The Digital Services Act: transparency as an efficient tool to curb the spread of disinformation on online platforms?

by **Alain Strowel and Jean De Meyere***

Abstract: The Digital Services Act (DSA), which aims at the creation of a safer online environment in Europe, addresses the lack of transparency in content moderation by online platforms. Therefore, the DSA imposes several new due diligence obligations. This article explores the implications of these transparency obligations on the spread of disinformation, in particular on the Very Large Online Platforms (VLOPs) that will be subject to additional scrutiny. The article highlights the potential benefits of the new regulatory framework that enables the access of vetted researchers to platforms' data, empowers users by reducing information asymmetry and mitigates certain risks. However, questions remain regarding the information overload for the regulators and the effectiveness of the future DSA enforcement. In view of the possible enforcement issues, the article proposes to go further, for example by adding a general principle of transparency (beyond the list of due diligences obligations) and by strengthening the co-regulatory and multistakeholder model of regulation (beyond what the DSA helpfully provides).

Keywords: disinformation; DSA (Digital Services Act); online platforms; transparency

Recommended citation: Alain Strowel, Jean De Meyere, The Digital Services Act: transparency as an efficient tool to curb the spread of disinformation on online platforms, 14 (2023) JIPITEC 66 para 1.

A. Introduction

1 Today, the role of platforms has become central in our life: to book a ride or a ticket, to organize travelling and accommodation, to access news or to exchange memories or thoughts, we constantly use online platforms[1]. Yet, they are notoriously opaque, in particular when ranking and propagating content and thus deciding about what we see and read (and buy, book as travel, etc.: the list is long!). While in the US, the Biden administration has announced principles to enhance platform accountability[2], the

* Alain Strowel, Professor, UCLouvain and USL-B, attorney and Jean De Meyere, PhD student, UCLouvain.

1 It is worth noting that, in 2023, the use of the term "platform" to designate, among others, the large social networks (Facebook, Instagram, TikTok, YouTube, Twitter...) is widely accepted, while, around 2015, the existence of those pivotal intermediaries, and the use of the term, were strongly opposed (for ex. by Google) and by certain researchers (see Thierry Pénard et Winston Maxwell, Réguler les plateformes: une fausse bonne idée, in L'Opinion, 23 avril 2015, at <https://www.lopinion.fr/economie/reguler-les-plateformes-une-fausse-bonne-idee>). In this paper, we focus on the very large online platforms (see below) as defined in the 2022 Digital Services Act.

2 On September 8, 2022, the White House released a statement containing some principles on platform accountability aiming, among others, to « increase transparency about platform's algorithms and content moderation decision [...] platforms are failing to provide sufficient transparency to allow the public and researchers to understand how and why such decisions [about content display] are made, their potential effects on users, and the very real dangers these

EU has recently adopted the Digital Services Act ("DSA")[3], an important piece of hard law which, among other things, imposes new transparency obligations on platforms.

2 In this contribution, we examine whether the transparency requirements of the DSA are adequate to fight the spread of online disinformation. We thus question whether the newly adopted rules are able to usefully highlight the platforms' mechanisms and (algorithmic) decisions about content prioritization and propagation, more commonly captured under the notion of 'content moderation'. Making those mechanisms and decisions more intelligible, in particular how the business choices on the platform's design influence information sharing, should facilitate the adoption of measures against some excesses in the spread of disinformation. We conclude that most of the new provisions are geared at reinforcing the 'reporting' requirements, with the risk of 'infobesity' and, in turn, of overwhelming the regulatory authorities. Some new provisions are, however, helpful in that they open the access to the content moderation mechanisms, for example to vetted researchers, but the possibility of online platforms to still hide their decisions, or to minimize their impact, behind the claimed protection of trade secrets or other concerns (as permitted by Article 40(5) DSA[4]) does not bode well for the implementation of the new rules. In the end, the efficiency of the new legal framework will mostly depend on how the enforcement mechanisms, including the Digital Services Coordinators (in particular, in the countries where the large platforms will be located) and the Commission, will put the rules into practice, and whether sufficient resources and skilled staff will be devoted to enforcement at the EU and national levels. This is not yet clear although it will be decisive for the DSA to be able to reach its objectives and to curb disinformation (and other unwanted content and behavior) on platforms.

