



**Australian Code of Practice on Disinformation and
Misinformation: Initial Report**

May 2021



Executive summary

Business and Content Context

TikTok is the fastest growing destination for short-form video. TikTok's mission is to inspire creativity and bring joy. We are building a global community where people can create and share, discover the world around them, and connect with others across the globe. As we grow, we are committed to maintaining a supportive environment for our community.

At TikTok, we prioritise safety, diversity, inclusion, and authenticity. We encourage creators to celebrate what makes them unique and viewers to engage in what inspires them; and we believe that a safe environment helps everyone express themselves openly. We deeply value the global nature of our community, and we strive to take into account the breadth of cultural norms where we operate. We also aim to cultivate an environment for genuine interactions by only allowing authentic content on TikTok.

Approach to Disinformation and Misinformation

At TikTok we do not allow misinformation on our platform. Misinformation is defined as content that is inaccurate or false. While we encourage our community to have respectful conversations about subjects that matter to them, we do not permit misinformation that causes harm to individuals, our community, or the larger public regardless of intent.

Our approach to misinformation (which includes disinformation) is set out in our [Community Guidelines](#). The specific measures that we employ within our end-to-end effort to prevent misinformation are outlined in the discussion about code outcomes below.

Our Community Guidelines apply to everyone and everything on TikTok. We proactively enforce them using a mix of technology and human moderation. We also encourage our community members to use the [tools we provide on TikTok](#) to report any content they believe violates our Community Guidelines.

We will remove any content – including video, audio, livestream, images, comments, and text – that violates our Community Guidelines. Creators are notified of our decisions and can appeal if they believe no violation has occurred. We will suspend or ban accounts and/or devices that are involved in severe or repeated violations; we



will consider information available on other platforms and offline in these decisions. When warranted, we will report the accounts to relevant legal authorities.

We allow content on our platform that counters misinformation. We therefore take a nuanced approach and allow reference to misinformation if the content could be in the public interest. Examples of this could include counter speech or satirical content.

Approach to monitoring performance

At TikTok we are committed to both addressing misinformation on our platform and being transparent about our performance in doing so. Our half-yearly global [Transparency Report](#) provides visibility into the volume and nature of content removed for violating our Community Guidelines, as well as insight into our work to counter misinformation related to COVID-19 and elections. It also includes how we respond to law enforcement requests for information, government requests for content removals, and intellectual property removal requests.

To further aid our efforts to identify and reduce the spread of misinformation, we are continuing our work with Agence France-Presse (AFP), our third-party fact-checking partner for Australia. AFP makes independent assessments about the veracity of claims on our platform. These contributions help us to strike a balance and prevent either under-moderating or over-moderating identified potential misinformation.

Data about trends

AFP provides us advice on misinformation trends that we can use to further limit the spread of misinformation on TikTok. We also work with a range of outside experts to understand what kind of misinformation trends are occurring on other platforms. By being alert to content trends outside of TikTok, we can issue guidance to our moderation teams to help them more easily spot and take action on violating content. The past 12 months have seen a particular focus of our efforts on combatting misinformation about COVID-19, including the effectiveness of vaccinations. Examples of content observed across platforms include:

- False claims that Prime Minister [Scott Morrison “faked” his Pfizer vaccination](#) despite clear video footage and photos showing that there was an exposed needle tip on the syringe.
- Health Minister Greg Hunt, who was diagnosed with cellulitis, was subject to the false claim that his [AstraZeneca vaccine was the cause of the skin infection](#).
- A false AstraZeneca vaccine claim after some members of the Australian Navy reported suffering mild side effects following their vaccination. The claim purported that [80 percent of Navy members had experienced “severe” side](#)



[effects](#) post-vaccination, and thus the Navy were down 50% of “healthy members”. The Department of Defence issued a statement disputing those claims, stating that it was simply “not true,” and that the side effects were mild and did not require medical care.

