



THE DSA, DUE DILIGENCE & DISINFORMATION: A DISJOINTED APPROACH OR A RISKY COMPROMISE?



BY
KATIE PENTNEY

DPHil Candidate in Law, University of Oxford: katie.pentney@law.ox.ac.uk.

THE DIGITAL SERVICES ACT - A LAUDABLE AMBITION, BUT WILL IT DELIVER?

By Michèle Ledger & Laura Sboarina



YOU MAY BE SUBJECT AS WELL!: DIGITAL SERVICES ACT - WHAT COMPANIES NEED TO KNOW

By Julia Apostle, Kelly Hagedorn, Christian Schroder & Adele Harrison



CONTENT MODERATION AND COMPETITION IN DIGITAL MARKETS

By Maciej Sobolewski & Néstor Duch-Brown



OPERATIONALIZING THE REGULATION OF ONLINE CONTENT UNDER A DEMOCRATIC DEFICIT: THE DIGITAL SERVICES ACT

By Dr. Joseph Downing



THE DSA, DUE DILIGENCE & DISINFORMATION: A DISJOINTED APPROACH OR A RISKY COMPROMISE?

By Katie Pentney



ALGORITHMIC SEARCH AND RECOMMENDER SYSTEMS IN THE DIGITAL SERVICES ACT

By Oliver Budzinski & Madlen Karg



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

THE DSA, DUE DILIGENCE & DISINFORMATION: A DISJOINTED APPROACH OR A RISKY COMPROMISE?

By Katie Pentney

The newly-introduced Digital Services Act (“DSA”) sets as its ambition ensuring a “safe, predictable and trusted online environment” by targeting the spread of illegal content, on the one hand, and the spread of harmful content, like disinformation, on the other. It imposes particular due diligence obligations on very large online platforms, like Facebook and Twitter, to achieve this end. But the vagueness of the provisions, the deference afforded to these platforms, and the disjointed approach to harmful content like disinformation specifically may hamper the DSA’s ability to fulfil its promise. This article sets out the key provisions of the heightened due diligence framework, the underlying compromises made during the negotiations, and the lingering challenges that lie ahead, particularly with a new leader – and self-proclaimed “free speech absolutist” – at the helm of Twitter.

Scan to Stay Connected!

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



01

INTRODUCTION

The long-awaited Digital Services Act (“DSA”) was finally signed into law by the European Union on October 19, 2022, after lengthy drafting and hard-fought negotiation processes.² The flagship Regulation harmonises existing rules applicable to internet intermediaries and imposes new transparency and accountability requirements on online platforms, as well as heightened due diligence obligations on so-called “very large online platforms” (“VLOPs”) like Facebook, Google (YouTube) and Twitter.³ The stated objective of the DSA is to ensure a “safe, predictable and trusted online environment” by addressing the dissemination of illegal content online, as well as “the societal risks that the dissemination of disinformation or other content may generate.”⁴

This was a long time coming for those concerned about the well-documented proliferation of illegal and harmful content online. The celebrations were, however, short-lived (or at least dampened): the day after the DSA was published in the EU Official Journal, marking the end of its adoption process (and the start of the 20-day countdown until its entry into force), Elon Musk completed his acquisition of Twitter. The takeover sparked concern that

the self-proclaimed “free speech absolutist” would roll-back existing content moderation practices and allow conspiracy theories, disinformation and hate speech to proliferate unabated on the platform.⁵ While it is still early days, at least some of these concerns appear to be well-founded: in the 48 hours following the takeover, Twitter’s Head of Safety & Integrity tweeted that “a small number of accounts post[ed] a ton of Tweets that include slurs and other derogatory terms,” before adding “To give you a sense of scale: More than 50,000 Tweets repeatedly using a particular slur came from just 300 accounts.”⁶ The entire human rights team at Twitter has since been fired,⁷ and Musk himself has since tweeted, and then deleted, an unfounded conspiracy theory regarding the attack on US Speaker of the House Nancy Pelosi’s husband, Paul.⁸ Before he had deleted the tweet, it had been retweeted 24,000 times and received more than 86,000 likes.⁹

