Community Standards Enforcement Report

HIGHLIGHTS - NOVEMBER 2019

Facebook builds services for people to express their voice. For people to express themselves, they need to know they are safe. That's why we have rules to determine what is and what isn't allowed.

The Community Standards Enforcement Report measures how we are doing at enforcing our policies.

What's new

1 DATA ON SUICIDE & SELF-INJURY

In addition to the 9 policy areas covered in the previous report, we now include our policy on suicide & self-injury.

2 DATA ON INSTAGRAM

We've added data related to 4 policies on Instagram: child nudity and sexual exploitation, regulated goods, suicide & self-injury, and terrorist propaganda.

3 DATA ON TERRORIST PROPAGANDA

To better reflect our efforts against terrorism worldwide, our data now covers enforcement against all terrorist organizations, and not just Al Qaeda, ISIS and their affiliates.

What we've done

PREVALENCE

Prevalence measures how often violating content is seen on Facebook. We have a better understanding of it in 7 of the 10 policy areas (up from 5 policy areas in the previous report).

7/10 POLICY AREAS

PROACTIVE ENFORCEMENT

In 8 of the 10 policy areas, we proactively detected over 90% of the content we took action on before someone reported it.
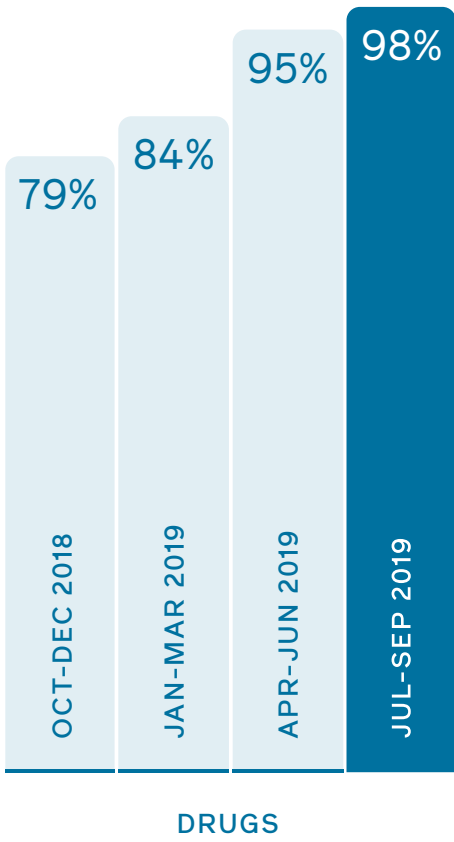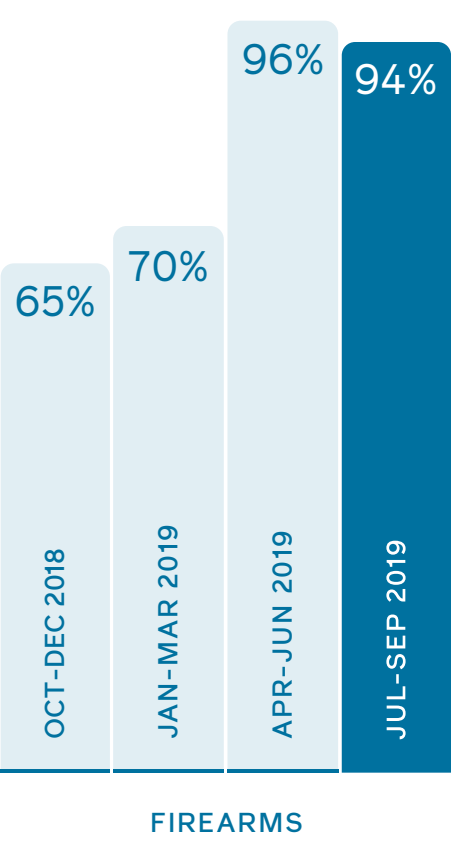
>90%

