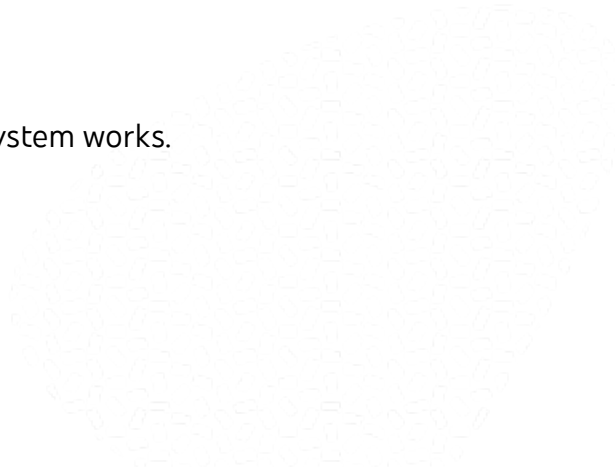INSIDE YOUTUBE

On YouTube's recommendation system

By Cristos Goodrow

VP of Engineering At YouTube

Sep.15.2021

A deeper look into how YouTube's recommendation system works.

When YouTube's recommendations are at their best, they connect billions of people around the world to content that uniquely inspires, teaches, and entertains. For me, that means diving into lectures exploring the ethical questions facing technology today or watching highlights from the University of Southern California football games I remember seeing as a kid. For my oldest daughter, it was finding laughter and community with the Vlogbrothers. And for my oldest son, recommendations brought about a better understanding of linear algebra through animated explainers by 3Blue1Brown—with breaks to watch KSI videos.

Recommendations drive a significa
amount of the o

YouTube, even more than channel subscriptions or search.