3 At the same time, beware: the role of public authorities should remain minimal to avoid encroaching on freedom of expression, thus the measures should be the least invasive and strictly necessary to reduce the (proved) harms linked to online disinformation. Therefore, we also plead in the conclusion for the development of 'middleware'[5], i. e. a new layer of software or content-curation services that give users more control over what they see and thus allow them to customize content moderation. To moderate the online conversation so as to improve the quality of exchanges requires all parties, the platforms of course—under the right incentives from the regulators—, but also the online users, whether speakers or receivers, to participate in this joint enterprise. The empowerment of users, through technology and other design measures, is thus a necessary complement to the regulatory measures adopted in the DSA.

4 First, we start this paper with an attempt to delineate which problematic situations are covered under the term "disinformation", and we distinguish this phenomenon from other information disorders (such as misinformation, fake news, malinformation, etc.). Three different criteria, based on their relation to truth, on the intentional element, and on the potential damage, should be used to identify disinformation cases.

5 In the second part, we briefly describe the evolving liability framework for online platforms and highlight some changes brought by the DSA. As a few online platforms concentrate a large number of Internet users, their impact on the online conversation is considerable, they are the source of the problem as well as the possible solution if they are adequately incentivized to take the right (self-regulatory) measures. In relation thereto we look into the EU Code of Practice against Disinformation, a self-regulatory instrument aimed at curbing the spread of online disinformation.

6 In the third part, we focus on the DSA and present the transparency obligations imposed in particular on a new category of online intermediaries, the Very Large Online Platforms ("VLOPs") as they are called under the DSA. (In brief, those are the online platforms having more than 45 million average monthly users in the European Union). We focus on four different types of transparency obligations:

decisions may pose. » (see <https://www.whitehouse.gov/briefing-room/statements-releases/2022/09/08/readout-of-white-house-listening-session-on-tech-platform-accountability/>)

3 Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (Text with EEA relevance) 2022 (OJ L).L277/1

4 This important Article 40, however, constitutes a major improvement over the DSA proposal whose initial Article 31 contained several loopholes.

5 Middleware has been defined in this context as "software and services that would add an editorial layer between the dominant internet platforms and internet users" (see the first Article that refers to this notion: Francis Fukuyama et alii, Middleware for Dominant Digital Platforms: Technological Solution to a Threat to Democracy, Stanford Cyber Policy Center, available, but not dated, at: <https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/cpc-middleware_ff_v2.pdf>, accessed 8 Sept. 2022; see also Daphne Keller, The Future of Platform Power: Making Middleware Work". Journal of Democracy, vol. 32, no. 3, July 2021, pp. 168-72).

transparency information-related obligations, transparency scrutiny-related obligations, reporting obligations and risk-assessment obligations. While reviewing those transparency obligations, we also look into the changes made from the initial DSA proposal of December 2020 to the regulation as adopted in 2022.

7 In the fourth and concluding part, we sketch three different paths to improve the overall framework for regulating harmful yet lawful content online: the implementation of a general transparency principle, the adoption of a co-regulatory model empowering users and third parties, such as vetted researchers and NGOs, and the creation of an independent authority in charge of regulating platforms and the conflicts arising from their use. (Indeed, we consider that the central role left to the Digital Services Coordinators constitutes the "weak link" in the new regulatory framework defined by the DSA; similarly, the role of national data protection authorities under the 2016 General Data Protection Regulation (GDPR) did not facilitate its enforcement.)

B. Disinformation: towards a definition

8 **An ancient issue.** Disinformation is not a new phenomenon. In a time of war, it takes the form of state-sponsored propaganda, as seen since the Ukraine war started[6]. Its usage can be traced as far as the battle of Actium in 31 BCE[7]—even though it is likely that disinformation was used before this. The evolution of disinformation closely follows the evolution of information itself; the more information spread, the more disinformation spread. The invention of the printing press in Europe in the 15th century and the wide development of the press during the industrial revolution allowed for a much larger dissemination of information—and disinformation—worldwide[8]. Of course, the invention of the Internet in the late 20th century caused an ever-growing dissemination of information, a phenomenon that was amplified by the emergence of the first social media platforms[9].