IN 8/10 POLICY AREAS

DRUG & FIREARM SALES CONTENT

Continued investments in our enforcement systems have enabled us to detect and remove more violating content, resulting in proactive rates above 90%.
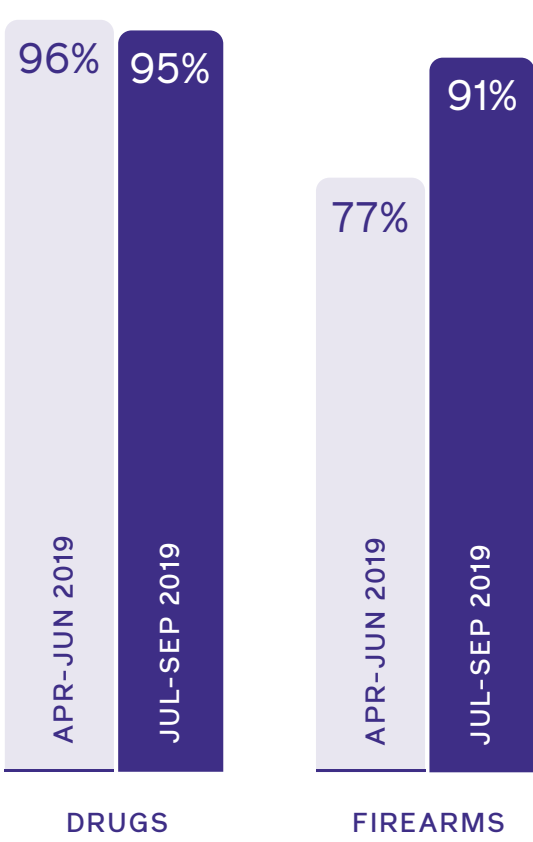
Facebook app

DRUGS
- OCT-DEC 2018: 79%
- JAN-MAR 2019: 84%
- APR-JUN 2019: 95%
- JUL-SEP 2019: 98%

FIREARMS
- OCT-DEC 2018: 65%
- JAN-MAR 2019: 70%
- APR-JUN 2019: 96%
- JUL-SEP 2019: 94%

Instagram

DRUGS
- APR-JUN 2019: 96%
- JUL-SEP 2019: 95%

FIREARMS
- APR-JUN 2019: 77%
- JUL-SEP 2019: 91%

PROACTIVE RATES %

SUICIDE & SELF-INJURY

Policy changes and improvements to our technology allowed us to remove more violating content.
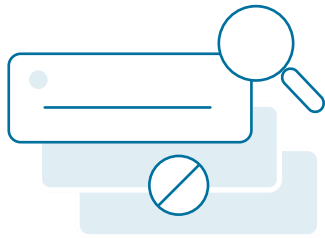
Facebook app

APR-JUN 2019	2.0M
JUL-SEP 2019	2.5M

Instagram

|	
APR-JUN 2019	835K
JUL-SEP 2019	845K

DATA TRANSPARENCY ADVISORY GROUP (DTAG)

We continue to work through the recommendations from DTAG, a group of international academic experts, to make the report more transparent.

What we're working on

HATE SPEECH

As AI systems have evolved and improved, we've started automatically removing hate speech in some cases where we have high confidence. As we try to be more proactive in combating hate speech, we continue to refine the accuracy of our systems and partner with our content review team to enforce our policy.
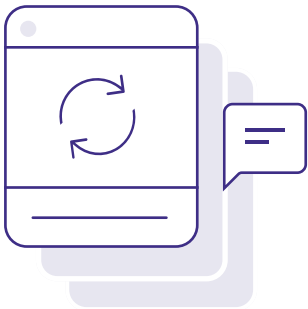
INSTAGRAM APPEALS AND RESTORES

In July, we started introducing the ability for people to appeal content decisions on Instagram. We will keep expanding these efforts. Appeals and restore metrics for Instagram will be included in future reports.

Continued efforts

We continually strive to improve our processes, refine our tools and technology, and find the right ways to measure enforcement. Transparency in our work will keep people informed and let the public hold us accountable for improving over time.

Learn more about the report at: **TRANSPARENCY.FACEBOOK.COM**

METRICS

PREVALENCE = estimate of how often content that violates our policies is seen

CONTENT ACTIONED = number of pieces of content (such as posts, photos, videos or comments) or accounts we take action on for going against our policies

PROACTIVE RATE = percentage of all content or accounts acted on that we found and flagged before users reported them to us

APPEALED CONTENT = number of pieces of content that people appeal after we take action on them for going against our policies

RESTORED CONTENT = number of pieces of content we restored after we originally took action on them