Understanding
the Facebook
**Community Standards
Enforcement Report**

WRITTEN BY

Alex Schultz VP, Data Analytics

Guy Rosen VP, Product Management

**Preliminary
Report**

Q4 2017 – Q1 2018

CONTENTS

Overview of the Facebook Community Standards Enforcement Preliminary Report

We want to protect and respect both expression and personal safety on Facebook. Our goal is to create a safe and welcoming community for the more than 2 billion people who use Facebook around the world, across cultures and perspectives.

To help us with that goal, we maintain a detailed set of Community Standards that define what is and isn't allowed on Facebook. We don't allow anything that goes against these standards, and we invest in technology, processes and people to help us act quickly so violations of standards affect as few people as possible.

We're sharing the Community Standards
Enforcement Preliminary Report publicly for the
first time to help people understand how we're
doing at enforcing our Community Standards.
The report measures how we help to minimize
the impact of standards violations on people using
Facebook by acting against those violations. It
covers the period October 2017 to March 2018.

This guide is a companion to that report.
It provides details about the processes we use
to enforce our Community Standards and how
we measure that enforcement.

While the report shares the metrics we use to measure our effectiveness over time, this guide provides a detailed explanation about topics such as:

How we enforce our standards

Our review and enforcement processes

How we label actions so we can measure our work

The methodologies we use to produce metrics

A first step

The teams at Facebook dedicated to keeping our community safe have always prioritized enforcing our standards and keeping the impact of violations low. We've focused for many years on building and scaling our standards enforcement operation: growing and training global teams of reviewers, refining policies and processes, and developing technology to help us find and take action on violations at scale.

Historically, we've been less focused on measuring and labeling these efforts cleanly and consistently in a way that would allow us to publish reports that were reliable and comparable over time. We've also used internal measurement that focused on metrics we previously used to define our success, such as operational efficiency.

Now, as we try to help our community better understand the decisions we make and actions we take, we've evolved how we measure our success — actual views of violations by Facebook users, an indicator of the impact violations may have on our community. Additionally, over the past year, we've put more emphasis on consistently categorizing and labeling the actions we take for different violation types.

Our metrics may not be perfect, and we still have a lot of work to do to refine them (which we'll explain in this guide), but we believe they are the best representation of the work we do. We're releasing them publicly for the first time in a preliminary report to provide a first look at how we're doing at enforcing our Community Standards. We want to help our community understand our decisions and actions and get feedback on how we're doing.

Metrics in development

All metrics in this document are still in development. The way we define and measure enforcement of these standards may change as we improve our methodologies. Historical comparisons may be imperfect as metric calculations evolve.

Maintaining Our Standards

Maintaining Our Standards

We maintain a comprehensive set of Facebook [Community Standards](#) that help us encourage free, personal expression while minimizing abuse in the Facebook community.

We recently released an updated version of these standards with more detailed explanations to help people understand where we draw the line on issues.

A global content policy team at Facebook is responsible for developing our Community Standards. The team employs people in regions around the world, including subject matter experts on issues such as hate speech, child safety and terrorism.

Our team regularly seeks input from experts and organizations outside Facebook so we can better understand different perspectives as well as the impact of our standards on different communities globally.

The evolving nature of our standards

Based on this expert input, as well as changes in social norms and language, our standards continuously evolve. What doesn't ever change are the underlying principles of safety, voice and equity, which we base our Community Standards on. We aim to make Facebook a place where people can express their opinions freely, even if others don't agree. But we also know that to be open and willing to express themselves, people need to feel safe. We work to apply these standards and maintain this balance between expression and safety consistently and fairly to all communities and cultures around the world.

Our efforts to improve and refine our Community Standards also depend on participation and input from people around the world. We've introduced feedback tools in Facebook so users can give us feedback on what they think shouldn't be on Facebook, and we recently announced a new appeals process and a series of forums that will help us continuously learn and improve our standards.

Changing standards, changing metrics

Changes we make to our standards, processes and methodologies can cause both metric calculations and metric results in our reporting to change. Learn more about our measurement methodologies in the "Measuring Results" section.

