

Threads and TikTok struggle to block misinformation and conspiracies

Article

The news: TikTok accidentally blocked searches for “WGA” and “WGA strike,” abbreviations for the **Writers Guild of America**, while attempting to filter out Qanon conspiracy theories, the

company confirmed following a [report](#) from Media Matters.

- As part of its sweep, TikTok blocked the prominent Qanon phrase, “Where we go one, we go all,” which is often abbreviated as, “WWG1WGA.” It accidentally flagged the final three letters for blockage.

Meta is caught up in a similar entanglement. It outright blocked any search related to COVID-19 on **Threads**, including the name of the disease and the word “vaccines.”

- Meta said these blockages were temporary and intended to stop users from viewing “potentially sensitive content” until it’s confident in the quality of search results.
- The decision drew criticism given Meta’s history with hosting COVID-19 misinformation and the current surge of the virus across the US.

Times are not a-changing: Social platforms still haven’t figured out how to manage misinformation and controversial topics on their platforms after all this time.

- There are more challenges than ever for platforms trying to curb the spread of harmful content. Artificial intelligence rapidly created a [wave of spam and misinformation](#) that platforms have yet to get a handle on.
- TikTok’s slipup with the WGA is less severe than blocking information about a pandemic, but it goes to show that something as seemingly simple as “banning Qanon content” is not so easily achieved, especially with the extraordinary volume of content uploaded to social platforms each day.

Figuring out Threads: Meta’s situation is trickier. If the company wants to turn Threads into a source of ad revenues, it has to put significant brand safety and content moderation measures in place to convince advertisers that it’s a safe place to spend.

- That means ensuring that health misinformation doesn’t have a place on the platform. But the timing of the move—in the middle of a COVID-19 surge—and lack of warning leaves a lot to be desired. Meta pulled a similar maneuver in Canada when it banned news [during a historic wildfire](#), drawing criticism from regulators—but that was because of a clash over new legislation.
- Releasing a statement clarifying why COVID-19 searches are blocked is a step in the right direction, but prior notice about when and why such changes are taking place would likely

help convince advertisers that Meta isn't making off-the-cuff decisions or banning crucial content.

Our take: Social platforms are still struggling to put systems in place to throttle harmful content, and the flow isn't stopping anytime soon. Blanket keyword bans might work as temporary solutions, but Meta and TikTok's missteps show a need for more sophisticated content moderation systems.

US Teens' Favorite Social Media Platform, Spring 2021-Spring 2023

% of respondents

	Spring 2021	Fall 2021	Spring 2022	Fall 2022	Spring 2023
TikTok	30%	30%	33%	38%	37%
Snapchat	31%	35%	31%	30%	27%
Instagram	24%	22%	22%	20%	23%
Discord	5%	5%	5%	4%	3%
Facebook	2%	2%	3%	2%	3%
Twitter	3%	2%	2%	2%	2%

Note: teens with an average age of 16.2; excludes "other" responses
Source: Piper Sandler, "Taking Stock With Teens," April 4, 2023

281162

eMarketer | InsiderIntelligence.com