Platform Responsibility in the Digital Services Act: Constitutionalising, Regulating and Governing Private Ordering

Giancarlo Frosio

(forthcoming in Andrej Savin and Jan Trzaskowski (eds), *Research Handbook on EU Internet Law*, Edward Elgar)

This chapter re-contextualises long-standing discussions about the emerge of private ordering online as a result of enhanced platform responsibility in light of a change in perspective brought about by the Digital Services Act. This reform updates online platform regulation in the European Union by emphasising the protection of fundamental rights online and providing a redress to the dystopian effects of private ordering. With a special emphasis on large platforms and the risks they might cause to users' rights online, the Digital Services Act has launched a process of constitutionalisation, regulation and institutional governance of private ordering practices that might serve as a bluprint for future reform in other jurisdictions.

I. INTRODUCTION

Recent events have revived long-standing concerns regarding private ordering and the moderating powers available to global platforms. In May 2021, the Oversight Board, a newly established body that should review selected Facebook content moderation decisions,[1] upheld former President Trump Facebook account's suspension following the infamous attack on the Capitol Building in Washington, D.C., while remanding the matter back to Facebook 'to determine and justify a proportionate response that is consistent with the rules that are applied to other users of its platform'.[2] Although, as the Oversight Board websites states, '[t]he purpose of the board is to promote free expression by making principled, independent decisions regarding content on Facebook and Instagram', questions abound on the independency—and adequacy—of the Board in issuing such sensitive decisions, with deep and unpredictable ramifications on political discourse and, finally, on the essential tenants of our democratic society. Why should Facebook take such decisions in the first place? Shouldn't such

[1] See Oversight Board <https://oversightboard.com> (composed by eminent independent figures, the Board should help Facebook answering most sensitive questions about freedom of expression online; the Board decisions to uphold or reverse Facebook's content decisions are binding and Facebook must implement them).

[2] See Oversight Board upholds former President Trump's suspension, finds Facebook failed to impose proper penalty (*Oversight Board*, May 2021) <https://oversightboard.com/news/226612455899839-oversight-board-upholds-former-president-trump-s-suspension-finds-facebook-failed-to-impose-proper-penalty> (the Oversight Board actually concluded that it was not appropriate for Facebook to impose the 'indeterminate and standardless penalty of indefinite suspension', while remanding the matter to Facebook to be reviewed).

'principled, independent decisions' be the domain of courts of law, and possibly, when they involve such sensitive questions, the domain of Constitutional Courts or, at least, administrative bodies that operate under State prerogatives and powers? In the words of the President of the European Commission von der Leyen,

> No matter how right it may have been for Twitter to switch off Donald Trump's account five minutes after midnight, such serious interference with freedom of expression should be based on laws and not on company rules. It should be based on decisions of parliaments and politicians and not of Silicon Valley managers.[3]

Instead, the 'privatization' of online enforcement and content moderation emerges as a troublesome reality of our platform society.

In contemporary information societies, the activities of online platforms and other service providers (hereinafter 'online intermediaries')[4] raise a number of ethical, legal, and social challenges. Some online intermediaries are in a unique position to influence their users' ability to access information and their interaction with it, thus fulfilling an important democratic function, acting as facilitators of users' speech, creativity and exchange of ideas.[5] These powers are so pervasive to be able to fundamentally affect the political discourse as in the case of a platform like Twitter permanently suspending the account of the President of the United States without any judicial nor independent review. This is even more preoccupying given that Twitter has been the major channel of communication between President Donald Trump and his electoral base during his mandate. Meanwhile, controversial entrepreneurs with a political agenda, like Elon Musk, stir new debate on a plan for acquiring Twitter and potential bias of content moderation practices that might follow that acquisition.[6] These events take platforms' liability, role, and regulation to a whole new level. platforms can influence elections or change the course of events by silencing speech according to alleged infringements of their Terms of Service. While certain harmful content posted online needs intervention, it has been increasingly

[3] See Speech by President of the European Commission von der Leyen at the European Parliament Plenary on the inauguration of the new President of the United States and the current political situation, Brussels, 20 January 2021, <https://ec.europa.eu/commission/presscorner/detail/en/speech_21_167>.

[4] For a taxonomical discussion of online intermediaries and service providers, see Graeme Dinwoodie, 'Who Are Internet Intermediaries?' in Giancarlo Frosio (ed), *The Oxford Handbook of Online Intermediary Liability* (OUP 2020) 37-56.

[5] See eg. Niva Elkin-Koren and Maayan Perel, 'Guarding the Guardians: Content Moderation by Online Intermediaries and the Rule of Law' in Giancarlo Frosio (Ed.), *The Oxford Handbook of Online Intermediary Liability* (OUP, 2020), pp. 669-678.

[6] See eg Max Zahn, 'A timeline of Elon Musk's tumultuous Twitter acquisition attempt' (*ABC News*, 13 July 2022) <https://abcnews.go.com/Business/timeline-elon-musks-tumultuous-twitter-acquisition-attempt/story?id=86611191>.

openly questioned if platform should alone be the one to decide what is available online and what is not.

Emerging innovation policy choices have been reversing an earlier approach that provided online intermediaries with substantial liability exemptions to incentivize their capacity to develop Internet infrastructure and applications. The regulatory and enforcement shift that is presently occurring depends on theoretical, market, policy, technological changes.[7] Today, due to changed market conditions, policy makers might want to switch the enormous transaction costs of online content regulation to online intermediaries as they might be the least cost avoiders, in particular given their economies of scale. Meanwhile, online content sanitisation is increasingly becoming the sole domain of private – and opaque – algorithmic technologies. Burdened by expansive enforcement obligations, online intermediaries resort to the widespread deployment of algorithmic enforcement tools, which remain the sole option at their disposal not to jeopardise their business models.

In this context, how are online intermediaries – gatekeepers of the Internet and masters of the algorithms – supposed to balance competing interests at stake? How can regulatory intervention help avoiding a dystopian future, for users in particular, in the 'black box society'?[8] Platform regulation is a complex conundrum in search of proportional balancing of fundamental rights and stakeholders' interests. In this context, it is of essence to clarify from the inception that rightholders' pressures leading to obligations imposed on intermediaries might negatively impact – and very often do – users' fundamental rights. Optimal regulation in the field of platform governance must attempt first to preserve users' and citizens' rights, thus democratization, as more online enforcement – and potential over-enforcement – equals to less access to information and less freedom of expression, thus less democratisation. The terms of the debate that online content moderation entails, via filtering and monitoring and the use of automated tools in particular, has been spelled out by the Court of Justice of the European Union (CJEU) multiple times. When imposing obligations on internet service providers a trifecta of interests must be taken into consideration, including the freedom of those service providers to conduct a business, guaranteed in Article 16 of the Charter, the fair balance between that freedom, the right to freedom of expression and information of the users of their services, enshrined in Article 11 of the Charter, and the right to intellectual property [and more broadly the private rights]

[7] See Giancarlo Frosio, 'Regulatory Shift in State Intervention: From Intermediary Liability to Responsibility' in Edoardo Celeste, Amélie Heldt, Clara Iglesias Keller (eds), *Constitutionalising Social Media* (Hart Publishing, 2022) 152-165.

