



PLATFORM CONTENT STANDARDIZATION



BY
MARTA CANTERO GAMITO

Associate Professor of IT Law, University of Tartu. Research Fellow, School of Transnational Governance, European University Institute.

THE TAMING OF INTERNET PLATFORMS - A LOOK AT THE EUROPEAN DIGITAL SERVICES ACT

By Natascha Just



WHY TECH REGULATORS NEED TO THINK LIKE GOOGLE

By Terry Flew



WHAT WE MEAN WHEN WE TALK ABOUT THE IMPACT OF DEPLATFORMING TRUMP

By Scott Babwah Brennen & Matt Perault



AESTHETICS, TECHNOLOGY, AND REGULATIONS

By Gregory Day



RANDOMIZED EXPERIMENTS FOR ONLINE CONTENT MODERATION POLICY

By Imanol Ramírez



PLATFORM CONTENT STANDARDIZATION

By Marta Cantero Gamito



PLATFORM CONTENT STANDARDIZATION

By Marta Cantero Gamito

As the enactment of the DSA is coming closer, anticipations around the new rules for curbing Big Tech power are mounting. A revision of the eCommerce Directive has been long due, but its modernization will indeed bring significant and necessary changes to make the online space a safer one, establishing procedural guarantees that protect fundamental rights and democracy online. However, it cannot be forgotten that the proposal relies on the harmonization of the internal market as a legal basis, and that this means that the rules will have an intense market regulation flavor. This policy choice will impact freedom of expression, as the new rules promote a sort of “standardization” of content moderation procedures. For example, compliance with regulatory obligations can be ensured by adopting recognized European and international standards. While, in principle, establishing similar guarantees for all platforms is needed, a corseted one-size-fits-all approach to content moderation could run the risk of compromising constitutional pluralism and to result in preventive cancellations across platforms. To avoid this, attention should be paid to discussions within standard-setting organizations following DSA’s adoption. Critical political decisions should not get lost into seemingly technical discussions.

Visit www.competitionpolicyinternational.com for access to these articles and more!

Scan to Stay Connected!

Scan here to subscribe to CPI’s **FREE** daily newsletter.



01

INTRODUCTION

Can (and should) moderation be standardized? Standardization can be defined as the development of consensus-based, often technical, non-binding guidelines to be followed by all the processes related to producing a product or performing a service. As part of market regulation, the legislator often delegates the definition of specific technical details to standardization. When that happens, voluntary compliance with the standards presumes conformity with the essential requirements established by legislation to place a product or to provide a service in the market.

By *platform content standardization* I refer here to the setting of standards that, not being necessarily technical, prescribe the specifications and procedural requirements to comply with regulatory obligations concerning content moderation. The emphasis on procedural steps in the forthcoming Digital Services Act (“DSA”) can also be seen as an example of standardization; a platform can provide services in the EU internal market provide that it abides by specific procedural guidelines. In this regard, it can be argued that standards are generally non-binding while the DSA establishes *mandatory* regulatory obligations. Underlying the claims to regulate content moderation there is a justified distrust of platform’s moderation policies and processes. However, by regulating procedures and not outcomes, the DSA opens the door to institutional variations, blurring the line between law and standards.

Administering speech is no easy task. But when faced with a high volume of cases, there is a trade-off between adjudicating in a rapid manner and compromising procedural guarantees in the protection of fundamental rights. While forthcoming due process provisions are welcome, there are already voices signaling the potential perils of the industrialization of content moderation.² Building on Keller’s insightful analysis, this article joins the critical voices on the DSA by offering a view on the shift from law to standards (and codes). In this short article I shall explain how the forthcoming EU rules promote and fuel platform content standardization, and why this might lead to suboptimal outcomes that counteract the (seemingly) desired policy and regulatory goals of the DSA. Moreover,

I reflect on how platform content standardization is leading a process of change in the way we perceive law and legal authority.

02

FROM LAW IS CODE TO CODE IS PROCEDURE

A platform in 2022 is something far removed from the law-maker’s concept of intermediaries when the safe harbors by the eCommerce Directive or Section 230 were drafted. Today’s platforms connect, entertain, employ, make richer, give voice, as well as censor, cancel, and generally govern users. As a result, the emphasis has been put on the private rules, practices and procedures through which platforms exercise their power in an attempt to understand how they regulate us.³ Platform (private) ordering plays an ever-increasing role in governing the conditions for freedom of expression and access to information. This is largely due to, first, the capacity and necessity of platforms to automate the administration of content and, second, the growing reliance of public regulation on private ordering and automated enforcement.

