Journal of Media Law

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rjml20

The Digital Services Act's red line: what the Commission can and cannot do about disinformation

Martin Husovec

To cite this article: Martin Husovec (07 Jul 2024): The Digital Services Act's red line: what the Commission can and cannot do about disinformation, Journal of Media Law, DOI: 10.1080/17577632.2024.2362483

To link to this article: https://doi.org/10.1080/17577632.2024.2362483

JOURNAL OF MEDIA LAW
https://doi.org/10.1080/17577632.2024.2362483

Routledge
Taylor & Francis Group

The Digital Services Act's red line: what the Commission can and cannot do about disinformation

Martin Husovec

London School of Economics and Political Science (LSE), London, UK

ABSTRACT

The Digital Services Act (DSA) creates a system of general risk management that is composed of two main obligations: risk assessment (Article 34), and risk mitigation (Article 35). The obligations are mandatory for very large online platforms and search engines (VLOPs/VLOSEs). The adoption of the risk-based approach to digital services make the law more future-proof. But inevitably it also makes the law very vague. This vagueness of the statutory language causes some to suggest that the European Commission will inevitably become the proverbial Ministry of Truth when tackling disinformation. This article argues that upon closer reading of the DSA, and its constitutional context, the worries that the Commission inevitably becomes a Ministry of Truth are misplaced. Suppressing incorrect or misleading lawful information is not the goal of the DSA. That is not to say that the DSA cannot be abused. But the law is not pre-programmed to do so.

KEYWORDS Risk mitigation; freedom of expression; DSA; Ministry of Truth; lawful but harmful speech

Introduction

The Digital Services Act[1] creates a system of general risk management that is composed of two main obligations: risk assessment (Article 34), and risk mitigation (Article 35). The obligations are mandatory for very large online platforms and search engines (VLOPs/VLOSEs). The periodical risk management exercise is overseen by the European Commission as the exclusive enforcer. It is aided by the official DSA Codes of Conduct that help to flesh out indicators, best practices, and industry-wide consensus. The adoption of the risk-based approach to digital services tries to make the law more future-proof. But inevitably it also makes the law very vague.

CONTACT Martin Husovec M.Husovec@lse.ac.uk

[1]Regulation (EU) 2022/2065 of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC [2022] OJ L277/1 (Digital Services Act), pp. 1–102.

The DSA says very little about the limits of the exercise. The scope of the risk assessment is extremely broad – touching upon everything we cherish. It covers most of the risks from legal and illegal content (or behaviour) that can transpire on digital services. As emphasised by Recital 84, VLOPs/VLOSEs 'should also focus on the information, which is not illegal, but contributes to the systemic risks identified in this Regulation' which includes 'misleading or deceptive content, including disinformation'. In addition, the measures that such providers are expected to take are explained only through examples (Article 35(1)) that mostly demand the adjustment of systems or processes but also of terms and conditions. This vagueness of the statutory language causes some[2] to suggest that the European Commission will inevitably become the proverbial Ministry of Truth when tackling disinformation.

This article argues that upon closer reading of the DSA, and its constitutional context, the worries that the Commission inevitably becomes a Ministry of Truth are misplaced. Suppressing incorrect or misleading lawful information is not the goal of the DSA.[3] That is not to say that the D SA cannot be abused. But the law is not pre-programmed to do so.

In the coming months and years, many will pressure the Commission to act against all sorts of social problems, including disinformation. However, even if these calls are justified by strong evidence, the Commission must stick to one important red line – it cannot invent new binding content rules. That is, it cannot tell providers what lawful expressions they must prohibit or suppress on their services. This still leaves room for many interventions against lawful disinformation whose goal is to improve the resilience of individuals against manipulation. As long as the Commission does not cross this red line, the arguments that it is becoming a Ministry of Truth are misplaced.