Finding Standards Violations

Finding Standards Violations

To keep abusive behavior off Facebook or catch it before it impacts the community, first we have to find it.

How do we identify violations of our standards in the midst of millions of pieces of content shared on Facebook every day?

We use a combination of technology, reviews by our teams and reports from our community to identify content that might violate our standards. While not always perfect, this combination helps us find and flag potentially violating content at scale before many people see or report it. It also helps us surface content that requires more human expertise and context so we can review accurately and with sensitivity.

Reports from the Facebook community

Each week, we receive millions of reports from Facebook users around the world asking us to review content, submitted through our reporting tools. Our users also use feedback tools to indicate when they feel content shouldn't be on Facebook. This helps us to identify new and emerging concerns quickly, as well as to improve the signals we use in our technology to detect and take action against content that goes against our standards.

For some violation types, we may rely more heavily on technology to find potential violations faster and at a larger scale than user reports can help us find. For others, where technology isn't as effective, we must rely on user reports to help us find potential violations. With 2.2 billion people using Facebook every month, the size and diversity of our community means people may disagree on what should be allowed on Facebook, and we often receive user reports for content that doesn't actually violate standards or that miscategorize why a piece of content doesn't meet standards.

Detection by Facebook teams

In limited circumstances, people on our trained teams also proactively identify potential violations, focusing on harmful types of content such as terrorist propaganda. Technology can sometimes aid this process as well. When we measure how much content we subsequently acted on that we detected and flagged before users reported it, we include our detection using technology and by our people in that measurement. See section "What We Measure" for more information.

Technology

Technology helps us efficiently and proactively enforce the Community Standards by:

Helping us find and take action on violating content faster.

Content uploaded to Facebook tends to get less attention the longer it's on the site. As we've improved our enforcement, we've prioritized identifying newly uploaded material. This is especially important in cases such as terrorism propaganda, which we want to remove quickly to try to prevent people from seeing it as much as possible.

Finding more violating content, including content that users don't report.

For example, in Q1 we took action on 1.9 million pieces of ISIS, al-Qaeda and affiliated terrorism propaganda, 99.5% of which we found and flagged before users reported them to us.

Increase the capacity of our review team to work on cases where human expertise is needed to understand the context or nuance of a particular situation.

While instrumental in our efforts, technology has limitations. We're still a long way off from it being effective for all types of violations. Our software is built with machine learning to recognize patterns, based on the violation type and local language. In some cases, our software hasn't been sufficiently trained to automatically detect violations at scale. Other violation types, such as hate speech or graphic violence, require us to understand context when we review reports and therefore require review by our trained teams. As emphasized above, it's important that people continue to report violations to help our enforcement efforts.

Our Review Processes

Our Review Processes

Once a reported piece of content is routed to our Community Operations team, they review the content and additional context to determine whether it violates our Community Standards. This section explains in more detail how we do this.

About our Community Operations team

How a review works

The Community Operations team includes content reviewers who review content in more than 50 languages, serving our global community with sites all over the world. We're committed to growing the teams focused on safety, security and content review from 10,000 to 20,000 people this year.

Our content reviewers are native language speakers who understand local context to review content and ensure that our Community Standards are enforced. Our team includes experts in areas such as child safety, hate speech, terrorism and law.

Reviewers undergo extensive training when they join the Community Operations team, including onboarding, hands-on practice, and ongoing support and training. We conduct regular training and testing to help our reviewers understand how to uphold the Community Standards and take the correct actions on content. We regularly audit the accuracy of reviewer decisions to help them improve if they make errors and address any gaps in tooling or questions about policy.

Our team reviews reports 24 hours a day, 7 days a week, and the vast majority of reports are reviewed within 24 hours.

Whether identified by our technology or reports from users, a flagged potential violation becomes a report in our system. We prioritize safety-related reports, including material related to terrorism and suicide. We use technology to try to detect imminent real-world harm, so that we can give top priority to these cases and to quickly get these reports to reviewers who have the right subject matter knowledge, language expertise and cultural context. In addition, we use automation to identify duplicate reports and stop the spread of such things as commercial spam.