The Twitter takeover by a self-proclaimed “free speech absolutist” illustrates the potential pitfalls of the EU’s chosen approach of “deferential regulating” – through which it imposes due diligence obligations on the likes of Twitter, Facebook and other VLOPs operating within the EU, but affords significant deference and leeway for internal decision-making by these online platforms. The battles to be waged are (somewhat ironically) best illustrated by a Twitter exchange between Musk and the EU’s Internal Market commissioner, Thierry Breton. Upon finalizing his acquisition, Musk tweeted,

2 Regulation (EU) 2022/2065 of the European Parliament and the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act) (“DSA”), Official Journal of the European Union L 277, Vol 65 (27 October 2022).

3 DSA, Recital 9. “VLOP” means, for the purposes of the Regulation, online platforms “which have a number of average monthly active recipients of the service in the Union equal to or higher than 45 million, and which are designated as [VLOPs...] pursuant to paragraph 4” (DSA, Article 33(1)). See also Natascha Just, *The Taming of Internet Platforms – A Look at the European Digital Services Act*, CPI TechREG CHRONICLE (June 15, 2022), <https://www.competitionpolicyinternational.com/the-taming-of-internet-platforms-a-look-at-the-european-digital-services-act/>.

4 DSA, *supra*, Recital 9.

5 Dan Milmo & Alex Hern, *Twitter takeover: fears raised over disinformation and hate speech*, THE GUARDIAN, Oct. 28 2022, <https://www.theguardian.com/technology/2022/oct/28/twitter-takeover-fears-raised-over-disinformation-and-hate-speech>; Guardian staff and agencies, *Elon Musk declares Twitter “moderation council” – as some push the platform’s limits*, THE GUARDIAN, Oct. 29, 2022 <https://www.theguardian.com/technology/2022/oct/28/elon-musk-twitter-moderation-council-free-speech>.

6 Yael Roth, Twitter, <https://twitter.com/yoyoel/status/1586542283469381632>.

7 Kate Conger, Ryan Mac & Mike Isaac, *Confusion and Frustration Reign as Elon Musk Cuts Half of Twitter’s Staff*, NEW YORK TIMES, Nov. 4, 2022, <https://www.nytimes.com/2022/11/04/technology/elon-musk-twitter-layoffs.html>; Sam Levin, Richard Luscombe & Graeme Wearden, *Twitter layoffs: anger and confusion as multiple teams reportedly decimated – as it happened*, THE GUARDIAN, Nov. 5, 2022 <https://www.theguardian.com/business/live/2022/nov/04/twitter-sued-layoffs-sizewell-nuclear-plant-uk-recession-us-jobs-business-live#:~:text=The%20human%20rights%20team%20has,in%20Ukraine%2C%20Afghanistan%20and%20Ethiopia>.

8 Julianne McShane, *Elon Musk, new owner of Twitter, tweets unfounded anti-LGBTQ conspiracy theory about Paul Pelosi attack*, NBC NEWS, Oct. 30, 2022 <https://www.nbcnews.com/news/us-news/elon-musk-new-owner-twitter-tweets-unfounded-conspiracy-theory-paul-pe-rcna54717>.

9 *Id.*

“the bird is freed”; shortly thereafter, Breton retorted (in Tweet form): “In Europe, the bird will fly by our [EU] rules.”¹⁰



This article offers some preliminary thoughts on the likelihood that these “EU rules” will achieve their stated aims of ensuring a “trusted online environment,” generally, and addressing the societal risks of online disinformation, specifically. While the DSA imposes transparency and other requirements on all internet intermediaries, the focus of this article is on the heightened due diligence framework imposed on VLOPs, in particular. It proceeds in two parts. First, I provide a brief overview of the key features of the risk-based due diligence framework, as well as some of the issues they raise. Second, I offer some reflections on the newly enacted DSA’s disjointed approach to disinformation, specifically, and the enforcement difficulties which seem poised to lie ahead, if Musk’s recent acquisition of Twitter is any indication.

