

The EU's Initiative to Extend its List of 'EU Crimes': A New Twist in the Saga of Policing Hate Speech

Ophélie Stockhem

On 25 November 2021, the European Commission published a [Communication](#), in which it recommends that the European Council adopt a decision to add hate speech and hate crime to the list of 'EU crimes' articulated in Article 83(1) of the founding 'Treaty on the Functioning of the European Union' (TFEU). If adopted unanimously by the European Council and after obtaining the consent from the European Parliament, this proposal would pave the way for the Commission to outline secondary legislation aiming to define and punish online and offline hate speech, and to push member states to adopt stronger measures against perpetrators. This short article analyses this initiative in light of the fragmented international legal framework and sometimes contradictory jurisprudence that exists concerning hate speech and how this interacts with the right to freedom of speech. In doing so, CDT highlights the significant fundamental rights concerns raised by the proposal to criminalise speech, as well as the difficulty the EU is likely to face in moving this politically untenable recommendation forward.

Combating 'Weaponised Words' in the Internet Ecosystem

There is no doubt that the Internet ecosystem has expanded our connectivity and definitively reshaped our democratic and societal spaces at a very rapid pace. Such an environment offers immeasurable benefits for free expression, social mobilisation—especially within oppressive regimes—and social progress. Changes that came with the advent of social networks are particularly hard to ignore: they have provided a new avenue to express opinions, space to give visibility to systemic and pressing structural inequalities, and space for collective mobilisation, which has led to marked social change.

However, this constantly evolving ecosystem has also raised growing fears about the proliferation of abusive speech such as hate speech and its impact on fundamental rights and democracy. As outlined in the [last report of the EU Fundamental Rights Agency \(FRA\)](#), the online and offline worlds have recently witnessed an increase in the targeting of ethnic and religious minorities, women, migrants, human rights defenders, and LGBTI+ people through messages of hate. Such phenomena not only negatively affect civic space and polarise society by preventing citizens from safely and meaningfully engaging in democratic processes, but also foster a general climate of intolerance, distrust and fear.

Online hate speech is, however, hard to define and renders the work of human moderators particularly challenging. Because of the volume of speech they must moderate, online platforms are increasingly resorting to automated filtering technologies to regulate their content. Despite [being controversially interpreted by the EU as authorised in the field of copyright](#), upload filters generally raise significant issues as they threaten free speech and the E-Commerce liability exemption. In the context of combating hate speech, these risks are heightened [due to difficulty](#)

in differentiating ‘illegal hate speech’ - a concept which is already highly controversial among member states - and speech that is lawful, but may be deemed inappropriate by some.

The variation on how EU member states view legislating on hate speech raises a fundamental question on how plausible the European Commission’s proposal to criminalise such speech at the EU level will actually be. The EU is already in the midst of updating and modernising its horizontal regulatory framework on content moderation, through the Digital Services Act, in a context in which a growing number of member states such as Germany, France, and Austria have proposed or adopted national laws which threaten to subvert the crucial general monitoring prohibition contained within the existing EU online liability regime.

For example, Germany’s Network Enforcement Act (NetzDG), adopted in 2017 and recently amended, probably best illustrates dubious attempts to regulate harmful content, including hate speech, by a Member State, as it compels removals and blocking of ‘violating content’ within short time periods, on the basis of provisions contained in the German Criminal Code and incentivises these actions through severe administrative penalties. Despite constitutional challenges, strong governmental opposition, and lawsuits filed by social media platforms, some EU countries are still contemplating the idea of adopting a similar regime, using the German law as a ‘model’. Those trends have been monitored closely by CDT and other civil society organisations, as both filtering mandates and intermediary liability laws significantly risk excluding legal and socially beneficial materials, exacerbating discrimination, enhancing biases and silencing dissent and minority voices.

These varying perspectives by EU Member States on platform governance and the intermediary liability framework already present a complex environment that the EU institutions are attempting to navigate. Though the growing concern on the proliferation of hate speech online is valid, we can only foresee that attempting to achieve unanimity within the Council to change the treaties to include hate speech will be similarly challenging, if not unfeasible. Not only because hate speech as a legal concept lacks a universal, undisputed and uncontroversial definition, but that it similarly maintains a history of legal uncertainty and contradictory jurisprudence, on which the EU simply cannot rely.

