

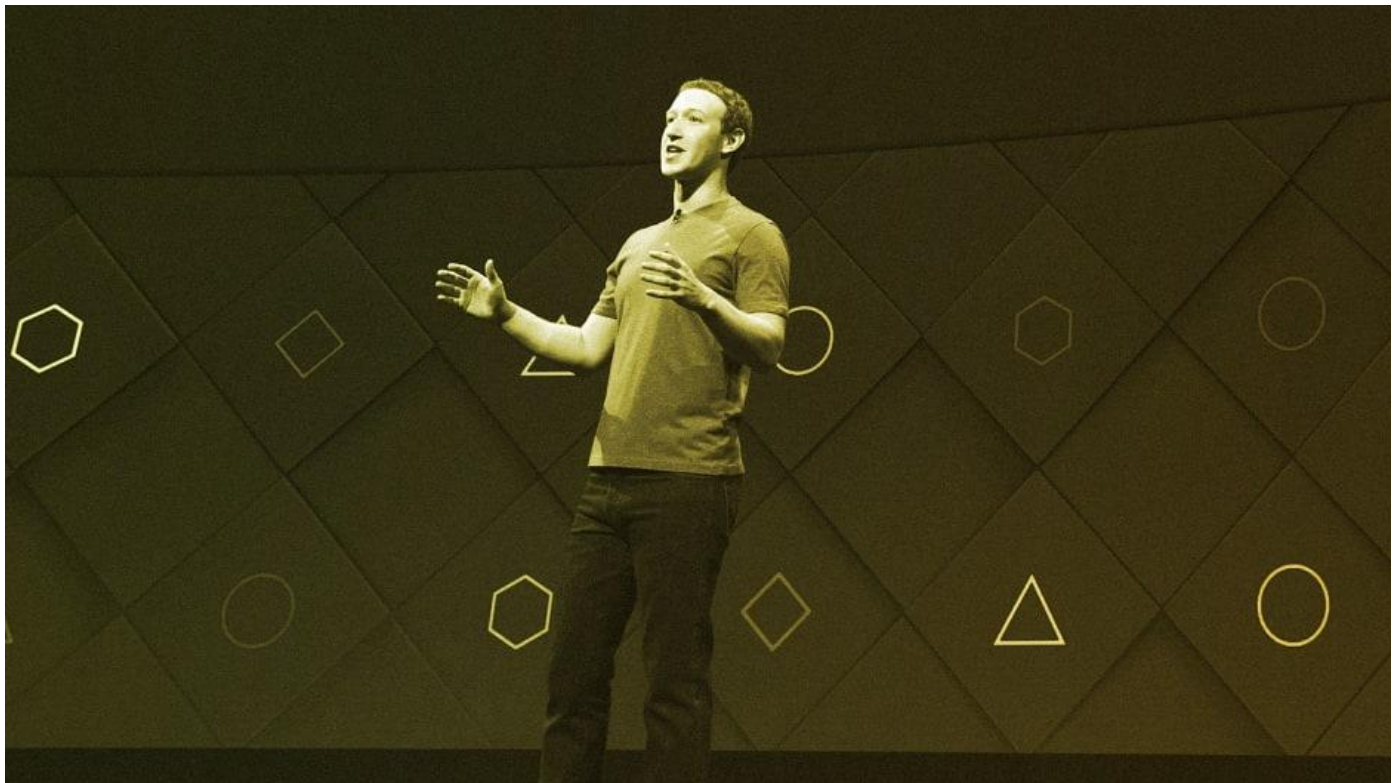


- Technology
- Leadership
- Entertainment
  - Ideas
  - Video
  - News

04.20.17 | MOST INNOVATIVE COMPANIES

## What Facebook Could Learn From Other Platforms About Content Moderation

While Facebook's size makes it hard to effectively filter out violent content, some other platforms have had more luck at keeping the demons at bay.



[Photo: courtesy of Facebook]



**BY CALE GUTHRIE WEISSMAN**

5 MINUTE READ

Facebook has a team of 60 engineers working on a [brain-computer interface that will someday scan your brain](#) a hundred times per second and translate your inner voice into text. Its Building 8 team is researching a [way for humans to hear through their skin](#). And it's reportedly developing eyeglasses and contact lenses [that will bring augmented reality to the masses](#).