02

THE DSA’S HEIGHTENED DUE DILIGENCE FRAMEWORK

The EU is not alone in expressing concerns about the societal risks that the proliferation of disinformation online may pose. To the contrary, such concerns are well documented and multifaceted, particularly when it comes to elections, public health emergencies or foreign invasions. The World Health Organization (“WHO”) has decried the “infodemic” that has accompanied – and at times, worsened – the COVID-19 pandemic: indeed, WHO notes that “In the first 3 months of 2020, nearly 6 000 people around the globe were hospitalized because of coronavirus misinformation” and during this same period, “research say at least 800 people may have died due to misinformation related to COVID-19.”¹¹ Carley notes that as COVID-19 spread around the world, so too did “an epidemic of disinformation and misinformation”:

Estimates suggest that there have been hundreds of thousands of distinct disinformation stories with respect to the pandemic. These stories included the innocuous—such as due to the lockdown pollution was lower in Venice and the swans and dolphins returned to the canals. Other stories were lethal—such as drink bleach to cure yourself of COVID-19. Still other disinformation stories were woven together to form larger conspiracy theories—such as Bill Gates invented the SARS-CoV-2 virus and the vaccine [...].¹²

Beyond COVID-19, the impact of so-called “information disorder”¹³ on elections in the US and France and on referenda in the United Kingdom and beyond has raised concerns about the effects of disinformation, misinformation

¹⁰ Thierry Breton, Twitter, https://twitter.com/ThierryBreton/status/1585902196864045056?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Cwtterm%5E1585902196864045056%7Ctwgr%5E1f36754db79be083c89e8995b46b97d9fff8f4ff%7Ctwcon%5Es1_&ref_url=https%3A%2F%2Fwww.theguardian.com%2Ftechnology%2F2022%2Foct%2F28%2Ftwitter-takeover-fears-raised-over-disinformation-and-hate-speech.

¹¹ World Health Organization, Fighting misinformation in the time of COVID-19, one click at a time (April 27, 2021) <https://www.who.int/news-room/feature-stories/detail/fighting-misinformation-in-the-time-of-covid-19-one-click-at-a-time>, citing Md Saiful Islam et al, *COVID-19-Related Infodemic and Its Impact on Public Health*, 103 Am. J. Trop. Med. Hyg. 4, 1621 (2020). See also European Commission, *Tackling coronavirus disinformation* https://ec.europa.eu/info/live-work-travel-eu/coronavirus-response/fighting-disinformation/tackling-coronavirus-disinformation_en.

¹² Kathleen Mary Carley, *A Political Disinfodemic*, in COVID-19 DISINFORMATION: A MULTI-NATIONAL, WHOLE OF SOCIETY PERSPECTIVE (Rita Gill & Rebecca Goosby eds, 2022) 1, 2.

¹³ This is the umbrella term used by Claire Wardle and Hossein Derakhshan to refer to three subcategories: disinformation, misinformation and malinformation. Claire Wardle & Hossein Derakhshan, *Information Disorder: Toward an interdisciplinary framework for research and policy making*, Council of Europe, DGI(2017)09 (2017) <https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-research/168076277c>.

and malinformation in public discourse and democratic processes.¹⁴ Similarly, the Russian state (and its affiliates) have weaponized disinformation to justify and perpetuate the war on Ukraine.¹⁵ Tim Wu notes the distorting effect of disinformation campaigns, which have “rapidly become the speech control technique of choice in the early 21st century.”¹⁶ He posits that “disinformation techniques are a serious threat to the functioning of the marketplace of ideas and democratic deliberation, and therefore, it has fallen upon other institutions—especially the press and sometimes others—to fight them.”¹⁷

It is against this backdrop that the EU has adopted the DSA – its flagship regulation imposing requirements on internet intermediaries to join the fight against the spread of illegal and harmful content online. For present purposes, the key feature of interest is the DSA’s imposition of a heightened due diligence framework on VLOPs in light of their scale, reach and importance in “facilitating public debate, economic transactions and the dissemination to the public of information, opinions and ideas and in influencing how recipients obtain and communicate information online.”¹⁸ There are three main pillars of the heightened due diligence approach: (i) a systemic risk assessment; (ii) mitigation of identified systemic risks; and (iii) an annual independent audit requirement.¹⁹ Each of these pillars is reviewed in turn.

