

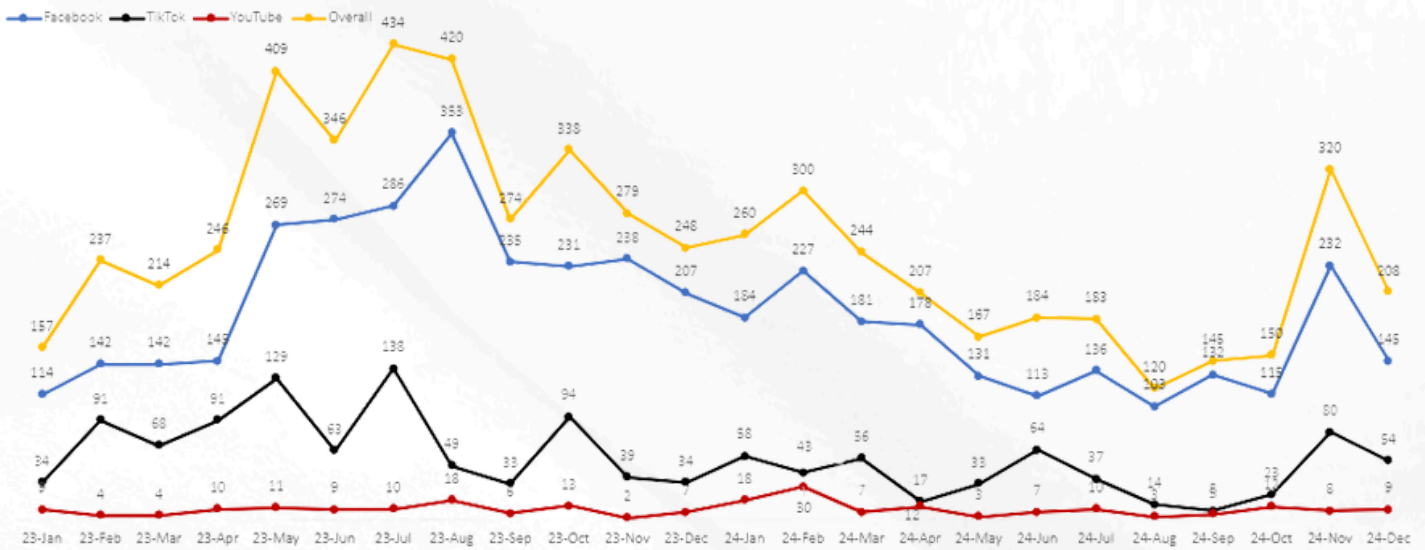
Harmful Speech in Sri Lanka

January 2025

“Hate speech may be used against women, refugees, migrants, gender-diverse and trans people, and minorities. It is massively amplified by the power of digital platforms and tools that enable it to spread across borders and cultures. States have an obligation under international law to prevent and combat incitement to hatred and to promote diversity, mutual understanding, and solidarity.”

UN Secretary General António Guterres

HARMFUL SPEECH TRENDS¹



Breakdown of identified month-on-month harmful speech by Social Media Platform.²

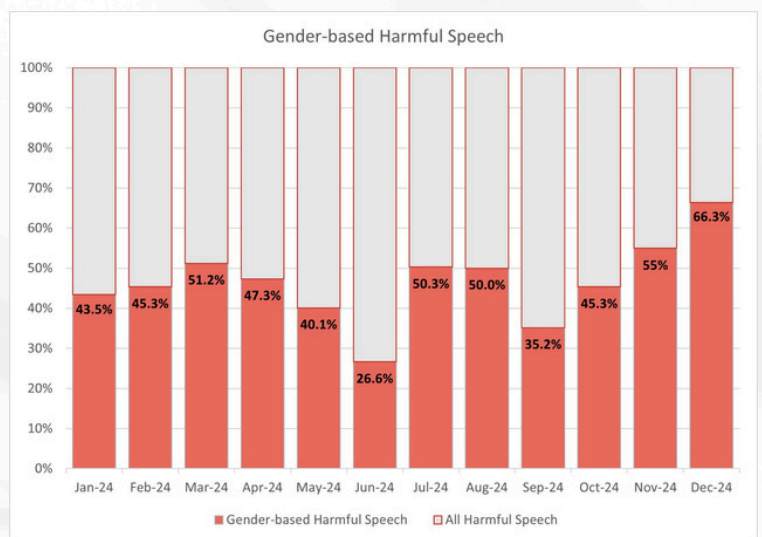
Harmful speech **increased by 113%** in November 2024, coinciding with the Sri Lankan parliamentary election held on 14 November.

This increase was primarily driven by a **159% surge in gender-based harmful speech**.

Over 35% of all gender-based harmful speech recorded in November was targeted at women candidates campaigning for the parliamentary elections.

Harmful speech saw a **35% decrease in December** following the conclusion of the November elections.

However, gender based harmful speech increased as a percentage against the total harmful content recorded, accounting for **66% of all harmful content recorded in December 2024**, up from 55% in November.



