Regulating the Cyber-Sphere: Impact and Challenges of the Digital Services Act

by Corbin Spang

KCL Policy Research Centre

Centre for European Affairs

Word Count: 1,727

May, 2025

The role and influence of internet platforms has come under increasing scrutiny in recent years. Regarding the issues of transparency, mis- and disinformation, the Digital Services Act (DSA) of the European Union is arguably the most impactful piece of regulation across the globe and has now been in full application since February 2024. This paper will evaluate the merits and demerits of the DSA, give an outlook on future digital regulation, and propose amendments to the overarching strategy of the EU in this field. The paper finds that the main areas for improvement of the DSA lie in the enabling of smooth and efficient compliance, coherent policy integration with other regulatory acts, and a clarification and operationalisation of terms, most importantly the term 'systemic risk'. Remaining challenges contingent on future developments are the advance of artificial intelligence, both in its capacity to support and undermine platform moderation, as well as tensions between the EU and the United States caused by US backlash to European regulation on American companies.

The Most Prominent Features of the DSA

- **Accountability and Liability**: The DSA implements a framework of liability for online intermediaries in the case that they fail to moderate harmful or illegal content (Turillazzi et al., 2023). Additional responsibility is assigned to platforms of a certain scale, the very large online Platforms (VLOPs), which are required to conduct additional, more sophisticated due diligence and transparency efforts to stem systemic risks (Nannini et al., 2024).
- **User's rights:**
    - **Freedom of Expression and Transparency**: In order to ensure that freedom of expression is nonetheless respected by online intermediaries and that moderation of content is restricted to a reasonable and adequate level (Tourkochoriti 2023), platforms are, under the DSA, obliged to provide statements of reason (SoRs) following the deletion or restriction of content (Trujillo, Fagni and Cresci 2025). For transparency purposes, VLOPs must upload these SoRs to a central database: the DSA Transparency Database (Kaushal et al., 2024; Trujillo, Fagni and Cresci 2025).
    - **The Dark Patterns Ban**: Expanding the Unfair Commercial Practices Directive and the General Data Protection Regulation, the DSA introduces a ban on manipulative website design (dark patterns) that specifically targets the online intermediaries (Nannini et al., 2024). Dark patterns are designs that serve

the purpose of influencing a user's behaviour, such as inducing them to share personal data.

- **Enforcement Mechanisms and Structure**: To ensure compliance, the DSA mandates a layered approach through controls on the VLOPs conducted by the European Commission (Trujillo, Fagni and Cresci 2025), while other, smaller online intermediaries operate under the oversight of national Digital Services Coordinators (Kaushal et al., 2024). These national agencies are instructed to cooperate and share data in transnational cases (DSA Articles 57, 58). VLOPs are further required to hire independent auditors to conduct annual compliance reports regarding issues such as transparency or the moderation of illegal content (Nannini et al., 2024; DSA Article 37). Summary findings of such reports must be made public. The DSA gives the European Commission the possibility of on-site inspections and allows penalties for non-compliance of up to 6% of the annual turnover of an online platform (DSA Articles 69, 74). Fines can be put into place daily to ensure quick adaptation, and, as a measure of last resort, the Commission can temporarily restrict access to the online service.

The DSA's Implementation Progress

Generally, the DSA has effectively supported the EU's digital regulation. Online intermediaries now use a unified reporting system, making data operable with much less effort. Takedown rates, i.e., the amount of content that the intermediaries identify and delete from their platforms, has risen significantly to roughly 9.5 billion submitted SoRs in 6 months prior to January 2025 (Mätzler and Jokic, 2025). As AI has grown considerably more powerful in the years since the DSA's implementation, it remains open to what extent AI is integrated into content moderation, with much of the screening still being carried out manually. The SoRs for content removal have taken a standardised form in the Transparency Database and are not written individually, leaving merely the screening itself as a task that is performed by humans. Considering the current and predicted progress of AI, as well as the fact that most VLOPs are in partnerships with many of the currently leading AI developers, it will be crucial to closely monitor the progress made in this field.