Our reviewers use our Community Standards and a step-by-step process to help them make decisions accurately and consistently for the appropriate violation type. We also provide our reviewers with tools to review the reported content and the available context required to identify the concern and determine whether a piece of content violates a standard.

Making the right decisions

We strive for accuracy, and we're working to improve our technology and decisions made by our review teams. This includes improving and expanding training, ongoing auditing and developing better review tools that provide more context and information. It also includes getting smarter about how to apply our Community Standards in different cultures and languages.

We've also introduced a way for people to appeal decisions we make about content. By appealing, people can let us know if they think we've made a mistake and request that we take another look, giving them a voice in the process that's essential to helping us build a fair system. Today, we offer appeals for posts that we took action on for nudity, hate speech or graphic violence. We're working to extend this process by supporting more violation types. Appeals will be available not just for content that we removed, but also for content that was reported but remains on Facebook.

The actions we take

If we determine that content or an account violates our Community Standards or may be disturbing for some audiences, we take action — removing the content, disabling the account or covering content with a warning. If we determine it doesn't violate standards or isn't disturbing, we leave the content as-is. We don't take action just because something has been reported multiple times — only if we identify a violation.

We respond differently depending on the severity of the violation, and we may take additional action against people who repeatedly don't follow our standards. In some cases, we involve law enforcement to prevent real-world harm.

One of our metrics answers the question "How many violations did we take action on?" for each violation type. "Taking action" can mean a number of different things. Some of the most common actions we take are detailed on the next page.

Removing content

When we determine a piece of content violates our standards, we remove it from Facebook. This makes the content inaccessible to anyone on Facebook. We may also remove some kinds of links, such as to sites created by spammers to install malware or imitate a Facebook login page to steal information. Though the websites' content doesn't live on Facebook, when we detect links to them on Facebook we delete those links.

Sometimes, a piece of violating content (such as a photo with nudity, for example) may be included in a post that contains other content that does follow standards. Depending on the circumstances, we might remove just the violating content (the photo), or we might remove the entire post.

Covering content with a warning

Some content may be disturbing to people even if it doesn't violate our Community Standards. For example, we may determine a piece of content with graphic violence in it was shared to condemn violence or spread awareness and doesn't violate our standards, but that it might be disturbing to sensitive audiences or underage viewers. We cover this content with a warning, and users can choose to uncover the content if they still want to see it. We don't show content covered with a warning to underage users.

Learn more about what kinds of content we remove or cover with warnings in our Community Standards on graphic violence.

Disabling accounts

When we identify a fake account, we disable it so it's no longer visible and its owner can't log in. When we suspect a fake account, we require the account owner to pass a series of challenges to prove the account is authentic before accessing Facebook again. If we can't verify authenticity, we disable the account. We also might disable accounts for repeated violations.

Escalations to external agencies

When we become aware of a specific, imminent and credible threat to human life, we notify law enforcement. We also suggest people contact law enforcement agencies or emergency services themselves if they see something about to happen that might require the authorities to intervene. We report all apparent instances of child exploitation appearing on our site from anywhere in the world to the National Center for Missing and Exploited Children (NCMEC).

Measuring Results

Measuring Results

We develop metrics to examine how effectively we enforce our Community Standards. With them, we try to understand and prioritize how we can do better and hold ourselves accountable to our community of more than 2 billion people. The first step in mitigating abuse is to fully understand how and when it occurs. Once we can measure this, we can test and measure detection and enforcement tactics.

The metrics reflected in the Community Standards Enforcement Preliminary Report are some of the same metrics we use to measure performance internally.

1 How prevalent are violations on Facebook?

2 How much content do we take action on?

What we measure

For each type of standards violation, we aim to answer 4 key questions:

We measure the estimated percentage of views that were of violating content, a metric we call prevalence. For fake accounts, we estimate the percentage of monthly active Facebook accounts that were fake. We consider prevalence to be a critical metric because it helps us measure how many violations impact people on Facebook.

These metrics are estimated using samples of content views and accounts from across Facebook. Learn more about how we estimate this metric and how we define views in the "How We Measure" section.

