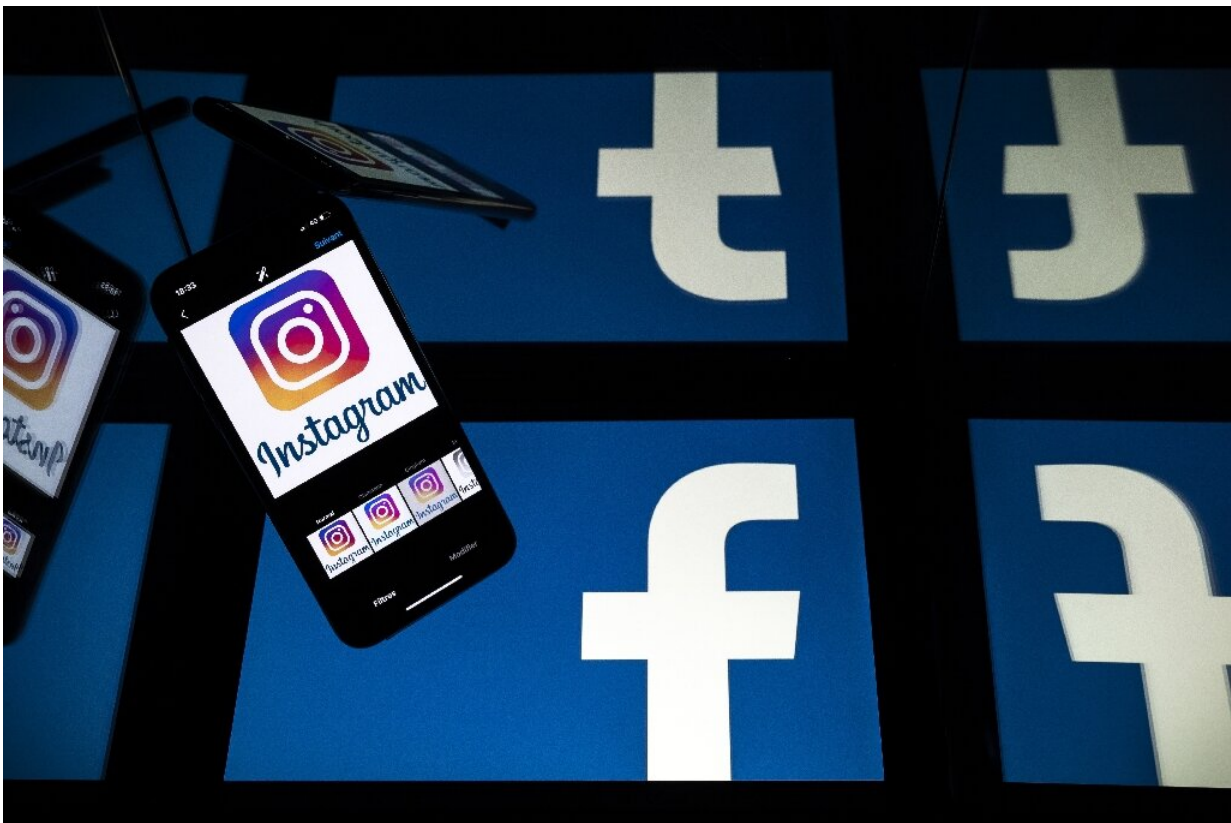


Facebook touts progress in curbing hate, violent content

May 19 2021



Facebook says its automated tools have helped it remove more hateful and abusive content on its core platform and on Instagram.

Facebook said Wednesday it was making steady progress in curbing hate speech and violent content on the massive social network as it released

its quarterly update on enforcing its standards.

The leading social media firm said refinements in [artificial intelligence](#) have helped it filter out more hateful and [violent content](#) along with bullying, inappropriate sexual material and fake accounts, among other things.

Facebook said the "prevalence" of hate speech in the first quarter of 2021 was between 0.05 and 0.06 percent, a drop of nearly half from mid-2020.

Facebook's [vice president](#) for integrity Guy Rosen told reporters using this metric of prevalence was appropriate because "it captures not what we took down but what we missed... hate speech prevalence on Facebook continues to decrease."

The update comes with Facebook under pressure from governments and activist groups to crack down on hateful and [abusive content](#), along with misinformation and incitements to violence.

The latest update showed violent and graphic content was cut to less than 0.04 percent on Facebook and under 0.02 percent on Instagram, and that the company had also reduced the prevalence of other inappropriate content including nudity and sexual material.

The report said Facebook in the quarter removed 25.2 million pieces of "hate speech content," 8.8 million pieces of "bullying and harassment content," and 9.8 million including "organized hate content."

Facebook said it was striving to enforce policies barring harmful misinformation about Covid-19 and had removed more than 18 million pieces of violating content from its platform and Instagram globally since the start of the pandemic.

Refinements to its automated technology have allowed Facebook to make progress in identifying inappropriate content not only in text, but in images and video or even memes which blend these elements, the company said.

"Today we proactively detect about 97 percent of [hate speech](#) content we remove," Rosen said.

© 2021 AFP

Citation: Facebook touts progress in curbing hate, violent content (2021, May 19) retrieved 23 April 2024 from

<https://techxplore.com/news/2021-05-facebook-touts-curbing-violent-content.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.