International and Regional Frameworks on Hate Speech: Legally Uncertain Examples ?

The concerns stemming from the dissemination of hate speech are not new, but as mentioned, the legal landscape that exists does not provide the EU with a solid foundation to build upon. Early in their emergence, international and regional human rights treaties attempted to regulate the issue as insurance against a return to Europe’s twentieth century totalitarian abuses. Within international legal standards, hate speech is addressed in the International Covenant on Civil and Political Rights (ICCPR),¹ which, among other things, prohibits the advocacy of national, racial or religious hatred. The Convention on the Prevention and Punishment of the Crime of

¹ Article 20§2.

Genocide, also condemns the incitement to commit genocide,² and the Convention for the elimination of racial discrimination (CERD), forbids propaganda and organisations based on racial hatred and discrimination.³

At the regional level, the Council of Europe (CoE) Convention on Human Rights (ECHR) prohibits hate speech through two different approaches. One broader approach is a categorical *exclusion* of certain expressions considered to be hate speech from the protection of freedom of expression granted by the Convention in case of ‘abuse of rights’⁴, which does not include a balancing of hate speech with freedom of expression. The second narrower approach involves *restriction* to the right to freedom of expression, which operates such a balancing process, with hate speech being a potentially valid legal ground to limit freedom of expression if the traditional legality, legitimate aim, and necessity criteria are fulfilled.⁵ The CoE additionally criminalises acts of a racist and xenophobic nature committed through computer systems in its Additional Protocol to the Convention on Cybercrime. Other regional systems similarly regulate harmful forms of speech, such as in the American Convention on Human Rights, as interpreted by the Inter-American Court of Human Rights.

Although these provisions have been essential developments in building legal standards against hate speech that also apply to the online world, they sometimes have been criticised for being hard to reconcile with the interpretation international courts have given to these norms.⁶ This is particularly true of the European Court of Human Rights (ECtHR), which, despite having ruled for decades that freedom of expression protects not only favourably received expressions but also those that ‘offend, shock or disturb’, has developed a significant body of case law which departs from this approach. The Strasbourg judicial body has relied instead on the prohibition of abuse of rights or so-called ‘guillotine effect’ clause, which excludes certain forms of expressions considered as hateful from the protection of the applicant’s right to freedom of expression, to the detriment of a more balanced analysis involving an in-depth reflection on the limitations of this right. Through this type of analysis, the ECtHR has often leaned heavily on national courts findings to hold that the (sometimes disproportionate)⁷ interference to free expression was necessary in a democratic society, establishing a precedent which might not adequately protect political speech on controversial issues such as criticism of public officials and government institutions.

² See Article 3.

³ Under Article 4 of the Convention.

⁴ Article 17 of the ECHR.

⁵ Article 10§2 of the ECHR.

⁶ Read, for instance, T. Mendel, “Does International Law Provide for Consistent Rules on Hate Speech?” In M. Herz & P. Molnar (Eds.), *The Content and Context of Hate Speech: Rethinking Regulation and Responses*, Cambridge University Press, pp. 417-429.

⁷ Read, in this regard, H. Cannie and D. Voorhoof, “The Abuse Clause and Freedom of Expression in the European Human Rights Convention: An Added Value for Democracy and Human Rights Protection?”, *Netherlands Quarterly of Human Rights*, Vol. 29/1, 2011, pp. 54–83.

The EU's 'Illegal Hate Speech' Puzzle

Within this context, an additional layer of complexity arises when considering that the EU's approach in this area has also been disharmonised and unclear across the Member States. To set the scene, it's important to take note of the specific EU legislative framework.

The commitment of the European Union to combat hate speech is closely related to other legal areas, such as non discrimination, as enshrined in the TFEU,⁸ the Charter of Fundamental Rights, and secondary legislation.⁹ Under this framework, the EU requires member states to combat discrimination based on protected characteristics such as sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation. The Union also directly requires member states to criminalise most severe forms of racism and xenophobia through a Framework Decision.¹⁰

In implementing this framework however, Member states have adopted significantly different strategies in updating these requirements into national laws. Some of them have failed to transpose the directly applicable Decision into their criminal legal system, while most of the others differ substantially in their interpretations of what constitutes illegal, and therefore punishable, speech. More so, the legal framework is heavily fragmented: most national legal systems include not only laws regulating hate speech directly, but also a broad array of related laws, such as non-discrimination, media regulation, and intermediary liability.¹¹ This can make the overall legal framework governing liability for hate speech difficult to interpret. In the Internet ecosystem, these difficulties add to the judicial uncertainty and inconsistency reflected in the way national courts deal with hate speech cases, especially in the context of content moderation, an area that the European Court of Justice also failed to clarify in its case law. In is also worth noting that, although indirectly, some national courts have interpreted national offences such as criminal defamation,¹² or "insult"¹³ as a way to criminalise certain types of online 'hate speech'.