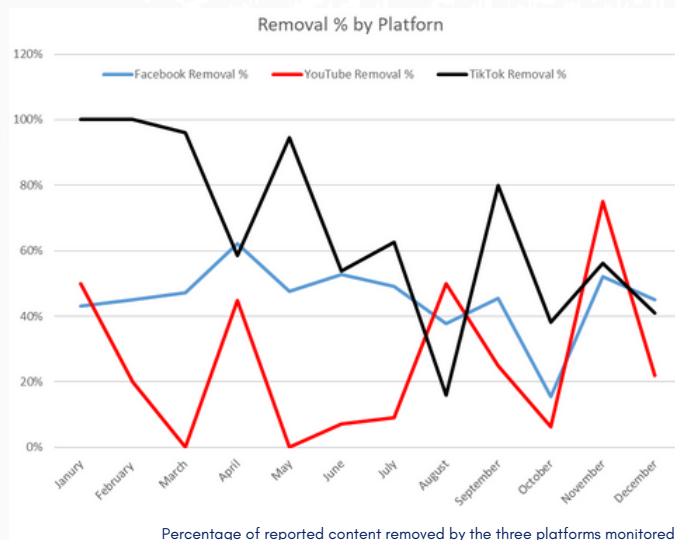
1. Harmful speech includes hate speech, dis/misinformation, harassment and incitement to or glorification of violence.
 2. The ability to monitor Facebook was negatively impacted by the retirement of Meta's CrowdTangle tool on 14 August 2024.
 3. While there is no universally accepted definition of disinformation, it is commonly understood as information that is not only inaccurate, but is also intended to deceive and is spread in order to inflict harm.
 4. The UN in Sri Lanka relies on trusted external fact-checkers to independently monitor dis- and misinformation. Actual levels of dis- and misinformation may be higher than reported due to monitoring capacity constraints.

Responsiveness by platforms to reported content

The responsiveness of Facebook, YouTube, and TikTok in removing reported harmful content **increased in November 2024**, following a sharp decline in October.

However, the removal rates for all three platforms **decreased in December**.

In parallel to the overall decrease, the removal rate for gendered harmful speech decreased from 54% in November to 41% in December.



Implications of changes to Meta's content policy and enforcement

In early January 2025, Meta (the parent company of Facebook) announced it was ending its third-party fact checking programme in the United States and also announced several major changes to content (moderation) policy, including changes to its:

- **Policy on hate speech**, now called "Hateful Conduct": This will allow some speech that was previously prohibited and removes some protections on topics such as immigration and gender.
- **Automated systems**: Meta will now focus on tackling illegal and high-severity content – such as terrorism, drugs, child exploitation, and frauds & scams – and will rely more on user reporting for other policy violations.
- **Approach to political content**: Meta will channel political content to users based on personalized signals, making it easier for people who want to see a particular type of content in their feeds to do so.

While the implications of these changes to content monitoring and reporting in Sri Lanka are as yet unclear, the reduction of safety guardrails intended to protect women, other sexual and gender minorities, and immigrants, among others, may result in less efficient removal of content targeting these groups on Meta-owned platforms.

Addressing gender-based hate speech and disinformation in Sri Lanka

On the International Day for Countering Hate Speech, the UN in Sri Lanka launched the [#UniteAgainstHate](#) campaign to raise awareness about the dangerous impact of harmful speech and ways to combat it.

The campaign includes a focus on gender-based harmful speech.

Forthcoming campaign material will focus on barriers that hinder the political participation of women in Sri Lanka.

HOW TO RESPOND TO **HARMFUL SPEECH**



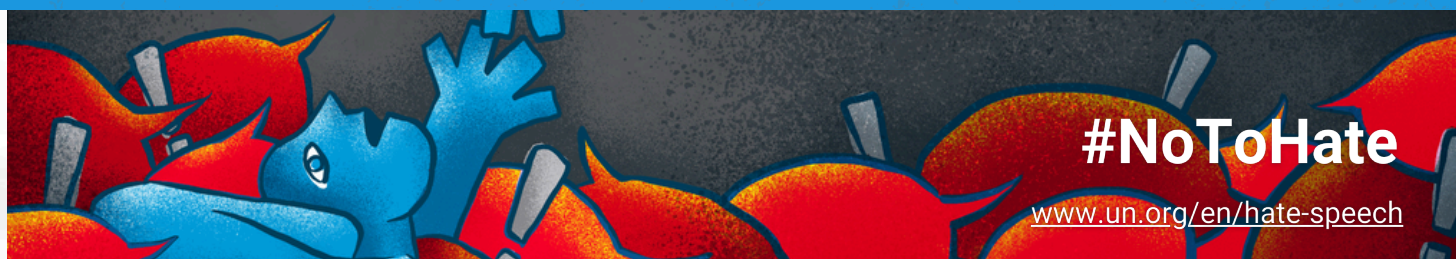
DISENGAGE when you know it's false and/or harmful to someone else



VERIFY authenticity before you share



AMPLIFY facts, call out mis/disinformation, and [#UniteAgainstHate](#)



#NoToHate

www.un.org/en/hate-speech