

# Anxieties over livestreams can help us design better Facebook and YouTube content moderation

March 19 2019, by Andrew Quodling

---



Livestream on Facebook isn't just a tool for sharing violence – it has many popular social and political uses. Credit: glen carrie / unsplash, CC BY

As families in Christchurch [bury their loved ones](#) following Friday's terrorist attack, global attention now turns to preventing such a thing ever happening again.

In particular, the role [social media](#) played in broadcasting live footage

and amplifying its reach is under the microscope. Facebook and YouTube face intense scrutiny.

New Zealand's Prime Minister Jacinta Ardern has [reportedly been in contact](#) with Facebook executives to press the case that the footage should not be available for viewing. Australian Prime Minister Scott Morrison has called for a [moratorium on amateur livestreaming services](#).

But beyond these immediate responses, this terrible incident presents an opportunity for longer term reform. It's time for [social media platforms](#) to be more open about how livestreaming works, how it is moderated, and what should happen if or when the rules break down.

## **Increasing scrutiny**

With the alleged perpetrator apparently flying under the radar prior to this incident in Christchurch, our collective focus is now turned to the [online radicalisation of young men](#).

As part of that, [online platforms](#) face increased scrutiny and Facebook and Youtube have drawn criticism.

After dissemination of the original livestream occurred on Facebook, YouTube became a venue for the re-upload and propagation of the recorded footage.

Both platforms have made public statements about their efforts at moderation.

YouTube noted the challenges of dealing with an "[unprecedented volume](#)" of uploads.

Although [it's been reported](#) less than 4000 people saw the initial stream

on Facebook, [Facebook said](#): "In the first 24 hours we removed 1.5 million videos of the attack globally, of which over 1.2 million were blocked at upload [...]"

Focusing chiefly on live-streaming is somewhat reductive. Although the shooter initially streamed his own footage, the greater challenge of controlling the video largely relates to two issues:

1. the length of time it was available on Facebook's [platform](#) before it was removed
2. the moderation of "mirror" video publication by people who had chosen to download, edit, and re-upload the video for their own purposes.

These issues illustrate the weaknesses of existing content moderation policies and practices.

## **Not an easy task**

Content moderation is a complex and unenviable responsibility. Platforms like Facebook and YouTube are expected to balance the virtues of free expression and newsworthiness with socio-cultural norms and personal desires, as well as the local regulatory regimes of the countries they operate in.

When platforms perform this responsibility poorly (or, utterly abdicate it) they pass on the task to others—like the New Zealand Internet Service Providers that [blocked access to websites that were re-distributing the shooter's footage](#).

People might reasonably expect platforms like Facebook and YouTube to have thorough controls over what is uploaded on their sites. However, the companies' huge user bases mean they often must balance the

application of automated, algorithmic systems for content moderation (like Microsoft's [PhotoDNA](#), and YouTube's [ContentID](#)) with teams of human moderators.

We know from [investigative reporting that the moderation teams](#) at platforms like Facebook and YouTube are tasked with particularly challenging work. They seem to have a relatively high turnover of staff who are quickly burnt-out by severe workloads while moderating the [worst content on the internet](#). They are supported with only meagre wages, and what could be viewed as [inadequate mental healthcare](#).

And while some algorithmic systems can be effective at scale, they can also be subverted by competent users who understand aspects of their methodology. If you've ever found a video on YouTube where the colours are distorted, the audio playback is slightly out of sync, or the image is heavily zoomed and cropped, you've likely seen someone's attempt to get around ContentID algorithms.

For online platforms, the response to terror attacks is further complicated by the difficult balance they must strike between their desire to protect users from gratuitous or appalling footage with their commitment to inform people [seeking news through their platform](#).

We must also acknowledge the other ways livestreaming features in modern life. Livestreaming is a [lucrative niche entertainment industry](#), with thousands of innocent users broadcasting hobbies with friends from [board games](#) to [mukbang \(social eating\)](#), to [video games](#). Livestreaming is important for [activists in authoritarian countries](#), allowing them to share eyewitness footage of crimes, and shift power relationships. A ban on livestreaming would prevent a lot of this activity.

## **We need a new approach**

Facebook and YouTube's challenges in addressing the issue of livestreamed hate crimes tells us something important. We need a more open, transparent approach to moderation. Platforms must talk openly about how this work is done, and be prepared to incorporate feedback from our governments and society more broadly.

A good place to start is [the Santa Clara principles](#), generated initially from a content moderation conference held in February 2018 and updated in May 2018. These offer a solid foundation for reform, stating:

- companies should publish the numbers of posts removed and accounts permanently or temporarily suspended due to violations of their content guidelines
- companies should provide notice to each user whose content is taken down or account is suspended about the reason for the removal or suspension
- companies should provide a meaningful opportunity for timely appeal of any content removal or account suspension.

A more socially responsible approach to platforms' roles as moderators of public discourse necessitates a move away from the black-box secrecy platforms are accustomed to—and a move towards more thorough public discussions about content moderation.

In the end, greater transparency may facilitate a less reactive policy landscape, where both public policy and opinion have a greater understanding around the complexities of managing new and innovative communications technologies.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

## Provided by The Conversation

Citation: Anxieties over livestreams can help us design better Facebook and YouTube content moderation (2019, March 19) retrieved 27 April 2024 from

<https://phys.org/news/2019-03-anxieties-livestreams-facebook-youtube-content.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.