (plus, for EEA users, Norwegian and Icelandic.

Where the banner is applied, the content will also become ineligible for recommendation into anyone's For You feed to limit the spread of information relating to unfolding events where details are still developing and which may potentially be misleading.

### 50 I

(Commitment 30,

Measure 30.1)

coverage

**Ensuring fact-checking** 

### Description of intervention

As part of our fact-checking program, TikTok works with 19 IFCN-accredited fact-checking organisations that support more than 50 languages, including Hebrew and Arabic, to help assess the accuracy of content in this rapidly-changing environment. In the context of the Conflict, our independent fact-checking partners are following our standard practice, whereby they do not moderate content directly on TikTok, but assess whether a claim is true, false, or unsubstantiated so that our moderators can take action based on our Community Guidelines. Fact-checker input is then incorporated into our broader content moderation efforts in a number of different ways, as further outlined in the 'indiciation of impact' section below.

In the context of the Conflict, we have also adjusted our information consolidation process to allow us to track and store Conflict related claims separately from our global repository of previously fact-checked claims. This facilitates quick and effective access to relevant assessments, which, in turn, increases the effectiveness of our moderation efforts.

Indication of impact (at beginning of action: expected impact) including relevant metrics when available



We see harmful misinformation as different from other content issues. Context and fact-checking are critical to consistently and accurately enforcing our harmful misinformation policies, which is why we have ensured that, in the context of the crisis, our fact-checking programme covers Arabic and Hebrew.

As noted above, we also incorporate fact-checker input into our broader content moderation efforts in different ways:

- Proactive insight reports that flag new and evolving claims they're seeing across the internet. This helps us detect harmful misinformation and anticipate misinformation trends on our platform.
- Collaborating with our fact-checking partners to receive advance warning of emerging misinformation narratives has facilitated proactive responses against high-harm trends and has helped to ensure that our moderation teams have up-to-date guidance.
- A repository of previously fact-checked claims to help misinformation moderators make swift and accurate decisions.

#### Relevant metrics:

- Number of fact checked tasks related to IL/Hamas 2.258
- Number of videos removed as a result of a fact checking assessment with words related to IL/Hamas 468
- Number of videos demoted (NR) as a result of a fact checking assessment with words related to IL/Hamas 463

# Description of intervention

#### Collaborating with our fact-checking partners in relation to emerging trends

Since the beginning of the Conflict, representatives from TikTok's misinformation policy and product teams have connected regularly with our external fact-checking partners, to receive advance insights into emerging misinformation narratives, facilitating proactive responses against high-harm trends.

(Commitment 31, Measure 31.1)

Indication of impact (at beginning of action: expected impact) including relevant metrics when available

#### Relevant metrics:

Number of videos removed because of violation of misinformation policy with a proxy (IL/Hamas) - 29,161





# Reporting on the signatory's response during an election

# 2024 European Parliament Elections

#### Threats observed during the electoral period: [suggested character limit 2000 characters].

As co-chair of the Code of Practice on Disinformation's Working Group on elections, TikTok takes our role of <u>protecting the integrity of elections</u> on our platform very seriously. We have comprehensive measures in place to anticipate and address the risks associated with electoral processes, including the risks associated with election misinformation in the context of the Elections for the European Parliament that took place across 27 Member States from 6 to 9 June 2024 (EU Elections).

TikTok has not observed major threats during the EU Elections:

- We removed over 2,600 pieces of content for violating our civic and election integrity policies, and over 43,000 pieces of content for violating our misinformation policies in the four weeks leading up to and during the EU Elections (between 6 May and 9 June 2024).
- We have not detected any covert influence operations solely focussed on the EU Elections during the reporting period.
- We received 5 notifications through the COPD Rapid Response System before the EU Elections, which were rapidly actioned and resulted in banning or geo-blocking of accounts and content removals for violation of Community Guidelines.

#### Mitigations in place during the electoral period: [suggested character limit: 2000 characters].

#### **Enforcing our policies**

#### (I) Moderation capabilities

We have over 6,000 people dedicated to moderating EU language content. As they usually do, our teams worked alongside technology to ensure that we were consistently <u>enforcing our rules</u> to detect and remove misinformation, covert influence operations, and other content and behaviour that can increase during an election period.

