

The EU's Digital Services and Digital markets act:

A new frontier against hate speech?

By Leandra Voss

The proliferation of social media platforms (SMPs) have enabled individuals to have widespread access to information available on the internet and a huge audience for their own online content. While SMPs such as Instagram and Tiktok have united many of us, many problems have surfaced as well, with perhaps the most pressing one being the spread of hate speech and the many dangers it poses. Many European democracies struggle with keeping hate speech under control while still respecting freedom of speech and plurality of opinion. With a rise in social media and online hate speech, this debate pertaining to human rights and their curtailment, when necessary, requires extensive examination. The EU is beginning to tackle this issue by adopting the Digital Services and Digital Markets act (DSA), which will be applied by the beginning of 2024 (Le Monde, 2022). Among other things, the act will hold social media companies accountable for content posted on their website and their promotion algorithms (Le Monde, 2022). Drawing clear, legal boundaries is especially important since countries are facing an up to 20% increase in online hate speech (Baggs, 2021). European countries are also affected by this: Finland, for example, saw severe anti-trans hate speech online after their parliament discussed laws on transgender rights, which culminated in homemade bombs detonated and pride benches being set on fire during pride (Eastham, 2023). Thus, creating these online protection laws is important to provide marginalised groups and vulnerable individuals with the freedom to live in safety, freedom of expression and the right to be legally protected in the same way everywhere, even online. However, in proposing these laws, the EU must avoid restricting everyone's right to expression. This paper will first examine the effects of online hate speech on human rights and review the existing hate speech regulations in the EU to evaluate how nations should best implement the

DSA. Finally, it argues that hate speech monitoring is absolutely necessary and tech companies should be heavily regulated and implicated in the monitoring and supervision system.

Problem definition: what is hate speech

Today, debates around online anti-hate speech laws are not new, they have just remained at the national-level. The tension in this debate is whether to maintain a completely open social media or restrict all hateful speech. To delve into this issue further, we need to clarify what hate speech is and why it's so important to regulate. According to the European charter for human rights, hate speech is "all forms of expression which spread, incite, promote or justify racial hatred, xenophobia, antisemitism or other forms of hatred based on intolerance" (Parmar, 2018). The online space is fertile ground for hateful narratives to spread and incite varying acts of hatred, therefore restricting an individual's right to exist without discrimination and to express themselves. Hate speech against minorities online needs to be legally controlled, as the UN council for minorities noted that hate speech online promotes and normalises the dehumanisation of minorities (United Nations, n.d.). These arguments are set starkly apart from ideas that restricting hate speech limits the plurality of opinion in democracies. This view considers hate speech as a matter of freedom of speech unless it is an explicit threat, a policy which is promoted by the US. Whilst creating the new set of laws, the EU must consider the effects that different hate speech has and should go about strategically interfering so as not to over-censor free speech.

Censorship law in action: The case of the NetzDG

Although the EU is only now creating EU-level regulations, several nations have gone through different iterations of hate-speech laws. Passed in 2017 and updated in 2020,

Germany has passed the NetzDG, similar legislation aiming to control and regulate hate speech more effectively on social media. It advises social media to make it easier to report hate speech and forces the content-reporting departments to send instances of hate speech to a designated branch of the German police, who could then evaluate if this was truly hate speech and take legal action. Using the new legislation, politician Renate Künast successfully reported 12 instances of hate speech towards her, as a politician and woman (Keilani, 2020). Through this legislation, Germany can take reports of threats online as seriously as in-person acts. However, there have been disagreements over NetzDG. Many, including Austrian Reporters without Borders, argue that there is an issue of *over-blocking* wherein non-hate comments are reported and needlessly censored (Reporter Ohne Grenzen, n.d).

Over-blocking: a new threat to freedom of expression

Over-blocking can either be caused by algorithms incorrectly judging hate speech or by attempts to restrict speech by governments with ulterior motives. Following the enactment of the NetzDG, several dictatorships including the Philippines, Turkey and Venezuela enacted their own hate speech laws (Delcker, 2020). These enactments allow nations to have needless control of social media and to censor opposing speech to maintain control. The free use of social media is especially sacred as it is the most accessible place to express freedom of speech and interact with a wide array of information.