[8] See Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Harvard University Press 2015).

of the rightholders, protected in Article 17(2) of the Charter.[9] The centrality of users' rights – and the overall goal of the EU legal system to preserve those rights against invasive proactive algorithmic enforcement – has been recently reiterated by the Grand Chamber of the CJEU in the Case C-401/19 of 26 April 2022, possibly acknowledging a fundamental right of users to share content online that cannot be limited by algorithmic content moderation.[10] As we will discuss later in this chapter, with a regulatory framework in constant flux, the Digital Services Act[11] might have started a process meant to formalize private ordering measures into legislatively mandated obligations, under a stricter application of the principle of proportionality and protection of fundamental rights, thus promoting constitutionalisation – as well as regulation and institutional governance – of platform responsibility, content moderation and related private ordering practices.

II. THE EMERGENCE OF PLATFORM RESPONSIBILITY AND PRIVATE ORDERING

Traditionally, intermediary and platform liability regulation have been struggling to find a proper balancing between competing rights that might be affected by online intermediaries' activities and obligations. Early regulatory approaches, reflected in the DMCA in the United States first[12] and the EU E-Commerce Directive (ECD), [13] have established limited liability frameworks including exemptions for online intermediaries, [14] which in turn provided strong safeguards for users' fundamental rights. In order to protect the ideal of a free internet—and maximise incentive for Internet entrepreneurs—policy makers decided initially to set up a legal framework that rejected any approach turning online intermediaries into Internet police.

With the mainstreaming of the Internet in the nineties, the legislators' decision that access and hosting providers should enjoy liability exemptions for unlawful acts of third parties on their networks was

[9] See C-314/12 *UPC Telekabel Wien GmbH v Constantin Film Verleih GmbH and Wega Filmproduktionsgesellschaft mbH* [2014] EU:C:2014:192, para 47.

[10] See C-401/19 *Republic of Poland v European Parliament and Council of the European Union* [2022] ECLI:EU:C:2022:297, para 80.

[11] See Proposal for a regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC, COM(2020) 825 (hereafter "DSA Proposal").

[12] See The Digital Millennium Copyright Act of 1998, 17 U.S.C. § 512 [hereafter DMCA].

[13] See eg Directive 2000/31/EC of the European Parliament and of the Council of 8 June 2000 on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market, 2000 OJ (L 178) 1-16 [hereafter 'eCommerce Directive' or 'ECD'] art 12-15. See alsoDirective 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the Harmonisation of Certain Aspects of Copyright and Related Rights in the Information Society, 2001 OJ (L 167) 10-19 [hereafter InfoSoc Directive] art 8(3); Directive 2004/48/EC of the European Parliament and of the Council of 29 April 2000 on the Enforcement of Intellectual Property Rights, 2004 OJ (L 195) 16 [hereafter Enforcement Directive], art 11.

[14] Similar legal arrangements have been, then, set up by several international jurisdictions. See for a review of safe harbour legislations, World Intermediary Liability Map (WILMap) (a project designed and developed by Giancarlo Frosio and hosted at Stanford CIS) <https://wilmap.law.stanford.edu>.

dictated by peculiar socio-economic conditions. It was a brave new world. The internet infrastructure was still to build. And, after completing that titanic task, the Internet would have remained an empty shell, absent the right incentives to make high-tech innovators develop millions of applications. It was a highly unpredictable business venture. Most companies were doomed to failure and just a few would have survived those early years. Thus, policy makers understood that both access providers, which had to set up the Internet infrastructure, and hosting providers, that had to populate that infrastructure with applications, had to be lured in the new market with the guarantee of liability exemptions for anything that could go wrong—and their users could do wrong. In the U.S., these market conditions justified the very broad safe harbours provided to access providers, for any content, and to hosting providers for the speech they carry, as well as the more limited safe harbours enjoyed by hosting providers for copyright infringing content.[15] Outside the U.S, other jurisdictions followed suit with very similar legal arrangements, although, at least in the EU, with more limited liability exemptions for hosting providers at large, regardless of whether they carry infringing speech or copyright infringing content.[16] In fact, most international jurisdictions were eager to attract investments and innovation from U.S. companies, largely dominating the infant Internet market, and concluded that the most beneficial policy option was to provide these companies with the same conditions they enjoyed at home, at least for the time being and as long as the Internet had to grow out of puberty. Therefore, given the great uncertainty regarding the future growth of the Internet, innovators were exempted from the obligation of policing its activities.[17]

As I have argued elsewhere,[18] the legal framework regulating online intermediaries has been changing considerably in the past few years. Today, market conditions have profoundly changed. The Internet has been established and has grown beyond any expectation, becoming the most pervasive medium in people's daily life. Rather than providing incentives to develop Internet infrastructure and applications, policy makers, governmental institutions, and law enforcement agencies are increasingly concerned with content moderation due to the relentless capacity of the Internet to disseminate content that might infringe upon multiple rights – and create social unrest at an unprecedented speed. At the same time, on one side, regulators face the conundrum of the enormous costs of online content moderation that have grown exponentially along with the growth of the Internet. On the other side, only large high-tech companies own the infrastructure and technology that can manage the massive

[15] See Communications Decency Act of 1996, 47 U.S.C. § 230; DMCA, 17 U.S.C. § 512.

[16] See ECD (n 13).

[17] Hence, the no general monitoring obligations principle enacted by most jurisdictions.

[18] See Giancarlo Frosio, 'Why Keep a Dog and Bark Yourself? From Intermediary Liability to Responsibility' (2018) 25 Oxford JILIT 1-33; Giancarlo Frosio and Martin Husovec, 'Intermediary Accountability and Responsibility' in Giancarlo Frosio (ed), *The Oxford Handbook of Online Intermediary Liability* (OUP 2020) 613-630.

content moderation volumes of the digital society. Thus, a sound policy solution might seemingly be to switch the enormous transaction costs of online content regulation to online intermediaries as they might serve as Calabresi-Coase's 'least-cost avoiders',[19] in particular given their economies of scale. Meanwhile, market geopolitics of the platform economy play a relevant role in this process, which is tightly attached to changes in incentives that might guide state intervention. For example, Europe has less concerns in switching transaction costs of enforcement on online intermediaries, mostly U.S. conglomerates.

Governments and rightholders have attempted for quite some time to enlist intermediaries to rid the internet of allegedly infringing and illicit materials. Public enforcement lacking technical knowledge and resources to address an unprecedented challenge in terms of global human semiotic behaviour would coactively outsource enforcement online to private parties. Recent legal developments can be situated within a trend towards the expansion of the liabilities — and responsibilities — of online intermediaries. The policy discourse is increasingly shifting from liability to enhanced 'responsibilities' for intermediaries under the assumption that OSPs' role is unprecedented for their capacity to influence the informational environment and users' interactions within it.[20] According to Shapiro, 'in democratic societies, those who control the access to information have a responsibility to support the public interest. […] these gatekeepers must assume an obligation as trustees of the greater good'.[21] This is happening via a number of policy developments, including private ordering and voluntary measures, which come as a consequence of an emergent emphasis on corporate social responsibility. At the same time, these developments also depend upon newly enhanced cooperation between States and private parties in online enforcement that often trespasses into public deal-making and jawboning.