Alongside the challenges seen in the implementation of these binding instruments, existing EU soft law measures, despite their useful flexibility and speed of adoption in times of emergencies such as the COVID-19 pandemic, may sometimes replicate decision-making patterns at the expense of rule of law standards.¹⁴ In 2016, CDT raised several concerns about the

⁸ See in particular Articles 10, 19 and 69(3) of the TFEU.

⁹ In particular, Directive 2000/43/EC6 (the Racial Equality Directive), Directive 2000/78/EC7 (the Employment Equality Directive) and Directives 2004/113/EC and 2006/54/EC8 (the Equal Treatment Directive).

¹⁰ Council Framework Decision 2008/913/JHA.

¹¹ See, for instance, at the EU level, the Audiovisual media service directive and the e-Commerce directive.

¹² See Austria, Poland, and Italy's legal frameworks.

¹³ E.g. In Poland and Germany.

¹⁴ See, in particular M. Eliantonio & O. Stefan, "The Elusive Legitimacy of EU Soft Law: An Analysis of Consultation and Participation in the Process of Adopting COVID-19 Soft Law in the EU", *Eur J Risk Regul.* 2021 Mar; 12(1): pp. 159–175.

Commission's Code of Conduct on countering illegal hate speech online, which creates voluntary commitments for IT companies to prevent and counter the spread of illegal hate speech online, including prioritising notifications of allegedly illegal hate speech received from so-called "trusted flaggers". Setting a dubious precedent for non-judicial authorities to be the arbitrators of illegal hate speech, the measure was also criticised by UN experts for the lack of independence of authorities responsible for removals of speech and the absence of adequate access to remedy for affected speakers.

Taking all of this into consideration, it is evident that though the aim of addressing hate speech online is a valid and necessary pursuit, the legislative landscape in which the EU operates is not conducive to the objective of adding hate speech to the EU criminal law frameworks. Much remains to be reconciled and alternative avenues would be more appropriate.

Conclusion

Despite undoubtedly demonstrating a high level of political will on the part of the European Union, the initiative to extend the list of EU Crimes will likely result in the EU being confronted with significant hurdles. While achieving unanimity within the European Council seems a politically untenable objective, so too does the prospect of being able to effectively implement secondary legislation uniformly and fairly across the EU Member States. Extending the list of 'EU crimes' to add 'hate speech', would add to the inconsistency and fragmentation of international, regional and national legal and judicial systems in the area of hate speech, and establish an additional framework which lacks the legal certainty, foreseeability, and clarity required by international human rights norms.

The Centre for Democracy & Technology Europe therefore strongly cautions the European Commission, and in light of the analysis presented here, we conclude that the proposal poses a threat to fundamental rights. Addressing the rise of offline and online hate speech and its root causes is crucial. However, giving considerable censorship powers to the European Commission and the 27 EU member states, some of which already prohibit forms of hate speech with varying degrees of compliance and significant differences in interpretation, risks fundamentally infringing on the right to freedom of expression of online users, especially in the face of efforts to automate content moderation online. CDT Europe also fears that, without a reconciliation of the issues we have presented here, which require monumental international efforts and clearer jurisprudence, the efforts by the EU as proposed would inadvertently lead to the amplification of hate speech, rather than reducing it, and to the silencing of the very voices that the Commission aims to protect in the first place with these measures.

For all these reasons, CDT recommends that the EU reconsider how it can effectively outline obligations to reduce the proliferation of hate speech through mechanisms it is already working on, such as the Digital Services Act. It further calls on EU institutions to reflect deeply on how



best to develop a holistic approach to combating hate speech, in respect to all fundamental rights and in close coordination with civil society and legal experts.