9 **"Fake news": too ambiguous.** The term "fake news" that was widely used by the press and the general public can cover a variety of situations, going from the honest mistake of a journalist to a campaign of invented news orchestrated by a foreign government with the goal of undermining democratic societies. It therefore appears justified to ban this term in scientific studies because it encompasses too many sorts of information disorders and speech acts (such as false statements, misdirection, biased allegations and outright propaganda) and cannot be relied on if one aims at designing effective counter-measures[10]. The weaponization of the term by various politicians, such as former US president Donald J. Trump, in order to discredit news-outlets sharing critical views, renders the term misleading[11].

10 **Constitutive elements of disinformation.** In order to correctly understand disinformation and to attempt to regulate it properly, we need a definition of disinformation. Unlike other nefarious content, such as pedo-pornography or apology for terrorism which are clearly illegal, disinformation involves what can be called "awful yet lawful" content[12]. Regulating this information disorder therefore could be incompatible with the requirements deriving from freedom of expression and of the press. The definition of disinformation at the same time must be comprehensive enough and well delineated in order to distinguish it from other disinformation

6 For a previous analysis of the Russian campaign orchestrating disinformation around the annexion of Crimea in 2014, see Sinan Aral, *The Hype Machine: How Social Media Disrupts Our Elections, Our Economy and Our Health - and How We Must Adapt* (HarperCollins Publishers Ltd 2020).

7 'Perspective | The Long History of Disinformation during War' *Washington Post* <https://www.washingtonpost.com/outlook/2022/04/28/long-history-misinformation-during-war/> accessed 26 July 2022.

8 Julie Posetti and Alice Matthews, 'A Short Guide to the History of 'fake News' and Disinformation' 20.

9 Carol A Watson, 'Information Literacy in a Fake/False News World: An Overview of the Characteristics of Fake News and its Historical Development' (2018) 46 International Journal of Legal Information 93.

10 W Lance Bennett and Steven G Livingston (eds), *The Disinformation Age: Politics, Technology, and Disruptive Communication in the United States* (Cambridge University Press 2021)., p. 193.

11 Content and Technology (European Commission) Directorate-General for Communications Networks, *A Multi-Dimensional Approach to Disinformation: Report of the Independent High Level Group on Fake News and Online Disinformation* (Publications Office of the European Union 2018) <https://data.europa.eu/doi/10.2759/739290> accessed 15 August 2022.

12 Miriam Buiten, 'Combating Disinformation and Ensuring Diversity on Online Platforms: Goals and Limits of EU Platform' (Social Science Research Network 2022) SSRN Scholarly Paper 4009079 <https://papers.ssrn.com/abstract=4009079> accessed 27 April 2022.

disorders and to avoid over-regulation of the information ecosystem.[13]

11 The assessment of disinformation must look at the nature of the content shared, at the intention or state of mind of the person circulating the content and at the effects of spreading it. First, the accuracy of the relevant information must be considered. In order to be defined as disinformation, information should be false, inaccurate or misleading[14]. But not all content lacking accuracy, or being plainly wrong, can be considered as disinformation. Second, it is important to look at the motives behind the production and distribution of the information. As the goal in the regulation of disinformation is to better protect our democracies and the public debate among citizens[15], only content that is intentionally fabricated or spread to undermine democratic values and the possibility of a reasonable debate should be qualified as disinformation. Third, disinformation supposes a will to cause public harm or to gain some advantage.[16] Quite often, the individuals who are propagating wrong information do not aim to induce harm, therefore such propagation does not involve disinformation, those persons just fall in the trap of misinformation (see below). Organizations or state-sponsored entities which disseminate false information for achieving some objectives are more likely to be involved in disinformation.