Moderation is intrinsic to platforms’ value propositions.⁴ Among the 2.5 quintillion bytes of content created every day there are millions of photos, videos, posts, and comments uploaded to the internet. In order to sustain their business model, largely based on advertising, platforms must organize and categorize information. This is true not only for the purposes of structuring of participation but also for preventing abuses. Assuming that each of the videos, photos or posts may be potentially illegal, harmful or infringe platforms’ Terms of Service, it seems almost inescapable that *effective* content moderation requires automation, or at least a certain level of it. Besides, the scale of moderation needed when platforms users’ pools expand can only be approached with the use of computerization techniques.

2 Daphne Keller, *The DSA’s Industrial Model for Content Moderation*, VERFBLOG (February 22, 2022), <https://verfassungsblog.de/dsa-industrial-model/>.

3 Kate Klonick, *The new governors: The people, rules, and processes governing online speech* 131 HARV. L. REV., 1598-1670 (2017).

4 See TARLETON GILLESPIE, CUSTODIANS OF THE INTERNET: PLATFORMS, CONTENT MODERATION, AND THE HIDDEN DECISIONS THAT SHAPE SOCIAL MEDIA (2018); James Grimmelman, *The virtues of moderation*. YALE JL & TECH., 17, 42 (2015).

Automated curation, also referred to as algorithmic commercial content moderation,⁵ allows not only a more efficient identification of inappropriate content, but also the possibility of taking immediate action consisting of removing or downgrading content and/or shadow banning users. By using digital hash technology or matching, filtering, and prediction (including the use of natural language processing or “NLP”), and image recognition tools, the detection of copyright-infringing, hate speech, extremist, and other types of unlawful content can be done in a matter of (micro) seconds.

Content moderation has thus become an ideal use case of AI for law enforcement purposes in digital environments. Aware of this, the legislator has been gradually supporting the use of algorithmic tools for regulatory compliance, evidencing the power of the Code and its capacity to constrain (online) behavior.⁶ In Europe, there are rules that require platforms to monitor and enforce rules against any behaviour infringing child rights (Audiovisual Media Services Directive and the Child Sexual Abuse and Exploitation Directive), copyright (Copyright Directive), exhibiting terrorist content (Counter-Terrorism Directive), or racist and xenophobic hate speech (Counter-Racism Framework Decision). There are also some soft law initiatives on disinformation, hate speech and illegal content online such as the Code of Practice on Disinformation or the EU Code of conduct on countering illegal hate speech online. The use of algorithmic tools for copyright enforcement was highly debated, especially Article 17 of the EU Copyright Directive, which requires the use of automatic recognition and filtering tools, which is at odds with the prohibition of general monitoring obligations contained in the eCommerce Directive (Article 15).⁷

Yet, despite criticism, the EU legislator has continued supporting the role of online intermediaries to voluntarily take the necessary measures to comply with the requirements of EU law, recognising the role of platforms as regulatory intermediaries.⁸ Now, due to legitimacy concerns regarding the adjudication of fundamental rights by private actors, the DSA will set a procedural safety net to ensure accountability and respect for fundamental rights. The threats over democracy escalate in highly concen-

trated markets where only a few dominant platforms are in charge of channeling public discourses, and the DSA stands as the promise for regulating platforms and the way they moderate.

“Content moderation has thus become an ideal use case of AI for law enforcement purposes in digital environments

As the EU's commissioner Thierry Breton has warned Elon Musk over the new direction in Twitter's content moderation policies,⁹ a more stringent approach towards content moderation by platforms seems to be not only a requirement for conditional immunity over user-generated and user-shared content, but also an entry condition to the EU's internal market. We should reflect upon the implications of governing fundamental rights with an (internal) market narrative. Seen from the perspective of market regulation, I argue that the legislator is standardizing content moderation by leaving the definition of the technical details to standard-setting organizations.

03 CAN CONTENT MODERATION BE STANDARDIZED?

Standards contribute to remove market barriers and to decrease compliance costs. Seen this way, there are good reasons why content moderation can be standardized. First, platform moderation operates in a one-to-many

5 Robert Gorwa, Reuben Binns & Christian Katzenbach, *Algorithmic content moderation: Technical and political challenges in the automation of platform governance*, 7 *BIG DATA & SOCIETY* 1-15 (2020).