Additionally, over-blocking can be used to silence specific people. This was the case for a trans woman who reported to Politico that right-wing trolls not only sent her death threats but also spam-reported her Twitter account, which led to Twitter shutting her account down for several months (Delcker, 2020). NetzDG regulations did not consider these kinds of limitations of social media censorship algorithms which have been subject to many similar controversies. Social media companies should not be entirely trusted in the censorship of

citizens when implementing anti-hate speech laws, as this can pose even more human rights abuses (in restricting people's freedom of speech). Generally, social media censorship regulations have heavily steered towards being advertiser friendly and have only recently been made public.

Recommendations

After this analysis, it is clear that in implementing the DSA, nations have to be cautious in their relationship with social media companies. The EU needs to carefully monitor and set healthy regulations for the censorship processes on social media platforms, including prioritising person moderators over AI. They should also carefully consider what precedents these laws set, and strengthen their foreign policy to advocate for freedom of speech online whilst also considering fair punishment of hate speech. For example, the EU has to be careful in monitoring its own increasingly autocratic nations, such as Poland, to ensure they are not abusing this censorship by targeting opponents unfavourably. Finally, the most important point is to avoid excessively limiting all negative speech online, but to find individuals and groups that are perpetuating and encouraging hate speech and take legal action against those people.

In conclusion, we are at a new frontier of speech, where we can communicate thoughts to millions of people at the push of a button. For this new frontier, we need to concretely set how far freedom of speech can go and how we can control hate speech to maximise plurality and safety in Europe and ensure and protect human rights for all.

Edited by Harshita Pugalia and the Short Paper Management Team.

References

- Baggs, M. (2021, November 15). Online hate speech rose 20% during pandemic: “We’ve normalised it.” *BBC News*. <https://www.bbc.com/news/newsbeat-59292509>
- Delcker, J. (2020, October 6). Germany’s Balancing Act: Fighting Online Hate While Protecting Free Speech. *POLITICO*. <https://www.politico.eu/article/germany-hate-speech-internet-netzdg-controversial-legislation/>
- Eastham, J. (2023, February 21). *ILGA-Europe reports spike in violence towards LGBTQI+*. [Www.euractiv.com; Euractiv. https://www.euractiv.com/section/non-discrimination/news/ilga-europe-reports-spike-in-violence-towards-lgbtqi/](https://www.euractiv.com/section/non-discrimination/news/ilga-europe-reports-spike-in-violence-towards-lgbtqi/)
- Keilani, F. (2020, March 25). Onlinehetze vor Gericht: Renate Künast gewinnt erneut in Hate-Speech-Verfahren. *Der Tagesspiegel Online*. <https://www.tagesspiegel.de/berlin/renate-kunast-gewinnt-erneut-in-hate-speech-verfahren-6058387.html>
- Le Monde. (2022, April 23). EU law to protect Internet users and their “fundamental rights online” against hate speech, disinformation. *Le Monde*. https://www.lemonde.fr/en/pixels/article/2022/04/23/eu-law-to-protect-internet-users-and-their-fundamental-rights-online-against-hate-speech-disinformation_5981333_13.html
- Parmar, S. (2018). *The Legal Framework for Addressing “Hate Speech” in Europe*. <https://rm.coe.int/opening-session-2-parmar-the-legal-framework-for-addressing-hate-speech/16808ee4bf>
- Reporter Ohne Grenzen. (n.d.). *Österreichisches Plattformgesetz: Fehler bei NetzDG vermeiden*. OTS.at. Retrieved March 1, 2023, from https://www.ots.at/presseaussendung/OTS_20200904_OTSO156/oesterreichisches-plattformgesetz-fehler-bei-netzdg-vermeiden

United Nations. (n.d.). *Hate speech and real harm*. United Nations. Retrieved from <https://www.un.org/en/hate-speech/understanding-hate-speech/hate-speech-and-real-harm>