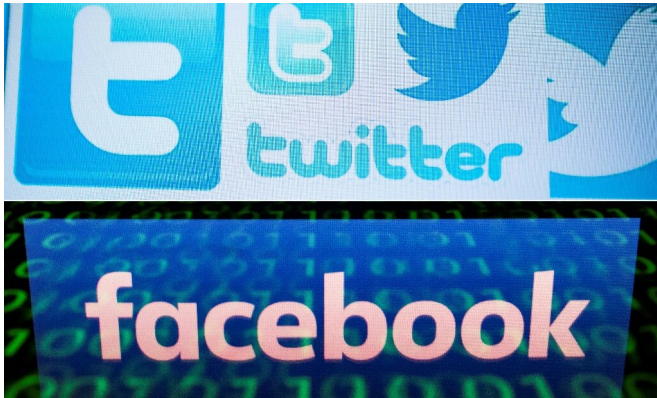


Tech giants strike deal with advertisers over hate speech

23 September 2020, by Glenn Chapman



Facebook and Twitter joined other online firms in an agreement with global advertisers on a common definition of hateful content

Web giants including Facebook struck a deal Wednesday with global advertisers to get on the same page on defining hate speech, a move aimed at helping companies steer clear of being associated with vile content.

The agreement—which also includes Twitter and YouTube—lays out for the first time a common set of definitions for hateful posts.

Facebook and others have long grappled with how to purge toxic content from online platforms while fending off accusations they stifle free expression in the process.

In July, hundreds of brands suspended advertising at Facebook as part of a #StopHateForProfit campaign, saying the social-media titan should do more to stamp out hatred and misinformation on its platform.

Earlier this month a group of celebrities—including Kim Kardashian, Leonardo DiCaprio and Katy Perry—stop using Facebook and Instagram for 24

hours, to push a similar message.

The announcement came from the Global Alliance for Responsible Media, a group which includes tech platforms and the major brands in the World Federation of Advertisers.

No mention was made of tightening rules regarding hateful posts at any of the social media platforms, just that they would have a standard by which to determine what that is.

The impact of the change, set to be implemented in 2021, was not clear amid a range of interpretations of what constitutes hate speech.

"We shouldn't get our hopes up though as there are other content issues that are already universally agreed to be harmful, such as terrorism and suicide, but social media companies continue to fail here," said Syracuse University assistant professor of communications Jennifer Grygiel.

Empty promises?

Key areas of the agreement were said to include applying common definitions of harmful content; developing reporting standards; establishing independent oversight; and rolling out tools for keeping advertisements away from harmful content.

"We welcome today's announcement that these social media platforms have finally committed to doing a better job tracking and auditing hateful content," said Anti-Defamation League chief executive Jonathan Greenblatt, a longtime critic of Facebook's content moderation efforts.

"These commitments must be followed in a timely and comprehensive manner, to ensure they are not the kind of empty promises that we have seen too often from Facebook."

The WFA said that defining online hate speech

solves the problem of different platforms using their own definitions, which it said makes it difficult for companies to decide where to put ads.

"As funders of the online ecosystem, advertisers have a critical role to play in driving positive change," said Stephan Loerke, chief executive of the WFA.

Luis Di Como, executive vice president of global media at Unilever, a major advertiser, sounded a note of cautious optimism.

"The issues within the online ecosystem are complicated, and whilst change doesn't happen overnight, today marks an important step in the right direction," Di Como said.

Facebook founder and chief executive Mark Zuckerberg has been adamant that the company did not want hate speech on the social network.

On Wednesday, the leading social network's vice president for global marketing solutions, Carolyn Everson, said the "uncommon collaboration" gave all parties "a unified language to move forward on the fight against hate online."

The deal is not enough to reduce societal risks posed by online social media platforms, contended Grygiel.

"We need social media companies to collectively engage in self-regulation at the industry level like the advertising industry does," Grygiel said.

© 2020 AFP

APA citation: Tech giants strike deal with advertisers over hate speech (2020, September 23) retrieved 25 November 2020 from <https://techxplore.com/news/2020-09-tech-giants-advertisers-speech.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.