6 LAWRENCE LESSIG, *CODE: AND OTHER LAWS OF CYBERSPACE* (1999).

7 Giancarlo Frosio, *To Filter, or Not to Filter-That Is the Question in EU Copyright Reform*, 36 *CARDOZO ARTS & ENT. LJ*, 331 (2018); João Quintais, *The new copyright in the digital single market directive: A critical look*, 1 *EUROPEAN INTELLECTUAL PROPERTY REVIEW* 28; Martin Senftleben & Christina Angelopoulos, *The Odyssey of the Prohibition on General Monitoring Obligations on the Way to the Digital Services Act: Between Article 15 of the E-Commerce Directive and Article 17 of the Directive on Copyright in the Digital Single Market* (2020), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3717022.

8 Christoph Busch, *Self-Regulation and Regulatory Intermediation in the Platform Economy* in *THE ROLE OF THE EU IN TRANSNATIONAL LEGAL ORDERING. STANDARDS, CONTRACTS AND CODES* (Marta Cantero Gamito & Hans W. Micklitz, eds. 2020).

9 <https://www.ft.com/content/22f66209-f5b2-4476-8cdb-de4befffebe5>.

environment. To date, every platform is governed by its own house rules, which determine the terms of use and access to information, including moderation policies and procedures. Therefore, standardizing platforms' moderation can contribute to preventing a situation whereby, within the internal market, what is allowed on one platform could be simultaneously prohibited on another. In this regard the DSA promotes the role of standards (and technological means) to facilitate the effective and consistent fulfillment of regulatory obligations.¹⁰ For example, in the proposed draft, the European Commission encourages the development of industry standards covering technical procedures for the submission of notices regarding violating content or for interoperable advertising repositories, among others. One of the most criticized aspects of the DSA is Article 14, dealing with notice and action mechanisms, as it empowers platforms to make decisions about the (il)legality of content. To assess content's legality, the DSA commends the value of industry standards for helping to "distinguish between different types of illegal content or different types of intermediary services, as appropriate."¹¹ Different standards to identify illegal content can also help to prevent fragmentation. Existing legislative initiatives, such as the German Network Enforcement Act ("NetzDG") or the French Law on Countering Online Hatred ("Avia Law"), set different guidelines on content moderation, and this was considered to pose a barrier to the free movement of services and prevent interoperability. In this regard, standardization would contribute to harmonize the internal market – let's not forget that the DSA is based on the internal market harmonization legal basis.¹²

“Standards contribute to remove market barriers and to decrease compliance costs

Secondly, standardization can contribute to reduce compliance costs. One of the downsides of the DSA is that, although not all, it establishes a set of procedures to be complied by all platforms regardless of size, significantly or even disproportionately increasing costs for small and medium online intermediaries and reinforcing the competitive disadvantage of these companies *vis-à-vis* incumbents.

Standardization can indeed minimize the required efforts to monitor online activity by setting more efficient content review systems. But if the EU wants to protect European-born innovation, it needs to be more flexible with those smaller intermediaries trying to reach a significant user base. This could perhaps explain why there are other areas for which the DSA opens the door to standardization, in addition to submission of notices, such as submission of notices by trusted flaggers, interfaces to comply with regulatory obligations (including APIs), standards for auditing, data transmissions, or concerning the interoperability of advertisement repositories.¹³ While this approach is a step towards levelling the playing field for big and small platforms, we should consider whether a one size fits all approach is suitable for content moderation.

04 SHOULD IT?

There are important problems associated with entrusting content moderation to standardization. Continuing with the example of Article 14 DSA, precise and adequately substantiated notices would constitute actual knowledge for the purposes of hosting liability. This means that standards are expected to define what content is illegal and what content is not. Provided that the submission of these notices would take away immunity, it is realistic to believe that platforms would likely remove notified content even when this is lawful.

This is problematic from the perspective of the assessment of content's legality. The impact and scale of platforms' action and power over the way we use our fundamental rights requires putting in place procedural safeguards but they can also reveal a darker side with undesired effects since content standardization can lead to automated and generalized cancellation. For example, while the DSA establishes due process obligations by regulating dispute resolution procedures in Articles 17 and 18, it does not provide any guidance on how platforms make decisions that generally affect the fundamental right to freedom of expression. This can be seen as a missed opportunity for the legislator to codify the existing (and abundant) case law concerning the limitation of fundamental rights and proportionality assessments. Any limitation to EU funda-

¹⁰ Article 34 and Recital 66.