As my family shows, there's an audience for almost every video, and the job of our recommendation system is to find that audience. Think about how hard it would be to navigate all of the books in a massive library without the help of librarians. Recommendations drive a significant amount of the overall viewership on YouTube, even more than channel subscriptions or search. I've spent over a decade at YouTube building our recommendation system and I'm proud to see how it's become an integral part of everyone's YouTube experience. But all too often, recommendations are seen as a mysterious black box. We want these systems to be publicly understood, so let me explain how they work, how they've evolved, and why we've made delivering responsible recommendations our top priority.
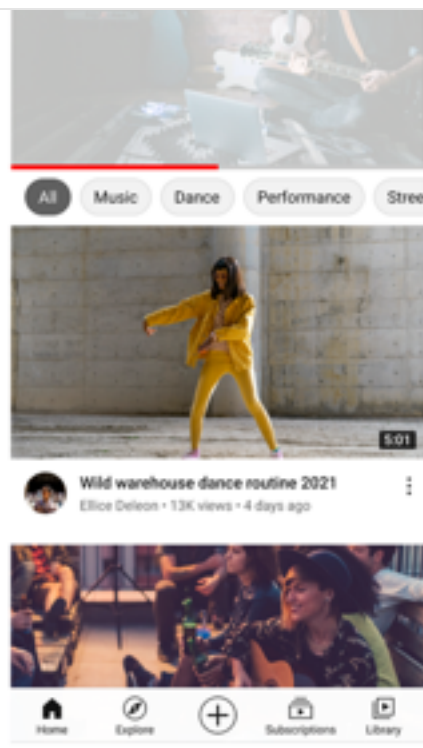
What is a recommendation system?
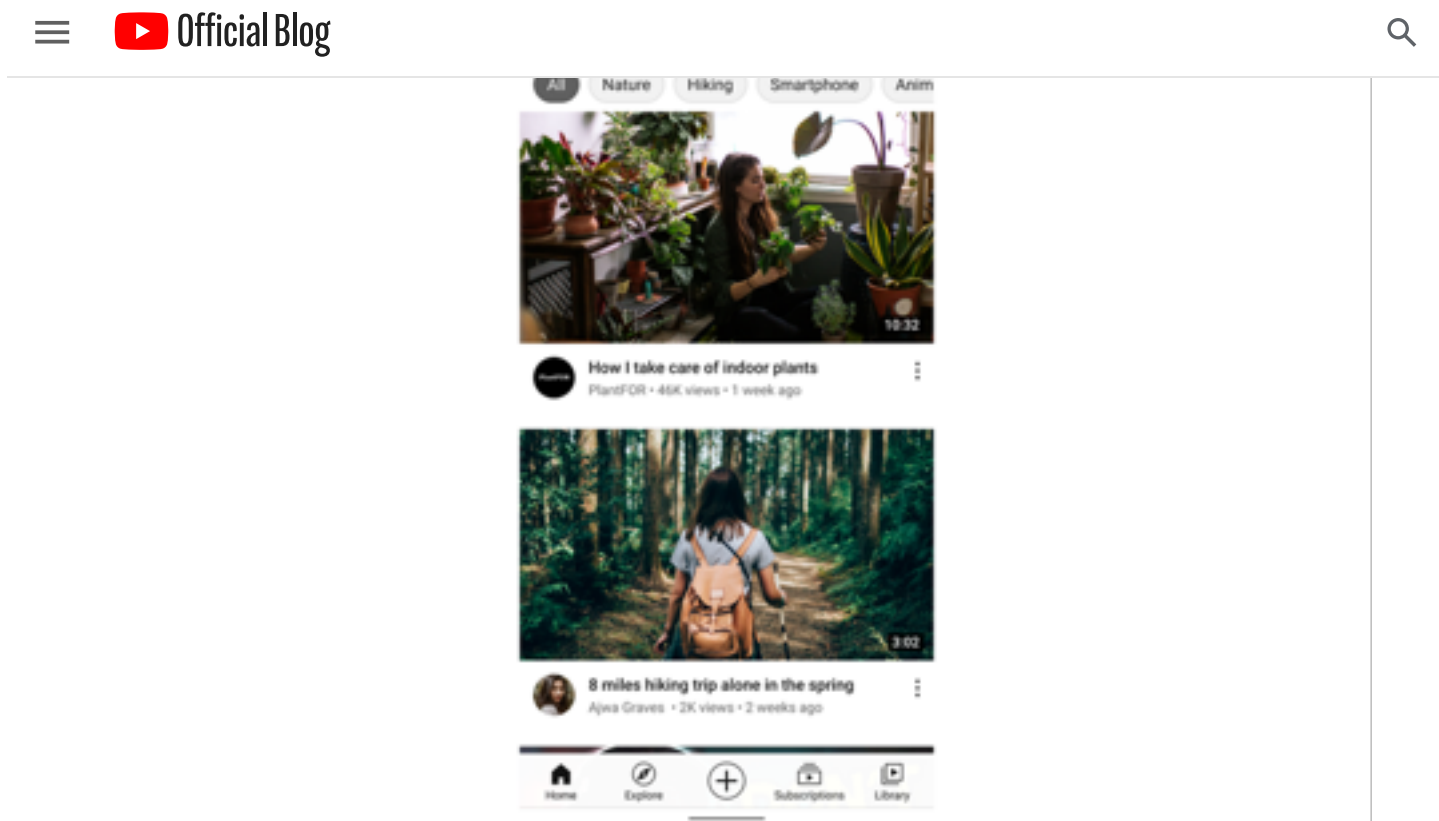
Recommendations on homepage.

Our recommendation system is built on the simple principle of helping people find the videos they want to watch and that will give them value. You can find recommendations at work in two main places: your homepage and the "Up Next" panel. Your homepage is what you see when you first open YouTube—it displays a mixture of personalized recommendations, subscriptions, and the latest news and information. The Up Next panel appears when you're watching a video and suggests additional content based on what you're currently watching, alongside other videos that we think you may be interested in.

Back in 2008, when we first started building our recommendation system, the experience was entirely different. Let's say you mostly watch cooking videos. Wouldn't it be frustrating if your homepage only recommended the latest sports and music videos to you because they had the most views? That was YouTube in the early days. The system ranked videos based on popularity to create one big "Trending" page. Not a lot of people watched those videos and the majority of YouTube's viewership came from searches or shared links off the platform.
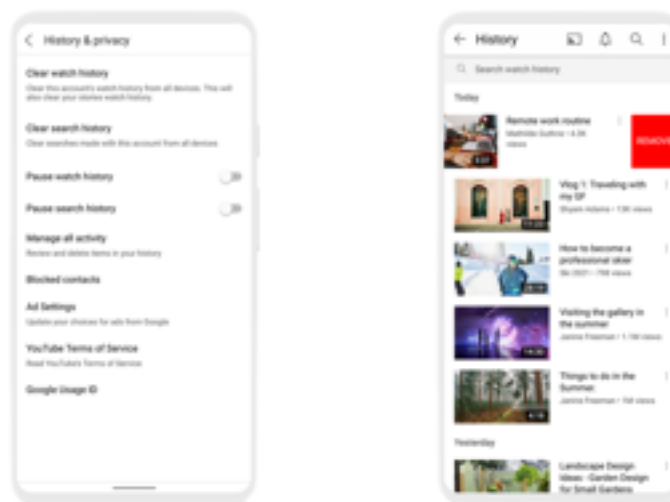
Recommendations on "Up next."