But they still can't figure out how to prevent users from abusing the platform by uploading and sharing violent videos, such as Steve Stephens's horrific murder of an elderly man in Cleveland last weekend (it took Facebook three hours to take down the video). And judging by the complications involved in such an effort, and similar attempts to tackle that problem by other social networks, it will take some time, as CEO [Mark Zuckerberg acknowledged at F8 this week](#).

## Why Facebook Should Amp Up Its Content Moderation

The platform could learn from other sites about how to filter out violent content.

Facebook's content moderation system relies on users who flag offending posts, which are then passed on to an army of moderators who make the determination of whether or not they're suitable for the site. Though it has thousands of human moderators, that's hardly enough to handle the swarm of content uploaded by close to 2 billion monthly active users. That is supplemented by artificial intelligence "to understand more quickly and accurately what is happening across our community," but that could take years to work effectively.

### OTHER PLATFORMS, OTHER SOLUTIONS

Other social-media platforms, though much smaller than Facebook, have had varying degrees of success tackling the problem. YouNow, a social live-streaming platform that's primarily targeted at teens, has a fairly robust content moderation system. The site has been around for quite a few years, and given its user demographic it needs to be extra careful with everything it uploads. One thing that YouNow has, which Facebook lacks, is a cohesive user culture and set of values that makes the site a "safe space." Those who use YouNow are part of a positive community that they don't want sullied. This was created, in part, thanks to strict content moderation that is quick to take down any abusive content. The founders used a mixture of community participation and technology to try and ensure that only kosher content was presented to the teens' eyes from the get-go.

Obviously Facebook can't adopt this solution since it lacks a unifying culture due to its sheer enormity as a global media juggernaut. (For context, last summer [YouNow said](#) that it had 150,000 live streams uploaded per day; Facebook [announced](#) in 2015 that it hit 8 billion daily video views per *day*.)

Still, YouNow does employ a [very thorough system](#) to make sure that inappropriate content isn't seen by minors. Like Facebook, users can flag offensive videos. The site also employs proprietary technology to filter out such content—both videos flagged by users as well as those detected via its algorithmic system. It does that partly by analyzing comments in real time, which helps detect content that should be flagged.

YouTube has been dabbling with a similar community-based moderation system. Called "YouTube Heroes," it empowers an elite group of users to police the site and help weed out the flood of horrific videos that you can imagine gets uploaded every second to the popular platform. Similarly, Reddit has been [working to curb](#) its well-known abuse and troll problem by rewriting offending posts and having moderators enforce a new content policy.

Instagram too is known for its heavy-handed content moderation, via such methods as blocking the hashtags for groups and communities that share potentially harmful content. Through this method, it has successfully blocked content that promotes eating disorders. How Instagram decides which hashtags to block remains a mystery. This is something many social media

researchers have [looked into](#), but given that hashtags are a prime way for users to connect, these blocks help reduce the flood of potentially abusive content.

On the other end of the spectrum there's Twitter, which has been plagued with abusive users for years. True, the company has been [doubling down on the problem](#), with a slew of new tools aimed at giving users more power to combat offending content and trolls. But when it comes to live-streaming video, users might have a harder time. According to Periscope's [terms of service](#):

*You understand that by using the Services, you may be exposed to Content that might be offensive, harmful, inaccurate or otherwise inappropriate, or in some cases, postings that have been mislabeled or are otherwise deceptive. All Content is the sole responsibility of the person who originated such Content. We may not monitor or control the Content posted via the Services and, we cannot take responsibility for such Content.*

## WAITING FOR AI TO GET SMARTER

The real question that stands out in this debate is how long it will take artificial intelligence to advance to the point that it's able to distinguish between a remake of *Pulp Fiction* filmed by some college freshmen and a real-life murder. YouNow is able to filter out such content through a blend of human moderation and technology, but that becomes much more difficult at the scale of Facebook. But the technology has advanced by leaps and bounds in recent years. Facebook told *Wired* earlier this week that [half of the content it flags comes from its AI program](#). The Cleveland video fell through the cracks since it wasn't flagged by humans immediately. Once it was flagged, says Facebook, it took a little over 20 minutes for it to be taken down.

