**Disguised Content**

Despite content regulation efforts, some individuals and organisations have used coercive and nefarious tactics to share prohibited imagery and videos online, especially aimed at those most vulnerable, children. Disguised content refers to harmful content that seeks to evade online content regulation or filtering mechanisms.

**TikTok, YouTube and ‘Splicing’**

In January 2021, users of TikTok, one of the world’s most popular social media sites, were exposed to a viral video containing a graphic depiction of bodily mutilation and the recorded death of an individual.[[1]](#footnote-1) Although TikTok strictly prohibits content of this nature, the video was able to be shared globally due to the original poster ‘splicing’ the clip of the violent crime behind a video of a girl dancing – allowing the upload to ‘trick’ the application’s artificial intelligence algorithm into thinking that it was a normal video.

In 2019, parents became aware of a similar issue when YouTube hosted popular videos that at first showed children’s content such as Peppa Pig that were found to be spliced with videos of a disfigured humanoid character called ‘Momo’ repeating inappropriate phrases and showing videos of cartoon characters being tortured.[[2]](#footnote-2) Another example involving YouTube involved children being exposed to spliced videos that gave explicit instructions on how to inflict self-harm.[[3]](#footnote-3)

Other manners of disguising content have begun to emerge. In August 2024, a new trend arose that involved nefarious creators uploading videos where the first few seconds appear relatively normal and fun before they begin to video themselves exposing themselves on camera or showing disturbing content.[[4]](#footnote-4)

To learn more, click [here](https://www.youtube.com/watch?v=o9p7XHkY6uE) to view a video essay on this topic by Visual Venture.

**Disguised content and the Online Safety Act**

In Australia the primary instrument that instructs online content regulation is the *Online Safety Act 2021* (Cth). This Act works to classify online content in conjunction with the National Classification Code into class 1 material and class 2 material under section 106 and 107. The Act utilises these classification systems as a way of directing enforcement, with sections 109-128 stipulating that the publication of harmful material against the regulations of the Act be followed by orders of deletion or the removal of content or platforms in extreme circumstances. However, these measures have been criticised for focusing too heavily on content removal rather than prevention of publication.[[5]](#footnote-5)

Disguised content is often class 1 material. Due to creators evading the regulatory algorithms on platforms like YouTube, these videos remain online until a substantial number of reports are made, and the videos are manually reviewed or flagged as class 1 material, by which point these videos have been widely circulated, causing harm. Exacerbating this issue is that the current laws are limited by geographical jurisdiction and therefore cannot effectively govern ‘cloud computing’ or the uploading or creation of disguised content from locations outside of Australia particularly with the more prevalent use of proxy servers such as VPNs that can hide the location of a poster.[[6]](#footnote-6)

Another issue around enforcement is that creators and uploaders of disguised content often are anonymous, undetectable and in viral circumstances, numerous due to re-uploads. This is also problematic for enforcing the new *Criminal Code Amendment (Sharing of Abhorrent Violent Material) Act* 2019 (Cth) provisions that criminalises the acts of uploading harmful material.

1. Brian Stieglitz, ‘Beheading Video Goes Viral on TikTok, Highlights Security Concerns’, Mail Online (9 June 2021) https://www.dailymail.co.uk/news/article-9669245/TikTok-apologizes-beheading-clip-tricks-AI-server-posing-dance-video-goes-viral.html. [↑](#footnote-ref-1)
2. Keza MacDonald, ‘Parents: Don’t Panic about Momo – Worry about YouTube Kids Instead | Keza MacDonald’, The Guardian (online, 28 February 2019) https://www.theguardian.com/commentisfree/2019/feb/28/parents-momo-scare-youtube-kids. [↑](#footnote-ref-2)
3. Doug Criss, ‘A Mom Found Videos on YouTube Kids That Gave Children Instructions for Suicide’, CNN (19 February 2019) https://edition.cnn.com/2019/02/25/tech/youtube-suicide-videos-trnd/index.html. [↑](#footnote-ref-3)
4. Anna Good, ‘What Is the Gorilla Mask Video on TikTok?’, The Daily Dot (5 August 2024) https://www.dailydot.com/memes/gorilla-mask-video-tiktok/. [↑](#footnote-ref-4)
5. Alex McIntosh, ‘Why Content Takedown Laws Aren’t the Only Answer to Reducing Harmful Content Online’, Medium (29 April 2021) https://medium.com/ausreset/why-content-takedown-laws-arent-the-only-answer-to-reducing-harmful-content-online-9da25dd293a8. [↑](#footnote-ref-5)
6. Dan Svantesson and Roger Clarke, ‘Privacy and Consumer Risks in Cloud Computing’ (2010) 26(4) *Computer Law & Security Review* 391 at [393]. [↑](#footnote-ref-6)