



# How the Code of Conduct helped countering illegal hate speech online



Factsheet | February 2019

**Věra Jourová**

Commissioner for Justice,  
Consumers and Gender Equality



Directorate-General for  
Justice and Consumers



Since the launch of the Code of Conduct in May 2016, this Commission initiative has effectively helped countering illegal hate speech online.

## Commitments of the Code of Conduct

Have rules and community standards that prohibit hate speech and put in place systems and teams to review content that is reported to violate these standards.

- ✓ All IT companies who signed the Code now have terms of service, rules or community standards prohibiting users from posting content inciting violence or hatred against protected groups. They have also significantly increased the number of employees monitoring and reviewing the content. Facebook reports having a global network of about 15,000 people working on all types of content review.



**Review the majority of the content flagged within 24 hours and remove or disable access to hate speech content, if necessary**

- ✓ On average, IT companies are now assessing 89% of flagged content within 24 hours, up from 81% one year ago. Removal rate remains stable at around 70%, which is satisfactory as hate speech is not easy to define. Its illegality has to be balanced with the right to freedom of expression and the context.

**Provide regular training to their staff**

- ✓ IT companies reportedly hold regular training, and provide coaching and support for their team of content reviewers.



**Engage in partnerships and training activities with civil society in order to enlarge their network of trusted reporters as well as work on promoting independent counter-narratives and educational programmes.**



- ✓ The IT companies reported a considerable extension of their network of ‘trusted flaggers’ in Europe and are engaging on a regular basis with them to increase understanding of national specificities of hate speech. In the first year after the signature of the Code of conduct, Facebook reported to have taken 66 EU NGOs on board as trusted flaggers; and Twitter 40 NGOs in 21 EU countries.

Out of a total of 38 training sessions provided in 2018 by YouTube to NGOs on their content policy and trusted flagger programme, 18 were focused on hate speech and abusive content.

IT companies also work together with their ‘trusted flaggers’ on campaigns for tolerance and pluralism online. For instance, between 2017 and 2019, three workshops took place at the headquarters of YouTube, Twitter and Facebook to facilitate such initiatives.

**Establish national contact points to communicate with competent national authorities.**

- ✓ All IT companies that subscribed to the Code of Conduct have established national points of contact to facilitate contact with the relevant competent authorities at national level. It is important to highlight that the work in the Code of Conduct complements legislation fighting racism and xenophobia, which requires authors of illegal hate speech offences - whether online or offline - to be effectively prosecuted.

**Promote transparency**

- ✓ In 2016, IT companies only made information available on the number of law enforcement request and rarely provided any detail on online hate speech as a specific ground for removal. Today, the removals of hate speech content are well presented, on a regular basis, on each of the IT Company transparency reports, for example see the transparency reports published by Facebook, Twitter and YouTube.
- ✓ Before the Code of Conduct was launched, users rarely received a response by IT companies when they notified hate speech content. In addition, the reporting or flagging function was often not very user-friendly. While there is scope for progress and the performance of the companies differs in this respect, now on average around two thirds of the notifications receive a systematic response detailing the outcomes and measures taken.

**Foster cooperation with other platforms with a view to promote the adherence to the Code**

- ✓ Since 2018, five new companies joined the Code of Conduct: Instagram, Google+, Snapchat, Dailymotion and jeuxvideo.com. This brings the total number of companies that are part of the Code of Conduct to nine and a coverage of 96% of the EU market share of online platforms which may be affected by hateful content. The Code has become an industry standard attracting a diverse range of platforms and both big and smaller businesses.