

Twitter adds more safety tools, will curb abusive accounts

1 March 2017, by Barbara Ortutay



This Nov. 4, 2013, file photo shows the sign outside of Twitter headquarters in San Francisco. Twitter announced Wednesday, March 1, 2017, it is adding more new tools to curb abuse, part of an ongoing effort to protect its users from hate and harassment. It is the second time in three weeks the company has released new features aimed at rooting out abusive content, signaling that it's getting more serious about the issue after being criticized for not doing enough in the decade since its founding. (AP Photo/Jeff Chiu, File)

Twitter is adding more tools to curb abuse on its service as part of its ongoing effort to protect users from hate and harassment. Among other things, the service will attempt to identify offenders on its own, even if no one has reported them first.

It is the second time in three weeks the company has released new ways to root out abusive material on its service. It's a sign Twitter is getting more serious about the issue; it's faced criticism for not doing more in the decade since its founding.

Twitter's effort to automatically flag abusive accounts is a new twist. Previously, the company would take no action until users reported

suspected abuse. Not that long ago, in fact, Twitter required the targets of abuse to make such reports themselves; it refused to take reports from friends, family or other third parties.

Once it identifies accounts that repeatedly engage in abusive behavior, Twitter will place temporary restrictions on them. For example, tweets from such accounts will be visible only to followers or other users who deliberately search them out.

"We aim to only act on accounts when we're confident, based on our algorithms, that their behavior is abusive," the company said in a blog post .

In addition, Twitter is also giving users the option to filter the notifications they get when someone mentions them in a comment. For example, people will be able to filter out accounts without a profile photo, or unverified email addresses or phone numbers. This can be helpful in screening out accounts that have been set up solely for the purpose of harassment or trolling.

© 2017 The Associated Press. All rights reserved.

APA citation: Twitter adds more safety tools, will curb abusive accounts (2017, March 1) retrieved 11 May 2021 from <https://phys.org/news/2017-03-twitter-safety-tools-curb-abusive.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.