¹¹ Recital 66.

¹² Article 114 of the Treaty on the Functioning of the European Union.

¹³ See Article 34 DSA.

mental rights must be provided for by law.¹⁴ Therefore, an assessment shall be made concerning whether potential limitations to freedom of expression made on the basis of standards' understanding of content legality would pass a proportionality test.

Resulting limitations to free speech do not result from any legal general obligations imposed on intermediaries to remove illegal content. The DSA does not require platforms to delete content but to take a decision with regard to the information considered to be illegal.¹⁵ Instead, potential limitations would come for instance from either platforms' policies concerning users' claims of illegality or by the threshold to be set in the standard that contains the technical specifications governing notice and action mechanisms – or both. The European Court of Human Rights has already recognized that the requirement that limitations to fundamental rights must be provided for by law can be interpreted expansively and in a flexible manner so as to avoid excessive rigidity, which would allow the law to keep pace with changing circumstances.¹⁶ From this perspective, standardization of content moderation seems suitable and compatible with the approach of the EU legislator supporting soft law and private regulatory initiatives to fight against hate speech or disinformation. However, this should not lead to intermediaries taking measures, or advocating for standards, that would affect the essence of the freedom of expression of users who share lawful content.¹⁷

From this perspective, platform content standardization would require the exploration of important questions. What does standardized and algorithmic moderation mean for the democratic control of the adjudication of fundamental rights? How does the regulatory reliance on standards and codes interplay with notions of authority? What does the recognized administrative power of platforms' private contracts and algorithms mean for our understanding of law, the legal system and, ultimately, the (digital) constitutional state? Can automated content moderation fix societal problems or is it the reason for societal problems to emerge? While these are questions not to be answered in this short article, some preliminary observations can be made.

Platform content standardization can affect fundamental rights on two levels: institutional (including procedural) and normative. Institutionally, “mandated” private ordering blurs the distinction between “standards” and “law.” On the one hand, standardization can be seen as a tool to depoliticize content moderation. However, standardization is ultimately about normative choices. Standards are inherently political, as they involve adherence to a particular decisional approach or understanding, representing critical value choices. This has been widely discussed with regard to internet governance from the perspective of the internet's infrastructure. Despite that, we are witnessing a growing reliance on multistakeholder standard-setting in other areas, such as artificial intelligence. On the other hand, the regulative power of standardization is often contested.¹⁸ However, inclusive standardization would allow the incorporation of multistakeholders' preferences in the regulatory process. The European Commission is currently trying to reinforce the democratic credentials of European standardization organizations (“ESOs”) by improving their decision-making processes and requesting them to “modernise their governance to fully represent the public interest.”¹⁹ Efforts need also to be made to prevent strategic participation and geopolitical opportunism. Therefore, it is important to keep an eye on how standardization following the DSA will occur, as it can be reasonably expected that small platforms would not have the means or the reach to be part of the standard-setting process, which would allow big platforms to hold the pen.

“*Platform content standardization can affect fundamental rights on two levels: institutional (including procedural) and normative*”

14 Article 52(1) Charter of Fundamental Rights of the European Union.

15 Article 14(6) DSA.

16 European Court of Human Rights, 16 June 2015, *Delfi AS v. Estonia* (64569/09), para. 121.

17 Cf. Court of Justice of the European Union, Judgment of the Court (Grand Chamber) 26 April 2022, Case C-401/19 *Poland v. Parliament and Council*. ECLI:EU:C:2022:297.

18 MARIOLINA ELIANTONIO & CAROLINE CAUFFMAN (EDS.). *THE LEGITIMACY OF STANDARDISATION AS A REGULATORY TECHNIQUE: A CROSS-DISCIPLINARY AND MULTI-LEVEL ANALYSIS* (2020).

19 EU Strategy on Standardization, EU Commission Communication “An EU Strategy on Standardisation. Setting global standards in support of a resilient, green and digital EU single market,” COM(2022) 31 final.