Disinformation as a signal of society's ills

Technology and public opinion are inextricably linked. It was the invention of the printing press that enabled the emergence of 'public opinion' by improving the ability of strangers to exchange texts and discuss ideas without being present.[4] It inevitably changed relationships between people and their rulers, and enabled democracy, a system of 'government by

[2]See Laurie Wastell, 'The EU's Orwellian Internet Censorship Regime' *The European Conservative* (Budapest, 24 August 2023) <https://europeanconservative.com/articles/commentary/the-eus-orwellian-internet-censorship-regime/> accessed 14 December 2023.

[3]Many national laws outlaw disinformation under some circumstances. See Ronan Ó Fathaigh, Natali Helberger and Naomi Appelman, 'The perils of legally defining disinformation' [2021] 10(4) Internet Policy Review https://policyreview.info/articles/analysis/perils-legally-defining-disinformation accessed 14 December 2023. I leave aside the problem of compliance of some of these laws with the European Convention on Human Rights.

[4]Robert Post, 'The Internet, Democracy and Misinformation' in Ronald Krotoszynski, Charlotte Garden and András Koltay (eds), *Disinformation, Misinformation and Democracy* (forthcoming, Cambridge

public opinion'.[5] As noted by Post, therefore, 'all modern democracies must allow for the free formation of public opinion'.[6] The highest form of human knowledge, science, equally relies on 'its ability to self-correct as new evidence is established'.[7] It constantly 'operates on the "edge of error"'.[8]

There is little doubt that the proliferation of disinformation in the current century is enabled by the human design of the newly invented digital ecosystem. The digital public sphere weakened the role of editors, such as newspapers or other media, who used to act as gatekeepers to information flows.[9] The ability of old-school editors to act as 'epistemological authorities' has been significantly undermined. That weakening of editors also brought many undisputed benefits to people, the public, and the health of democracy, but it introduced new challenges. Humanity undoubtedly needs to invent new ways to determine what and whom to trust.

The proliferation of disinformation potentially challenges society's ability to agree upon facts on which individuals base their views. Those views co-determine the public opinion by which we are then in turn ruled. As noted by Laufer and Nissenbaum, '[t]he endpoint of this downward trend is that societies with weakened, fragmented epistemic processes might lose the capacity to distinguish between reliable reporting and disinformation and fail to find common ground among believers of opposing facts – a modern-day Tower of Babel'.[10] Thus, regardless of whether disinformation is a sickness or only a symptom of other society's ills,[11] the fact is that its significant proliferation does not foretell anything good for society at large.

However, the term 'disinformation' covers many types of expressions whose level of risk to society can range from negligible to significant. The simple fact alone that people are patently wrong in their beliefs does not give the state sufficient ground for suppressing their expressions. The human rights law conditions interventions by the state upon proper justification. As explained by Article 19, '[t]he falsity of information is not a legitimate basis for restricting freedom of expression under international and

University Press 2024); Yale Law School, Public Law Research Paper <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4545891> accessed 19 December 2023, pp. 1, 4.

[5]Ibid, fn. 10.

[6]Ibid, p. 5.

[7]The Royal Society, *The online information environment Understanding how the internet shapes people's engagement with scientific information* (The Royal Society 2022) <https://royalsociety.org/-/media/policy/projects/online-information-environment/the-online-information-environment.pdf?la=en-GB&hash=691F34A269075C0001A0E647C503DB8F> accessed 14 December 2023, p. 92.

[8]Ibid, p. 92.

[9]Brian D. Loader, and Dan Mercea, 'Networking Democracy? Social Media Innovations and Participatory Politics' (2011) 14 Information, Communication & Society 757.

[10]Benjamin Laufer and Helen Nissenbaum, 'Algorithmic Displacement of Social Trust' (Knight First Amendment Institute 2023) <https://knightcolumbia.org/content/algorithmic-displacement-of-social-trust> accessed 19 December 2023.

[11]For the discussion, see Sascha Altay, Manon Berriche and Alberto Acerbi, 'Misinformation on Misinformation: Conceptual and Methodological Challenges' (2023) 9(1) Social Media + Society <https://journals.sagepub.com/doi/full/10.1177/20563051221150412> accessed 6 February 2024.

regional human rights law'.[12] Disinformation can be war propaganda that is illegal under international law[13] or someone's belief that the Earth is flat. It can include politicians questioning, without justification, the validity of elections that they lost, or someone promoting the health benefits of late-night eating of crisps relying on questionable studies.