Current challenges to the DSA

Various issues have arisen and continue to restrict the DSA's effectiveness. The data-sharing by the intermediaries is limited and SoRs are occasionally incomplete - despite the high

potential fines (Trujillo, Fagni and Cresci 2025). This issue is complicated by intransparency around the algorithms' usage and training patterns, which has virtually created a black box (Kaushal et al., 2024). As these artificial intelligence systems make content moderation decisions on a massive scale (Drolsbach and Pröllochs 2023), individual case review is rendered futile.

Further, the broad economic landscape of the digital realm has been shifted through the DSA. Adherence requires extensive compliance efforts, and these resource-intensive processes have led to a strongly increased barrier to entry into the market (Turillazzi et al., 2023). This is, in return, counterproductive to the European Union's overarching goal of promoting European companies and reducing dependence on non-European actors. Platforms whose business models are based on advertisement as sole revenue source might face additional pressure should the DSA's stricter guidelines on privacy complicate the personalisation of advertisement - therefore reducing its profitability.

Lastly, there have been concerns raised about definitional issues regarding the terms disinformation, harmful content and systemic risk, potentially making compliance with the requirements of the DSA more difficult (Nannini et al., 2024), or, at the minimum, posing a loophole for online intermediaries to avoid implementing the intended meaningful changes to their platforms. Similarly, scholars suggest a clear distinction between content deemed harmful and content that is 'merely' controversial (Turillazzi et al., 2023).

Potential Future Challenges

The rapid advance of artificial intelligence has the potential to significantly enhance content moderation in line with the regulation of the DSA. Scholars highlight previous examples of its success: the videotaped terrorist attack on a mosque in Christchurch, Australia, was deleted 1.5 million times, and taken down automatically roughly 80% of the time before reaching users' feeds (Sonderby, 2019). Other experiments show that artificial intelligence can detect toxic language even more accurately than a human. However, it is frequently pointed out that artificial intelligence is still far from being able to operate such systems autonomously (Vargas Penagos, 2024). Neither has consensus been reached on AI's effectiveness in content moderation, nor does there appear to be a framework to understand whether the content moderation algorithms will outpace generative artificial intelligence used to create illegal content (Nannini et al., 2024; Fisher et al., 2024), or vice versa.

Internationally, the DSA as well as other European regulation attempts are likely to result in frictions with the United States and the US-based very large online intermediaries (Tourkochoriti 2023). This is especially true for the Trump administration, as demonstrated on the one hand by J.D. Vance's speech at the Munich Security Conference and the unwillingness of the United States to accept the recent fines of 700 million Euros imposed on Apple and Meta for violating terms of the Digital Markets Act of the European Union (Holland and Singh, 2025), but also by Meta's Mark Zuckerberg referring to the DSA as "institutionalising censorship" (Sullivan, 2025). While there is little action taken by the administration so far, there is a likelihood that this issue moves higher up on the agenda should more extensive regulation come into place, or perhaps in the scenario that such fines affect Elon Musk's X (formerly known as Twitter).

Policy recommendations

Complaints about regulation burdens are not uncommon. Given the major potential threat that unregulated online intermediaries pose to privacy, digital safety and entire democracies, ensuring that such spaces cause no harm is paramount. Concerns about heightened barriers to entry into the platform market, on the other side, must be taken very seriously, especially in the face of – from a European perspective - desirable independence from non-European actors. By itself, through scaling and network effects, the digital environment in which the intermediaries find themselves is one that tends towards a monopoly (Barwise and Watkins, 2018). While the Digital Services Act is crucial to reigning in the power of the VLOPs, it is imperative to acknowledge that overreaches of the same are merely a symptom of the underlying problem that is the dominance of these platforms in a non-competitive environment. Reevaluating this framework through case studies and analyses of compliance costs for small intermediaries ought to be conducted to ensure that the EU does not sabotage the possibility of more socially sustainable alternatives to the current VLOPs entering the market.

Artificial intelligence is sure to be one of the next large challenges to democracies around the world. Once again championing regulation, the European Union's AI Act has implemented checks on this rapidly advancing technological field. In extension to the new regulation and confronted with what is still frequently referred to as a 'black box' when talking about platform algorithms, it is crucial to ensure that independent, experienced software

researchers gain complete access to the code, usage, and training patterns of artificial intelligence as part of the annual risk assessment duties (Nannini et al., 2024).

