

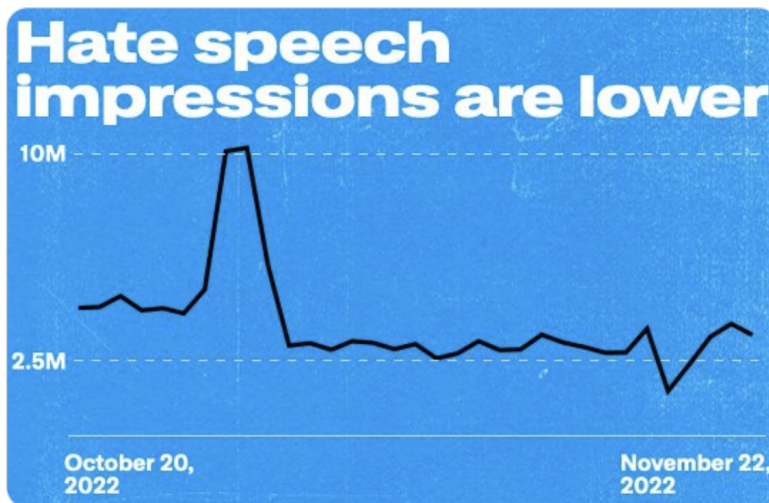
Tweets containing hate speech are still rising despite Elon Musk's claims

On November 24, Elon Musk tweeted to claim that “hate speech impressions” are down by a third from “pre-spike levels”. His tweet was accompanied by a graph showing the number of “hate speech impressions” declining in recent weeks after a spike when he first took over the platform.



Elon Musk
@elonmusk

Hate speech impressions down by 1/3 from pre-spike levels. Congrats to Twitter team!



11:07 PM · Nov 23, 2022 · Twitter for iPhone

However, new research conducted by the Center for Countering Digital Hate (CCDH) shows that the week up to his claim saw a substantially higher number of daily tweets containing hateful slurs when compared to the norm before he took over the platform on October 27.

In the week running up to Musk's tweet, there were:

- 30,546 tweets containing “ni**er”, triple the rate on the month before Musk's takeover
- 30,068 tweets containing “fa**ot”, up 31% on the month before Musk's takeover
- 35,938 tweets containing “tra**y”, up 3% on the month before Musk's takeover
- 2,366 tweets containing “k*ke”, up 24% on the month before Musk's takeover

- 916 tweets containing “sp*c”, up 58% on the month before Musk’s takeover

The data shows that hateful language on Twitter is only rising under Musk. Notably, his tweet claims to analyse “impressions”, meaning the number of times a given tweet has been seen. This emphasis obscures the fact that the number of hateful tweets being hosted by Twitter is actually rising under his leadership, indicating a failure to enforce platform community standards. Musk also failed to explain what his definition of “hate speech” is.

Imran Ahmed, CEO of the Center for Countering Digital Hate said:

“Elon Musk has again been exposed as misleading users and advertisers, claiming mission accomplished despite his clear failure to meet his own self-proclaimed standards to clamp down on vicious bigotry.

“Community standards ensure users feel welcome and advertisers’ brands are safe.

“Advertisers can either enable Mr Musk’s gaslighting, or send a message and stop their marketing dollars enabling the spread of deadly hate. It’s time for advertisers to do their bit to counter hate in our societies.”

– **END**–

Notes

- Elon Musk’s full tweet reads: “Hate speech impressions down by 1/3 from pre-spike levels. Congrats to Twitter team!”

Elon Musk, Twitter, 24 November 2022,
<https://twitter.com/elonmusk/status/1595630109116989440>

- Elon Musk completed the Twitter deal on 27 October 2022

New York Times, 27 October 2022,
<https://www.nytimes.com/2022/10/27/technology/elon-musk-twitter-deal-complete.html>

- The analysis compares the average daily tweets from the month preceding Musk’s takeover (27th September – 26 October) to the week up to Musk’s claim about hate speech impressions (18th November – 24th November).

Summary Table

Term	Volume in week before tweet	% inc. over Pre-Musk Month	% inc. over 2022 average
N****r	30,546	197%	260%
T****y	35,938	3%	63%
F****t	30,068	31%	91%
K**e	2,366	24%	12%
S**c	916	58%	64%

Methodology

- Data was collected using Brandwatch, which includes original tweets, retweets and quote retweets.
- Tweets analyzed were from all locations worldwide but were restricted to the English language
- The underlying dataset can be found at the link below

CCDH, 221125 Slurs under Musk vs preceding Month,
<https://docs.google.com/spreadsheets/d/1ZnOgaD3szi1vqanhTtyxxIfY7uQQ2m1IYVJ9HFGsmYg/edit#gid=0>