The lack of internal differentiation is the key reason why the concept of disinformation, even if defined as 'false or misleading content that is spread with an intention to deceive or secure economic or political gain',[14] without any further qualification, is hardly useful for lawyers trying to design more restrictive policy responses. While social scientists might derive insights about trust and institutions from the proliferation of disinformation, lawyers trained to endlessly balance the interests and think of the worst possible scenarios must inevitably struggle to use 'disinformation' as an operational concept.

This starting point is crucial also for the EU's Digital Services Act. Disinformation is indeed mentioned many times by the DSA in recitals.[15] It constitutes one of the risks to which the providers must pay attention. While the Code of Practice on Disinformation is not yet a DSA official Code of Conduct,[16] it can become one in the future. But even if it does, the key question will *not* be what is required by the Codes of Conduct. The key question will be: what interventions are required by Articles 34 and 35 DSA?

Mandated risk mitigation is narrower than risk assessment

The Digital Services Act in Article 34 obliges very large providers of online platforms and search engines to annually assess risks stemming from the design, functioning and use of their services. These risk assessments must be informed by research, views of individuals and civil society. The assessments are reviewed by auditors and regulators who oversee the entire industry. The scope of the risk assessments is incredibly broad. It covers risks posed by illegal content, or to fundamental rights, civic discourse, electoral processes, public security, or people's physical and mental well-being.

Article 35 then asks companies to act upon these risk assessments. They are asked to 'put in place reasonable, proportionate and effective mitigation

[12]Article 19, 'Response to the consultations of the UN Special Rapporteur on Freedom of Expression on her report on disinformation' (Article 19 2021) <https://www.article19.org/wp-content/uploads/2021/02/SR-report-submission-on-disinformation-ARTICLE-19.pdf> accessed 10 January 2024, p. 4.

[13]International Covenant on Civil and Political Rights, art 20(1).

[14]Commission, 'Communication from the Commission to the European Parliament, the Council, and Social Committee and the Committee of the Regions on the European democracy action plan' COM (2020) 790 final, s 4.

[15]Digital Services Act, recitals (2), (9), (69), (83), (84), (88), (95), (104), (106), (108).

[16]European Commission, 'The Strengthened Code of Practice on Disinformation 2022' (European Commission 2022) <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation> accessed 19 December 2023.

measures, tailored to the specific systemic risks identified pursuant to Article 34, with particular consideration to the impacts of such measures on fundamental rights' (Article 35(1)). However, the mandated interventions are left to the double discretion – firstly that of the companies, and secondly that of the regulators. The interventions that can be imposed are also circumscribed by the DSA itself, such as limiting them by fundamental rights, including the principle of legality, and prohibition of general monitoring.

Two recent studies related to disinformation show the pitfalls of using broad analytical insights for DSA enforcement. The TrustLab 2023 study measured the prevalence of mis/disinformation and sources of disinformation in selected EU states.[17] The Reset 2023 study measured the prevalence of pro-Russian disinformation and mapped the behaviour of its actors. Reset went as far as to conclude that 'Article 35 standards of effective risk mitigation were not met in the case of Kremlin disinformation campaigns'.[18] Both studies try to come up with benchmarks that could then serve as a basis for measuring compliance with the obligation to mitigate risks under the DSA. The envisaged interventions range from empowerment through labelling, and fact-checking, to demonetisation, and possibly also demotions or removal of content.[19]

Such broadly designed studies looking at disinformation, its proliferation, potential causes, or risk factors can be helpful in many respects. However, we need to be careful about the *legal* implications that we draw from such observations.