“It is against this backdrop that the EU has adopted the DSA – its flagship regulation imposing requirements on internet intermediaries to join the fight against the spread of illegal and harmful content online

A. The Risk Assessment

The first and foundational element of the heightened due diligence framework is the requirement that VLOPs undertake a risk assessment in which they “diligently identify, analyse and assess any systemic risks in the Union stemming from the design or functioning of their service and its related systems, including algorithmic systems, or from the use of their services.”²⁰ The risk assessment must be “specific to their services and proportionate to the systemic risks, taking into consideration their severity and probability” and must include the following identified “systemic risks”:

- (a) the dissemination of illegal content through their services;
- (b) any actual or foreseeable negative effects for the exercise of fundamental rights, including human dignity, respect for private and family life, data protection, freedom of expression, and non-discrimination;
- (c) any actual or foreseeable negative effects on civic discourse and electoral processes, and public security; and
- (d) any actual or foreseeable negative effects in relation to gender-based violence, the protection of public health and minors and serious negative consequences to the person’s physical and mental well-being.²¹

This provision indicates the two strands of content identified to pose a “risk” and therefore targeted by the Regulation: illegal content, on the one hand, and “lawful but awful” content, on the other. However, while one of the stated objectives of the DSA is to address the “societal risks that the dissemination of disinformation or other content may generate,” disinformation is not included as a specific systemic risk of which VLOPs must be aware. This may be because disinformation traverses the systemic risks iden-

14 See generally Max Bader, *Disinformation in Elections*, 29 Sec. and Hum. R. 24 (2018); SANDRINE BAUME ET AL. (eds) MISINFORMATION IN REFERENDA (1st ed., 2021).

15 Olivia B Waxman, *What Putin Gets Wrong About ‘Denazification’ in Ukraine*, TIME, Mar. 3, 2022, <https://time.com/6154493/denazification-putin-ukraine-history-context/>; Brian Klaas, *Vladimir Putin Has Fallen Into the Dictator Trap*, THE ATLANTIC, Mar. 16, 2022) <https://www.theatlantic.com/ideas/archive/2022/03/putin-dictator-trap-russia-ukraine/627064/>. See also *Allegations of Genocide under the Convention on the Prevention and Punishment of the Crime of Genocide (Ukraine v. Russian Federation)*, Order of the International Court of Justice (March 23, 2022) §§ 28-47.

16 Tim Wu, *Disinformation in the Marketplace of Ideas*, 51 Seton Hall L.R. 169, 169 (2020).

17 *Id.* 170.

18 DSA, *supra*, recital 75. See generally DSA, Section 5.

19 For a more in-depth review of the (draft) provisions, see Tarlach McGonagle & Katie Pentney, *From risk to reward? The DSA’s risk-based approach to disinformation* in UNRAVELLING THE DIGITAL SERVICES ACT PACKAGE (IRIS Special, European Audiovisual Observatory, M. Cappello ed., 2021) 43.

20 DSA, *supra*, Article 34(1).

21 *Id.*

tified – from negatively affecting civic discourse and electoral processes, to public security, to protection of public health. Yet it is a notable divergence with the approach to “illegal content,” which is explicitly identified as a systemic risk, and included without further elaboration of particular kinds of illegal content.²² But there is very little guidance or direction about what kinds of “actual or foreseeable negative effects” on fundamental rights, civic discourse/elections, public security, or protection of public health VLOPs would fit the bill, what threshold must be reached in order for the risk to be “systemic,” or how proximate such effects must be. Read broadly, this provision could capture much of what happens in the online ecosystem, given the scope of the fundamental rights included in the Regulation and the breadth and vagueness of the systemic risks listed. This could have serious repercussions for the flow of information and ideas online – particularly those which might “offend, shock or disturb”²³ – when read together with the second element of the due diligence framework: the requirement of mitigation.