Since such technological advancements could take years to implement, Facebook will have to adopt more human moderation and other methods—including the idea of making it more difficult to download videos to prevent users from quickly sharing them on third-party sites like Live Leak. But maybe it can adopt some of the solutions that have been used by other platforms or even partnering with them to collaborate on technologies that tackle the problem. If Zuckerberg is all about “connecting the world,” maybe one place to start is by partnering with allies in a shared campaign to keep out violent and hateful content.

For now, Facebook will be faced with a choice that cuts to the heart of its purpose: Does it want to create a safer space for users by implementing more rigorous safeguards, or does it favor a free space where there are only limited controls so that users can still upload their content in just seconds? Whatever the company says, it's going to be years until it can have both.

---

### ABOUT THE AUTHOR

Cale is a Brooklyn-based reporter. He writes about business, technology, leadership, and anything else that piques his interest. [More](#)

---

## Technology Newsletter

YOUR EMAIL ADDRESS

**SIGN UP**

☒ Receive special Fast Company offers.

[See All Newsletters](#)

## VIDEO

This ad will end in 0

### La La Anthony On How A Fake ID (And Hard Work) Launched Her Career



### La La Anthony On How A Fake ID (And Hard Work) Launched Her Career

ADVERTISEMENT

01.08.18 | 12:19 PM

## NBC wants you to know it hasn't endorsed Oprah for president



Last night, NBC took a beating from conservative Twitter users after a tweet during the Golden Globes was perceived as an endorsement of a speculative presidential run by Oprah Winfrey. (Let's all take a moment to absorb that sentence and how completely un-crazy it sounds in 2018.)

Winfrey [gave a rousing speech](#) last night as she accepted the Cecil B. DeMille Award for lifetime achievement, fueling a flurry of social media chatter about a possible 2020 White House bid. **But not everyone thinks it's a great idea, and some of those critics tweeted their disapproval at the folks at NBC for ostensibly declaring themselves Team Oprah.** The offending tweet—which came from the Peacock Network's [main account](#)—said, “Nothing but respect for OUR future president.” It included a GIF of the former queen of daytime talk.

NBC has since deleted the tweet, saying it was sent by a third-party social media agency and “not meant as a political statement.”



**NBC**  
@nbc

Yesterday a tweet about the Golden Globes and Oprah Winfrey was sent by a third party agency for NBC Entertainment in real time during the broadcast. It is in reference to a joke made during the monologue and not meant to be a political statement. We have since removed the tweet.

9:09 PM - Jan 8, 2018

4,322

1,067

2,250

Make of that explanation what you will, but let's not forget this is the same network that took a lot of flak for letting Donald Trump host *SNL* in 2015. Also, it's probably worth mentioning that Oprah's talk show was distributed by CBS. —cz

### News Newsletter

YOUR EMAIL ADDRESS

**SIGN UP**

☒ Receive special Fast Company offers.

[See All Newsletters](#)

## IDEAS

### IDEAS

If You Find Yourself Automated Out Of A Job, Blame The New Tax Plan

IDEAS

The Growing Movement To Create City-Run Public Banks

---

IDEAS

This New Blockchain Protocol Wants To Create Accountability For Social Impact

---

## ENTERTAINMENT

---

ENTERTAINMENT

Nine #MeToo Moments That Defined The Golden Globes

---

ENTERTAINMENT

Why A Small Minneapolis Fashion Brand Is Making Climate Change Part Of Its Business

---

ENTERTAINMENT

Political Statements, Black Dresses, And Surprise Wins: What To Expect At The Golden Globes This Sunday

---

## CO.DESIGN

---

GRAPHICS

Coke's New Font Is Design At Its Worst

---

CITIES & SPACES

These Enchanting Fireflies Are Actually A Robot Swarm

---

CITIES

Google Street View Now Has A Soundtrack, Thanks To AI

---

## FAST COMPANY

---

NEWS

Facebook just signed another music licensing deal—this time with Sony/ATV

---

IDEAS

If You Find Yourself Automated Out Of A Job, Blame The New Tax Plan

---

NEWS

Here's why GoPro finally pulled the plug on its disastrous drone

---