A move from intermediary liability to platform responsibility has been occurring first with special focus on intermediaries' corporate social responsibilities and their role in implementing and fostering human rights,[22] so that online intermediaries are increasingly expected to act according to current social and cultural values.[23] Thus, responsible behaviour beyond the law finds justification in intermediaries'

[19] See Guido Calabresi, 'Some Thoughts on Risk Distribution and the Law of Torts' (1961) 70 Yale L J 499; Ronald Coase, 'The Problem of Social Cost' (1960) 3 J L & Econ. 1.

[20] See eg Communication from the Commission to the European Parliament, the Council, and the Economic and Social Committee, and the Committee of the Regions, Tackling Illegal Content Online. Towards an enhanced responsibility of online platforms, COM(2017)555final, 28 September 2017, § 6 (noting "the constantly rising influence of online platforms in society, which flows from their role as gatekeepers to content and information, increases their responsibilities towards their users and society at large).

[21] Andrew Shapiro, *The Control Revolution: How the Internet is Putting Individuals in Charge and Changing the World We Know* (Public Affairs 2000) 225.

[22] See Frosio (n 7) 156-159.

[23] See Emily Laidlaw, *Regulating Speech in Cyberspace: Gatekeepers, Human Rights and Corporate Responsibility* (CUP 2015); Dennis Broeders and others, 'Does Great Power Come with Great Responsibility? The Need to Talk

corporate social responsibilities as derived from those pertaining to States, and endorsed by the UN Human Rights Council declaration of Internet freedom as a human right,[24] or those directly applied to corporations in the preamble of the UN Norms on the Responsibilities of Transnational Corporations and Other Business Enterprises,[25] or the UN Guiding Principles on Business and Human Rights. [26]

In light of the widespread acceptance of corporate social responsibility, States have been pushing new forms of State-driven cooperation with private parties, such as online intermediaries, to alleviate the burden on the public of online sanitisation of allegedly infringing—or merely dangerous– content without deploying prescriptive norms and mandatory obligations. For quite sometimes, the European Commission's regulatory efforts have been emphasising voluntary and proactive measures to remove certain categories of illegal content and allow intermediaries to fulfil enhanced duties of care reasonably expected from them,[27] with special emphasis on enforcement of intellectual property rights' (hereinafter IPR) infringement.[28] For States, coercing intermediaries into "voluntary" measures has doubtless advantages by allowing to circumvent the EU Charter on restrictions to fundamental rights, avoiding the threat of legal challenges, and taking a quicker reform route. In the EU, the 'Communication on Online Platforms and the Digital Single Market' puts forward the idea that 'the responsibility of online platforms is a key and cross-cutting issue.'[29] Again, in 'Tackling Illegal Content Online', the Commission made this goal even clearer by openly pursuing 'enhanced responsibility of online platforms' on a voluntary basis and noting that 'the constantly rising influence of online platforms in society, which flows from their role as gatekeepers to content and information, increases their responsibilities towards their users and society at large'.[30] This terminology plainly evokes the

About Corporate Responsibility' in Mariarosaria Taddeo and Luciano Floridi, The Responsibilities of Online Service Providers (Springer 2017) 315-323; Mariarosaria Taddeo and Luciano Floridi, 'New Civic Responsibilities for Online Service Providers' in Mariarosaria Taddeo and Luciano Floridi, *The Responsibilities of Online Service Providers* (Springer 2017) 1; Mariarosaria Taddeo and Luciano Floridi, 'The Debate on the Moral Responsibility of Online Service Providers' (published online November 27, 2015) Sci Eng Ethics 1, 1.

[24] See Human Rights Council of the United Nations, Resolution on the Promotion, Protection and Enjoyment of Human Rights on the Internet (2012).

[25] See UN Norms on the Responsibilities of Transnational Corporations and Other Business Enterprises (August 13, 2003).

[26] See United Nations, Human Rights, Office of the High Commissioner, Guiding Principles on Business Human Rights: Implementing the United Nations "Protect, Respect, and Remedy" Framework (2011) [hereinafter UN GPBHRs].

[27] See eg European Commission, Public consultation on the regulatory environment for platforms, online intermediaries, data and cloud computing and the collaborative economy (24 September 24 2015) 21-23.

[28] See European Commission, Public Consultation on the Evaluation and Modernisation of the Legal Framework for the Enforcement of Intellectual Property Rights (9 December 2015) D.1.

[29] European Commission, 'Online Platforms and the Digital Single Market: Opportunities and Challenges for Europe' (Communication) COM(2016) 288 Final, 9.

[30] See European Commission, 'Tackling Illegal Content Online. Towards an enhanced responsibility of online platforms' COM(2017) 555 final, s 6.

mentioned move from intermediary liability to intermediary responsibility in platform governance. Under this framework of enhanced responsibilities, online intermediaries and platforms would be invested by a duty to ensure a safe online environment for users,[31] hostile to criminal and other illegal exploitation, which should be deterred and prevented by (1) allowing prompt removal of illegal content, (2) adopting effective *proactive* measures to detect and remove illegal content online, [32] rather than limiting themselves to reacting to notices, and (3) encouraging the use of automated monitoring and filtering technology.[33] In particular, 'online platforms must be encouraged to take more effective voluntary action to curtail exposure to illegal or harmful content, such as incitement to terrorism, child sexual abuse and hate speech'.[34] This should occur, in particular, by setting up a privileged channel with 'trusted flaggers', competent authorities and specialized private entities with specific expertise in identifying illegal content,[35] regardless of 'being required to do so on the basis of a court order or administrative decision, especially where a law enforcement authority identifies and informs them of allegedly illegal content'.[36] In sum, online platforms should be able to prevent and contrast online illegal activities by allowing rapid content take-down, setting up proactive intervention, deploying automated filtering technologies, and adopting effective notice and actions mechanisms, which might not necessarily require users to identify themselves when reporting content that they consider illegal.[37] The 2019 UK Government's *Online Harms White Paper* reinforces this approach of adding responsible duties beyond those provided in the liability framework by proposing a new duty of care towards users, holding companies to account for tackling a comprehensive set of online harms, ranging from illegal activity and content to behaviours which are harmful but not necessarily illegal.[38] The goal of the proposal is to set out 'high-level expectations of companies, including some specific expectations in relation to certain harms'.[39] The violation of the duty of care would be assessed separately from liability for particular items of harmful content. This 'systemic form of liability' essentially superimposes a novel duty to cooperate on the providers and turns it into a separate form

[31] ibid s 3.

[32] Communication (n 30) s 3.3.1.

[33] ibid s 3.3.2. See also Recommendation from the Commission on Measures to Effectively Tackle Illegal Content Online, C(2018)1177final, 1 March 2018, para 37.

[34] Communication (n 29) 9.

[35] Communication (n 30) § 3.2.1.

[36] ibid § 3.1.

[37] ibid § 3.2.3.