This provision indicates the two strands of content identified to pose a “risk” and therefore targeted by the Regulation

B. The Mitigation of Risk Requirement

The second pillar is the requirement that VLOPs put in place “reasonable, proportionate and effective mitigation measures” which are “tailored to the specific systemic risks identified” and “with particular consideration to the impacts of such measures on fundamental rights.”²⁴ The Regulation sets out a list of illustrative examples of such mitigation measures, including adapting the design, features or functioning of their platforms, taking awareness-raising measures to give users more information, and ensuring that false or inauthentic information “is distinguishable through prominent markings when presented on their online interfaces.”²⁵ While the Regulation requires mitigation measures that are tailored to the systemic risks identified, it once again defers to VLOPs with respect to how best to do so, and provides little guidance about what would fulfill the qualitative requirements that the measures be reasonable, proportionate and effective.

The more generalized mitigation measures are supplemented by the “crisis response mechanism” particularized in Article 36, which is triggered (somewhat unhelpfully) and imprecisely “[w]here a crisis occurs.”²⁶ The preamble notes that a crisis “should be considered to occur when extraordinary circumstances occur that can lead to a serious threat to public security or public health in the Union or significant parts thereof” and further provides that such crises “could result from armed conflicts or acts of terrorism, [...] natural disasters [...] as well as from pandemics and other serious cross-border threats to public health.”²⁷ The crisis response mechanism was a late addition to the DSA: it did not appear in earlier drafts, but was added in response to the Russian war on Ukraine.²⁸ It was the subject of significant criticism from civil society organizations when it was introduced late in the process on the basis that it was “an overly broad empowerment of the European Commission to unilaterally declare an EU-wide state of emergency” and would “enable far-reaching restrictions of freedom of expression and of the free access

22 Recital 12 does provide that “the concept of ‘illegal content’ should be defined broadly to cover information relating to illegal content, products, services and activities.” (DSA, *supra*, Recital 12). It further states that “Illustrative examples include the sharing of images depicting child sexual abuse, the unlawful non-consensual sharing of private images, online stalking” and so on.

23 *Handyside v. United Kingdom*, App no 5493/72 (Plenary, December 7, 1976) § 49.

24 DSA, *supra*, Article 35(1).

25 *Id.*

26 DSA, *supra*, Article 36.

27 *Id.*, Recital 91.

28 38 organizations called on DSA negotiators to “stop negotiating outside their respective mandates and respect the democratic process of the EU”: see Press Release, European Digital Rights (**EDRI**), A new crisis response mechanism for the DSA (April 12, 2022) <https://edri.org/our-work/public-statement-on-new-crisis-response-mechanism-and-other-last-minute-additions-to-the-dsa/>. See also Press Release, Access Now, Civil society to EU: don’t threaten rights with last-minute ‘crisis response mechanism’ in DSA (April 13, 2022) <https://www.accessnow.org/crisis-response-mechanism-dsa/>.

to and dissemination of information in the Union.”²⁹ Some of the specific concerns were addressed in the Regulation as adopted, including requiring that the actions taken in line with this provision are “strictly necessary, justified and proportionate, having regard in particular to the gravity of the serious threat referred to in paragraph 2, the urgency of the measures and the actual or potential implications for the rights and legitimate interests of all parties concerned.”³⁰