If the studies serve to teach us about media literacy problems, the need for more user agency, or fact-checking as types of interventions under the DSA, the rigorous legal classification of the underlying expressions is less important. However, to substantiate interventions that are content-specific, i.e. where the meaning of information (e.g. the existence of 'pro-Kremlin view') triggers the intervention, the Commission cannot be equally lax in reliance on definitions and methodology. To justify suppressive interventions, such as demonetisation, deamplification, or removal, distinctions

[17]TrustLab, 'Code of Practice on Disinformation: A Comparative Analysis of the Prevalence and Sources of Disinformation across Major Social Media Platforms in Poland, Slovakia, and Spain' (TrustLab,Transparency Centre 2023) <https://disinfocode.eu/wp-content/uploads/2023/09/code-of-practice-on-disinformation-september-22-2023.pdf> accessed 19 December 2023.

[18]European Commission, *Digital Services Act: Application of the risk management framework to Russian disinformation campaigns* (Publications Office of the European Union 2023) <https://op.europa.eu/en/publication-detail/-/publication/c1d645d0-42f5-11ee-a8b8-01aa75ed71a1/language-en> accessed 19 December 2023.

[19]The two studies argue in this regard the following: "Note that platform actions for mis/disinformation will not necessarily be content removal but could include demonetization, applying warning labels or adding pointers to fact-check articles." TrustLab (n 17), p. 58; "While it is important to document the ecosystem-level effects of platform policies, we do not mean to suggest that companies necessarily should have imposed bans or demotions on all Kremlin-aligned accounts. However, in the context of systemic risk mitigation – as defined under Article 35 of the DSA – it is clear that mitigation requirements are not defined by the actor but rather by the severity of risk" European Commission (n 18), p. 46.

between lawful and unlawful disinformation must be firmly part of the methodology. Otherwise, there is a risk of imposing content-specific interventions on lawful modes of expression.

Let me explain this point.

While Article 34 DSA imposes obligations of merely a procedural nature – what steps are to be followed when identifying risks, Article 35 undoubtedly implies more substantive obligations. However, the scopes of the two obligations are not the same. The companies do not have to act upon all risks. And when they do act, they are not required to act against all risks in the same way. Unlawful disinformation (e.g. war propaganda) is likely to justify more stringent treatment than lawful disinformation (e.g. flat earthers) already because the legislature said one is unlawful while the other is not. Thus, while disinformation as such can be a relevant source of risk that companies must periodically assess, this says nothing about what sort of actions can be substantively required by regulators supervising compliance with Article 35 DSA. What exactly can be imposed on companies depends on the powers that the European Commission can derive from Article 35 DSA. And those powers must differ for risks posed by lawful and unlawful disinformation.

The powers of the European Commission

The European Commission is heavily constrained by the principle of legality (Article 52 of the EU Charter), and the absence of formal content rule-making powers under the DSA, when enforcing the mandatory risk mitigation measures.

The DSA's design has an implicit division of powers: parliaments make content rules and digital regulators only give the effect to them. This is why national and EU law are the sources of illegality, and why some of the DSA provisions explicitly only apply to illegal content (e.g. Article 16). The idea that the Commission could simply make on-the-go new content rules that would be specifically applicable to very large digital services has no support in the text of the Regulation.

This division of powers has not only consequences for the separation of the executive from the legislative authority but also for the vertical division of powers between the Union and its Member States. If the European Commission were allowed to ad hoc regulate the illegality of content, it would encroach on the competencies that the Member States did not ask to be delegated through the Regulation. Thus, the attempt by the Commission to become a surrogate legislature would lead to an excess of its powers in two ways: an overreach of the executive, and an abuse of competence.

To be clear, to maintain this division it is not enough to simply engage in the proportionality exercise because the key flaw is the lack of legality and not

the compliance with the outer limits of the fundamental rights. Thus, even relying on the proportionality principle in the human rights framework would not justify the action of the Commission to create de facto content rules.[20] Restricting expressions proportionately does not make up for the flaw that the action is taken by the institution not empowered to restrict it in the first place.

Having said that, it is easier to say this in the abstract than it is to demarcate the line that the Commission cannot cross. The reason is that *any* risk mitigation restriction imposed on content that is lawful, including fact-checking, inevitably means some type of restriction on the freedom of expression of someone. Thus, there seems to be a tension between the two starting points. On one hand, the DSA intends to entrust the European Commission to act against a wide range of risks as they dynamically unfold in society, including lawful disinformation. At the same time, the principle of legality says that Article 35 cannot amount to granting the Commission the power to be a surrogate legislature of content on VLOPs and VLOSEs. How to reconcile these two positions?