It assumes that the impact caused is proportional to the number of times the content is viewed within a given violation category. We also think of this metric as *how many views of violating content we didn't prevent*. It's the number to which we hold ourselves accountable.

With fake accounts, the metric assumes the impact on Facebook users is proportional to the number of active fake accounts on Facebook, even if people don't ever see or experience these accounts.

We measure the number of pieces of content (such as posts, photos, videos or comments) or accounts we take action on for going against standards. We use this metric because it shows the scale of activity. "Taking action" could include removing a piece of content from Facebook, covering photos or videos that may be disturbing to some audiences with a warning, or disabling accounts. In the event we escalate content to law enforcement, we don't additionally count that.

It might be tempting to read this metric as an indicator of how effectively we find violations or the impact of those violations on our community. However, the volume of content we take action on is only part of the story. It doesn't reflect how long it took to detect a violation or how many times users saw that violation while it was on Facebook, which other metrics show.

This metric can go up or down due to external factors that are out of our control. As an example, consider a cyberattack during which spammers share 10 million posts featuring the same malicious URL. After we detect the URL, we remove the 10 million posts. The metric would report 10 million pieces of content acted on, an enormous spike. This number doesn't necessarily reflect that we got better at acting on spam; it reflects more that spammers decided that month to attack Facebook with unsophisticated spam that was easy to detect. It also doesn't indicate how much of that spam actually affected users: people might have seen it a few times, or a few hundred or thousand times. The 10 million pieces of content acted on metric doesn't tell us that; the prevalence metric does. After the cyberattack, the content actions metric might decrease dramatically, even if our detection moving forward improves.

3 How much violating content do we find before users report it?

4 How quickly do we take action on violating content? (Metric not yet available)

This metric shows the percentage of all content or accounts acted on that we found and flagged before users reported them to us. We use this metric as an indicator of how effectively we detect violations.

We use detection technology and people on our trained teams to help find potentially violating content and accounts and flag them for review. Then, we review them to determine if they violate standards and take action if they do. We take action on the remaining percentage of content and accounts because users report them to us first.

That being said, as with content we took action on, this metric only tells part of the story. It doesn't reflect how long it took to detect a violation or how many times it was viewed during that time. It also doesn't reflect how many violations we failed to detect altogether or how many times people viewed them. The percentage of content we proactively detected can be very high, but even the remaining small percentage can cause significant impact to users.

The metric can go up or down due to external factors. For example, in the cyberattack example above, if we detected the malicious URL before any user reported it to us, this metric would go up during the cyberattack and go down afterward, even if our detection technology didn't change during the period. This metric can also increase or decrease based on how our processes and tools change—for example, it might go up if our detection technology gets better, but it might go down if our user reporting improves and we rely less on proactive detection.

We try to act as quickly as possible against violating content to minimize its impact on users. One way that we're considering answering this question is by measuring views before we can take action. We're developing this metric to help us understand how well we handle the violations that we find. We're finalizing our methodologies for how we measure this across different violation types, and we'll make these metrics available in future versions of this report.

How we measure:
what's included in our metrics

Constructing metrics requires us to make decisions about what and what not to count in our metrics and how to categorize actions and activity. As we use these metrics for our own internal tracking, they represent our best attempt to fairly represent how effectively we enforce our Community Standards.

In this section, we explain in more detail the practices around how we measure.

1. How prevalent are violations on Facebook?

2. How much content do we take action on?

3. How much violating content do we find before users report it?

1. How prevalent are violations on Facebook?

Our goal is to minimize impact on people using Facebook caused by violations of our Community Standards. We measure prevalence of violating content to gauge how we're performing against that goal.

Prevalence represents how much violating content people may have experienced on Facebook. We determine how often violating content that we haven't taken action on was viewed on Facebook, either because we haven't caught it yet or we missed it altogether. We want to make this number as low as possible.

We calculate this metric as: the estimated number of views that showed violating content, divided by the estimated number of total content views on Facebook. If the prevalence of adult nudity and sexual activity was 0.07% to 0.09%, that would mean of every 10,000 content views, 7 to 9 on average were of content that violated our standards for adult nudity and sexual activity.