[38] See Department for Digital, Culture, Media & Sport and Home Department, *Online Harm* (White Paper, Cp 59, 2019).

[39] ibid 67.

of responsibility, which is enforceable by public authorities by means of fines.[40] At the same time, it leaves the underlying responsibility for individual instances of problematic content intact.

In addition, online platforms' increased responsibilities might result from market-driven private ordering and voluntary measures, either implemented in the intermediary's self-interest or because of rational business decisions based on dealings with the right holders. As per the first category, a number of factors might lead online intermediaries to expand their own responsibility.[41] First, it might be the need of improving users' experience as users might be misled or defraud by illegal content, such as in the case of abusive or spam content.[42] Second, online platforms might want to enhance their credibility and reputation, which helps attracting advertising or other investments, as in the case of Yelp that incorporated a right-to-reply into its review service after public pressure from the business community.[43] On the other side, increased platforms' responsibility might result from market dynamics, when right holders cut deals with platforms, or leverage their existing business relationships. For example, voluntary enforcement schemes were usually initiated when intermediaries such as Internet access providers tried to vertically integrate into markets where they had to do business with major right holders and license their content.[44] In exchange the right holders might ask to implement some sort of enforcement strategy.

In practice, miscellaneous forms of 'responsible' behaviours beyond the law have been recently emerging via multiple enforcement tools applied by different categories of online service providers.[45] They reflect a globalised, ongoing move towards privatisation of law enforcement online through proactive actions and algorithmic tools that spans all subject matters relevant to intermediary liability online. Their common denominator is that that they go beyond the baseline legal expectations created by the legal liability framework. Inherently, their common trajectory is towards more proactive tackling of illegal or otherwise objectionable content. They are usually applied via codes of conducts and standardisation, while targeting a vast array of online intermediaries, such as (1) DNS providers in the

[40] ibid 59.

[41] See Marin Husovec, *Injunctions Against Intermediaries in the European Union: Accountable But Not Liable?* (CUP 2017) 13.

[42] For an overview of industry practices and their corresponding business reasons, see Emma Goodman, *Online Comment Moderation: Emerging Best Practices* (WAN-IFRA 2013) <https://www.wan-ifra.org/reports/2013/10/04/online-comment-moderation-emerging-best-practices>.

[43] See Claire Miller, 'Yelp Will Let Businesses Respond to Web Reviews' *New York Times* (New York, 10 April 2009).

[44] See Frosio and Husovec (n 18) 625-626.

[45] In fact, these enforcement strategies are actually deployed through a larger spectrum of policy options that spans from legally-mandated obligations to private ordering. Even the same type of enforcement arrangements, can be the result of a private ordering scheme, ad-hoc governmental policy administered by agencies, or the application of legislatively-mandated obligations.

case of private DNS content regulation, (2) access providers for website blocking and graduate response; (3) search engines, in the case of online search manipulation; (4) hosting providers, deploying proactive monitoring and filtering strategies; and (5) online advertisement and payment providers, in the case of payment blockades and follow-the-money strategies.[46] Of course, such private solutions are agreed upon secretly by the parties involved. Implementation terms are customarily confidential. Therefore, they are not subject to any public scrutiny, even though such measures might limit fundamental rights of users affected, causing obvious tensions with due process.

III. THE RISE OF THE MACHINE

Meanwhile, private ordering and enforcement practices have been mainly looking for an 'answer to the machine in the machine'.[47] By enlisting online intermediaries as watchdogs, governments would *de facto* delegate online enforcement to algorithmic tools.[48] Given the unsustainable transaction costs related to manual checking of infringing content, online intermediaries are forced to deploy algorithmic tools to perform monitoring and filtering, and limit their liability.[49] Content moderation online, including the moderation of intellectual property infringing content, defamatory content, dangerous and hate speech, child pornography and abuse, or online disinformation – and the adjudication of disputes with and between users – have been increasingly dealt through automated filtering and other algorithmic means. As Calabresi-Coase's 'least-cost avoiders',[50] OSPs will inherently try to lower transaction costs of adjudication and liability and, in order to do so, might functionally err on the side of overblocking. This process might likely bring about pervasive over-enforcement, which becomes invisible and leads users to get accustomed to a highly censored online ecosystem.[51]

As per the effects of this policy development, intermediary responsibilities privately enforced via automated tools challenges the rule of law and a vast array of fundamental rights. First, algorithmic

[46] See Frosio (n 7) 165-174.

[47] See Charles Clark, 'The Answer to the Machine is in the Machine' in P. Bernt Hugenholtz (ed), *The Future of Copyright in a Digital Environment* (Kluwer Law International 1996) 139 (discussing the application of digital right management systems to enforce copyright infringement online).

[48] See, in general, Giancarlo Frosio, 'Algorithmic Enforcement Online' in Paul Torremans (ed), *Intellectual Property Law and Human Rights* (Wolters Kluwer 2020) 709-744.

[49] See, for a discussion of this very issue in connection to the enactment of Art. 17 of Directive 2019/790/EU, Giancarlo Frosio, 'Reforming the C-DSM Reform: a User-Based Copyright Theory for Commonplace Creativity' (2020) 52(6) IIC 709-750.

[50] See Guido Calabresi, 'Some Thoughts on Risk Distribution and the Law of Torts' (1961) 70 Yale L J 499; Ronald Coase, 'The Problem of Social Cost' (1960) 3 J L & Econ. 1.

[51] But see Niva Elkin-Koren and Husovec arguing that technologies might be the only way how we address the concerns of over-blocking on scale and with necessary speed,: Niva Elkin-Koren, 'Fair Use by Design' (2017) 64 UCLA L Rev 22 (2017); Martin Husovec, 'The Promises of Algorithmic Copyright Enforcement: Takedown or Staydown? Which is Superior? And Why?' (2019) 42(1) Columbia J of Law & the Arts 53-84.

enforcement leads to privatisation of enforcement and delegation of public authority. Automated and algorithmic content moderation implies a number of policy choices, which are operated by platforms and rightsholders and subject freedom of expression and online cultural participation to a new layer of 'private ordering'.[52] Private ordering—and the retraction of the public from online enforcement—does push an amorphous notion of responsibility that incentivizes intermediaries' self-intervention to police allegedly infringing activities on the Internet.

Second, expansive online intermediaries' algorithmic enforcement obligations will result in architectural changes leading to sanitisation of allegedly infringing content—possibly legit freedom of expression—by design. Therefore, enforcement can become implicit by modifying the architecture of the Internet,[53] as, in Lessig words, the code is the law of cyberspace.[54]

Third, algorithms take decisions reflecting policy's assumptions and interests that have very significant consequences to society at large, yet there is limited understanding of these processes. Socially relevant choices online increasingly occur through automated private enforcement run by opaque algorithms, which creates a so-called 'black-box' society.[55] Algorithmic enforcement finds its primary Achilles' heel in algorithms' transparency and accountability, which remains an issue challenging from the outset efficient and democratic semiotic regulation online.[56]

Finally, this process highlights unescapable tensions with fundamental rights—such as freedom of information, freedom of expression, freedom of business or a fundamental right to Internet access—by limiting access to information, causing chilling effects, or curbing due process.[57] In this context, tensions have been highlighted between algorithmic enforcement and the European Convention of Human Rights (ECHR) and the Charter of Fundamental Rights of the European Union (EU Charter).[58] In

---

[52] See Matthew Sag, 'Internet Safe Harbors and the Transformation of Copyright Law' (2017) 93 Notre Dame L Rev 1.