In my upcoming monograph,[21] I argue that one way to draw the boundary between overreach and justified supervision is to confine the exercise of authority to two conceptually distinct situations: risks posed by solely illegal expressions, and other risks, which inevitably means risks posed also by legal expressions. When risks are posed by illegal content, the Commission can demand interventions that are content-specific. As long as its requests comply with the fundamental rights, and the prohibition of general monitoring (Article 8 DSA), it can try to quantify compliance, including by aggregate quotas, targeted de-amplification, removals, etc. In this case, the Commission is only using the DSA as a tool to suppress the distribution of content that was determined illegal by parliament(s).

However, when the targeted class of expressions is not solely illegal, such as in the case of many types of disinformation that are largely lawful, the Commission must shy away from demanding any content-specific restrictions. It either isolates the illegal disinformation (e.g. foreign election interference, or war propaganda) and acts more harshly against it, or remains *content-neutral* if it is unable to do so. This still empowers the Commission to act against disinformation in general, however, in ways that mostly involve empowerment of users, or redesign expectations that apply to services in general, such as circuit breakers, limits on authentic behaviour or super-users.

[20]European Commission (n 18), p. 63, for instance, relies on the UN Rabat Plan to scope of the risks.
[21]Martin Husovec, *Principles of the Digital Services Act* (forthcoming, Oxford University Press 2024), ch. 13, 15.

The same limit applies to the obligations that are imposed through a crisis response mechanism. The DSA does not say that a crisis justifies the Commission overriding the principle of the separation of powers. The mechanism envisaged in Article 37 provides no special legal basis. It does *not* vest the executive with deeper powers of any kind. Recital 91 clarifies that the added value is in allowing the executive to move faster by breaking the annual cycle and focusing the attention of VLOPs and VLOSEs on a specific ongoing crisis. These 'additional' measures are thus not extraordinary in substance, but only in speed.

This red line should also have consequences for the drafting of the DSA Codes of Conduct. Even though these Codes are not binding, they have some legal effects. It is questionable whether they should include substantive commitments that the Commission cannot enforce based on Article 35 DSA. The Codes of Conduct expand and detail the scope of periodical risk assessment (Article 37(1)(b)), but their value is more than simply procedural. Several DSA provisions suggest certain spill-over effects into substantive obligations (Articles 45(3), 75(2), 75(3)). Because such Codes build an evidence base for the follow-up compliance with Articles 34–35 DSA, it is highly questionable whether they should include content-specific KPIs linked to lawful content at all (e.g. demonetise all 'pro-Kremlin disinformation'), if such commitments are not enforceable by Article 35 DSA.

Inevitably, the Commission will stumble upon cases where content-neutral interventions are exhausted and fail. That in itself does not mean that the DSA fails. The DSA is largely a procedural tool that creates a lot of new evidence about what is going on in the digital ecosystem. If content-neutral interventions, such as user empowerment, cannot solve risks posed by lawful content, the DSA delivers tools to study such risks that can inform the legislatures who need to take the lead. Through the DSA, parliaments can better learn and improve the content rules too. However, it is them who need to take the action. In liberal democracies, parliaments have the legitimacy to make content rules and are accountable to the judiciary and the public.

To be clear, the red line suggested above only applies to the European Commission as the sole enforcer of Article 35 DSA. It says nothing about what tech companies can do. Since the companies have the freedom to design their policies to fit their business, they can naturally exceed the obligations that the European Commission can impose on them. Thus, for instance, even if the Commission might be unable to limit some type of disinformation, they can decide to contractually restrict it anyway. As a result, companies do not have to make as strict distinctions between legal and illegal content because they can redefine it contractually for their purposes. The key limit of their rule-making powers is Article 14(4) DSA and risk mitigation strategies that devolve some decision-making to individuals and groups

(e.g. an obligation to give choice in recommender systems or allow user-level moderation).

