JEWISH TELEGRAPHIC AGENCY

EST 1917

United States

Jewish groups pushed Meta to crack down on hate speech. Now, the company is reversing course.

The company offered "Jews are flat out greedier than Christians" as an example of newly allowed speech.

By Asaf Elia-Shalev January 13, 2025 5:30 pm



Meta founder and CEO Mark Zuckerberg, left, and Joel Kaplan, named in January 2025 as Meta's president of global affairs, descend a staircase after leaving a meeting with a Republican senator in Washington, D.C., Sept. 19, 2019. (Samuel Corum/Getty Images)

Six months ago, Jewish groups celebrated a policy win when Meta <u>banned the use of "Zionist" as a coded slur</u> against Jews and Israel. Now, the same organizations are condemning the company for dramatically loosening restrictions on speech across its social media platforms.

"It is mind-blowing how one of the most profitable companies in the world, operating with such sophisticated technology, is taking significant steps back in terms of addressing antisemitism, hate, misinformation and protecting vulnerable and marginalized groups online," the Anti-Defamation League's CEO, Jonathan Greenblatt, said in a statement.

Greenblatt was responding to Meta CEO Mark Zuckerberg's <u>announcement</u> that the company would do away with its fact-check program and broaden the terms of what can be said on Facebook and Instagram. It will also stop using automation to detect and remove hate speech — limiting its use to catching terrorism, child sexual exploitation and other offenses.

"This is a trade-off," Zuckerberg said in a video message. "It means that we're going to catch less bad stuff, but we'll also reduce the number of innocent people's posts and accounts that we accidentally take down."

Zuckerberg, who is Jewish, said the moves respond to what he called a "cultural tipping point" following the U.S. presidential election. Conservatives have long alleged that Meta suppresses right-wing speech, and Zuckerberg's advisors and detractors both explained the move as a <u>strategic pivot toward accommodating a political era</u> increasingly shaped by Donald Trump and the Republican Party.

To replace its fact-checking program, Meta said it would introduce a "community notes" feature similar to the one on X, the social media platform owned by Trump ally Elon Musk, and otherwise favor allowing users to rebut posts they view as objectionable over removing them entirely.

"We want to undo the mission creep that has made our rules too restrictive and too prone to over-enforcement," the company said in its announcement. "We're getting rid of a number of restrictions on topics like immigration, gender identity and gender that are the subject of frequent political discourse and debate. It's not right that things can be said on TV or the floor of Congress, but not on our platforms."

The changes could significantly affect what users see on social media. Last year, Meta's automated systems detected approximately 95% of hate speech violations on Facebook and 98% on Instagram, according to <u>data self-reported by the company</u>. Millions of posts were removed as a result.

That was a positive sign for advocates who have battled the proliferation of antisemitism and other hate speech online. Meta has historically invited outside input when faced with content questions, and Jewish groups, such the ADL, the World Jewish Congress and a nonprofit focused on online antisemitism called CyberWell, have lobbied the company for years hoping to rein into online antisemitism.

Now, those groups are balking at the social media giant's retreat from policing the content on its platforms, which it announced after <u>a six-week post-election policy</u> overhaul reportedly conducted by Zuckerberg and just a handful of confidentes.

The World Jewish Congress, for example, criticized Meta's new reliance on user-generated "community notes" to combat misinformation, arguing that it shifts the burden of addressing hate speech onto marginalized groups.

"In an online environment already rife with hostility, reducing protections and clear guidelines will open the floodgates to content that fuels real-world threats, including violent acts targeting Jewish communities," said Yfat Barak-Cheney, executive director of the organization's Technology and Human Rights Institute.

The debate of how to police misinformation is a matter of enormous consequence for Jews given that <u>studies show a strong link</u> between the spread of conspiracy theories and antisemitic attitudes.

"This change means one thing, very in line with the trend of both the quantity and quality of content that we have seen on X since Musk acquired Twitter – more hate speech, more politicized content, more silos and less effective responses from the platforms," Tal-Or Cohen Montemayor, who heads the Israel-based CyberWell, said in a statement.

"This change particularly undermines the safety of all marginalized communities, including the Jewish community which is currently experiencing one of the worst onslaughts of widespread Jew-hatred in both online and offline spaces," Montemayor added.

Soon after Meta announced the policy shift — first in an Instagram Reel by Zuckerberg, then in a Fox & Friends appearance by Joel Kaplan, a Republican operative who has long worked for Meta and was recently named president of global affairs — leaked training materials surfaced with examples of derogatory statements that will now be permitted on the company's platforms.

The materials, obtained by the left-wing news organization <u>The Intercept</u> and not disputed by Meta, include antisemitic examples, such as "Jews are flat out greedier than Christians." Other examples of permitted hate speech target other groups, with statements like "Gays are freaks!" and "Immigrants are grubby, filthy pieces of shit."

Some hateful statements about Jews remain prohibited under the new policy. For instance, content moderators are instructed to remove comments such as "Jewish women are slutty" under guidelines addressing "insults about sexual immorality." Similarly, Meta's rules continue to ban the use of certain curse words directed at protected groups; for example, "Ugh, the fucking Jews are at it again" would not be allowed.

Meta did not immediately respond to a request about the fate of the ban on using "Zionist" as a proxy for Jews and Israel in hate speech.

The company also did not signal any changes to its ban on Holocaust denial in its announcement, and the leaked materials also did not address the ban, which <u>Zuckerberg announced in 2020</u>, reversing a previous position under pressure from Jewish groups. Some education and advocacy groups <u>say automatic detection of hate content has ensnared their efforts to create educational content about the Holocaust; nearly all "algorithmic overreach" of that nature would be a thing of the past under Meta's new approach.</u>