How prevalent are violations on Facebook? (continued)

Measuring prevalence based on impact

For prevalence, we estimate how often content is seen rather than the amount of content published as a way to determine how much that content affected people on Facebook. A violation could be published once but seen 1,000 times, 1 million times or not at all. The prevalence metric reflects the number of views to show the impact on the community. A small prevalence number can still correspond to a large amount of impact on Facebook, due to the large number of overall views of content on Facebook.

Definition of a content view

We record a content view when a piece of content appears on a user's screen. Specifically, a view happens when someone:

Views a post, even if there are multiple pieces of content in that post. The view is assigned to the post.

Clicks to enlarge a photo or video player.
The view is assigned to the photo or video.

+1

+1
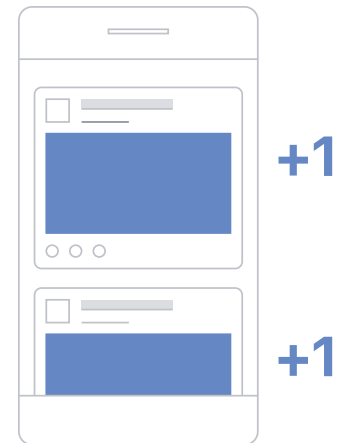
How prevalent are violations on Facebook? (continued)

Estimated metrics using sampling

We estimate prevalence by sampling content views on Facebook. To do this, we manually review samples of views and the content shown in them. Then we label the samples as violating or not violating according to our Community Standards. The teams who do this sampling review the entire post for violations, even if the sampled view didn't expose all the content in the post.

Using the portion of these samples that were of violating content, we estimate the percentage of all views that were of violating content. Note that our current prevalence sampling methodology doesn't include sampled views of comments, search results or content from some less prominent areas of Facebook.

We use stratified sampling, which increases the sample rate if the context indicates the content view is more likely to contain a violation. For example, if violations were viewed more frequently in Groups than in News Feed, we would sample views in Groups with a higher probability than we sample views in News Feed. One reason we do this is to reduce the uncertainty due to sampling. We express this uncertainty by quoting a range of values, for example by saying 7 to 9 out of every 10,000 views are on violations for adult nudity and sexual activity. This range reflects a 95% confidence window. This means that if we performed this measurement 100 times using different samples each time, we expect the true number to lie within the range 95 out of the 100 times.

How prevalent are violations on Facebook? (continued)

Caveats

The people who apply labels to our samples sometimes make mistakes, including labeling a violation as non-violating or labeling a non-violation as a violation. The relative rate of these mistakes could impact the prevalence measurement. We use audits to measure error and then adjust the prevalence calculation to account for it. The difference due to this adjustment is within the uncertainty range for the metrics we quote.

Prevalence for fake accounts

Our process to estimate the prevalence of fake accounts is slightly different. We sample monthly active users and label them as fake or not. We define a monthly active user (MAU) as a registered Facebook user who logged in and visited Facebook through our website or a mobile device, or used our Messenger application (and is also a registered Facebook user), in the last 30 days as of the date of measurement. Fake accounts can be especially difficult to discern: reviewers often must look at a profile that contains very little information and determine whether or not they're actually fake.

2. How much content do we take action on?

Content can refer to text, images, videos, links, live videos or comments on posts. This content can exist in multiple places across Facebook, such a user's Timeline, News Feed, Pages or within Groups.

The metric can go up or down due to external factors and can be enormously affected by events like cyberattacks, where we find and remove the content quickly. As explained earlier, this metric doesn't reflect how long it took to detect a violation or how many times users saw that violation while it was on Facebook.