#### (II) Mission Control Centre: internal cross-functional collaboration

Ahead of the EU Elections, we established a dedicated Mission Control Centre (MCC) space in our Dublin office, bringing together in-person employees from multiple specialist teams within our safety department. Through the MCC, our teams were able to provide consistent and dedicated coverage of potential election-related issues in the run-up to, and during, the EU Elections.

# (III) Countering misinformation

Our misinformation moderators receive enhanced training and tools to detect and remove misinformation and other violative content. We also have teams on the ground who partner with experts to ensure local context and nuance is reflected in our approach.



In the four weeks leading up to and including the EU Elections, we removed over **2,600 pieces of content** for violating our civic and election integrity policies, and over **43,000 pieces of content** for violating our misinformation policies. We removed over **96%** of violative misinformation content before it was reported to us, and over 80% before receiving a single view. Some of the misinformation narratives we observed included narratives around migration, climate change, security and defence and LGBTQ rights.

#### (IV) Fact-checking

Our global fact-checking programme is a critical part of our layered approach to detecting harmful misinformation in the context of elections. The core objective of the fact-checking program is to leverage the expertise of external fact-checking organisations to help assess the accuracy of potentially harmful claims that are difficult to verify.

Across Europe, we worked with 12 local fact-checking partners and covered at least one official language of every Member State during the EU Elections. These partners supported our work through proactive surfacing and flagging of potential misinformation; verification of content and contribution to a repository of fact-checked claims; and in-app interventions. We also collaborated with fact-checkers to create media literacy content and produce Election Speaker Series.

#### (V) Deterring covert influence operations

Deceptive actors do sometimes try to target online platforms during elections, and we have dedicated experts working to detect, disrupt, and stay ahead of deceptive behaviours. To further increase transparency, accountability, and cross-industry sharing we introduced, earlier this year, <u>dedicated covert influence operations reports</u>.

During the reporting period we have not detected any covert influence operations (CIOs) focussed on the EU Elections. Based on our covert influence operations transparency reporting between January and June 2024, we did, however, disrupt 2 CIOs that we assessed to have targeted EU Member States with broader narratives:

- a network assessed to be targeting political discourse in Germany consisting of 46 accounts with 12,837 followers; and
- a network assessed to be targeting political discourse in Sweden consisting of 35 accounts with 109,537 followers.

We also continued to remove accounts from previously disclosed CIO networks when they try to reestablish a presence on the platform and our efforts there are disclosed monthly on our CIO transparency page.

# (VI) Tackling misleading Al-generated content

In addition to requiring creators to label any realistic Al-generated content (AIGC) and launching, in September 2023, an Al-generated content label to help people do this, ahead of the EU Elections we further strengthened our approach to tackling risks associated with AIGC:



- TikTok updated its existing 'Synthetic and Manipulated Media' policy and renamed it the 'Edited Media and Al-Generated Content (AIGC)' policy, which became effective in May 2024. Through these updates, we adopted more commonly used and easily understood language when referring to AIGC, and clarified our existing prohibitions on AIGC showing fake authoritative sources or crisis events, or falsely showing public figures in certain contexts including being bullied, making an endorsement, or being endorsed.
- As AI evolves, we continue to invest in combating harmful AIGC by evolving our proactive detection models, consulting with experts, and partnering with peers on shared solutions.
- TikTok joined the Content Authenticity Initiative and the Coalition for Content Provenance and Authenticity, and became the first video sharing platform to put Content Credentials into practice. We launched the ability to read Content Credentials that attach metadata to content, which we can use to instantly recognise and label AIGC. This helped us to expand auto-labelling to AIGC created on some other platforms.
- TikTok also joined industry partners as a party to the Tech Accord to Combat Deceptive Use of AI in 2024 Elections to work together to safeguard communities against misleading or deceptive AIGC in the election year.

#### (VII) Government, Politician, and Political Party Accounts (GPPPAs)

Many European leaders, ministers, and political parties are joining TikTok on a regular basis - many of them did ahead of the EU Elections, including approx. 30% of the sitting Members of European Parliament (MEPs) and the EU Parliament itself. These politicians and parties play an important role on our platform - we believe that verified accounts belonging to politicians and institutions provide the electorate with another route to access their representatives, and additional trusted voices in the shared fight against misinformation. Before the EU Elections we engaged with both sitting MEPs and candidates from across the EU 27 to ensure that those using TikTok had the necessary information to do so safely and authentically.