Number of Australian violation videos with the term “Coronavirus” or “Covid” (removed from the platform)

Month	Count
October 2020	138
November 2020	94
December 2020	91
January 2021	149
February 2021	86
March 2021	93

Number of Australian medical misinformation videos (removed from the platform)

Month	Count
October 2020	11
November 2020	24
December 2020	42
January 2021	24
February 2021	61
March 2021	60

More information about our efforts in removing COVID-19 misinformation can be found [here](#).

Future Initiatives

In consultation with relevant stakeholders, we [update our Community Guidelines](#) at regular intervals to evolve alongside new behaviours and risks, as part of our commitment to keeping TikTok a safe place for creativity and joy.

We are also continuing to expand our fact-checking network globally and to build partnerships with local stakeholders, including NGOs, regulators, and state and federal electoral commissions.



Objective 1: Safeguards against Disinformation and Misinformation

Outcome 1a: Signatories contribute to reducing the risk of Harms that may arise from the propagation of Disinformation and Misinformation on digital platforms by adopting a range of scalable measures.

A robust policy framework

In response to global concern about the propagation of disinformation and misinformation, we have updated [our policies](#) on misleading content to provide further clarity on what is and isn't allowed on TikTok. Our Community Guidelines reflect our ongoing commitment to providing a safe platform for creative self-expression and fostering a trust-based community that promotes authentic behaviour. We want to ensure that people understand these guidelines, when and why we take action and what kind of content is not allowed on the platform.

In the past 12 months we have added a policy which prohibits synthetic or manipulated content that misleads users by distorting the truth of events in a way that could lead to real world harm. Our intent was to explicitly protect users from things like shallow or deep fakes, even though that kind of content was already broadly covered by our guidelines.

We have also made our policy around influence campaigns more transparent. Our Community Guidelines already prohibit content around disinformation campaigns and this addition makes our stance against coordinated inauthentic behaviour unambiguous: *"Do not engage in coordinated inauthentic activities (such as the creation of accounts) to exert influence and sway public opinion while misleading individuals, our community or the larger public about the account's identity, location or purpose."*

It is also notable that we do not allow paid [political advertising](#) on our platform in Australia.



Appropriate resources to address misinformation and disinformation

We continue to [build our teams](#), processes and technology to address the challenges faced by the spread of misinformation. Our 24/7 computer and human moderation capability is supported by Agence France-Presse (AFP): global fact-checkers who review and verify reported content across 16 languages. If the information is found to be false or misleading, we take proactive steps to remove the content in line with our Community Guidelines.

We use hashtag identification within our search function to help refer our users to verifiable information, including on upcoming elections and as a means of taking [action against Covid-19 and vaccine misinformation](#) on the platform. When searching for terms related to COVID-19, users are directed to our COVID-19 information hub, which was viewed 2,625,049,193 times globally in the second half of 2020.

Australian page views and user views for COVID-19 Information Hub		
<i>Month</i>	<i>Page views</i>	<i>User views</i>
November 2020	86,034	72,191
December 2020	78,962	58,690
January 2021	112,823	73,765
February 2021	80,512	59,990
March 2021	45,703	27,899

We are also rolling out specific product measures to limit the spread of potentially misleading information. Our [Know The Facts](#) campaign was implemented in Australia in February 2021 to prompt users to think twice before sharing content with unsubstantiated claims.

Fact checks are sometimes inconclusive or content is not able to be confirmed, especially during unfolding events. In these cases we need to strike a careful balance between countering misinformation and avoiding overreach. Our recommendation system on TikTok is also designed with safety as a key consideration and can be used in these cases to prevent unverified content from entering our users' [For You feed](#). We also inform viewers when we identify a video with unsubstantiated content in an effort to reduce sharing.



Number of Australian videos tagged with COVID-19 information notices	
<i>Month</i>	<i>Count</i>
October 2020	3,479
November 2020	3,148
December 2020	3,521
January 2021	2,715
February 2021	2,394
March 2021	1,628

Outcome 1b: Users will be informed about the types of behaviours and types of content that will be prohibited and/or managed by Signatories under this Code.