[53] See Joel Reidenberg, 'Lex Informatica: The Formulation of Information Policy Rules Through Technology' (1998) 76 Texas L. Rev 553.

[54] See Lawrence Lessig, *The Code and Other Laws of Cyberspace* (Basic Books 1999) 3-60; William Mitchell, *City of Bits: Space, Place, and the Infobahn* (MIT Press 1995) 111. See also Lawrence Lessig, *Code: Version 2.0* (Basic Books 2006).

[55] Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Harvard U Press 2015).

[56] See, for some proposals promoting algorithmic transparency and accountability, Frosio (n 77) 740-743.

[57] See, for detailed overview of fundamental rights and content moderation, Giancarlo Frosio and Christophe Geiger, 'Taking Fundamental Rights Seriously in the Digital Services Act's Platform Liability Regime' (2022) European Law Review (forthcoming) <https://ssrn.com/abstract_id=3747756>.

[58] See Convention for the Protection of Human Rights and Fundamental Freedoms (European Convention on Human Rights, as Amended) [1950] (hereafter ECHR); Charter of Fundamental Rights of the European Union 2012 OJ (C 326) 391 (hereafter "EU Charter"). See also eg *Akdeniz v Turkey*, App no 20877/10, 11 March 2014.

particular, as the CJEU made clear in *Scarlet* and *Netlog*, automated filtering tools cannot properly balance intellectual property rights and other competing fundamental rights, such as freedom of expression, freedom of business and privacy.[59] For example, present algorithmic filtering tools impinge upon freedom of expression as they might be unable to identify uses made by applying exceptions and limitations or might not discern the public domain status of works.[60] More broadly, algorithmic enforcement appears to be inherently unsuitable to apply any equity-based—thus inherently human—judgement.[61]

IV. INSTITUTIONALISING PRIVATE ORDERING IN THE DIGITAL SERVICES ACT

Private ordering runs counter any constitutionalisation process of online regulation. State intervention have been so far increasingly delegating online enforcement and regulation to private parties, thus fragmenting responses and standards, rather than consolidating and 'constitutionalising' Internet governance. However, there are counter-posing forces at work in the present internet governance struggle. A centripetal move towards digital constitutionalism for Internet governance alleviates the effects of the centrifugal platform responsibility discourse. Efforts to draft an "Internet Bill of Rights" can be traced at least as far back as the mid-1990s. [62] Two full decades later, aspirational principles have begun to crystallize into law.

A move towards digital constitutionalism is partially occurring at multiple levels. Gill, Redeker and Gasser have described more than thirty initiatives spanning from 1999 to 2015 that can be labelled under the umbrella of 'digital constitutionalism'.[63] These initiatives show great differences—and range from advocacy statements to official positions of intergovernmental organisations to proposed legislation—but belong to a broader proto-constitutional discourse seeking to advance a relatively comprehensive set of rights, principles, and governance norms for the Internet.[64] Governments and

---

[59] See eg C-360/10 *Belgische Vereniging van Auteurs, Componisten en Uitgevers CVBA (SABAM) v Netlog NV* [2012] ECLI:EU:C:2012:85, para 51; Case C-70/10 *Scarlet Extended SA v Société belge des auteurs, compositeurs et éditeurs SCRL (SABAM)* [2011] ECLI:EU:C:2011:771, para 53. See also C-314/12, *UPC Telekabel Wien* [2014] EU:C:2014:192 (hereafter *UPC Telekabel*).

[60] *Netlog* (n 59) para 50.

[61] See Dan Burk and Julie Cohen, 'Fair Use Infrastructure for Copyright Management Systems' (2000) Georgetown Public Law Research Paper 239731/2000 <ssrn.com/abstract=239731>. See also, Amélie Pia Heldt, 'Upload-Filters: Bypassing Classical Concepts of Censorship?' (2019) 10(1) JIPITEC <https://www.jipitec.eu/issues/jipitec-10-1-2019/4877>.

[62] See Lex Gill, Dennis Redeker, and Urs Gasser, Towards Digital Constitutionalism? Mapping Attempts to Craft an Internet Bill of Rights (Berkman Center Research Publication No. 2015-15, November 9, 2015).

[63] See Gill, Redeker and Gasser (n 62) 1; *see also* Oreste Pollicino and Graziella Romeo (eds), *The Internet and Constitutional Law* (Routledge 2016); Edoardo Celeste, 'Digital Constitutionalism: A New Systematic Theorisation' (2019) 33 International Review of Law, Computers & Technology 76.

[64] See Gill, Redeker and Gasser (n 62) 1.

policymakers are heavily pressuring companies to take action and a few jurisdictions have already responded with new regulatory initiatives or proposals for future reform, including the Online Harm White Paper, the Digital Services Act or the Brazilian "Fake News" bill.[65]

Recently, the European Parliament might be strengthening this process of constitutionalisation of Internet governance, *de facto* institutionalising previous approaches that might have instead emphasised private ordering solutions. The Commission made a proposal for a Digital Services Act (hereinafter DSA) in December 2020 and on 23 April 2022 a political agreement was reached.[66] The European Parliament approved, then, the DSA on 5 July 2022.[67] The DSA Proposal would move in the direction of a substantive and procedural 'constitutionalisation' of Internet governance in the EU by structuring, standardising, and homogenising content moderation processes that have been recently promoted under a private ordering approach, while confirming that the *ex-post* knowledge-and-take-down mechanism of the ECD should be preferred to *ex ante* proactive moderation as it provides better guarantees for the fundamental rights of users.[68]

First, the proposed DSA constitutionalises private ordering practices by seeking a more balanced approach in digital policies where all competing fundamental rights and interests of stakeholders involved receive equal safeguards. Starting with the resolution on the 'Digital Services Act and fundamental rights issues posed',[69] the DSA Proposal highlights from the inception the central role of fundamental rights in online regulation, and for the DSA in particular. Specific reference to the key role of fundamental rights' protection in digital services' regulation abounds in the DSA Proposal, starting from the Explanatory Memorandum, which ensures safeguards for fundamental rights and reminds that the ECD prohibition of general monitoring obligation is crucial to the fair balance of fundamental

[65] See PL (Bill) 2630/20 'Brazilian Law of Freedom, Responsibility and Transparency on the Internet' (2020) < https://www25.senado.leg.br/web/atividade/materias/-/materia/141944>.

[66] See Proposal for a regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC, COM(2020) 825 (hereafter "DSA Proposal").

[67] See the EP legislative resolution of 5 July 2022 on the proposal for a regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC (COM(2020)0825 – C9-0418/2020 – 2020/0361(COD). At the time of writing, the DSA is still awaiting for formal adoption by Council of the European Union, after which the act will come into force twenty days after publication in the Official Journal of the European Union. The rules will begin to apply fifteen months after coming into force or on 1 January 2024.