12 **Disinformation in the EU texts.** There is currently no legal definition of disinformation, and the DSA does not define what it covers—even though some of its recitals address the rise of online disinformation.[17] However, the European Action Plan for Democracy defines disinformation as: "false or misleading content that is spread with an intention to deceive or secure economic or political gain and which may cause public harm".[18] Under this definition, which we rely on in this paper, three conditions must be met for a circulating content to be considered disinformation:

- The information must be inaccurate: the truth condition;
- There must be an intent to gain economic or political gains behind the diffusion of the information: the intentionality condition;
- There must be a potential for the information to cause public harm: the public harm condition.

13 It is important to note that the third condition, the potentiality to cause public harm, is not always explicitly mentioned in the literature defining disinformation.[19] We believe the inclusion of such a condition is important as restricting lawful content without significant negative consequences on the public, for example on the cohesion of our societies, would not be proportional and therefore risks to be an unlawful restriction on freedom of expression and of the press.

14 **Disinformation v. misinformation.** Disinformation is to be distinguished from misinformation, which is defined in the European Democracy Action Plan as: "false or misleading content shared without harmful intent" but whose "effects can be still harmful".[20] With misinformation, the false/misleading content requirement and the public harm condition are met, while the condition of intent is not: the person sharing the information did not share the content with the intention to deceive or to secure economic or political gain. This is the case when a person unknowingly shares false information.

15 The remedies to misinformation partly differ from the responses to disinformation. The European Commission points out that misinformation could be more easily countered than disinformation, mostly through better communication strategies, awareness raising and increased media literacy.[21] Furthermore, overregulating speech which was not shared or produced with a malicious intent might pose an excessive risk to freedom of expression.[22] This justifies a stronger response to disinformation, in particular when orchestrated by powerful (State) actors.

13 'European Democracy Action Plan' (*European Commission - European Commission*) <https://ec.europa.eu/info/strategy/priorities-2019-2024/new-push-european-democracy/european-democracy-action-plan_en> accessed 26 July 2022.

14 Directorate-General for Communications Networks (n 11).

15 'European Democracy Action Plan' (n 13).

16 Directorate-General for Communications Networks (n 11).

17 See DSA recital 2, recital 9, recital 69, etc.

18 'European Democracy Action Plan' (n 13).

19 'Information disorder: Toward an interdisciplinary framework for research and policy making' (*Council of Europe Publishing*) <https://edoc.coe.int/fr/medias/7495-information-disorder-toward-an-interdisciplinary-framework-for-research-and-policy-making.html> accessed 16 May 2022.

20 'European Democracy Action Plan' (n 13).

21 COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS On the European democracy action plan 2020.

22 Noémi Bontridder and Yves Poullet, 'The Role of Artificial Intelligence in Disinformation' (2021) 3 Data & Policy e32.

16 Disinformation v. parody and satire. Satire or parody, if wrongly perceived and shared, without the necessary second-degree humor and understanding, could create some information disorder. In that case, while the person sharing it does not realize that the shared content—if taken at face value—is false, or at least exaggerated, the information is generally not communicated with a malicious intent nor has the potential to cause public harm.[23] However, there have been cases where parodical or satirical content were not clearly identified as such by its author, causing confusion.[24] Politicians and public figures have also been known for sharing parodical articles from websites such as The Onion or Le Gorafi, well-known parodical websites.[25] Although the line between disinformation and parody/satire is not always clear (at least for the persons ignoring the context), it is important to keep the irreverent expression immune from legal interference, thus regulating disinformation must be adequately finetuned to preserve the room of parodical speech.

17 Disinformation v. malinformation. Malinformation is "genuine information that is shared to cause harm".[26] In that case, the truth condition is respected while the intentionality condition is not. Malinformation is not illegal *per se* but could in some circumstances constitute an illegal behavior such as harassment.[27]

18 Regulations touching upon illicit disinformation. Content that commonly qualifies as disinformation can also fall under the scope of prohibitions, for example misleading advertising.[28] Another example is the negation of the Holocaust, which is illegal under the laws of certain European countries.[29] Prohibition of these forms of disinformation is usually justified because they pose a serious threat to customers or democratic societies. The DSA will help to curb the spread of those *illicit* types of content as the DSA permits a better online enforcement of the laws banning such content.

