



Official Blog

Broadcast Yourself

Expanding our work against abuse of our platform

Monday, December 4, 2017

As the CEO of YouTube, I've seen how our open platform has been a force for creativity, learning and access to information. I've seen how activists have used it to advocate for social change, mobilize protests, and document war crimes. I've seen how it serves as both an entertainment destination and a video library for the world. I've seen how it has expanded economic opportunity, allowing small businesses to market and sell their goods across borders. And I've seen how it has helped enlighten my children, giving them a bigger, broader understanding of our world and the billions who inhabit it.

But I've also seen up-close that there can be another, more troubling, side of YouTube's openness. I've seen how some bad actors are exploiting our openness to mislead, manipulate, harass or even harm.

In the last year, we took actions to protect our community against violent or extremist content, testing new systems to combat emerging and evolving threats. We tightened our policies on what content can appear on our platform, or earn revenue for creators. We increased our enforcement teams. And we invested in powerful new machine learning technology to scale the efforts of our human moderators to take down videos and comments that violate our policies.

Now, we are applying the lessons we've learned from our work fighting violent extremism content over the last year in order to tackle other problematic content. Our goal is to stay one step ahead of bad actors, making it harder for policy-violating content to surface or remain on YouTube.

More people reviewing more content

Human reviewers remain essential to both removing content and training machine learning systems because human judgment is critical to making contextualized decisions on content. Since June, our trust and safety teams have manually reviewed nearly 2 million videos for violent extremist content, helping train our machine-learning technology to identify similar videos in the future. We are also taking aggressive action on comments, launching new comment moderation tools and in some cases shutting down comments altogether. In the last few weeks we've used machine learning to help human reviewers find and terminate hundreds of accounts and shut down hundreds of thousands of comments. Our teams also work closely with NCMEC, the IWF, and other child safety organizations around the world to report predatory behavior and accounts to the correct law enforcement agencies.

We will continue the significant growth of our teams into next year, with the goal of bringing the total number of people across Google working to address content that might violate our policies to over 10,000 in 2018.

At the same time, we are expanding the network of academics, industry groups and subject matter experts who we can learn from and support to help us better understand emerging issues.

Tackling issues at scale

We will use our cutting-edge machine learning more widely to allow us to quickly and

[Labels](#)
[Archive](#)
[Feed](#)
YouTube on
[Follow @youtube](#)
[Follow](#)
[Learn more in the YouTube Help Center.](#)

efficiently remove content that violates our guidelines. In June we deployed this technology to flag violent extremist content for human review and we've seen tremendous progress.

- Since June we have removed over 150,000 videos for violent extremism.
- Machine learning is helping our human reviewers remove nearly five times as many videos than they were previously.
- Today, 98 percent of the videos we remove for violent extremism are flagged by our machine-learning algorithms.
- Our advances in machine learning let us now take down nearly 70 percent of violent extremist content within eight hours of upload and nearly half of it in two hours and we continue to accelerate that speed.
- Since we started using machine learning to flag violent and extremist content in June, the technology has reviewed and flagged content that would have taken 180,000 people working 40 hours a week to assess.

Because we have seen these positive results, we have begun training machine-learning technology across other challenging content areas, including [child safety](#) and hate speech.

Greater transparency

We understand that people want a clearer view of how we're tackling problematic content. Our [Community Guidelines](#) give users notice about what we do not allow on our platforms and we want to share more information about how these are enforced. That's why in 2018 we will be creating a regular report where we will provide more aggregate data about the flags we receive and the actions we take to remove videos and comments that violate our content policies. We are looking into developing additional tools to help bring even more transparency around flagged content.

A new approach to advertising on YouTube

We're also taking actions to protect advertisers and creators from inappropriate content. We want advertisers to have peace of mind that their ads are running alongside content that reflects their brand's values. Equally, we want to give creators confidence that their revenue won't be hurt by the actions of bad actors.

We believe this requires a new approach to advertising on YouTube, carefully considering which channels and videos are eligible for advertising. We are planning to apply stricter criteria, conduct more manual curation, while also significantly ramping up our team of ad reviewers to ensure ads are only running where they should. This will also help vetted creators see more stability around their revenue. It's important we get this right for both advertisers and creators, and over the next few weeks, we'll be speaking with both to hone this approach.

We are taking these actions because it's the right thing to do. Creators make incredible content that builds global fan bases. Fans come to YouTube to watch, share, and engage with this content. Advertisers, who want to reach those people, fund this creator economy. Each of these groups is essential to YouTube's creative ecosystem—none can thrive on YouTube without the other—and all three deserve our best efforts.

As challenges to our platform evolve and change, our enforcement methods must and will evolve to respond to them. But no matter what challenges emerge, our commitment to combat them will be sustained and unwavering. We will take the steps necessary to protect our community and ensure that YouTube continues to be a place where creators, advertisers, and viewers can thrive.


Susan Wojcicki, CEO of YouTube



No comments :

Post a Comment

Enter your comment...

 Comment as: Google Accou ▾

[Publish](#) [Preview](#)

Links to this post

[Create a Link](#)



YouTube

- [About YouTube](#)
- [Press & Blog](#)
- [Copyright](#)
- [Creators & Partners](#)
- [Advertising](#)
- [Developers](#)
- [Help](#)

More Blogs

- [YouTube Creator Blog](#)
- [YouTube Engineering and Developers Blog](#)
- [YouTube Trends Blog](#)
- [Google Blog](#)
- [More blogs from Google](#)