To do this, we start with the knowledge that everyone has unique viewing habits. Our system then compares your viewing habits with those that are similar to you and uses that information to suggest other content you may want to watch. So if you like tennis videos and our system notices that others who like the same tennis videos as you also enjoy jazz videos, you may be recommended jazz videos, even if you've never watched a single one before (for categories like news and information, this might function differently - more on that later). A few years ago, our system recommended videos from Tyler Oakley to my oldest daughter, because that's who many of the people who watched Vlogbrothers also watched at the time. She ended up becoming a big fan, so much so that we later took her to see him at a meet-up.

Today, our system sorts through billions of videos to recommend content tailored to your specific interests. For example, our system recognized that I watched a classic USC football highlight and found for me other sports highlights from my youth. Without recommendations, I would never have known these videos were available. Unlike other platforms, we don't connect viewers to content through their social network. Instead, the success of YouTube's recommendations depends on

But of course, we also know not everyone wants to always share this information with us. So we've built controls that help you decide how much data you want to provide. You can pause, edit, or delete your YouTube search and watch history whenever you want.

How we personalize recommendations

Recommended for You

Key Moments in YouTube's Recommendation System

2008

Launched recommendation system, ranking videos based on popularity

2011

Built classifiers to identify videos that were racy or violent and prevented them from being recommended
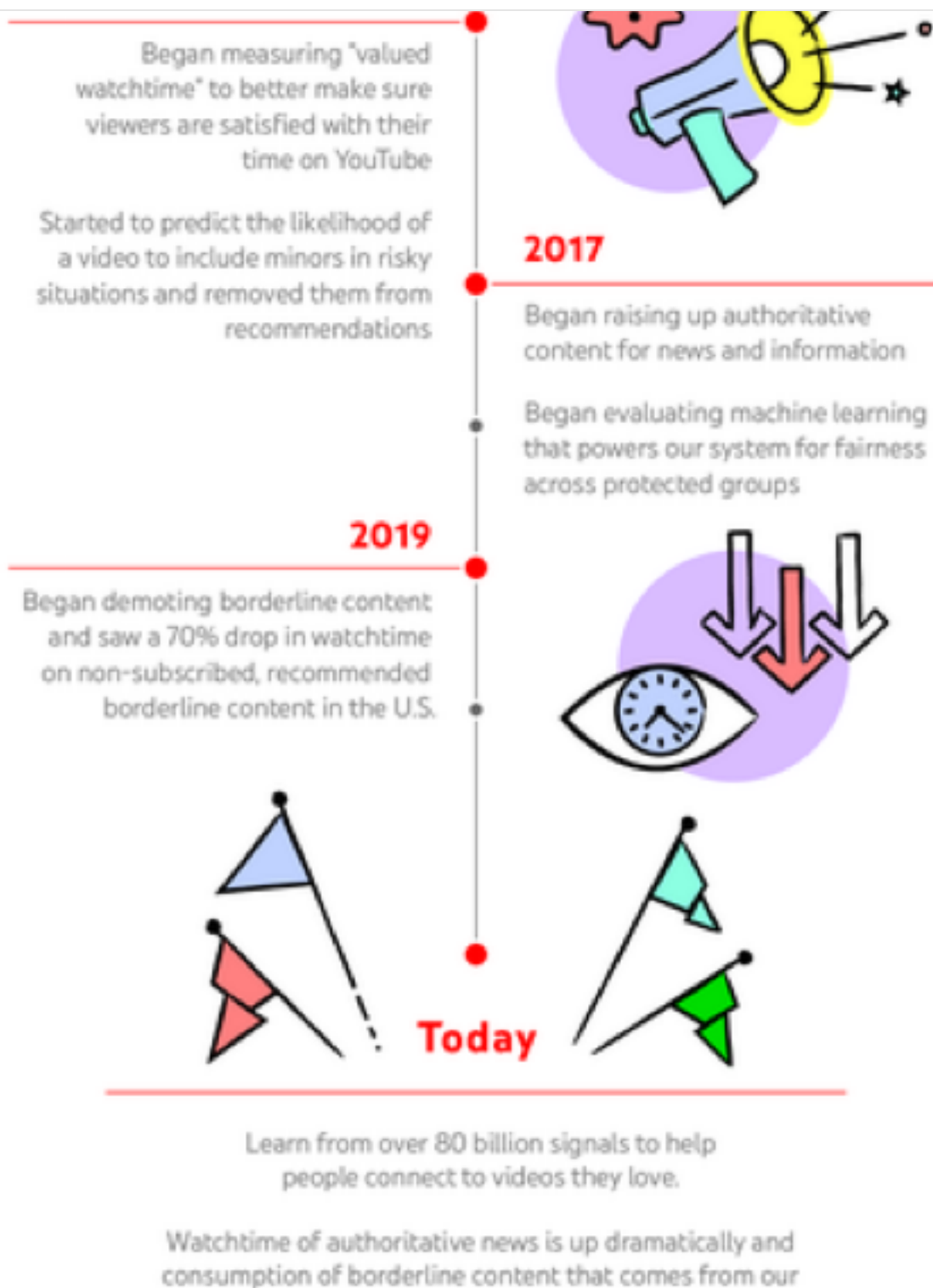
2012