Conclusion

Whenever the Commission tackles problems that are posed also by *lawful* expressions, the mitigation measures that it requires from companies must remain strictly content-neutral. Any attempt by the Commission to prescribe content-specific measures for legal content, such as forcing companies to ban some specific lawful content in terms and conditions, would mean that the Commission assumes the role of a surrogate legislature regarding content. The DSA offers no empowerment for such formal rulemaking. Doing so would mean that the Commission oversteps its competencies. It crosses a red line.

We should be reminded of Article 1(1) of the DSA which states that the main goal of the law is to create 'a safe, predictable and trusted online environment'. In policy debates, there is a tendency to emphasise safety. And often it is for a good reason. Safety from illegal content fosters trust in the ecosystem and thus improves the freedom of individuals to express themselves. For instance, citizens who are safe from harassment and hate speech are more likely to express their views publicly.

But safety can also undermine trust. Let me illustrate this with the example of suicide reporting. Social science research shows that sensationalist coverage of suicides can lead to copycat behaviour among vulnerable individuals.[22] Thus, in many countries, the media are encouraged to follow guidance about how to report on the suicides, which can range from not mentioning them at all to not reporting on methods and avoiding any sensationalism. There is no doubt that such interventions can be justified by the safety of vulnerable individuals. However, as countries that have tried to legislate on the issue have learned,[23] overbroad rules can undermine trust because individuals feel under-informed. Such suppression can undermine trust.

Thus, it is not too difficult to imagine that even well-intentioned and evidence-based limits on speech can instil distrust in people's minds because they feel that they are not being told the whole truth. For this reason, it is important to balance safety and trust. Often this means not abandoning the effort to regulate such a situation but doing so very carefully, and paying attention to who has legitimacy to undertake what intervention.

[22] See for instance, Thomas Niederkrotenthaler and others, 'Association between suicide reporting in the media and suicide: systematic review and meta-analysis' [2020] 368 British Medical Journal <https://www.bmj.com/content/368/bmj.m575.full> accessed 19 December 2023.

[23] Coroners Act 2006 (New Zealand), s 71. For analysis, see James Hollings, 'Reporting suicide in New Zealand: Time to end censorship' (2023) 19 Pacific Journalism Review 136.

The superior goal of the DSA enforcement should be to build people's resilience when they encounter disinformation, and not to entirely prevent them from encountering it. There are many content-neutral interventions, such as media literacy tools, user empowerment tools, labelling or fact-checking, that can help in that regard.

Let me end with a speech given by the Commission's Vice-President Jourová on the occasion of 2022 Václav Havel European Dialogues in Prague shortly before the DSA's adoption:[24]

> But we also see that digitalisation, the social media, has been massively used to spread and amplify disinformation. It is increasingly difficult to see who is saying what and why. Yet, new technologies should be tools for emancipation, not for manipulation. When I say this, some people argue, or rather shout: "Jourová – you want to censor the Internet!" So, let me repeat: Freedom of speech is the most cherished value of democracy. I said many times: "I don't want to see a Ministry of Truth". To bring in Vaclav Havel again: "Follow the man who seeks the truth; run from the man who has found it". This is why our actions don't focus on assessing the content.

If the Commission keeps Havel's message in mind when enforcing the DSA, Europeans will demonstrate to the world that a risk-based approach can be applied to digital services in ways that uphold the principles of liberal democracy.[25] The DSA is pre-programmed to help people find the truth rather than find it for them. But it is the actions, not words, that will determine what the law is.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributor

Martin Husovec is an Associate Professor of Law at The London School of Economics and Political Science (LSE). His scholarship deals with questions of innovation policy and digital liberties, in particular, regulation of intellectual property and freedom of expression.

[24] See European Commission, 'Speech by Vice-President Jourová on the occasion of 2022 Václav Havel European Dialogues' (European Commission 10 May 2022) <https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_22_2983> accessed 19 December 2023.

[25] More on this, Martin Husovec, 'Rising Above Liability: The Digital Services Act as a Blueprint for the Second Generation of Global Internet Rules' (2023) 38 Berkeley Technology Law Journal (forthcoming).