[68] See Frosio and Geiger (n 57) 6-13.

[69] European Parliament, 'Resolution on the Digital Services Act and fundamental rights issues posed' [2020/2022(INI)] 20 October 2020, Provisional edition, P9_TA-PROV(2020)0274 <https://www.europarl.europa.eu/doceo/document/TA-9-2020-0274_EN.html> (hereafter 'EP Resolution on the DSA and FRs').

rights in the online world.[70] Again, Recital 3 of the DSA stresses that a responsible behaviour of Digital Service Providers ('DSPs') is essential for allowing the exercise of the fundamental rights guaranteed in the EU Charter, 'in particular the freedom of expression and information and the freedom to conduct a business, and the right to non-discrimination'.[71] Finally, the DSA should be interpreted and applied in accordance to the fundamental rights recognised by the EU Charter with an obligations for public authorities exercising the powers provided by the DSA to achieve a fair balance of the conflicting fundamental rights, in accordance with the principle of proportionality.[72] The DSA also includes at least two major specific prescriptive obligations for DSPs to enforce fundamental rights. First, by defining its scope, the DSA states that the aim of the Regulation is to regulate an online environment 'where fundamental rights enshrined in the Charter are *effectively* protected'.[73] Second, the DSA has included the impact of digital services on the exercise of fundamental rights protected by the EU Charter as a category of systemic risks that should be assessed in depth—and mitigated— by *very large online platforms*,[74] a new category of online platform to which special obligations apply.[75] In particular, the risk assessment of platforms' services must regard the impact of digital services on (i) the freedom of expression and information, (ii) the right to private life, (iii) the right to non-discrimination and (iv) the rights of the child.[76]

Second, the DSA regulates and institutionalises content moderation practices via the introduction of procedural guarantees that comply with the broad framework of fundamental rights.[77] On one side, the DSA associates effective protection of user rights with the possibility for users to obtain more information about content moderation practices, including transparency of the content removal process and human oversight.[78] On the other side, procedural guarantees for notice, counter notice

[70] See DSA Proposal (n 66) p 12 (according to the Explanatory Memorandum, general monitoring obligations 'could disproportionately limit users' freedom of expression and freedom to receive information, and could burden service providers excessively and thus unduly interfere with their freedom to conduct a business').

[71] ibid, Recital 3.

[72] ibid, Recital 105.

[73] ibid, art 1(2)(b) (emphasis added).

[74] ibid, Recital 56-57, art 26(1)(b).

[75] ibid, Recital 56, art 25 (including ‚online platforms which provide their services to a number of average monthly active recipients of the service in the Union equal to or higher than 45 million').

[76] ibid., Recital 57.

[77] ibid, art. 12 (2) (stating that in applying and enforcing any restrictions to the use of their services, DSPs shall act with due regard to the fundamental rights of the users).

[78] ibid, art. 12 (1) (providing that DSPs' terms and conditions must include ‚information on any policies, procedures, measures and tools used for the purpose of content moderation, including algorithmic decision-making and human review').

and complaint procedures must be ensured.[79] In particular, the DSA regulates procedural rights under a notice-and-actions mechanism by providing safeguards to fundamental rights.[80] In order to do so, the DSA is meant to provide guidance regarding: (i) the right to issue a notice and the form of notices;[81] (ii) procedural safeguard for processing notices and for final decision-making,[82] with a special regime that might be applied to trusted flaggers;[83] (iii) safeguards against the abuse of the system allowing to sanction parties that systematically and repeatedly submit wrongful notices[84] (or manifestly illegal content); (iv) transparency reports;[85] (v) access to internal complaint mechanisms that should be transparent, effective, fair and expeditious;[86] as well as (vi) the possibility to resort to out-of-court dispute settlement mechanisms and judicial redress.[87] In this regard, Article 18 of the DSA Proposal would like to set up a mechanism of certification of out-of-court dispute settlement bodies. Such bodies would be certified by the Digital Service Coordinator (a new entity to be established by the DSA) of the Member State of establishment of the body upon meeting certain requirements, such as impartiality, independence, expertise, technological capability, cost-efficiency, and procedural fairness.[88]

Third, the DSA includes some—scattered and partial—regulation of algorithmic enforcement. In this context, unlike art 17 of the Copyright in the Digital Service Market (C-DSM) Directive,[89] the DSA does

[79] As per the adequacy of the notice procedure to fundamental rights, the DSA Proposal categorizes as a risk, subject to systemic risk assessment and mitigating measures, the misuse of very large online platforms" service "through the submission of abusive notices or other methods for silencing speech or hampering competition". DSA Proposal (n 66) Recital 57.

[80] See DSA Proposal (n 66) Recital 41.

[81] See DSA Proposal (n 66) art 14(1) (stating that the submission mechanism should be user-friendly and by electronic means), (2) (clarifying that notices should be sufficiently precise and adequately substantiated and defining the elements that a notice should include) and (3) (noting that if the requested formalities of the notice are fulfilled, the notice "shall be considered to give rise to actual knowledge").

[82] See DSA Proposal (n 66) art 14(6). According to art. 15 of the DSA Proposal, the hosting provider must inform the user of the content removal or blocking and provide "a clear and specific statement of reasons of that decision", including a number of information listed in art. 15(2) of the DSA Proposal.

[83] ibid, art 19.

[84] See DSA Proposal (n 66) Recital 57, art 14(2)(d) and art 20(1)-(2) ((1)-(2). In particular, the status of "trusted flagger" can be revoked, following an investigation, if, *inter alia*, the trusted flagger "submitted a significant number of insufficiently precise or inadequately substantiated notices". DSA Proposal (n **Error! Bookmark not defined.**) art. 19(5-6).

[85] ibid, art. 13, 23, and 33.

[86] See DSA Proposal (n 66) art 17 (providing that DSPs must make available an online, free-of-charge, internal complaint-handling mechanism for at least six months following decisions to (a) remove or disable access to the information; (b) suspend or terminate the provision of the service, in whole or in part, to the recipients; (c) suspend or terminate the recipients" account).

[87] See DSA Proposal (n 66) art 18.

[88] See DSA Proposal (n 66) art 18(2).

