

Reddit CEO says racism is permitted on the platform, and users are up in arms



Reddit CEO Steve Huffman has found himself once again embroiled in a controversy surrounding his website's policy on moderation. In a Reddit thread announcing the platform's 2017 transparency report findings, in which Reddit identified and listed close to 1,000 suspected Russia-linked propaganda accounts that have been banned, Huffman replied to a straightforward question about the company's rules around hate speech, which is a verbal attack based on race, religion, or another protected class.

"I need clarification on something: Is obvious open racism, including slurs, against reddit's rules or not?" asked Reddit user chlomyser. "It's not," Huffman, who operates on Reddit under his original handle "spez," responded.

Huffman elaborated on his point, adding:

"On Reddit, the way in which we think about speech is to

separate behavior from beliefs. This means on Reddit there will be people with beliefs different from your own, sometimes extremely so. When users actions conflict with our content policies, we take action."

Our approach to governance is that communities can set appropriate standards around language for themselves. Many communities have rules around speech that are more restrictive than our own, and we fully support those rules.

It's a controversial approach, to say the least, and it has many Reddit users outraged that communities like the Trump-centric r/The_Donald are allowed to walk up to and over the line of racism repeatedly without any site action. Many Reddit users responded to Huffman by pointing out that hate speech does constitute behavior in a way, and that communities like r/The_Donald directly participated in the conversation and organizing of events like the Charlottesville, Virginia, white supremacist rally that resulted in Heather Heyer's death. This conversation around Reddit's light moderation has been simmering for quite some time, boiling over most recently last month when the company discussed its approach to Russia propaganda.

Huffman's position here is an evolving one. Nearly a decade ago, Huffman's approach to hate speech mirrored that of other major social media platforms today, which is to ban it except in extremely narrow or uniquely circumstantial situations. For instance, Facebook's policies on hate speech are well-documented, and saying something racist will typically lead to some type of disciplinary action. Other platforms like Twitter, YouTube, and Instagram all have hate speech policies as well that can result in suspensions or bans.

Huffman's approach to hate speech has evolved over the last 10 years

"I guess I'm a little late to the party, but I banned him. We

rarely ban non-spammers, but hate-speech used in that context is not something we tolerate,” Huffman wrote in a thread nine years ago about banning a user over hate speech. “This isn’t any change in policy: we’ve always banned hate speech, and we always will. It’s not up for debate. You can bitch and moan all you like, but me and my team aren’t going to be responsible for encouraging behaviors that lead to hate,” Huffman wrote in response to another user in the same thread.

Yet, when Huffman took over in 2015 for interim CEO Ellen Pao, who was pushed out of her position in part due to the platform’s toxic and vehement opposition to Pao’s leadership, his approach to hate speech had shifted. “While my personal views towards bigotry haven’t changed, my opinion of what Reddit should do about it has,” Huffman wrote on the topic nearly three years ago, a few weeks after he had returned to lead the company. “I don’t think we should silence people just because their viewpoints are something we disagree with. There is value in the conversation, and we as a society need to confront these issues. This is an incredibly complex topic, and I’m sure our thinking will continue to evolve.”

Reddit still takes hardline stances on calls to violence, threats, doxxing, and other activities that may lead to real-world harm. But Huffman has often been wishy-washy on moderating the more complex gray areas in between innocuous content and those extreme examples. That’s where hate speech, which is not illegal in the US, thanks to First Amendment protection, typically falls. For instance, in 2015, Reddit banned the fat-shaming community r/fatpeoplehate and the openly racist community r/coontown. Infamous situations prior to that included the banning of the community sharing leaked celebrity nude photos and a community dedicated to sharing so-called “creepshots” of underage girls.

Reddit takes action against some communities only when it seems it has to

More recently, Reddit took action against the artificial

intelligence-generated fake porn community r/deepfakes as well as a handful of alt-right subreddits and Nazi boards. But each time it does this, Reddit cites a specific rule like the use of violent speech, doxxing, or the sharing of non-consensual pornography.

When it comes to raw speech, however, Huffman seems to be more permissive, which stands in stark contrast to other tech industry platforms, nearly all of which are grappling with hard questions about moderation these days. Just this week, Facebook CEO Mark Zuckerberg was grilled by Congress over the ongoing Cambridge Analytica data privacy scandal, and he was asked numerous questions about how the company plans to handle hate speech on its platform. The matter for Facebook is especially pressing, as ethnic violence in Myanmar has erupted, thanks in part to organizing and the spreading of propaganda on the social network.

Facebook's approach seems largely centered on AI. Zuckerberg says his company is increasingly looking to automated algorithms that parse text, photos, and videos to do the work even tens of thousands of human moderators cannot. That work decides whether involving a piece of content breaks the company's policies around fake news, hate speech, obscenity, and other inadmissible forms of content.

Huffman thinks banning speech won't make it go away

Reddit's approach, on the other hand, seems to be focused less on sweeping rules and more on case-by-case evaluations. That won't do very much to calm critics who want it to ban communities like r/The_Donald or make the use of racial slurs a punishable offense. Huffman seems to take the free speech absolutism approach of letting sunlight disinfect the world of extremist viewpoints and bigotry, or to offload the work to subreddit admins and site moderators when applicable.

Yet that approach falls apart when it becomes inconvenient for Reddit as a company, like in the presence of a legal and PR

nightmare resulting from letting neo-Nazis or illegal pornography run rampant on the site. As many Reddit users pointed out to Huffman in responses to the thread, a study published last year on the banning of r/fatpeoplehate and r/coontown, titled "You Can't Stay Here: The Efficacy of Reddit's 2015 Ban," showed clear positive effects of banning hateful communities.

"Many more accounts than expected discontinued their use of the site; and, among those that stayed active, there was a drastic decrease (of at least 80 percent) in their hate speech use," the study authors concluded. "Though many subreddits saw an influx of r/fatpeoplehate and r/coontown 'migrants,' those subreddits saw no significant changes in hate speech use. In other words, other subreddits did not inherit the problem."

Banning hateful groups helps clean up communication platforms
Whatever Huffman's evolving approach on the topic, it's Reddit users that seem to be the most directly affected by the proliferation of hate speech on the platform.

"Spez, what qualifies as bannable hate speech to you? Because I kinda wonder if you'd be able to justify allowing some of the things on your platform that you do allow on your platform in front of Congress," wrote user PostimusMaximus. "Zuckerberg is sitting over here getting grilled for not removing hate-speech fast enough due to AI limitations and yet you find yourself passing hate speech off as okay because you think it's not a dangerous thing to allow on your platform or because you expect t_d [r/The_Donald] to self-moderate and hopefully if they troll long enough they'll die out on their own."

"I think aside from Russian interference you need to give a thorough answer explaining what the logic is here," the user added, linking to specific Reddit threads filled with anti-Muslim hate speech on the Trump-centric subreddit. "You are literally letting users spread hate speech and pretend it's

politics in some weird sense of free speech as if it's okay and nothing bad is happening."

<https://www.theverge.com/2018/4/11/17226416/reddit-ceo-steve-huffman-racism-racist-slurs-are-okay>