We strongly recommended GPPPAs to get <u>verified</u>. Verified badges help users make informed choices about the accounts they choose to follow. It is also an easy way for notable figures to let users know they're seeing authentic content, and it helps to build trust among high-profile accounts and their followers.

#### **Directing people to trusted sources**

# (I) Investing in media literacy

We invest in media literacy campaigns as a counter-misinformation strategy. In 2024, ahead of the EU Elections, we collaborated with fact-checkers to launch localised media literacy campaigns for each of the 27 EU Member States. This localised approach helped to ensure that messages were relevant to our community and encouraged more engagement.

From 28 March 2024 we launched an individual in-app Election Centre for each of the 27 EU Member States covering 24 official EU languages. Working with electoral commissions and civil society organisations, these Election Centres connected people with reliable voting information, including when, where, and how to vote; eligibility requirements for candidates; and, ultimately, the election results.



All Election Centres contained a media literacy section that provided tips and resources on how to identify misinformation and improve media literacy skills. This section included videos from our trusted fact-checking partners and local media literacy bodies: Agence France-Presse (AFP), Deutsche Presse-Agentur (DPA), Demagog.cz, Facta, Faktograf, Logically Facts, Newtral, Poligrafo, Delfi.lt, The Journal, Nieuwscheckers, Funky Citizens, DigiQ, Ostro.

We also worked closely with the European Parliament to incorporate their own media literacy videos into the Election Centres, with a 'follow' button to the Parliament's own TikTok account as a further link to authoritative election-related information.

We directed people to the Election Centres through prompts on videos, LIVEs and searches related to elections, as well as through push notifications. These Election Centres were visited more than 7.5 million times, and search banners were viewed more than 63 million times.

#### External engagement at the national and EU levels

#### (I) Rapid Response System: external collaboration with COPD Signatories

TikTok actively participated in the set up of the COPD Rapid Response System (RRS) launched in April 2024 to streamline the exchange of information among civil society organisations, fact-checkers, and online platforms. The RRS was a time-bound dedicated framework for cooperation among signatories during the EU Elections, which allowed non-platform signatories to flag time sensitive content, accounts or trends that may present threats to the integrity of the electoral process.

TikTok received 5 notifications through the RRS before the EU Elections, which were rapidly actioned and resulted in banning or geo-blocking of accounts and content removals for violation of Community Guidelines.

# (II) Engagement with local experts

To further promote election integrity, and inform our approach to the EU Elections, we organised Election Speaker Series with local and regional external experts, particularly from the fact-checking community, who shared their insights and market expertise with our internal teams.

Our fact-checking partners and local media literacy bodies have also supported our launch of the 27 localised Election Centres that included videos from them.

## (III) Engagement with national authorities and stakeholders

In order to prepare for the EU Elections, our Government Relations teams engaged with relevant authorities, including local Digital Services Coordinators, in many EU Member States.



#### (IV) Engagement at the EU-level

#### Engagement with the European Parliament and EU political parties

Our teams have also been closely engaging with the European Parliament since mid-2023 in order to discuss our preparations for the EU Elections and answer questions about our policies.

Our teams continued to actively engage with all European political groups and political parties, to explain our approach to elections, to help with account verification when requested, and answer their questions about our policies. To assist these efforts, our teams developed specific resources to help explain our policies and tools to interested politicians, and developed a new guide to account verification. In H1 2024 our teams also organised dedicated briefings for political parties at national level. At EU level, our teams were in close contact with all EU political groups and parties, and were available in both Brussels and Strasbourg for bilateral meetings in the lead up to the EU Elections.

#### Engagement with the EU authorities

Ahead of the EU Elections, our teams engaged with the European Commission (EC), including regarding the Guidelines for VLOPs on the mitigation of systemic risks for electoral processes pursuant to the Digital Services Act (DSA). Our teams also participated in events conducted by the EC. In particular, they took part in a simulation exercise to test our election readiness in April 2024 that was attended by civil society organisations and a number of European authorities, dedicated EU Elections roundtables in May and June 2024, and a post-election roundtable in July 2024.