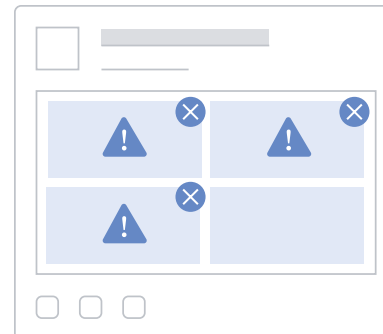# How much content do we take action on? (continued)

Counting content and actions

How we count an individual piece of content in our measurement can be complicated. We only count what we explicitly take action on. The number of pieces of content acted on that we show in reporting may vary depending on the situation.
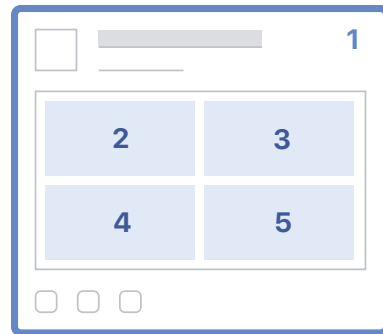
post = 5 pieces of content

3 pieces of violating content removed

1 piece of violating content removed

As a basic example, if a Facebook user publishes a post with 4 photos and text, that creates 5 pieces of content in our system: 1 for the post, 4 for the images.

If we later determine that 3 of the 4 photos violate standards and remove them, that would count as 3 pieces of content acted on.
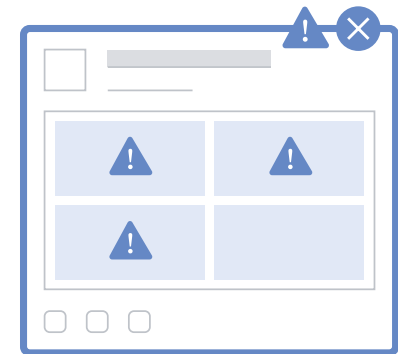
However, if we determine that the entire post is violating, in which it would count as only 1 piece of content acted on. Some of this counting becomes more complex depending on context, but this is a basic illustration of how the process works.

How much content do we take action on? (continued)

Counting content and actions (continued)

Counting fake accounts disabled is simpler: It's the number of accounts we disable for being fake.

Note that we don't count when we block content or accounts from being created in the first place, as we do when we detect spammers attempting to post with high frequency or the creation of a fake account. If we included these blocks as if they were content or account actions, it would dramatically increase the numbers (likely by millions a day) for fake accounts disabled and spam content removed.

At times, a piece of content violates multiple standards. For the purpose of measuring, we attribute the action to only one primary violation. Typically this will be the violation of the most severe standard. In other cases, we ask the reviewer to make a decision about the primary reason for violation. Learn more about how we label content in the "Labeling Content by Abuse Type " section.

How much content do we take action on? (continued)

How we measure actions on accounts, Pages, Groups and events

Large volumes of content can live within user accounts, Pages, Groups or events on Facebook. One of these as a whole can violate standards, based on content or behavior within it. We can usually determine that it violates standards without reviewing all the content within it.

If we disable an account, Page, Group or event, all the content within it automatically becomes inaccessible to users. However, in our metrics, we only count the content we determined to be violating during our review and that we explicitly took action on. We don't count the content automatically removed if we disable the account, Page, Group or event that contained that content.

In some cases, our technology may identify violations and remove content when the parent object that contained it (such as a Group) was already acted on. This additional content acted on may also be counted in our metrics.

Except for fake accounts, we don't currently include in this report any metrics for objects we took action on.

3. How much violating content do we find before users report it?

This metric shows the percentage of all content or accounts acted on that we found and flagged before people using Facebook reported them. Its complementary metric shows the percentage of content or accounts we acted on that users reported to us first. Together, the two percentages make up 100%.

How much violating content do we find before users report it? (continued)

This percentage is calculated as: the number of pieces of content acted on that we found and flagged before people using Facebook reported them, divided by the total number of pieces of content we took action on.

We aim to reduce violations to the point that our community doesn't regularly experience them. We use technology, combined with people on our teams, to detect and act on as much violating content as possible before users see and report it.

The rate at which we can do this is high for some violations, meaning we find and flag most content before users do. This is especially true where we've been able to build artificial intelligence technology that automatically identifies content that might violate our standards. Such technology is very promising but is still years away from it being effective for all kinds of violations, for example due to limitations in the technology's ability to understand context and nuance. In those areas, the rate will be lower.

