Commentary

Self-regulation 2:0? A critical reflection of the European fight against disinformation

In presenting the European Democracy Action Plan (EDAP) in 2020, the European Commission pledged to build more resilient democracies across the EU. As part of this plan, the Commission announced intensified measures to combat disinformation, both through the incoming Digital Services Act (DSA) and specific measures to address sponsored content online. Ostensibly, these reforms would end the era of haphazard self-regulation that has characterized the EU response to disinformation. However, purported changes in this area are vaguely framed, and fail to address critical issues such as the regulation of harmful but lawful content. While instruments like the DSA show signs of improvement, shortcomings in this evolving framework represent a continuation of the EU’s piecemeal approach to disinformation.

Author: Ethan Shattock
Affiliation: Department of Law, Maynooth University, Ireland
Received: March 4th, 2021. Accepted: May 15th, 2021. Published: May 31st, 2021.

Decoding the Codes of Practice (CPD)

Since 2018, the European Commission has driven the legislative agenda for online disinformation through the Code of Practice on Disinformation (CPD). In a nutshell, technological signatories including Facebook, Twitter, and more recently TikTok, voluntarily commit to minimizing disinformation and coordinated election interference. Platforms agree to secure services against inauthentic behavior, encourage transparent “issue-based advertising,” and share relevant data to the “research community” (European Commission, 2020).

As a self-regulatory framework, platforms are not obliged to implement any specific practice but agree to report their activities to the Commission. Ultimately, incentives for implementing the CPD are largely predicated on reputation and evading regulation (Bertolini et al., 2021). Accordingly, there is no binding framework “specifically designed to tackle disinformation online,” (De Cock Buning et al., 2018). The CPD review system audits five key areas. These are the scrutiny of advertising, issue-based advertising, integrity of services, empowering consumers, and empowering the research community. The monitoring process itself is somewhat fractured. Numerous bodies are responsible for assessing the CPD. The European Regulators Group for Audiovisual Media Services (ERGA), the Commission, the signatories, and

1 A publication of the Shorenstein Center on Media, Politics and Public Policy at Harvard University, John F. Kennedy School of Government.
Disinformation in the Digital Services Act (DSA): Light touch liability?

The CPD sits within a broader evolving framework in the EU. Since 2000, the flagship instrument for regulating digital services in the single market has been the Electronic Commerce Directive. Arguably, the defining feature of the Directive is not its obligations, but its liability exemptions. Articles 12-15 exempt providers from liability for third-party unlawful content, conditional in that those providers must act “expeditiously” to remove such content. Once providers take steps to remove illegal content, they are effectively absolved from secondary liability for illegal content disseminated by users.

The consensus now is that this Directive is obsolete. Much has changed in Europe’s communication landscape since 2000, and new platforms yield unprecedented opportunities and threats (Borgesius et al., 2018). Unsurprisingly, these shifts have prompted calls for reform in order to align legislation with social
media (De Streel, 2018). In response to these calls, the Commission pledged to modernize platform regulation in its “Shaping Europe’s Digital Future” commitment. In 2020, the Commission unveiled the long-awaited Digital Services Act (DSA). The DSA, in a nutshell, is an attempt to recalibrate intermediary responsibilities for curbing illegal content, in light of contemporary technologies.

The DSA differentiates between different classifications of service providers. These are broken down into:

- Intermediary services (internet access providers, domain name registrars)
- Hosts (cloud and web hosting services)
- Online platforms (app stores and social media platforms)
- Very large online platforms (platforms reaching more than 10% of 450 million monthly European consumers)

Certain provisions of the DSA apply to all providers. For example, all providers are subjected to new transparency obligations, and all providers must establish a “single point of contact” (SPoC) in the EU. However, certain obligations are tailored according to the classification of providers. “Hosts” must furnish transparent rules related to notice and takedown mechanisms, and disclose reasons underlying decisions to disable access to illegal content. This information must be made available in a database controlled by the Commission. Online platforms and very large online platforms submit to a “trusted flagger” regime, where flaggers notify platforms of illegal content and platforms can act on this “with priority and without delay.” This formalizes Notice and Takedown (NTD) procedures into binding EU law. NTD procedures have been inculcated into domestic law in Europe, for example in Germany’s Network Enforcement Act (NetzDG). The DSA necessitates that large platforms establish trusted flaggers to initiate this process as part of an internal complaints system. Larger platforms also have specific obligations to verify the identity of advertisers and disclose relevant information related to profiling in advertising procedures. This is a notable development in light of the role of tailored advertisements in the Cambridge Analytica scandal (Cadwalladr, 2018).