[Note: Signatories are requested to provide information relevant to their particular response to the threats and challenges they observed on their service(s). They ensure that the information below provides an accurate and complete report of their relevant actions. As operational responses to crisis/election situations can vary from service to service, an absence of information should not be considered a priori a shortfall in the way a particular service has responded. Impact metrics are accurate to the best of signatories' abilities to measure them].

Policies and Terms and Conditions			
Outline any changes to your policies			
Policy	Changes (such as newly introduced policies, edits, adaptation in scope or implementation)	Rationale	



# Integrity and Authenticity (I&A) Policies

In May 2024 we updated our policies related to 'Misinformation', 'Civic and Election Integrity', 'Synthetic and Manipulated Media' (renamed 'Edited Media and Al-Generated Content (AIGC)'), and 'Spam and Deceptive Behavior'.

TikTok updated its Community Guidelines by adding new policies and enforcement regarding unverified elections claims, election misconduct, distribution of hacked materials, and significantly misrepresented authoritative civic information.

To ensure more robust policy coverage, TikTok also updated its existing policies related to synthetic media and AIGC to adopt more commonly used and easily understood language when referring to AIGC, and clarify existing prohibitions on AIGC showing fake authoritative sources or crisis events, or falsely showing public figures in certain contexts including being bullied, making an endorsement, or being endorsed.

#### **Scrutiny of Ads Placements**

Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.

#### **Description of intervention**

Scrutiny of Ad Placements, including prohibition on monetisation and fundraising campaigns for GPPPAs (Commitment 1 and Measure 1.1) We prohibit political advertising, which is outlined in more detail in the Political Advertising section below. In order to do this, we classify certain TikTok accounts as Government, Politician, and Political Party Accounts (GPPPA). GPPPAs can include government-run entities, elected officials or candidates, political parties, current/former heads of state or government. Where accounts are designated as GPPPAs, they are banned from placing ads on TikTok and from monetisation features.

Solicitations for campaign fundraising by GPPPAs are not permitted either on the platform. That includes content like a video from a politician asking for donations, or a political party directing people to a donation page on their website. We provide GPPPAs account holders with information on our policies and moderation systems as well as actively encourage them to verify their profiles. Read more <a href="here">here</a>.

In order to enforce these policies, our local Government Relations teams ensured that as many EU GPPPAs as possible were identified and labelled correctly ahead of the EU Elections.



#### Indication of impact including relevant metrics when available

Our GPPPA-specific policies helped us recognise the public interest associated with these accounts, while also putting guardrails in place (for example by restricting advertising and monetisation features).

#### **Political Advertising**

Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.

TikTok prohibits political advertising on its platform.

We recognise that some members of our community may want to create or engage with political content on TikTok and so sharing political beliefs and engaging in political conversation is allowed as organic content provided it does not violate our <a href="Community Guidelines">Community Guidelines</a> or <a href="Branded Content policy">Branded Content policy</a>. However, users cannot pay to advertise or promote this content. Furthermore, in an effort to maintain both the integrity of the platform and safety of the people who use it, accounts belonging to <a href="politicians or political parties">political parties</a> cannot advertise or make money on TikTok.

# Prohibition on Political Advertising

(Commitment 5, Measure 5.1)

We do not allow paid political promotion, political advertising, or fundraising by politicians and political parties (for themselves or others). Our political ads policy prohibits both traditional paid ads and creators being paid to make branded political content. This has been a <u>long-standing policy</u> at TikTok since 2019. We do not allow paid or branded content that promotes or opposes a candidate, current leader, political party or group, or issue at the federal, state, or local level – including election-related ads, advocacy ads, or issue ads. This means that accounts we identify as belonging to politicians and political parties (i.e. GPPPAs) have their access to advertising features turned off.

All ads are reviewed before being displayed on our platform. Upon review, if an ad is deemed to violate our strict ads policies including our prohibition of political ads, it will not be permitted on our platform. We also review ads reported to us and upon review, violating ads will be removed. Our community can also report ads, through our reporting tools, that they believe break our rules, which our teams will review and remove if they are found to be violative.

We take extremely seriously our responsibility to ensure electoral integrity. We strive to combat election misinformation and promote fair elections. We have instituted practices to mitigate the risk of harmful misinformation impacting election integrity, such as investing in media literacy campaigns as a counter-misinformation strategy. Earlier this year there were media reports in Ireland relating to political ads appearing in our Commercial Content Library. We promptly investigated the cause of this





content appearing on our platform and removed ads that we deemed to be in breach of our political advertising policy. We continue to regularly review and improve our policies and processes in order to combat increasingly sophisticated attempts to spread disinformation and to further strengthen our systems.

