Harmful Speech in Sri Lanka
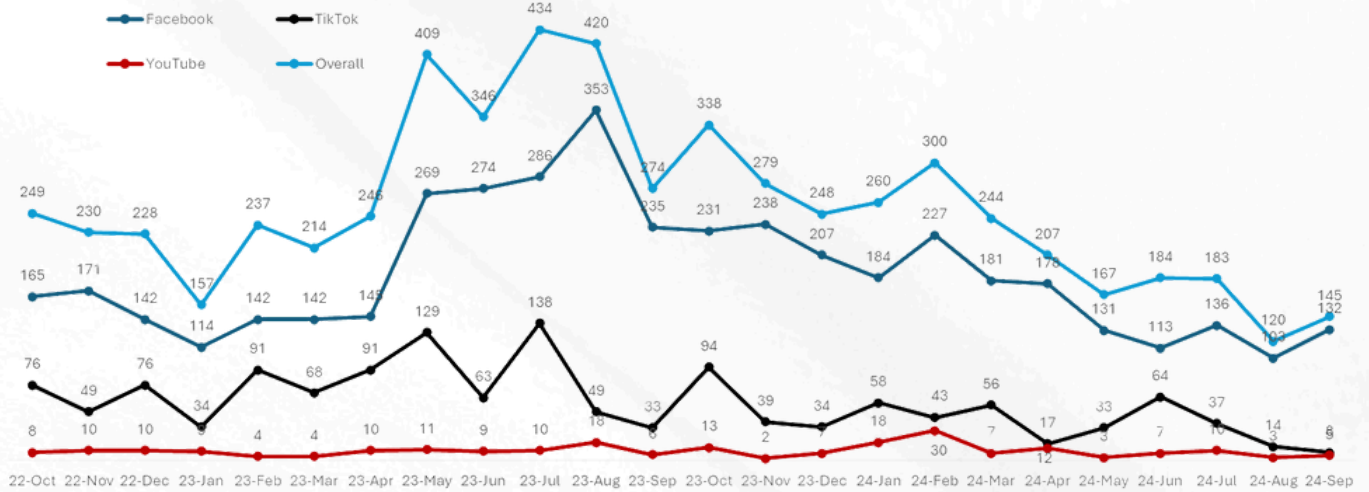
October 2024

> *Hate speech refers to any kind of communication in speech, writing or behaviour, that attacks or uses pejorative or discriminatory language with reference to a person or a group on the basis of who they are, in other words, based on their religion, ethnicity, nationality, race, colour, descent, gender or other identity factor.*
>
> **UN Strategy and Plan of Action on Hate Speech**

HARMFUL SPEECH TRENDS

According to the methodology employed by the United Nations in Sri Lanka, harmful speech includes hate speech, mis/disinformation, harassment and incitement to or glorification of violence.

Breakdown of identified month-on-month harmful speech by Social Media Platform.[1]

Incidences of harmful speech have been on a broad downward trend since August 2023.

Harmful speech increased by 21% month-on-month from 120 posts in August 2024 to 145 posts in September 2024.

This increase was mainly caused by an 8-fold rise in hate speech targeting ethnic and/or religious minorities that followed the declaration of results of the presidential election on 21 September.

Post-election harmful content included derogatory references to minority communities in the Northern and Eastern provinces, as well as the estate sector, in response to their voting preferences.

While marking a 15% month-on-month decrease, disinformation remained elevated in September.[2]

Responses from Platforms to content reported via trusted partner/flagger channels (September 2024)

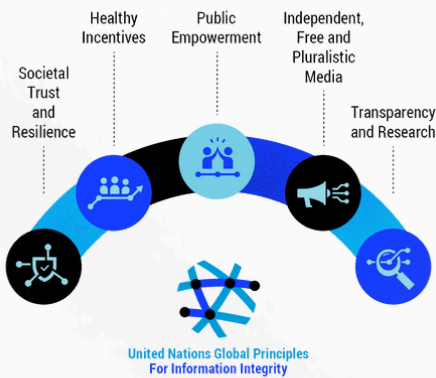|	Response Rate	Removal Rate
YouTube	60%	40%
Facebook	96%	45%
TikTok	100%	75%

1. The ability to monitor Facebook was negatively impacted by the retirement of Meta's CrowdTangle tool on 14 August 2024.
2. 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.

UN RESOURCES TO HELP COMBAT HARMFUL SPEECH

United Nations Global Principles
For Information Integrity

United Nations Global Principles for Information Integrity

The United Nations Global Principles for Information Integrity present a vision of a future in which power imbalances are redressed so that a small group of actors, including technology companies based in a handful of countries, no longer monopolize control over global information flows.

The five principles envision an information ecosystem that delivers choice, freedom, privacy and safety for all, in which people everywhere can express themselves freely and make informed and independent decisions.

A Comprehensive Methodology for Monitoring Social Media

UNOSAPG, with support from the United Kingdom, recently launched a Comprehensive Methodology for Monitoring Social Media. Building upon a long-term engagement with tech and social media companies, the methodology sets out a systematic approach to monitoring hate speech based on the UN Strategy and Plan of Action on Hate Speech and in full respect of international human rights law and standards.

Adressing Harmful Speech in Sri Lanka

On the International Day for Countering Hate Speech (18 June), the UN in Sri Lanka launched the #UniteAgainstHate campaign to raise awareness about the dangerous impact of harmful speech and ways to combat it. Check out some of the campaign resources below.

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