

New Twitter policy aims to pierce fog of war misinformation

May 19 2022, by David Klepper



In this April 26, 2017, photo is a Twitter app icon on a mobile phone in Philadelphia. Twitter is stepping up its fight against misinformation with a new policy cracking down on posts that spread potentially dangerous false stories. Under the new rules, which take effect Thursday, May 19, 2022, Twitter will no longer automatically recommend posts that mischaracterize conditions during a conflict or make misleading claims about war crimes or atrocities. Credit: AP Photo/Matt Rourke, File

Twitter is stepping up its fight against misinformation with a new policy cracking down on posts that spread potentially dangerous false stories. The change is part of a broader effort to promote accurate information during times of conflict or crisis.

Starting Thursday, the platform will no longer automatically recommend or emphasize posts that make misleading claims about the Russian invasion of Ukraine, including material that mischaracterizes conditions in conflict zones or makes false allegations of war crimes or atrocities against civilians.

Under its new "crisis misinformation policy," Twitter will also add warning labels to debunked claims about ongoing humanitarian crises, the San Francisco-based company said. Users won't be able to like, forward or respond to posts that violate the new rules.

The changes make Twitter the latest social platform to grapple with the misinformation, propaganda and rumors that have proliferated since Russia invaded Ukraine in February. That misinformation ranges from rumors spread by well-intentioned users to Kremlin propaganda amplified by Russian diplomats or fake accounts and networks linked to Russian intelligence.

"We have seen both sides share information that may be misleading and/or deceptive," said Yoel Roth, Twitter's head of safety and integrity, who detailed the new policy for reporters. "Our policy doesn't draw a distinction between the different combatants. Instead, we're focusing on misinformation that could be dangerous, regardless of where it comes from."

The new policy will complement existing Twitter rules that prohibit digitally manipulated media, false claims about elections and voting, and health misinformation, including debunked claims about COVID-19 and

vaccines.

But it could also clash with the views of Tesla billionaire Elon Musk, who has agreed to pay \$44 billion to acquire Twitter with the aim of making it a haven for "free speech." Musk hasn't addressed many instances of what that would mean in practice, although he has said that Twitter should only take down posts that violate the law, which taken literally would prevent action against most misinformation, personal attacks and harassment. He has also criticized the algorithms used by Twitter and other social platforms to recommend particular posts to individuals.

The policy was written broadly to cover misinformation during other conflicts, natural disasters, humanitarian crises or "any situation where there's a widespread threat to health and safety," Roth said.

Twitter said it will rely on a variety of credible sources to determine when a post is misleading. Those sources will include humanitarian groups, conflict monitors and journalists.

A senior Ukrainian cybersecurity official, Victor Zhora, welcomed Twitter's new screening policy and said that it's up to the global community to "find proper approaches to prevent the sowing of misinformation across social networks."

While the results have been mixed, Twitter's efforts to address misinformation about the Ukraine conflict exceed those of other platforms that have chosen a more hands-off approach, like Telegram, which is popular in Eastern Europe.

Asked specifically about the Telegram platform, where Russian government disinformation is rampant but Ukraine's leaders also reaches a wide audience, Zhora said the question was "tricky but very

important." That's because the kind of misinformation disseminated without constraint on Telegram "to some extent led to this war."

Since the Russian invasion began in February, [social media platforms](#) like Twitter and Meta, the owner of Facebook and Instagram, have tried to address a rise in war-related misinformation by labeling posts from Russian state-controlled media and diplomats. They've also de-emphasized some material so it no longer turns up in searches or automatic recommendations.

Emerson Brooking, a senior fellow at the Atlantic Council's Digital Forensic Research Lab and expert on social media and disinformation, said that the conflict in Ukraine shows how easily misinformation can spread online during conflict, and the need for platforms to respond.

"This is a conflict that has played out on the internet, and one that has driven extraordinarily rapid changes in tech policy," he said.

© 2022 The Associated Press. All rights reserved. This material may not be published, broadcast, rewritten or redistributed without permission.

Citation: New Twitter policy aims to pierce fog of war misinformation (2022, May 19) retrieved 25 April 2024 from <https://techxplore.com/news/2022-05-twitter-policy-aims-pierce-fog.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.