The DSA has been labeled as the EU’s “most ambitious plan yet to rein in online platforms” (Milo & Kreko, 2021). However, an appraisal of the DSA as a watershed moment for European intermediary liability could be premature. There are undeniably positive signs in the DSA pertaining to disinformation. There is a shift of focus from the narrow category of “political advertisements” to the broader scope of “paid for” content. There is also scope for the implementation of the CPD to be assessed as part of systemic risk management, which could carve room for robust oversight in how anti-disinformation measures are implemented. This has led to the DSA being characterized as a “co-regulatory backstop” for disinformation (Tambini, 2021). However, important aspects of platform responsibilities remain unclear. The cornerstone of liability exemption remains fundamentally unchanged in the DSA, as the regulation does not envisage general monitoring obligations to “actively . . . seek facts or circumstances indicating illegal activity.” In its currently proposed format, the DSA introduces transparency requirements for political advertising. While language has shifted, a major misconception is the interchangeable association between political ads and disinformation. Political advertising is only a fraction of this problem. European disinformation campaigns often spring from organic users, and counterfeit news sites (Bennett & Livingston, 2018). Attempts to focus on countering disinformation through political advertising will also run into the problem that domestic legislation for political advertisements lacks uniformity across EU Member states. As highlighted by the European Court of Human Rights (ECHR) in Animal Defenders v United Kingdom, there is “no European consensus between the contracting states on how to regulate paid political advertising” (Kleinlein, 2017). Furthermore, an over-emphasis on political advertising is in itself problematic, as this represents only a subset of disinformation as a wider problem, not the problem in its entirety.
The quagmire of harmful but lawful content

A problem unlikely to be resolved by the DSA is the regulation of harmful but lawful content. Unlike child pornography and copyright infringement, disinformation is often not illegal per se. While other online harms are subjected to binding rules in Europe, disinformation is relegated to piecemeal soft law. A pervasive concern often raised when considering harmful but lawful content is that regulating such content could undermine freedom of expression, and this argument has cropped up when positing disinformation regulation (Smith, 2019).

These concerns are somewhat misplaced. If disinformation is not addressed with recursive legislation, the governance of disinformation will be reserved for commercial platforms. While the removal of unlawful content entails clear-cut responsibilities, the lack of concrete obligations to remove disinformation leaves wide discretion. Accordingly, a chief concern is that "notice and takedown" regimes allow platforms to indirectly regulate expression. This was crystallized with the NetzDG in Germany. Under this law, platforms must remove unlawful content within a 24-hour period, at the risk of financial sanctions. This elicited criticism that characterized the NetzDG as a “vague” and “overbroad” mechanism that “turns private companies into overzealous censors to avoid steep fines, leaving users with no judicial oversight or right to appeal” (Human Rights Watch, 2018). In this connection, an absence of legal oversight on account of the lawful nature of disinformation will not necessarily safeguard fundamental rights, and it is conceivable that a continued regulatory vacuum could further exacerbate freedom of expression concerns.

The gap for harmful but lawful content is evidenced in the DSA. The trusted flaggers notify platforms of content that is unlawful. Reporting mechanisms for notice and takedown procedures are directed towards unlawful content. While the risk assessment obligations for very large online platforms are not strictly limited to unlawful content, the scope of systemic risks is widely defined, emboldening anxieties surrounding how content moderation decisions could adversely affect digital speech (Kuczerawy, 2019). Many domestically sourced websites and blogs will not be subjected to these risk assessments, despite the reality that smaller platforms remain influential in diffusing disinformation in European Parliamentary elections (Pierri & Ceri, 2021). The penetration of domestic websites through organic user interaction is an important way in which “radical right” parties have capitalized on waning “institutional legitimacy” to influence elections such as the Brexit referendum (Bennett & Livingstone, 2018). Even the European Commission has acknowledged the DSA’s limitations in this respect, noting that the regulation “will not explicitly address some of the very specific challenges related to disinformation.” Instead, more tailored changes are reserved to the updated Codes of Practice on Disinformation (European Commission, 2020).

The current European legal framework for disinformation remains linked to a model that has already proved ineffective, that of self-regulation (Madiega, 2020). This reality exists in spite of the illusory reform under the DSA. While tempting to posit direct regulation, this prospect is fraught with legal, political, and territorial challenges. The EU is a collection of 27 Member States, each possessing unique political environments wherein different country-specific disinformation campaigns diffuse (Bayer, 2019). An aggressive top-down attempt to harmonize strict rules for harmful but lawful content could destabilize political cohesion in the Union. In addition, there is a need for further debate on how to reconcile the regulation of harmful but legal content with fundamental rights to freedom of expression.

