

# Facebook turns to real people to fix its violent video problem

May 4 2017, by Jennifer Beckett

---



Credit: AI-generated image ([disclaimer](#))

Facebook has recently been [under fire](#) for not doing enough to keep disturbing content out of our newsfeeds. It hopes a hiring spree will fix the problem.

In [a Facebook post](#) on Wednesday, company founder and CEO Mark

Zuckerberg announced plans to add an additional 3,000 moderators to its community operations team. These new employees will help review posts that are flagged as troubling by the community and "improve the process for doing it quickly".

Given the recent spate of news stories about Facebook Live being used to stream everything from [rapes](#) to [murders](#) in real time, it's fair to say the [social media](#) giant has some work to do.

## Algorithms fall short

So why are people necessary when algorithms can do the job?

Facebook does use software to filter some types of explicit or illegal content, but there are limits to its capabilities. That's where humans come in.

At the 2015 SWARM conference for community managers, Mia Garlick, Facebook's director of policy for Australia and New Zealand, noted the social media [company uses Microsoft's PhotoDNA software](#) to weed out known images of child pornography posted onto the platform. And therein lies the rub – the program must already "know" the images to delete them.

Algorithms are only as good as we teach them to be. In the case of something like PhotoDNA, someone – a human – has to find and add any additional images to the database that sits behind the software.

These automated systems can also be blunt tools. They see the world as black and white, so they're not always good at deciding when something falls into a grey area – say, making decisions on whether that nude photo is a piece of art or just a nude photo.

This issue also arises when cultural context or local law is at odds with a Facebook post.

In [a 2015 blog post](#) explaining Facebook's community guidelines, Monika Bickert, Facebook's head of global policy management, and Chris Sonderby, deputy general counsel, touched on the vexed notion of blasphemy to demonstrate how hard it can be to moderate language and imagery in different countries.

The update followed [a news story](#) about Facebook blocking a page in Turkey after a court there ruled it was insulting the Prophet Mohammed.

"While blasphemy is not a violation of the Community Standards, we will still evaluate the reported content and restrict it in that country if we conclude it violates local law," Bickert and Sonderby explained in the post.

In other words, we need humans to catch anything that falls between the cracks. That's why Facebook relies on its human community to report content and flagged posts are still screened by its team of human moderators.

## **Moderators are people, too**

Good moderation is at the heart of building a safe space online, and Zuckerberg's [stated commitment](#) to building a better Facebook community is an important first step.

Still, one thing is often forgotten in this discussion: the people whose labour ensures the rest of us have a pleasant experience online.

Take Microsoft, creators of that PhotoDNA software used by Facebook, and its online safety team. In December 2016, two former employees of

Microsoft's team [sued the technology company](#) after allegedly developing post traumatic stress disorder.

According to the suit, the pair were responsible for not only reviewing video and imagery that would end up being part of the algorithm's "learning", but also for working with law enforcement.

In particular, the lawsuit alleges the two employees were not told of the possible impacts of their work, did not receive sufficient training despite the traumatic nature of their work, and did not receive enough support from Microsoft.

If the lawsuit is successful, it could have serious ramifications for any company employing content moderators. That could be anything from banks to the media and, of course, Facebook and its 3,000 new moderators.

These companies have a moral duty to support employees who may be routinely exposed to rape and torture imagery, though recent news reports suggest that's not always the case.

In December, Germany's Süddeutsche Zeitung newspaper [found Facebook's Berlin moderators facing dismaying conditions](#). The journalists reported employees were viewing violent and explicit material daily with little support while making "barely over minimum wage". Facebook declined to comment to the publication.

Whether Facebook's new influx of moderators will create the "global community that works for everyone" Zuckerberg [outlined in his recent manifesto](#) remains to be seen. With billions of pieces of content being posted to Facebook everyday, it's an almost impossible task.

This article was originally published on [The Conversation](#). Read the

[original article](#).

Provided by The Conversation

Citation: Facebook turns to real people to fix its violent video problem (2017, May 4) retrieved 21 July 2024 from <https://phys.org/news/2017-05-facebook-real-people-violent-video.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.