## Indication of impact (at beginning of action: expected impact) including relevant metrics when available

By prohibiting political advertising, we help ensure our community can have a creative and authentic TikTok experience, and we reduce the risk of our platform being used to advertise and amplify narratives that may be divisive or false.

Number of ads removed for our political advertisement policies during the 4 weeks leading up to and including the days of the EU Elections (6 May-9 June 2024) - 265 for the EEA, 263 for EU

#### **Integrity of Services**

Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.

## **Description of intervention**

During the EU Elections we continued our work to detect and disrupt covert influence operations (CIOs) that attempt to establish themselves on TikTok and undermine the integrity of our platform. To further increase transparency, accountability, and cross-industry sharing we introduced <u>dedicated covert influence operations reports</u>.

# Identifying and removing CIO networks

(Commitment 14, Measure 14.1)

# Indication of impact (at beginning of action: expected impact) including relevant metrics when available

During the reporting period we have not detected any CIOs focussed on the EU Elections. Based on our covert influence operations transparency reporting between January and June 2024 we did, however, disrupt, 2 CIOs that we assessed to have targeted EU Member States with broader narratives:

- a network assessed to be targeting political discourse in Germany consisting of 46 accounts with 12,837 followers; and
- a network assessed to be targeting political discourse in Sweden consisting of 35 accounts with 109,537 followers.

We also continued to remove accounts from previously disclosed CIO networks when they try to reestablish a presence on the platform and our efforts there are disclosed monthly on our CIO transparency page.



#### **Description of intervention**

Our Edited Media and Al-Generated Content (AIGC) policy <u>makes it clear</u> that we do not want our users to be misled about political issues. For the purposes of our policy, AIGC refers to content created or modified by artificial intelligence (AI) technology or machine-learning processes, which may include images of real people, and may show highly realistic-appearing scenes, or use a particular artistic style, such as a painting, cartoons, or anime.

We do not allow misleading AIGC or edited media that falsely shows:

- Content made to seem as if it comes from an authoritative source, such as a reputable news organisation.
- A crisis event, such as a conflict or natural disaster,
- A public figure who is:
  - o being degraded or harassed, or engaging in criminal or anti-social behaviour
  - taking a position on a political issue, commercial product, or a matter of public importance (such as an elections)
  - o being politically endorsed or condemned by an individual or group.

# Tackling misleading AIGC and edited media

(Commitment 15, Measures 15.1 and 15.2) In addition, all AIGC or edited media, including that which depicts public figures, such as politicians, must be clearly labelled as AI generated, and can not be used for endorsements.

In September 2023, we launched a new <u>Al-generated content label</u> for users to easily inform their community when they post AIGC. The label can be applied to any content that has been completely generated or significantly edited by AI which makes it easier to comply with the obligation to disclose AIGC that shows realistic scenes. Creators can do this through this label or through other types of disclosures, like a sticker, watermark or caption.

TikTok is also proud to have joined, earlier this year, the Content Authenticity Initiative and the Coalition for Content Provenance and Authenticity, and became the first video sharing platform to put Content Credentials into practice. We launched the ability to read Content Credentials that attach metadata to content, which we can use to instantly recognise and label AIGC. This helped us to expand auto-labelling to AIGC created on some other platforms.

Lastly, acknowledging that it is crucial for tech companies to work together to safeguard communities against misleading and deceptive AI in this historic election year, TikTok joined industry partners as a party to the Tech Accord to Combat Deceptive Use of AI in 2024 Elections. Through this Accord, we pledged to work collaboratively on tools to detect and address online distribution of such AIGC and acknowledge the importance of tracking the origin of deceptive election-related content and the need to raise public awareness about the problem.