[89] See Directive 2019/790/EU of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC [2019] OJ L 130/92,

not contain any obligation for hosting service providers or other technical intermediaries to use automated tools in content moderation. Instead, the DSA includes provision to promote transparency of algorithmic decision-making when deployed by platforms. First, when hosting providers use automated means for processing or decision-making, they shall include information on such use in a notice to be sent to the party that submitted the notice.[90] Second, that information on the use of automated means must be included in the statement of reasons that must accompany removal/blocking decisions.[91] Third, the DSA provides a systemic risk assessment obligation in relation to the design of algorithmic systems, coupled by the obligation of introducing mitigating measures, for *very large online platforms* in connection to the impact of their services on the exercise of fundamental rights.[92] Finally, the DSA imposes upon DSPs an obligation to submit annual transparency reports, which should also be published on a publicly accessible database. [93]

Fourth, new obligations in the DSA must be proportionate to the scale of reach as well as technical and operational capacities of DSPs. In sharp contrast, again, with the solution endorsed by art 17 C-DSM Directive, a proportional approach to obligations of DSPs might better safeguards fundamental rights.[94] The DSA endorses this graduate obligation approach by distinguishing between (i) providers of intermediary services, (ii) micro or small enterprises, (iii) providers of online platforms, and (iii) very large online platforms. Intermediary services are the broader category including access, caching, and hosting providers,[95] while 'online platforms' are hosting providers 'which, at the request of a recipient of the service, [store] and [disseminate] to the public information'.[96] Among online platforms, the DSA Proposal distinguishes between online platforms proper, micro or small enterprises,[97] which are excluded from additional obligations applicable to online platforms,[98] and very large online platforms,

art 17 (providing that a subset of online platforms, known as Online Content Sharing Service Providers (OCSSP'), are now directly liable for communicating to the public copyright infringing content that their users might make available on their networks, thus, under certain conditions, OCSSP have an obligation either to (1) conclude licencing agreements to make available that protected content on the platforms or (2) make 'best efforts' to make unavailable any infringing content on their networks, which might turn *de facto* in an obligation to set up proactive monitoring and filtering or so-called 'upload filters'). See also Giancarlo Frosio, 'Reforming the C-DSM Reform: a User-Based Copyright Theory for Commonplace Creativity' (2020) 52(6) IIC 709-750.

[90] DSA Proposal (n 66) Recital 57, art. 14(6)

[91] ibid, art 15(2)(c).

[92] See DSA Proposal (n 66) Recital 56-57.

[93] See, for further information on the nature of these reports, DSA Proposal (n 66) art 13 and 23.

[94] Directive 2019/790/EU (n 89) art 17. See also Frosio (n 89) 721, 727-728 (discussing negative externalities on start-ups and smaller platforms of the new obligations included in the C-DSM Directive).

[95] See DSA Proposal (n 66) art. 2 (f).

[96] ibid, art. 2(h) ("unless that activity is a minor and purely ancillary feature of another service and, for objective and technical reasons, cannot be used without that other service").

[97] Within the meaning of the Annex to Recommendation 2003/361/EC.

[98] See DSA Proposal (n 66) art 16.

which are subject to stricter obligations than mere online platforms.[99] In particular, very large online platforms, as mentioned already, must assess risks in connection to (i) dissemination of illegal content though their services, (ii) negative effects on fundamental rights, and (iii) intentional manipulation of their services affecting public health, minors, civic discourse, elections, and public security.[100] Afterwards, if it is the case, very large online platforms must put in place mitigation measures as a result of their risk assessment,[101] and independent audit.[102]

Finally, the DSA would like to create EU-wide coordinated mechanisms to oversee compliance by online intermediaries with the DSA regulation, with special emphasis on standards for content management and algorithmic technologies deployed to that end.[103] The DSA is meant to implement an arrangement similar to that governing the privacy domain. Member States should designate one or more authorities competent for the enforcement of the regulation, one of which should serve as Digital Service Coordinator (DSC) in the Member States.[104] The DSCs shall have broad powers spanning from ordering the cessation of infringements, imposing remedies to bring infringements to an end, to imposing fines, periodic penalty payments, and adopting interim measures to avoid serious harm.[105] The DSCs should make up an "European Board for Digital Services", similar to the European Data Protection Board, formerly the Article 29 Working Party, an independent body tasked with advising the DSCs and the Commission on the application of the DSA and the supervision of very large online platforms.[106]

V. CONCLUSIONS

In the last few years, enhanced intermediary responsibilities have broadened private ordering practices to the extent that, as mentioned at the very beginning of this chapter, platforms might influence, for example, critical political speech with their own voluntary measures. This trend might have brought about a marked democratic deficit in content moderation online with considerable repercussions on the enjoyment of multiple fundamental rights and the proper balancing of competing interests online. The forthcoming DSA attempts to govern this widespread private ordering practices via their constitutionalisation, regulation and institutional governance.

[99] ibid, art 25.

[100] ibid, art 26.

[101] ibid art 27.

[102] ibid, art 28.

[103] ibid, art 38-48.

[104] Ibid, art 38.

[105] ibid, art 41.

[106] ibid, arts 47-48.

First, the DSA lunches a process of constitutionalisation of online regulation by emphasizing the role of fundamental rights online and taking the safeguard of the fundamental rights of users, platforms, and other stakeholders very seriously. In doing so, the DSA redeploys the constitutional legal framework set up by ECHR and the EU Charter as primary references, and the DSA serves, inter alia, as a tool to promote the application of the EU Charter online. The DSA highlights the central role of the EU Charter in online regulation in Recital 3 by stressing that a responsible behaviour of DSPs is essential for allowing the exercise of the fundamental rights guaranteed in the EU Charter, 'in particular the freedom of expression and information and the freedom to conduct a business, and the right to non-discrimination'. Again, according to Recital 105, the DSA should be interpreted and applied in accordance with the fundamental rights recognised by the EU Charter with an obligation for public authorities exercising the powers provided by the DSA to achieve a fair balance of the conflicting fundamental rights, in accordance with the principle of proportionality.

Second, the DSA introduces provisions meant to regulate content moderation online in a proportionate and user-friendly manner. Differing regulatory approaches should be applied to illegal, manifestly illegal, and harmful content. The DSA choses the right policy approach by not regulating harmful content, but rather harmonizing rules for tackling illegal content.[107] Potentially harmful content should not be specifically addressed from a DSA perspective as controversial content cannot be censored just because it makes the audience uncomfortable.[108] Rather than subjecting potentially harmful (but not illegal) content to removal obligations, it would be preferable from a freedom of expression perspective to consider alternative remedies to content removal, such as DSPs and users flagging of inappropriate content, and other counter-speech approaches, such as like or dislike, or availability of more users' choices in what they want to see.[109] Meanwhile, the DSA should have made a distinction between illegal content[110] and manifestly illegal content, which could be removed in a simplified manner.[111] Along with apology of acts constituting an offence against human dignity, war crimes, crimes against humanity, human trafficking, incitement to or apology of violence, acts of terrorism and child abuse content, manifestly illegal content could include manifestly IP infringing

[107] See DSA Proposal (n 66) Recital 12. See also EP Resolution on the DSA and commercial and civil rules (n 125) point 3; European Parliament, 'Resolution on the Digital Services Act and fundamental rights issues posed' [2020/2022(INI)] 20 October 2020, Provisional edition, P9_TA-PROV(2020)0274 <https://www.europarl.europa.eu/doceo/document/TA-9-2020-0274_EN.html> (hereafter 'EP Resolution on the DSA and FRs') point 5.

[108] See Richard Wingfield, The Digital Services Act and Online Content Regulation: A slippery slope for human rights? (*Medium*, *The GNI Blog,* 15 July 2020) <https://medium.com/global-network-initiative-collection/the-digital-services-act-and-online-content-regulation-a-slippery-slope-for-human-rights-eb3454e4285d>.