The TikTok Community Guidelines define a set of norms and common code of conduct for TikTok; they provide guidance on what is and is not allowed to make a welcoming space for everyone. Under the Integrity and Authenticity header, we set out clear guidelines for our users to help prevent misinformation on our platform:

Misinformation

Misinformation is defined as content that is inaccurate or false. While we encourage our community to have respectful conversations about subjects that matter to them, we do not permit misinformation that causes harm to individuals, our community, or the larger public regardless of intent.

Do not post, upload, stream, or share:

- Misinformation that incites hate or prejudice
- Misinformation related to emergencies that induces panic
- Medical misinformation that can cause harm to an individual's physical health
- Content that misleads community members about elections or other civic processes
- Conspiratorial content that attacks a specific protected group or includes a violent call to action, or denies a violent or tragic event occurred
- Digital Forgeries (Synthetic Media or Manipulated Media) that mislead users by distorting the truth of events and cause harm to the subject of the video, other persons, or society



Do not:

- Engage in coordinated inauthentic behaviours (such as the creation of accounts) to exert influence and sway public opinion while misleading individuals and our community about the account's identity, location, or purpose.

Outcome 1c: Users can report content or behaviours to Signatories that violates their policies under section 5.10 through publicly available and accessible reporting tools.

TikTok users can [report content](#) that they believe violates our Community Guidelines, at any time within the app. Misinformation is a distinct reporting category within our platform and appears as the first suggested issue when making a report about content.

Non-users can make a report to us about content on our platform using our out-of-app [reporting form](#).

We are also building relationships with regulators, NGOs and academics and establishing systems that enable direct reports to us of potential misinformation for further consideration and action.

We have published blogposts about publicly available and accessible reporting tools. These are available on our website [here](#).

Outcome 1d: Users will be able to access general information about Signatories' actions in response to reports made under 5.11.

Information about TikTok's approach to violations of our Community Guidelines, which include violations of our misinformation policies, is available on our [website](#).

All creators are notified when their content has violated our Community Guidelines. Our goals are to enhance the transparency and education around our Community Guidelines to reduce misunderstandings about content on our platform.



Our [Global Transparency Report](#) also provides information on our platform-wide efforts to address misinformation, particularly as it relates to COVID-19. The report details that in the second half of 2020, 2.4% of videos removed from our platform globally related to Integrity and Authenticity concerns. The Integrity and Authenticity category captures, among other things, videos containing misinformation as well as spam and impersonation.

This global figure was up from 1.2% of all videos in the [first half of 2020](#). 91.3% of these videos were removed within 24 hours. 70.5% were removed before a user made a report to us.

Objective 2: Disrupt advertising and monetisation incentives for disinformation.

Outcome 2: Advertising and/or monetisation incentives for Disinformation are reduced.

TikTok is a new entrant into the Australian advertising market as a platform offering marketing solutions for brands and small and medium-sized businesses.

We have strict [advertising policies](#) and do not allow political advertising. We also do not allow the making of deceptive or misleading claims in paid content.

While we believe advertising and monetisation are low risk insofar as it relates to misinformation given our strict policies and processes, we will continue to actively monitor this space and ensure our policies continue to reflect risks as they emerge.

Objective 3: Work to ensure the integrity and security of services and products delivered by digital platforms.

Outcome 3: The risk that Inauthentic User Behaviours undermine the integrity and security of services and products is reduced.

TikTok does not allow accounts that pose as another person or entity in a deceptive manner. When we confirm a report of impersonation, we will ask the user to revise the profile or suspend or ban the account. We do allow parody, commentary, or fan



accounts, as long as the user indicates in the bio and username that it is fan, commentary, or parody and not affiliated with the subject of the account.

We also do not permit coordinated behaviour that can mislead or cause harm to individuals, our community, or the larger public regardless of intent.

Under our Community Guidelines, users must not:

- pose as another person or entity by using someone else's name, biographical details, or profile picture in a misleading manner
- engage in coordinated inauthentic behaviours (such as the creation of accounts) to exert influence and sway public opinion while misleading individuals and our community about the account's identity, location, or purpose.

We are working with a number of industry-leading threat assessment platforms to bolster our ability to detect inauthentic activity and improve our safeguards against it.

Objective 4: Empower consumers to make better informed choices of digital content.