Indication of impact (at beginning of action: expected impact) including relevant metrics when available



	Number of videos removed because of violation of synthetic and manipulated media policy during the 4 weeks leading up to the EU Elections and including the days of the EU Elections (6 May-9 June 2024): 1,176		
Empowering Users  Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.			
(Commitment 17, Measure 17.2)	From 28 March 2024 we launched an individual in-app Election Centre for each of the 27 EU Member States covering 24 official EU languages. Working with electoral commissions and civil society organisations, these Election Centres connected people with reliable voting information, including when, where, and how to vote; eligibility requirements for candidates; and, ultimately, the election results.		
	All Election Centres contained a media literacy section that provided tips and resources on how to identify misinformation and improve media literacy skills. This section included videos from our trusted fact-checking partners and local media literacy bodies: Agence France-Presse (AFP), Deutsche Presse-Agentur (dpa), Demagog.pl, Demagog.cz, Facta, Faktograf, Logically Facts, Newtral, Poligrafo, Delfi.lt, The Journal, Nieuwscheckers, Funky Citizens, DigiQ, Ostro.		
	We also worked closely with the European Parliament to incorporate their own media literacy videos into the Election Centres, with a 'follow' button to the Parliament's own TikTok account as a further link to authoritative election-related information.		
	We directed people to the Election Centres through prompts on videos, LIVEs and searches related to elections, as well as through push notifications.		
	Indication of impact (at beginning of action: expected impact) including relevant metrics when available		
	The 27 in-app Election Centres launched before the EU Elections were visited more than 7.5 million times, and search banners were viewed more than 63 million times.		
Engagement with local and regional experts	Description of intervention		



# (Commitment 17, Measure 17.2)

To further promote election integrity, and inform our approach to the EU Elections, we organised Election Speaker Series with local and regional external experts, particularly from the fact-checking community, who shared their insights and market expertise with our internal teams. The Election Speaker Series were held with the following organisations:

- German Press Agency (dpa): Austria, Germany, the Netherlands
- Agence France-Presse (AFP): France, Belgium, Greece, Cyprus
- Facta: ItalyNewtral: SpainPoligrafo: Portugal
- Faktograf: Croatia, Slovenia
- Demagog.pl: Poland
- Funky Citizens: Romania, BulgariaDemagog.cz: Czechia, Slovakia
- LeadStories: HungaryThe Journal: Ireland
- Delfi.lt: Latvia, Lithuania, Estonia
- Logically Facts: Sweden, Denmark, Finland

Our fact-checking partners and local media literacy bodies have also supported our launch of the 27 localised Election Centres that included videos from them: Agence France-Presse (AFP), Deutsche Presse-Agentur (dpa), Demagog.pl, Demagog.cz, Facta, Faktograf, Logically Facts, Newtral, Poligrafo, Delfi.lt, The Journal, Nieuwscheckers, Funky Citizens, DigiQ, Ostro. This localised approach helped to ensure that messages were relevant to our community and encouraged more engagement.

Indication of impact (at beginning of action: expected impact) including relevant metrics when available

This engagement with external regional and local experts allowed us to inform our country-level approaches because, although the EU Elections represented a unique transnational democratic process, in practice, they consisted of 27 national elections.

# **Empowering the Research Community**

Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.

# Providing access to our Research API (Commitment 26 and Measures 26.1 and 26.2)

# **Description of intervention**





Through our Research API, academic researchers from non-profit universities in the US and Europe can apply to study public data about TikTok content and accounts. This public data includes comments, captions, subtitles, and number of comments, shares, likes, and favourites that a video receives, and comments from our platform. More information is available <a href="here">here</a>.

Indication of impact (at beginning of action: expected impact) including relevant metrics when available

Number of Research API applications related to the EU Elections that have been approved from January to June 2024: 8

#### **Empowering the Fact-Checking Community**

Outline approaches pertinent to this chapter, highlighting similarities/commonalities and differences with regular enforcement.

# Ensuring fact-checking coverage (Commitment 30, Measure 30.1)

#### **Description of intervention**

We ensured fact-checking coverage by a combination of permanent and temporary partnerships covering at least one official language of each of the EU Member States, for the EU Elections, and expanded our misinformation moderation teams' coverage accordingly.

We onboarded Faktograf, the first Croatian media specialised in fact-checking, thus expanding our fact-checking network in Europe to 12 fact-checking partners.

Indication of impact (at beginning of action: expected impact) including relevant metrics when available

Number of videos reviewed by fact-checking partners during the 4 weeks leading up to the EU Elections and including the days of the EU Elections (6 May-9 June 2024): approx. 1000