The last issue that appears to require addressing is the one of definitions. While a concise definition of misinformation appears moderate in the challenges it poses (see Nannini et al., 2024), there seems to be considerable uncertainty about what constitutes a 'systemic risk' (Sullivan and Pielemeier, 2023), which poses a challenge to the DSA's success. It seems sensible to improve definitions to prevent loopholes and ensure compliance as intended by the online intermediaries.

Conclusion

Ensuring smooth and resource-efficient compliance, coherent policy across different regulation acts and operationalisable terms is essential for regulatory success. The Digital Services Act by the European Union has induced substantive progress, while at the same time leaving room for improvement regarding transparency and terminology. The rapid advancement of generative artificial intelligence, while certainly posing risks itself, should not detract from the existing threats posed by unregulated online platforms, and it is crucial to address flaws in policy in a timely manner - even against the backdrop of US pushback - to limit the damage done to individuals' privacy rights and democratic stability.

References

- Barwise, P. and Watkins, L. (2018) 'The Evolution of Digital Dominance: How and Why We Got to GAFA', in Moore, M., Tambini, D. and Freedman, D. (eds) *Digital Dominance: The Power of Google, Amazon, Facebook, and Apple*. New York: Oxford University Press, pp. 21–50.
- Drolsbach, C. and Pröllochs, N. (2023) Content Moderation on Social Media in the EU: Insights From the DSA Transparency Database. Unpublished manuscript.
- Fisher, S.A., Howard, J.W. and Kira, B. (2024) 'Moderating synthetic content: the challenge of generative AI', Philosophy & Technology, 37, 133. Available at: https://doi.org/10.1007/s13347-024-00818-9.
- Holland, S. and Singh, K. (2025) 'US calls EU fines on Apple, Meta "economic extortion"', *Reuters*, 23 April. Available at: https://www.reuters.com/sustainability/boards-policy-regulation/us-calls-eu-fines-apple-meta-economic-extortion-2025-04-23/
- Kaushal, R., van de Kerkhof, J., Goanta, C., Spanakis, G. and Iamnitchi, A. (2024) Automated Transparency: A Legal and Empirical Analysis of the Digital Services Act Transparency Database. In *ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT '24)*.
- Mätzler, A. and Jokic, K. (2025) 'The EU Digital Services Act: Ready to meet reporting obligations?', *IAPP*, 20 February. Available at: https://iapp.org/news/a/the-eu-digital-services-act-ready-to-meet-reporting-obligations.
- Nannini, L., Bonel, E., Bassi, D. and Maggini, M.J. (2024) Beyond phase-in: assessing impacts on disinformation of the EU Digital Services Act. *AI and Ethics*. (forthcoming).
- Sonderby, C. (2019) 'Update on New Zealand', Facebook Newsroom, 18 March. Available at: https://perma.cc/ZA85-2Y3X
- Sullivan, D. (2025) 'Systemic risk assessments hold clues for EU platform enforcement', *Lawfare*, 11 February. Available at: https://www.lawfaremedia.org/article/systemic-risk-assessments-hold-clues-for-eu-platform-enforcement.
- Sullivan, D. and Pielemeier, J. (2023) 'Unpacking "Systemic Risk" under the EU's Digital Service Act', *Tech Policy Press*, 19 July. Available at:

https://www.techpolicy.press/unpacking-systemic-risk-under-the-eus-digital-service-act/.

- Tourkochoriti, I. (2023) The Digital Services Act and the EU as the Global Regulator of the Internet. *Chicago Journal of International Law*. 24(1), 129–146.
- Trujillo, A., Fagni, T. and Cresci, S. (2025) The DSA Transparency Database: Auditing Self-reported Moderation Actions by Social Media. *Proceedings of the ACM on Human-Computer Interaction*. 9(2), Article CSCW187.
- Turillazzi, A., Taddeo, M., Floridi, L. and Casolari, F. (2023) The digital services act: an analysis of its ethical, legal, and social implications. *Law, Innovation and Technology*. 15(1), 83–106. Available at: https://doi.org/10.1080/17579961.2023.2184136
- Vargas Penagos, E. (2024) 'ChatGPT, can you solve the content moderation dilemma?', International Journal of Law and Information Technology, 32, eaae028. Available at: https://doi.org/10.1093/ijlit/eaae028