Since this metric is based on counting pieces of content acted on, it has the same caveats as that metric. In addition, we compute this metric using a strict attribution of user reports to content. For example, if someone reports a Page and, while reviewing the Page, we identify and act on some violating content, we would report flagging that content proactively unless there were specifically additional user reports of it. This strict way of attributing user reports is not ideal, but we haven't yet determined a better method.

We can have a very high proactive rate for a type of violating content, and people still might see a lot of that kind of content on Facebook. For example, even if we find and flag 99% of content that we subsequently take action on, people may still see a lot of that content before we're able to remove it.

For fake accounts, we calculate this metric as the percentage of accounts disabled for being fake that we found and flagged before users reported them to us. It's calculated as the number of disabled accounts we found and flagged before users reported them, divided by the total number of accounts disabled for being fake.

Getting better at measurement

As we've stated, the metrics we published in the Community Standards Enforcement Preliminary Report aren't perfect. We're continually assessing them to learn how we can improve our methodologies for measuring.

Labeling
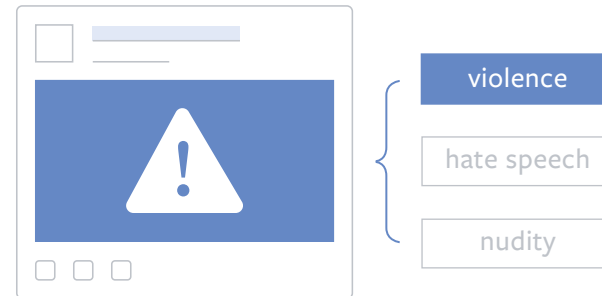
To count the content acted on for a specific standard violation, we must label the violation each time we take an action. When reviewers look at reports, they first select whether the material violates our policies or not. If they select yes, they then label with the violation type. Note that this labeling process is used to count the amount of content acted on and doesn't factor into how we measure prevalence.

Labeling (continued)

In the past, we didn't require our reviewers to label the violations when they made decisions. Instead, we relied on information that users gave us when they submitted the report. We've upgraded our review process, as of last year, to record more granular information about why reviewers remove a piece of content. While this has slowed the reviewer process down, it's allowed us to establish more accurate metrics. We also updated our detection technology so it labels violations as they're found, flagged or removed using the same labels as our reviewer decisions.

This is the reason we only report metrics as of Q4 2017. In fact, our records of these violations started part way through October, but we've prorated the data for Q4 to provide a meaningful quarter-over-quarter comparison between Q4 and Q1.

Continuing to improve our metrics

Our metrics, which are still in development, have a lot of limitations. For example, we enforce our Community Standards across Facebook, but our measurement of prevalence doesn't sample from all kinds of content views—it doesn't yet include sampled views of comments, search results or content from some less prominent areas of Facebook.

We continue to review our standards and processes and the methodologies behind them. Changes to these standards, processes and methodologies will change the metrics calculations themselves. These methodology or process changes may be in addition to trends indicating that we're getting better or worse at mitigating violations.

Conclusion

Conclusion

As we said earlier, this preliminary report is just a first step for sharing with our community how we uphold the Facebook Community Standards to keep people safe while maintaining an open platform for personal expression. These metrics aren't perfect, and we have a lot of work to do to improve our internal processes, refine our tools and technology, and find the right ways to measure our enforcement reliably.

One reason we want to share this report publicly is to begin a dialog with our community about the Facebook Community Standards and our enforcement of them.

As announced when we released our internal guidelines for enforcing Community Standards in April 2018, we plan to hold a series of public events around the world this year to get people's direct input on our standards and enforcement of them.

The Community Standards Enforcement Preliminary Report is a move toward holding ourselves accountable—and letting others in our community hold us accountable—for doing better at enforcing our standards. This guide explains our methodology so the public can understand the benefits and limitations of the numbers we share, as well as how we expect these numbers to change as we refine our methodologies. We're committed to doing better, and communicating more openly about our efforts to do so, going forward.

See details on our Community Standards at:

https://www.facebook.com/communitystandards

See the Community Standards Enforcement Preliminary Report at:

**https://transparency.facebook.com/
community-standards-enforcement**