Incorporated watchtime into recommendations to augment for personalization, which results in a short term drop in watchtime, but helps improve recommendation quality

2015

Demoted sensationalistic tabloid content on homepages

Began measuring "valued
watchtime" to better make sure
viewers are satisfied with their
time on YouTube

Started to predict the likelihood of
a video to include minors in risky
situations and removed them from
recommendations

2017

Began raising up authoritative
content for news and information

Began evaluating machine learning
that powers our system for fairness
across protected groups

2019

Began demoting borderline content
and saw a 70% drop in watchtime
on non-subscribed, recommended
borderline content in the U.S.

Today

Learn from over 80 billion signals to help
people connect to videos they love.

Watchtime of authoritative news is up dramatically and
consumption of borderline content that comes from our

To provide such custom curation, our recommendation system doesn't operate off
of a "recipe book" of what to do. It's constantly evolving, learning every day from over

understanding all the data that feeds into our system. A number of signals build on each other to help inform our system about what you find satisfying: clicks, watchtime, survey responses, sharing, likes, and dislikes.

- **Clicks:** Clicking on a video provides a strong indication that you will also find it satisfying. After all, you wouldn't click on something you don't want to watch.

But we learned back in 2011 that clicking on a video doesn't mean you actually watched it. Let's say you were searching for highlights from that year's Wimbledon match. You scroll through the page and click on one of the videos, which has a thumbnail and title suggesting it shows footage of the match. Instead, it's a person in their bedroom talking about the match. You click on a video our system recommends in your Up Next panel, only to find another fan talking about the match. Again and again you click through these videos until finally you're recommended a video with footage of the match that you want to watch. That's why we added in watchtime in 2012.

- **Watchtime:** Your watchtime—which videos you watched and for how long—provides personalized signals to our system about what you most likely want to watch. So if our tennis fan watched 20 minutes of Wimbledon highlight clips, and only a few seconds of match analysis video, we can safely assume they found watching those highlights more valuable.

When we first incorporated watchtime into recommendations, we saw an immediate 20% drop in views. But we believed that it was more important for us to deliver more value to viewers. Still, not all watchtime is equal. Sometimes I've stayed up late, watching random videos, when I could have instead been learning a new language on YouTube or refining my cooking skills through a tutorial. We don't want viewers regretting the videos they spend time watching and realized we needed to do even more to measure how much value you get from your time on YouTube.

- **Survey Responses:** To really make sure viewers are satisfied with the content they're watching, we measure what we call "valued watchtime"—the time spent watching a video that you consider valuable. We measure valued watchtime through user surveys that ask you to rate the video you watched from one to five stars, giving us a metric to determine how satisfying you found the content. If you rate a video one to two stars, we ask why you gave such a low rating. Similarly, if you give the video four to five stars,

Of course, not everyone fills out a survey on every video they watch. Based on the responses we do get, we've trained a machine learning model to predict potential survey responses for everyone. To test out the accuracy of these predictions, we purposely hold back some of the survey responses from the training. This way we're always monitoring how closely our system tracks with the actual responses.