What can be done in the interim is to smoothen the transition from self-regulation to co-regulation in a manner that rectifies gaps in the EU fight against disinformation. There needs to be a reassessment as to whether disinformation can realistically be subsumed into the DSA in its current format. At present, the Commission has committed to revamping the CPD in 2021. However, the Commission has simultaneously recognized that many of the CPD’s failures stem from their self-regulatory nature (European Commission, 2020). Therefore, it would be preferable to tackle disinformation through the
attempted co-regulation in the DSA. If it is deemed too late to reshape the DSA in a manner that puts more focus on disinformation, other instruments should be considered. The focus should be on firstly rectifying inconsistencies in the CPD. Consistent definitions and approaches need to be encouraged, and sanctions for non-compliance need to be considered. Oversight should scrutinize how co-regulation can improve anti-disinformation efforts while safeguarding fundamental rights. A natural starting point is that the scope of this problem, and its effect in the EU, need to be accurately understood and transparently communicated. Accordingly, the obscured access to relevant data for researchers should be lifted, respecting the contours of the General Data Protection Regulation (GDPR). Platforms should not use such instruments as a superficial rationale to avoid compliance with requests for data that can help ascertain their amplification of disinformation.

**Conclusion**

For disinformation in Europe, the era of self-regulation appears to be nearing a close. The CPD, while establishing important commitments, lacks enforcement and gives too much discretion to digital platforms to implement practices (Colliver, 2020). This discretion, as recognized by the European Commission, has led to severe implementation gaps that have prevented a coordinated response to online disinformation across Europe (Madiega, 2020). As the announcement of the incoming Digital Services Act (DSA) would seem to suggest, there is a need to inculcate greater transparency and due diligence in the way that digital actors are accountable for harmful content on their platforms, and the European Commission has responded by bringing EU rules in line with the realities of contemporary digital engagement. However, with respect to disinformation, the devil is in the detail. Important gaps still continue to plague the response to disinformation, and the DSA fails to address many of these gaps.

A key question that should shift debates forward is whether harmful but lawful content should continue to eschew regulation purely because it is not illegal per se. As the European Union’s “Democracy Action Plan” points out, the right to free and fair elections and the strengthening of “media freedom” should be front and center of the evolving agenda in this area. Accordingly, a number of important points must be addressed going forward. Firstly, while not all forms of disinformation are unlawful, some are. Some aspects of discriminatory and racist disinformation may well run contrary to both domestic legislation and EU law. Secondly, the imposition of binding rules for disinformation does not necessarily have to manifest through mandates for content removal. Tangible and concrete sanctions for platforms that fail to minimize disinformation could be a reasonable backstop. Obligations do not necessarily have to involve concrete take down mechanisms, and could focus on systemic approaches to minimizing disinformation, disincentivizing inauthentic behavior, and collaborating with researchers to identify persistent disinformation campaigns at an EU and EU Member State level. This could be achieved with the introduction of a new Directive that independently addresses grey areas and systemic risks associated with disinformation while still maintaining room for discretion as to how Member States implement harmonized rules at a domestic legislative level. Crucially, it must be pointed out that, under European human rights law, scrutiny of legal interferences with free speech is not merely focused on whether content was lawful or unlawful. As evidenced in a wide array of case law, both in EU and non-EU-related courts, legal interferences with harmful content are often centered on the legitimate democratic aim of restrictions, the legal precision of sanctions, and the proportionality of actions. Just because content may not be strictly illegal does not preclude any degree of binding regulatory scrutiny. This recognition has already been expressed by the European Commission, as current proposals in the DSA provide a scope for oversight of not only unlawful content but also content that poses risks to “public interests” and “fundamental rights” (European Commission, 2020).
At a bare minimum, there is a clear need for stronger oversight, harmonized approaches, and greater access to important knowledge on how and where disinformation crops up in the EU. Because of continuing shortcomings in the current framework, longstanding questions surrounding the trade-off between disinformation legislation and freedom of expression will, for now, continue to go unanswered (Helm & Nasu, 2021; Posetti & Bontcheva, 2020). Increasingly, it is clear that disinformation, in being a unique legal problem in Europe, requires a designated legislative agenda that finally graduates beyond self-regulation. Attempts to shoehorn reform into the DSA should not be mistaken for concrete progress.

Bibliography


Acknowledgements
The author would like to thank the editorial committee for reviewing this commentary. A special thanks is given to Costanza Sciubba Caniglia for providing helpful feedback on earlier drafts of this commentary.

Funding
Funding supplied by Maynooth University.

Copyright
This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original author and source are properly credited.

Competing interests
The author has no potential conflicts of interest.