[109] See Frosio and Geiger (n 57) 44-46.

[110] EP Resolution on the DSA and FRs (n **Error! Bookmark not defined.**) point 5.

[111] See Frosio and Geiger (n 57) 51.

content for which no equity-based assessment is needed.[112] In this latest regards, any content that might encompass a privileged use[113] or whose protected status could be otherwise questioned should never be considered 'manifestly illegal'. Simplified forms of removal, including via automated tools, should be limited to content that is manifestly illegal or has been found illegal by an independent (be it judicial or quasi-judicial) authority.[114] Instead, should any doubts exist as to the content's 'illegal' nature, such content should be subject to human review and not be removed without further investigation, either carried out internally by the DSPs or run by external, independent entities.[115] Of course, legal content should not be subject to any legal removal or blocking obligations.[116] This would include misinformation that does not meet the notion of illegal content in Recital 12 of the DSA.[117]

However, in updating the content moderation regulatory framework, the DSA misses more specific and granular obligations to address algorithmic opacity and "black box society" broader challenges. In particular, the DSA should include safeguards against their abuse that might impinge on fundamental rights. First, should any automated decision-making tools be used on DSPs' own initiative to restrict content, the DSA should establish transparency and non-discrimination requirements for algorithms, data sets used to train them and the decision-making process.[118] Secondly, the decision-making process should be subject to 'human-in-command' principle, which is the capability to oversee the overall activity of the AI system, including its impact, and the ability to decide when and how to use the system.[119] Thirdly, automated removal and blocking mechanisms should be subject to external audits by independent entities or by an EU oversight structure to be specifically created or empowered for DSA enforcement purposes,[120] who should be authorised to carry out a case-by-case oversight and

[112] Ibid. EP Resolution on the DSA and commercial and civil rules (n 125) suggests to define it as the content which 'unmistakably and without requiring in-depth examination [is] in breach of legal provisions regulating the legality of content on the internet'.

[113] A good example for such a privileged use in the field of copyright are exceptions and limitations which have been qualified as users' right and need to be safeguarded from a fundamental rights perspective (see *supra*).

[114] See DSA Proposal (n 66) art 9 (1-2) (defining conditions meant to safeguard rights of the affected parties that orders issued by a judicial or administrative authority to act against illegal content must meet).

[115] See Frosio and Geiger (n 57) 39-40.

[116] Ibid, 39-40.

[117] ibid (with a broader discussion on content moderation of misinformation).

[118] Ibid, 43-44 (noting that platforms should flag at which step of the process algorithmic tools were used, be transparent in the processing of algorithms and of the data used to train them, explain the logic behind the automated decisions taken, and also explain how users can contest their decisions).

[119] See European Parliament, Resolution on the Digital Services Act: Improving the functioning of the Single Market [2020/2018(INL)] 20 October 2020, Provisional edition, P9_TA-PROV(2020)0272, <https://www.europarl.europa.eu/doceo/document/TA-9-2020-0272_EN.html> (hereafter 'EP Resolution on the DSA and the Single Market') point 41 and Annex to the Resolution, 28. See also High-Level Expert Group on Artificial Intelligence, 'Ethics Guidelines For Trustworthy AI' (European Commission, 8 April 20219) 15 <https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419>.

[120] See EP Resolution on the DSA and commercial and civil rules (n 125) point 12.

recurrent risk assessment.[121] Fourthly, there should be clear accountability, liability and redress mechanisms to deal with potential harm resulting from the use of AI applications, automated decision-making and machine learning tools.[122] Measures to ensure non-discrimination and transparency of algorithms and data sets should be the cornerstone of an effective DSA reform,[123] while the use of automated tools should be 'proportionate, covering only justified cases'.[124]

Finally, the DSA institutionalises private ordering by tasking new specialized institutional oversight entities with monitoring DSPs' compliance with the DSA. However, in order to provide a more harmonized outlook and avoid forum shopping and political biases, the establishment of "an independent EU oversight structure" with power to adopt decisions directly against DSPs, rather than national DSCs coupled with an EU advisory board, might better promote regulatory harmonisation as well as the proper enforcement of the DSA.[125] Also, the proposals on the table seem to construe this entity closer to an antitrust authority that oversees compliance of DSPs with DSA's obligations, rather than serving as guarantor of fundamental rights. Instead, for this purpose, this entity could work closely with the European Union Agency for Fundamental Rights.[126] Working in strict collaboration with the European Union Agency for Fundamental Rights, this EU entity should identify difficulties in implementation and issuing policy guidelines on how to remedy them, taking duly into account the fundamental rights impacted. Finally, an Ombudsperson may be empowered to join proceedings on behalf of users to protect their rights online.[127] The Ombudsperson could act when a content hosting platform has decided to remove, take down or disable access to content, or otherwise to act in a manner that is contrary to the interest of the uploader or infringes fundamental rights. Recently, also the Parliamentary Assembly of Council of Europe called on the European Union to consider whether

[121] See EP Resolution on the DSA and the Single Market (n 118) point 41 and Annex to the Resolution, 28.

[122] ibid Annex to the Resolution, 29.

[123] Transparency obligations should also apply to users profiles' content curation.

[124] See EP Resolution on the DSA and commercial and civil rules (n 125) point 'H'.

[125] See Frosio and Geiger (n 57) 44-46 (in particular also re forum shopping mentioned above). See European Parliament, Resolution on the Digital Services Act: adapting commercial and civil law rules for commercial entities operating online [2020/2019(INL)] 20 October 2020, Provisional edition, P9_TA-PROV(2020)0273, https://www.europarl.europa.eu/doceo/document/TA-9-2020-0273_EN.html (hereafter 'EP Resolution on the DSA and commercial and civil rules') point 40. See also Jan Nordemann and others, 'Annex III: Digital services act: Improving the functioning of the single market' in Niombo Lomba and Tatjana Evas, *Digital services act: European added value assessment* (European Parliament, October 2020) <https://www.europarl.europa.eu/RegData/etudes/STUD/2020/654180/EPRS_STU(2020)654180_EN.pdf>, 303-307 (highlighting positive and negative impacts of a new regulatory agency).

[126] See Frosio and Geiger (n 57) 44-46.

[127] ibid

to set up an Internet Ombudsman to foster harmonisation of legislation on internet content, while facilitating swift deletion of manifestly illegal content.[128]

The DSA has launched a process of constitutionalisation, regulation and institutional governance of online content moderation and private ordering practices. With a special emphasis on large platforms and the risks they might cause to users' rights online, the DSA pushes a model for platform responsibility and governance that takes fundamental rights's protection very seriously and might serve as a bluprint for future reform in other jurisdictions. Still, algorithmic enforcement remains a serious threat for user's rights online. Making more transparent the black boxes where the true governance of our online experience remains hidden is a high call that the DSA has not fully aswered and should instead be a major focus of any platform responsibility regulation.

[128] See Standing Committee of the Parliamentary Assembly of the Council of Europe, Resolution 2334 (2020), Provisional version, 'Towards an Internet Ombudsman institution', 15 September 2020, point 6, <https://pace.coe.int/en/files/28728/html>.