- **Sharing, Likes, Dislikes:** On average, people are more likely to be satisfied by videos that they share or like. Our system uses this information to try to predict the likelihood that you will share or like further videos. If you dislike a video, that's a signal that it probably wasn't something you enjoyed watching.

Like your recommendations though, the importance of each signal depends on you. If you're the kind of person to share any video that you watch, including the ones that you rate one or two stars, our system will know not to heavily factor in your shares when recommending content. All of this is why our system doesn't follow a set formula, but develops dynamically as your viewing habits change.

A focus on responsible recommendations

Clicks, views, watchtime, user surveys, shares, likes and dislikes work great for driving recommendations for topics like music and entertainment—what most people come to YouTube to watch. But over the years, a growing number of viewers have come to YouTube for news and information. Whether it's the latest breaking news or complex scientific studies, these topics are where the quality of information and context matter most. Someone may report that they're very satisfied by videos that claim "the Earth is flat," but that doesn't mean we want to recommend this type of low-quality content.

That's why recommendations play such an importa in how we maint

They connect viewers to high-quality information and minimize the chances they'll see problematic content.

That's why recommendations play such an important role in how we maintain a responsible platform. They connect viewers to high-quality information and minimize the chances they'll see problematic content. And they complement the work done by our robust Community Guidelines that define what is and isn't allowed on YouTube.

We've used recommendations to limit low-quality content from being widely viewed since 2011, when we built classifiers to identify videos that were racy or violent and prevented them from being recommended. Then in 2015, we noticed that sensationalistic tabloid content was appearing on homepages and took steps to demote it. A year later, we started to predict the likelihood of a video to include minors in risky situations and removed those from recommendations. And in 2017, to ensure that our recommendation system was fair to marginalized communities, we began evaluating the machine learning that powers our system for fairness across protected groups—such as the LGBTQ+ community.

The rise of misinformation in recent years led us to further expand the ways we use our recommendation system to include problematic misinformation and borderline content—that is content that comes close to, but doesn't quite violate our Community Guidelines. This includes conspiracy theory videos ("the moon landing was faked") or other content that spreads misinformation ("orange juice can cure cancer").

We're able to do this by using classifiers to identify whether a video is "authoritative" or "borderline". These classifications rely on human evaluators who assess the quality of information in each channel or video. These evaluators hail from around the world and are trained through a set of detailed, publicly available rating

To determine authoritativeness, evaluators answer a few key questions. Does the content deliver on its promise or achieve its goal? What kind of expertise is needed to achieve the video goal? What's the reputation of the speaker in the video and the channel it's on? What's the main topic of the video (eg. News, Sports, History, Science, etc)? Is the content primarily meant to be satire? These answers and more determine how authoritative a video is. The higher the score, the more the video is promoted when it comes to news and information content.

Any video classified borderline is demoted in recommendations.

To determine borderline content, evaluators assess factors that include, but aren't limited to, whether the content is: inaccurate, misleading or deceptive; insensitive or intolerant; and harmful or with the potential to cause harm. The results are combined to give a score for how likely the video contains harmful misinformation or is borderline. Any video classified borderline is demoted in recommendations.

These human evaluations then train our system to model their decisions, and we now scale their assessments to all videos across YouTube.

Addressing common questions about recommendations

Recommendations play a pivotal role across our entire community, introducing

▶ Official Blog   🔍

the spread of harmful misinformation. Because while clicks, watchtime, user surveys, shares, likes and dislikes are important signals that inform our system, they can be overruled by our commitment to meeting our responsibility to the YouTube community and to society.

Top creator questions around recommendations answered.

There are a few remaining questions that I'm commonly asked about our recommendation system that I think are important to address:

1. Does borderline content get the most engagement?

Actually, through surveys and feedback, we've found that most viewers do not want to be recommended borderline content, and many find it upsetting and off-putting. In fact, when we demoted salacious or tabloid-type content we saw that watchtime actually increased by 0.5% percent over the course of 2.5 months, relative to when we didn't place any limits.