At the normative level, platform content standardization is based on a non-traditional philosophical foundation of law and law enforcement. Automated moderation brings the technology to the forefront of law enforcement, with algorithms massively filtering and removing content. Governing mandated content moderation practices by non-legislative actors raises a whole set of issues related to the incorporation of value choices into privately designed algorithms with enforcement capabilities, the interplay between the code and the legal systems, and their inevitable impact on the ethos of fundamental rights adjudication. The protection of fundamental rights online is replacing the law-based enforcement discourse with a new set of code-based enforcement values consisting of access (to platform services), neutrality, and algorithmic transparency.

From this perspective, standardized moderation can solve some problems (e.g. procedural fairness) while it exacerbates others such as lack of pluralism. What constitute harmful content can vary and mean different things depending on the geographical, cultural, substantive, and subjective context. Content moderation is not neutral, and every moderation model involves a trade-off between competing interests and values. For example, moderation obligations under DSA may empower platforms to use public regulation to pick winners and losers. There are indeed important concerns related to the ambitions of protecting democracy through standards and algorithms. The most important one is that standardizing content moderation has a fundamental design flaw. Can we rely on AI and standards to oversee public interests? With more than 500 million tweets per day, there are just so many and mutually inconsistent moderation demands to be met that renders impossible to please everyone. The solution to the “wrong speech” should not be “standardized speech.” However, this triggers the following question: does the definition of values belong to the law?

“*At the normative level, platform content standardization is based on a non-traditional philosophical foundation of law and law enforcement*”

Lastly, technosolutionism is leading a process of change from states based on the rule of law to states increasingly centered on the rule of code and system-level bureaucracy.²⁰ This raises questions as to, first, whether attempts to constitutionalize the cyberspace are compatible with the EU’s constitutional pluralism and, second, whether technical solutions improve or instead worsen citizens’ fundamental rights.

05 CONCLUSION

The internet once favored a remarkable development towards greater democratization. It facilitated a mechanism for anonymous users to voice and amplify their opinions. Today’s sentiment is the opposite; internet’s governance has the power to manipulate and shape public opinion, eroding democracy.

The broad recognition of the role of platforms in the functioning of democracy prompted lawmakers to investigate platforms’ activities more closely. As a result, there is much hope in the initiatives by the European Commission to curb Big Tech power. From this perspective, the DSA can be seen as an *in-progress* political building process for internet governance. Yet, in my opinion, it would be incorrect to think that the DSA would solve all the existing problems and that business models based on users’ data extraction would not find alternative ways to sustain their revenue models. Most importantly, the DSA should not be a victim of its own success. The public expectation of increased responsibility taken by platforms is overshadowing an underlying process of institutional innovation and the use of alternative regulatory techniques, which includes an excessive reliance on less accountable codes and standards.

This article has outlined the pros and cons of legitimizing the regulative power of non-legislative regulatory tools. Not much consideration has been given regarding the protection fundamental rights with (internal) market regulation narratives, although this can be seen as a legacy problem in the historical constitutional configuration of the EU. In this regard, it is argued that, while improvements can be made, an EU-level approach towards platform regulation

20 Mark Bovens & Stavros Zouridis, *From street-level to system-level bureaucracies: how information and communication technology is transforming administrative discretion and constitutional control*, 62 *PUBLIC ADMINISTRATION REVIEW* 2, 174-184 (2002).

is indeed welcome and necessary for effectively and consistently protecting fundamental rights online.²¹ However, discussions should not terminate with the final approval of the DSA. Instead, attention should be paid to standardization, where actual value choices affecting fundamental rights are to be made. Moreover, it is important to monitor that standardization is not resulting in a level of substantial harmonization that would compromise constitutional pluralism. The focus should be on advocating for a process of greater institutionalization and accountability of standard-setting that aims at reproducing the more participatory decision-making structures of the early days of the internet to counterbalance the existing centralization of platform power. ■

“*The internet once favored a remarkable development towards greater democratization*”

21 The proposal for a Regulation laying down harmonized rules on artificial intelligence (“AI Act”) and the forthcoming initiative for protecting media freedom, the European Media Freedom Act (“EMFA”) are also using legal harmonization in the internal market (Article 114 TFEU) as their legal basis.

CPI SUBSCRIPTIONS

CPI reaches more than **35,000 readers** in over **150 countries** every day. Our online library houses over **23,000 papers**, articles and interviews.

Visit competitionpolicyinternational.com today to see our available plans and join CPI's global community of antitrust experts.

