

**From:** [REDACTED]  
**To:** [REDACTED]  
**Cc:** [BIISA Lucrezia \(CAB-REYNERS\)](#); [REDACTED]  
**Subject:** RE: Our commitment to the Code EU Code of Conduct on Illegal Hate Speech Online - some proposals for your consideration  
**Date:** Wednesday 5 May 2021 12:45:00  
**Attachments:** [image001.png](#)  
[image002.jpg](#)  
[image003.jpg](#)

---

Dear [REDACTED], dear all,

Many thanks for your email of last 23 April, and for your continuous engagement on fighting against hate speech online. We are grateful to receive this joint message from all of you, reiterating your commitments to progress together on the implementation of the Code of conduct.

Regarding your specific proposals in view of the publication of the results of the 6<sup>th</sup> monitoring: we are very open to consider options to integrate the factsheet with additional information that you may provide. At this stage, we would have the following comments/feedback:

- On the first idea: as you know, we have always been interested in getting from you any information you had available on hate speech content detected (including automatically), and removed, other than the flags sent to you during the monitoring period. We would therefore encourage you to discuss and agree on a possible set of common indicators on which you all have data, and which would allow to make comparisons over the (coming) years. This will allow to have a more structured approach to this reporting. Indicatively we could think of adding a box of half a page per company, highlighting the main results.
- Regarding your second proposal, and in line with the message we conveyed to you in the past monitoring rounds, we would be very open to include figures on the number of cases where your assessment diverged from the one by the NGOs. What we can't unfortunately do is to enter into each specific case, categorise them and arbitrate on what is "right" to flag, or "illegal", as this is not our role. What we find interesting however is to duly reflect that there is a certain number of discrepancies which mirror the complexity of making assessments on hate speech content in the different (legal) national contexts.

We would like to propose a meeting altogether towards the end of May if OK for you, where we could discuss in depth about the proposals below as well as how we see, from our respective perspectives, the next steps on the Code, including on the engagement with trusted flaggers. Grateful if you could propose already few dates that would be convenient for all of you in the week starting on 24<sup>th</sup> May.

We look forward to discussing this with you soon,

Best regards,

[Redacted]



**European Commission**  
DG Justice and Consumers  
Fundamental Rights Policy  
[Redacted]  
B-1049 Brussels/Belgium  
[Redacted]

-



-

**From:** [Redacted]

**Sent:** Friday, April 23, 2021 4:09 PM

**To:** [Redacted]

**Cc:** BUSA Lucrezia (CAB-REYNDERS) [Redacted]; [Redacted]

[Redacted]  
[Redacted]  
[Redacted]  
[Redacted]  
[Redacted]

**Subject:** Our commitment to the Code EU Code of Conduct on Illegal Hate Speech Online - some proposals for your consideration

Caro [Redacted],

Thank you for sharing that the 6th monitoring round has recently concluded and for DG JUST's commitment to work with us and the NGO partners to consider the results.

We are proud to be signatories of the EU Code of Conduct on Illegal Hate Speech Online and thank DG JUST for its leadership in this area. We believe that the Code has proven so far to be a fruitful process and we are pleased with the positive collaboration with the trusted partners and the continued progress to quickly action reported content.

Since the Code was first published in 2016, the broad landscape and challenges associated with tackling hate speech have constantly been evolving. We have continued to invest heavily in scaling up our efforts against hate speech. In particular, we continue to invest in machine learning and automatic detection so that we can find and remove hate speech content before people report it to us, and sometimes before anyone sees it.

We also continue to analyze how hate speech manifests itself on our platforms, to keep pace with the way it can evolve and inform our enforcement decisions.

As part of our commitment to the Code, we have also been working on developing further our relationship with our trusted partners and building their capacity to help us tackle this kind of content online. We have also dedicated resources to support their efforts to challenge hateful and extremist narratives, including through counterspeech campaigns.

In light of these developments and in order to keep the Code of Conduct relevant, we believe it would be worth updating how we communicate publicly about the Code's achievements and to better frame the progress made in effectively preventing and countering the spread of illegal hate speech online.

In particular, we think it would be important to:

- **Have the opportunity to communicate the breadth of our efforts to tackle hate speech alongside the results of the Code of Conduct's yearly monitoring exercise.** For many companies, the vast majority of removals now happen before content is even flagged. Therefore, the results of the monitoring exercises (including only data about removals based on reports submitted by civil society organizations) do not depict fully the reality of our efforts to curb the spread of hate speech online, nor represent a sufficient indicator to measure the success of those efforts. Therefore, we believe it could be valuable to use the opportunity of the publication of the yearly test results to also shine a light on the broader context, which could be done by also referring to our transparency reporting efforts.
- **Identify separate categories focused on cases where platforms' legal teams and NGOs disagree on the illegality of the content and exclude reports that were not illegal content from the results**

We think it is important for the test results to register the complexity of assessing illegal hate speech in cases where the applicable legislative thresholds are somewhat vague or the intended meaning of the reported content itself is open to different interpretations. To this end, the monitoring exercise results' factsheet could highlight these cases in a separate category and report on their number to help shed light on the complexity of such content.

We are committed to remaining constructive partners and playing our part in making the Code of Conduct successful. We would like to discuss with you opportunities for DG JUST to promote further collaboration between platforms and NGOs outside of the monitoring exercise. We have found that discussing our approaches to tackling hate speech and sharing insights has been extremely useful for all involved.

We remain at your disposal to further discuss these ideas.


 (on behalf of YouTube, Facebook, Twitter and TikTok)

--



Director Public Policy and Government Affairs EMEA

YouTube | Google, Inc.



This email may be confidential or privileged. If you received this communication by mistake, please don't forward it to anyone else, please erase all copies and attachments, and please let me know that it went to the wrong person. Thanks.