C. The Independent Audit

The third and final pillar of the due diligence scheme is the independent audit, to which VLOPs shall be subjected on an annual basis to assess compliance with the transparency and due diligence obligations set out in Chapter III and with any commitments they’ve undertaken pursuant to codes of conduct and crisis protocols.³¹ The audit must result in a report which includes an opinion on whether the VLOPs complied with their obligations and commitments.³² Where the opinion is not “positive,” the report must also include operational recommendations on the specific measures to achieve compliance and the recommended timeframe for doing so.³³ The report may be redacted as necessary to protect confidential information.³⁴ Upon receipt of the audit report, providers of VLOPs must “take due account of the operational recommendations addressed to them with a view to take the necessary measures to implement them.”³⁵ They have one month from receiving the recommendations to adopt an “audit implementation report” setting out implementation measures.³⁶ Given the scope of the obligations set out in the DSA, it may be impractical – if not impossible – for VLOPs to respond to the audit report within this timeframe, or to do so in more than a cursory way. Moreover, while this third and final piece brings in the independent oversight needed to peer behind the veil, the requirement that VLOPs “take due account of” the recommendations provided “with a view to take the necessary measures to implement them” seems to leave significant leeway to

VLOPs about how quickly, and how thoroughly, they must make changes.

The third and final pillar of the due diligence scheme is the independent audit, to which VLOPs shall be subjected on an annual basis to assess compliance with the transparency and due diligence obligations set out in Chapter III and with any commitments they’ve undertaken pursuant to codes of conduct and crisis protocols

D. A Disjointed Approach or a Risky Compromise?

The risk-based approach thus attempts to balance the competing interests and calls from interested sectors of the population, including the public and regulators, civil society, and online platforms. It responds to regulators’ (and members of the public’s) desire to combat the proliferation of harmful and illegal content online by requiring VLOPs to play ball in addressing the problem. At the same time, it takes on board the concerns raised by civil society organizations within (and beyond) Europe relating to the lack of transparency about how content moderation decisions are made by large online platforms like Facebook, Twitter and Google (YouTube) and the absence of oversight as to whether such decisions comply with fundamental rights under the EU Charter.³⁷ Finally, the approach aims to appease the tech sector by deferring to online platforms and affording significant leeway in identifying the systemic risks that most affect their services and users, and selecting the best options to mitigate them. But how this negotiat-

29 EDRI, *Public Statement: ON NEW CRISIS RESPONSE MECHANISM AND OTHER LAST MINUTE ADDITIONS TO THE DSA* (April 12, 2022) <https://edri.org/wp-content/uploads/2022/04/EDRI-statement-on-CRM.pdf>.

30 DSA, *supra*, Article 36(3).

31 *Id.*, Article 37.

32 *Id.*, Article 37(3) and (4). The audit opinion must indicate whether it is “positive,” “positive with comments” or “negative” (per Article 37(4)(g)).

33 *Id.*, Article 37(4)(h).

34 *Id.*, Article 37(2).

35 *Id.*, Article 37(6).

36 *Id.*

37 See, for instance, the Santa Clara Principles on Transparency and Accountability in Content Moderation, <https://www.santaclaraprinciples.org/>; Rikke Frank Jørgensen (ed.), *HUMAN RIGHTS IN THE AGE OF PLATFORMS* (2019).

ed compromise will work in practice remains a significant question mark, particularly in responding to the proliferation of so-called “lawful but awful” content like disinformation. The next section offers some broader context about how the DSA came to address disinformation at all and outlines a few of the lingering questions that remain in respect of implementing and enforcing the risk-based approach as against disinformation.

03

IMPLEMENTING & ENFORCING THE RISK-BASED APPROACH VIZ. DISINFORMATION

The DSA’s approach to disinformation can be described as ambiguous, uneasy or disjointed – terms that legislative drafters should seek to avoid. Whatever qualifier one chooses, the upshot is that VLOPs’ internal compliance and human rights teams are left in the unenviable position of having to make sense of these newly-imposed, but imprecisely drafted, requirements in rather short order.