Also, we haven't seen evidence that borderline content is on average more engaging

average, flat earth videos get far fewer views. Surveys show that borderline content is satisfying to only a very small portion of viewers on YouTube. We've invested significant time and money toward making sure it doesn't find its way to broader audiences through our recommendations system. Today, borderline content gets most of its views from sources other than non-subscribed recommendations.

In 2019 we first began demoting borderline content in recommendations, resulting in a 70% drop in watchtime on non-subscribed, recommended borderline content in the U.S.

2. Does borderline content grow watchtime for YouTube?

For the vast majority of people, borderline content doesn't meet that bar of time well spent on YouTube. That's why in 2019 we first began demoting borderline content in recommendations, resulting in a 70% drop in watchtime on non-subscribed, recommended borderline content in the U.S. Today, consumption of borderline content that comes from our recommendations is significantly below 1%.

3. Do recommendations drive viewers to increasingly extreme content?

As I've explained, we actively demote low-quality information in recommendations. But we also take the additional step of showing viewers authoritative videos about topics that may interest them. Say I watch a video about the COVID-19 vaccine. In

(to the extent that our system can detect them).

Alongside those COVID-19 news and explainer videos, I'll also get personalized recommendations from other topics based on my watch history—a sketch from Saturday Night Live or a TEDx Talk about the Super Mario Effect. This personalized diversity helps viewers access new subjects and formats versus the same type of video over and over again.

A growing number of independent researchers have been looking into how tech platforms impact the consumption of borderline content—and while ongoing study continues—recent published papers conclude YouTube recommendations aren't actually steering viewers towards extreme content. Instead, consumption of news and political content on YouTube more generally reflects personal preferences that can be seen across their online habits.

4. Does borderline content make money?

To start, our advertiser-friendly guidelines already prohibit lots of borderline content from monetizing. Many advertisers have told us that they don't want to be associated with this type of content on YouTube and often choose to opt out of advertising against it. This means each borderline video watched is a lost opportunity to monetize, leading to real lost revenue to YouTube. Likewise, this kind of content breeds distrust and raises concern not just with advertising partners, but with the public, press, and policy makers. The reality is that as our work on responsibility has grown, so has our company and the entire creator economy. Responsibility is good for business.

With all that, why don't we simply remove borderline content? Misinformation tends to shift and evolve rapidly, and unlike areas like terrorism or child safety, often lacks a clear consensus. Also, misinformation can vary depending on personal perspective and background. We recognize that sometimes, this means leaving up controversial or even offensive content. So we continue to heavily focus on building responsible recommendations and take meaningful steps to prevent our system from widely recommending this content.

views of borderline content from recommendations below 0.5% of overall views on YouTube.

Taken together, all of our responsibility work around recommendations has shown real impact. Watchtime of authoritative news is up dramatically and borderline viewing is down. This doesn't mean we've solved the issues—it just means we'll need to continue refining and investing in our systems to keep improving. Our goal is to have views of borderline content from recommendations below 0.5% of overall views on YouTube.

YouTube's mission is to give everyone a voice and show them the world. It's made a tremendous difference in my own family's life. Videos that brought lessons of tolerance and empathy had a profound and positive impact on my oldest daughter's character. My son made it through some tough moments in his linear algebra class. I've learned a meaningful amount of context and nuance from lectures by leaders in technology ethics. And our commitment to openness has given rise to new voices and ideas that otherwise wouldn't have a platform. Creators like Marques Brownlee, MostlySane, or NikkieTutorials have inspired millions with their expertise, advocacy, and honesty.

Our recommendation system is getting better everyday thanks to feedback from all of you, but it can always be better. My team and I are committed to keeping that work going, and delivering you the most helpful and valuable experience possible.

Related Topics

≡   ▶ Official Blog   🔍

▶ **Official Blog**   Q

Explore the latest company news, creator and artist profiles, culture and trends analyses, and behind-the-scenes insights on the YouTube Official Blog.

Our Channels   ⌄

Twitter   ⌄

Connect   ⓘ   f

About YouTube   ⌄

YouTube Products   ⌄

For Business   ⌄

For Creators   ⌄

Our Commitments   ⌄

▶ YouTube

Policy & Safety   Copyright   Brand Guidelines   Privacy   Terms

? Help   English   ⌄

≡ ▶ Official Blog

🔍