No. 2.9
Version 01

PRIVATISED ENFORCEMENT AND THE RIGHT TO FREEDOM OF EXPRESSION

The purpose of this paper is to explore the risks of privatised enforcement in the field of terrorism propaganda, **stemming from the EU Code of conduct on countering illegal hate speech online**. The Code merely focuses on the removal of 'illegal hate speech' whereas the countering of terrorism propaganda formed one of the main incentives for its adoption. The author argues that implementation of it may undermine the rule of law and give rise to private censorship. In order to outweigh these risks, IT companies should improve their transparency, especially towards users whose content have been affected. Where automated means are used, the companies should always have in place some form of human intervention in order to contextualise posts. At the EU level, the Commission should provide IT companies with clearer guidelines regarding their liability exemption under the e-Commerce Directive. This would help prevent a race-to-the bottom where intermediaries choose to interpret and apply the most stringent national laws in order to secure at utmost their liability. The paper further articulates on the fine line that exists between 'terrorist content' and 'illegal hate speech' and the need for more detailed definitions.

The Dangers of Practice of 'privatised enforcement'

When placing private companies at the frontline of law enforcement online, the risk may arise that our right to freedom of expression is merely guided by their terms of service, which may not always be in accordance with the level of protection guaranteed under human rights instruments. The EU Commission issued, on 12 September 2018, a proposal for a Regulation on the prevention of terrorist content online.

Taking into account the primary profit-making nature of platforms, it is questionable in how far delegation of such large-scale public functions, which are fundamental to the proper function of our democracy, may be at odd with their business objectives and thereby result in a conflict of interests. **Taking into account the intermediaries' data-driven business model, placing them at the frontline of law enforcement may be dangerous from a legal point of view but also for democracy in general.**

Besides the proposal's general requirement that hosting service providers should remove or disable access to terrorist content within one hour after receipt of a removal order, it also encourages the use of 'referrals', whose content should be assessed against the companies *own terms and conditions*. In that respect, it makes no reference to the law.

The EU Code of conduct is non-binding instrument encourages companies to assess the legality of a post within 24 hours after being notified and to remove or block access to it in case of unlawfulness. Importantly, it explicitly stipulates that the notified posts have to be primarily reviewed against the company's rules and community guidelines and only *'where necessary'* (emphasis added) against national laws. Through these means, specifically encouraging the companies to 'take the lead' and initiative in tackling illegal hate speech online, the Code stimulates the occurrence of privatised enforcement (as a practice in which private companies undertake 'non-law based "voluntary" enforcement measures').

Different ways to balance the dangers of privatized enforcement on the right to freedom of expression

One way to counterbalance the issue of overly broad terms of service through which the rule of law may be threatened would be to provide **legal safeguards to end users**. In this regard, it is important for IT companies to be transparent and accountable and to take into account due process principles.

Whilst the Code states that it promotes transparency, it only does so by encouraging publication of transparency reports. In the two latest periodical reviews, no attention was paid to the existence of transparency measures towards end users whose post had been notified and/or removed. The main focus was whether the companies had provided feedback to *notifying* users. Whilst the Commission did stress, in its communication, the importance of transparency reports, it also stressed the importance of being transparent towards users whose post had been notified and that information shall be provided about received counter-notices. Intrinsically related to this point the companies should have in place a system of counter-notices. This would help uphold due process principles in notice-and-actions procedures.

Another way to secure respect for the rule of law online would be through the States' **positive obligations.** Discussions should find place in order to 'operationalize relevant positive obligations of States in the context of self-regulatory or privatised law enforcement measures by online intermediaries'.

Concerning the countering of private censorship, IT companies should have **more legal certainty about their liability exemption** provided for under the e-Commerce Directive. Indeed, when taking into account that internet intermediaries could potentially be subject to the laws of all countries in which their content is accessible, the safest way for them to act would be to take a restrictive approach and treat the harshest laws as threshold for content removal.

Another possible way to achieve a higher level of legal certainty would, yet again, be through positive state obligations. Importantly, the ECtHR established in *Dink v. Turkey* (para. 137) that one of these obligations consists in ensuring that individuals can express themselves without fear. In light of this, legal scholars have held that such a positive obligation could include **the duty to reduce internet intermediaries' fear of being held liable**, which would be a 'promotional obligation'.

Conclusion

From an EU-perspective, a shift from the focus on 'speed' to 'legality' should take place. Whereas the Code adopted a 24-hour framework for removal of illegal content, the recommendation on tackling illegal content online and the recently proposed regulation (COM(2018) 641 final) encourages removal of terrorist content within one hour. Such short time frame, paired with the unclear definition attributed to 'terrorist content', will undoubtedly magnify the risks of over-removal of content. Moreover, the EU should clarify the liability exemption under the e-Commerce Directive by giving clear guidance on what the terms contained therein entail. This would help prevent a race-to-the bottom where intermediaries choose to interpret and apply the most stringent national laws in order to secure at utmost their liability. IT companies should always provide counter-notices and provide feedback. Human intervention should also be a *conditio sine qua non* in cases where there is no human in the loop and thus not only 'where appropriate' as stipulated in the recommendation.

Compiled by SCM (2019) from
Coche, E. (2018). Privatised enforcement and the right to freedom of expression in a world confronted with terrorism propaganda online. *Internet Policy Review*, *7*(4). DOI: 10.14763/2018.4.1382