For starters, the term “disinformation” is used, but nowhere defined, in the Regulation. In light of the variation in definitions – within and beyond the EU – this seems a glaring oversight (at best) or an intentional omission (at worst).³⁸ In either case, it leaves online platforms in the unenviable position of having to sort it out for themselves, which may result in inconsistent approaches between platforms, and over-regulation of con-

tent, with all of the corresponding human rights issues that entails.³⁹ In addition, each of the thirteen references to “disinformation” are found in the DSA’s preambular recitals, rather than its substantive provisions setting out the risk-based approach, and many are sandwiched between the companion focuses of “illegal content” (which is defined) and “other societal risks” (which appears to be a catch-all for the negative impacts of the online ecosystem in the offline realm).⁴⁰

Of course, the DSA is but one piece of a broader and complex regulatory and policy landscape governing disinformation within the EU. Though the DSA’s stated objective refers to the proliferation of disinformation, the Regulation is not primarily concerned with disinformation: it operates in parallel with other (more targeted) efforts to combat disinformation, including co-regulatory efforts like the Strengthened Code of Practice on Disinformation 2022, which was negotiated alongside the DSA and adopted earlier this year.⁴¹ Whether the EU’s intention was to take a soft-touch with the DSA to allow the 2022 Strengthened Code of Practice to do the heavy lifting in respect of disinformation remains unclear. However, the resulting “piecemeal” approach to disinformation has been the subject of criticism,⁴² and its omission from the “systemic risks” identified in Article 34 leaves lingering uncertainty about whether and to what extent the DSA enables or requires VLOPs to address its spread on their platforms, separate and apart from any obligations they have agreed to under the 2022 Strengthened Code of Practice on Disinformation.⁴³

The disjointed approach to disinformation – perhaps best illustrated by the preamble’s frequent references to the problem and the total exclusion of the concept from the DSA’s substantive provisions – may in fact be a by-product of the hard-fought drafting and negotiation processes within the EU. Indeed, the question of whether disinformation ought to be addressed by the DSA at all was a fundamental issue throughout the negotiations. The

38 For the definitional dilemmas, see McGonagle & Pentney, *supra*, 44-47; Ronan Ó Fathaigh, Natali Helberger & Naomi Appelman, *The Perils of Legally Defining Disinformation*, 10 *Internet Pol. Rev.* 4, 1-25 (2022).

39 See generally Jørgensen (2019), *supra*; Jillian C. York, *SILICON VALUES: THE FUTURE OF FREE SPEECH UNDER SURVEILLANCE CAPITALISM* (2021).

40 See e.g. DSA, *supra*, Recitals (2) and (9). Recital 84, by contrast, refers to disinformation within the broader category of “misleading or deceptive content.” Tambini has characterized the DSA as a “co-regulatory backstop” for disinformation: Damien Tambini, *Media policy in 2021: As the EU takes on the tech giants, will the UK?* LONDON SCHOOL OF ECONOMICS, Jan. 12, 2021 <https://blogs.lse.ac.uk/media-else/2021/01/12/media-policy-in-2021-as-the-eu-takes-on-the-tech-giants-will-the-uk/>.

41 Strengthened Code of Practice on Disinformation (June 2022), <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation>.

42 Ethan Shattock, *Self-regulation 2:0? A critical reflection of the European fight against disinformation* (Harvard Kennedy School Misinformation Review, May 31, 2021) <https://misinforeview.hks.harvard.edu/article/self-regulation-20-a-critical-reflection-of-the-european-fight-against-disinformation/>.

43 For instance, signatories agreed to take action in “demonetising the dissemination of disinformation; ensuring the transparency of political advertising; empowering users; enhancing the cooperation with fact-checkers; and providing researchers with better access to data.” (2022 Strengthened Code of Practice, *supra*).

Committee on Civil Liberties, Justice and Home Affairs of the European Parliament (“LIBE Committee”) thought not: its Draft Opinion, released in May 2021, put forward a number of amendments, most crucially for present purposes the deletion of the provisions setting out the risk-based due diligence approach (discussed above).⁴⁴ The LIBE Committee justified these amendments on the basis that they were necessary to protect freedom of expression and to ensure the DSA was tailored to address the dissemination of *illegal* rather than *harmful* content.⁴⁵ The LIBE Committee expressed concern that the requirements in Article 26 (setting out the risk-based approach) went “far beyond illegal content where mere vaguely described allegedly “negative effects” are concerned.”⁴⁶ Similar concerns were raised regarding the independent audit requirements set out in Article 28.⁴⁷ The LIBE Committee’s suggested amendments illustrate the disconnect between the broad aims sought to be achieved by the drafters, and the more circumscribed scope preferred by the LIBE Committee, which would have effectively removed from the DSA’s purview “lawful but awful” speech, such as disinformation.