Outcome 4: Users are enabled to make more informed choices about the source of news and factual content accessed via digital platforms and are better equipped to identify Misinformation.

TikTok supports a number of initiatives globally, and within Australia, to boost media literacy and help users make informed choices about content on our platform. Much of this work focuses on our younger users, including through the establishment of a dedicated TikTok [Youth Portal](#) where teens and their families can learn about [internet safety](#) and the tools and controls built into TikTok.

In collaboration with the [Alannah & Madeline Foundation](#), TikTok has rolled out a series of videos that shine a spotlight on key, practical tips to help navigate misinformation and use online platforms safely. The *Get Media Smart* video series translates internet safety education in a digestible and engaging way for TikTok users of all ages. The six TikTok videos feature popular local creators, [@rifenstine](#) and [@sakshammagic](#), and playfully bring to life scenarios that everyday Australians may encounter as they use the internet and platforms like TikTok.



The video series encourages smart and safe online behaviour, builds confidence to seek help when needed and empowers users to stop the spread of misinformation by reporting content and by skipping the 'like' or 'share' buttons.

The *Get Media Smart* series covers a range of topics and guides:

- '[Question the Source](#)' asks users to find out who is authoring the content that they see.
- '[Not Interested](#)' shows users that they can control what they see.
- '[Fact vs. Opinion](#)' helps users determine whether statements are expressing truths or beliefs.
- '[Verify the visuals](#)' challenges the TikTok community to analyse the visual images they come across.
- '[Challenge Your Bias](#)' reminds users to consider how their experiences influence what they believe.
- '[Calling Out Fakes](#)' encourages users to reflect on whether the content they want to share, should be shared.

Since their launch in February 2021, these videos have been promoted to TikTok users across Australia, reaching hundreds of thousands of views. The videos are also available through the Alannah and Madeline Foundation's [eSmart Media Literacy Lab](#), a gamified education resource for teachers and students aged 12-16.

Objective 5: Improve public awareness of the source of political advertising carried on digital platforms.

Outcome 5: Users are better informed about the source of Political Advertising.

As mentioned above, TikTok does not allow paid political advertising on our platform.

Objective 6: Strengthen public understanding of Disinformation and Misinformation through support of strategic research.



Outcome 6: Signatories support the efforts of independent researchers to improve public understanding of Disinformation and Misinformation.

TikTok's [Asia Pacific Safety Advisory Council](#) was formed in 2020 to formally bring together a group of independent online safety experts to advise on issues related to trust and safety at TikTok. The members of the council are thought leaders from academia, law and government from various sectors in the region.

They hold extensive knowledge across shared topics of interest including cyber wellness, mental well-being, minor protection and countering falsehoods. We made a deliberate choice to bring on board experts with strong independent opinions and the courage to challenge our thinking. Hearing from a diverse set of voices is valuable for TikTok when crafting products, policies and processes to serve the evolving needs of our growing community.

Objective 7: Signatories will publicise the measures they take to combat Disinformation.

Outcome 7: The public can access information about the measures Signatories have taken to combat Disinformation and Misinformation.

We understand that a copy of this report will be made available online, and we will continue to publicise our efforts in identifying and removing misinformation from our platform.

The H2 2020 TikTok Transparency Report, including our COVID-19 Safety Centre is available at: <https://www.tiktok.com/safety/resources/transparency-report-2020-2>

A full range of articles, blogposts and reports relating to our safety initiatives are available on the TikTok website, through our newsroom at: <https://newsroom.tiktok.com/en-au>.



Concluding remarks

Although TikTok is only a recent entrant to the Australian market, we have established robust policies and tools to address the spread of misinformation. But we know this work never stops.

Our most important commitment at TikTok is to keep our community safe. Content that is intended to mislead or deceive any of our users endangers our trust-based community.

We're tackling misinformation by investing heavily in technology and review teams, introducing in-app features, promoting trusted information from authoritative sources and developing policies to prevent its spread.

We place great importance on the measures we can take to support platform integrity so TikTok continues being a home where compelling, authentic content can thrive.