

REPORTS ON REMOVING PROBLEMATIC CONTENT BY PLATFORMS

Social media platforms are getting faster at responding to hate speech online. The major social media platforms removed 72 percent of illegal hate speech on their platforms during 2018. The response rate is a big increase from two years ago, when tech companies removed just 28 percent of content.

10,000 people are now working to fight hate speech on Google's platforms only.

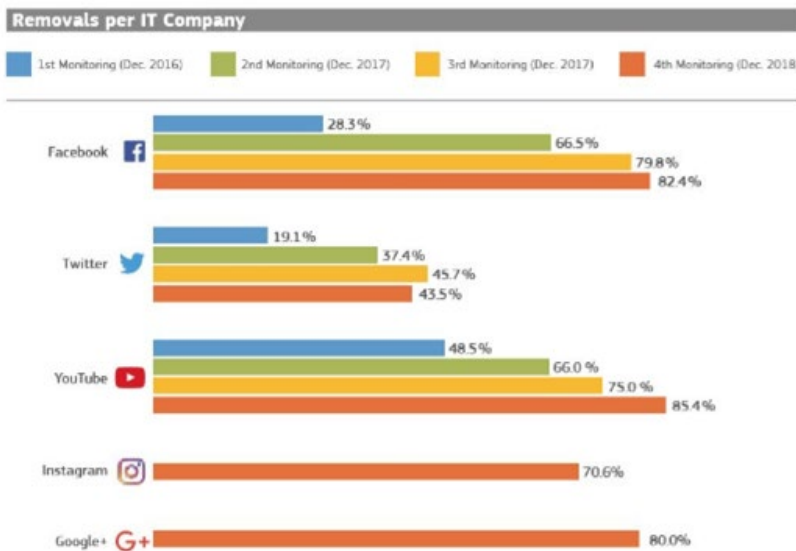
Eighty-nine percent of content flagged as hate speech was reviewed by platforms within 24 hours, up from 40 percent in 2016.

Facebook now removes 82 percent of illegal hate speech, up from 28 percent in 2016.

However, 43.5 percent of hate speech flagged on Twitter was removed in that time frame, down from 45 percent in December 2017. Twitter is now reviewing 88 percent of all notifications received within 24 hours.

It might be useful to compare this data with situation in Germany after the NetzDG. As can be seen in Table 1, noticeably, all platforms responded to most of the reported content that required action within one day, and deleted or blocked access. However, by a large majority platforms rejected the complaints they received from users. Especially Twitter was flooded with complaint notices, but only identified a need to take action in roughly 11 percent of reported pieces of content.

Figure 1: Removals per Platform



² YouTube has also limited the features of an additional 23 videos: this implies that while not being removed, a video may not be liked, commented, or shared and does not appear in searches.

Compiled by SCM (2019) from

Elizabeth Schulze (2019, February 4). EU says Facebook, Google and Twitter are getting faster at removing hate speech online

Kirsten Gollatz, Martin J. Riedl and Jens Pohlmann (2018). Removal of online hate speech in numbers. Media Policy Project Blog.

Table 1: Reported Data for Germany for the first Half of 2018

Platform	Total items reported	Reports resulted in action (removal rate)	Removal rate within 24h
Facebook	1,704	362 (21.2 percent)	76.4 percent (of reports)
YouTube	241,827	58,297 (27.1 percent)	93.0 percent (54.199)
Google+	2,769	1,277 (46.1 percent)	93.8 percent (1.198)
Twitter	264,818	28,645 (10.8 percent)	97.9 percent (28.044)