Where, then, does that leave VLOPs when it comes to identifying and mitigating the risks posed by disinformation? Several points appear (relatively) clear even at this early stage. First, the DSA is focused on particular *contexts* rather than specific *content*: the proliferation of disinformation that has actual or foreseeable negative effects on civic discourse, electoral processes, public security or the protection of public health must be included in VLOPs’ risk assessments and mitigated accordingly. As a threshold, this at least appears straightforward. However, from there, issues arise: how can one establish that particular (knowingly false and intentionally shared) content has had actual negative effects on public health or civic discourse? What level of causation is necessary, or sufficient, for VLOPs to take action? What level of foreseeability is required in order to identify, assess and mitigate a systemic risk posed by disinformation in relation to electoral processes? Is the proliferation of disinformation in previous elections sufficient to foresee a similar risk arising in future? And even where such a risk has been identified in

the risk assessment, how can it be mitigated in a manner that accords sufficient protection for political speech or debates of questions of public interest, for which few restrictions are permitted?⁴⁸ More broadly, will the heightened due diligence framework have any (micro) effect on specific disinformation that is shared on the platforms, or will it simply result in broader design and “system” changes on a macro level, for instance changes to algorithmic content moderation at scale?

Of course, the DSA is but one piece of a broader and complex regulatory and policy landscape governing disinformation within the EU

Finally, and most fundamentally, a large question remains

about whether the deference afforded to VLOPs in identifying, analysing, assessing and mitigating systemic risks stemming from the design or functioning of their service is a gamble that will pay off. Elon Musk’s Twitter acquisition, and subsequent firing of the entire human rights team, casts this in stark relief, but the problem goes deeper still. Facebook, Twitter and Google (YouTube) are based in the US, with a free speech tradition that diverges significantly from that of the EU.⁴⁹ Leaving it to the likes of Elon Musk and Mark Zuckerberg (or their chosen executives) to not only balance competing rights and interests, but to *decide what to weigh on the scales*, may prove an unwise choice. It may also severely limit the potential of the DSA to achieve its stated objective of ensuring a safe, predictable and trusted online environment. Just how freely the bird will fly in Europe – and how far the EU succeeds in clipping VLOPs’ wings – remains to be seen. ■

44 Committee on Civil Liberties, Justice and Home Affairs for the Committee on the Internal Market and Consumer Protection on the proposal for a regulation of the European Parliament and of the Council Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC (COM(2020)0825) (May 19, 2021) Amendments 21-24, 28, 29, 91-93, https://www.europarl.europa.eu/doceo/document/LIBE-PA-692898_EN.pdf.

45 *Id.* Amendment 91, “Justification” p. 64/84.

46 *Id.* pp. 64-65/84.

47 *Id.* Amendment 102, pp. 69-70/84.

48 *Castells v. Spain*, App no 11798/85 (Chamber, April 23, 1992) § 43; *Wingrove v. United Kingdom*, App no 17419/90 (Chamber, November 25, 1996) § 58.

49 See e.g. Jared Schroeder, *Meet the EU Law That Could Reshape Online Speech in the U.S.*, SLATE, Oct. 27, 2022 <https://slate.com/technology/2022/10/digital-services-act-european-union-content-moderation.html>; Mark Scott, *Musk vs. Europe: The upcoming battle over free speech*, POLITICO, April 26, 2022 <https://www.politico.eu/article/elon-musk-europe-online-content-free-speech/>.

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](https://www.competitionpolicyinternational.com) today to see our available plans and join CPI's global community of antitrust experts.