19 Currently, the day-to-day control of online disinformation remains in the hands of private, profit-oriented actors, i.e. the social media platforms such as Meta and Google.[30] Those platforms have been accused of encouraging, by their design and decisions, the rise of disinformation.[31] In part 2 below, we briefly summarize how their business models favor the rise of disinformation. This is why some specific regulatory measures should target those online platforms with regard to disinformation, and this should be distinguished from the liability rules and processual tools for reducing illicit content online.

C. The new liability framework for platforms and some self-regulatory measures to fight disinformation

20 The conditional exemptions of liability for intermediaries still in place with the DSA. The Internet we know today is much different than that of the (early) 1990s, when the Internet was still made of a large number of small communities, for instance researchers, journalists or professionals, who were accustomed to self-regulating their expression (e.g., due to the ethical rules known and shared by them, while the "netiquette" rules never achieved the same moderating effect on the social networks' most aggressive participants). The Internet was a decentralized network without powerful

23 Christine Sinclair, 'Parody: Fake News, Regeneration and Education' (2020) 2 Postdigital Science and Education 61.

24 'Bye Bye Belgium: en 2006, le docu-fiction de la RTBF créait un électrochoc' (*RTBF*) <https://www.rtbf.be/article/bye-bye-belgium-en-2006-le-docu-fiction-de-la-rtbf-creait-un-electrochoc-9479103> accessed 16 August 2022.

25 'Quand Christine Boutin cite sans sourciller le site parodique Le Gorafi' (*LEFIGARO*, 4 February 2014) <https://www.lefigaro.fr/politique/2014/02/04/01002-20140204ARTFIG00255-quand-christine-boutin-cite-sans-sourciller-le-site-parodique-le-gorafi.php> accessed 16 August 2022.

26 'Information disorder: Toward an interdisciplinary framework for research and policy making' (n 19).

27 ibid.

28 Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No 2006/2004 of the European Parliament and of the Council ('Unfair Commercial Practices Directive') (Text with EEA relevance) 2005.

29 For ex. the Loi n° 90-615 du 13 juillet 1990 tendant à réprimer tout acte raciste, antisémite ou xénophobe.

30 Daphne Keller and Paddy Leerssen, 'Facts and Where to Find Them: Empirical Research on Internet Platforms and Content Moderation' (16 December 2019) <https://papers.ssrn.com/abstract=3504930> accessed 16 August 2022.

31 Christian Stöcker, 'How Facebook and Google Accidentally Created a Perfect Ecosystem for Targeted Disinformation' in Christian Grimme and others (eds), *Disinformation in Open Online Media* (Springer International Publishing 2020).

intermediaries dealing with the content (contrary to intermediaries such as telecom operators dealing with the network infrastructure). "The Web of the 1990s could arguably be thought of as a neutral marketplace of ideas, one in which anyone with a dial-up connection and a bit of training in HTML could write online and potentially find a modest audience".[32] Of course, it does not mean that disinformation was not already present online. But the relatively small audience at the time made online disinformation a marginal issue affecting probably only the people actively looking for this type of content.[33]

21 This situation led regulators, first in the United States and then in Europe, to take measures in order to preserve the neutrality of the Internet. Webhosts could be considered neutral actors in the digital world, as they did not interfere with the content on their networks. In the US, Section 230 of the Communications Decency Act[34] (the "Safe Harbor" clause) made websites non-liable for content posted by their users.[35] Article 14 of the eCommerce Directive contains a similar liability exception.[36] Although the online world has fundamentally changed since the 1990s, this last provision has now been inserted in Article 6 DSA showing that the same regulatory approach remains in place (the other liability exemptions have also been imported in the DSA). Nevertheless the DSA also takes into account new realities and innovates[37]: there is, for instance, a new special rule (Article 6(3)) on the hosting provider liability under consumer law (in particular distance selling); also, the new Good Samaritan provision (Article 7) will clearly encourage platforms to take voluntary measures to tackle illicit content or to comply with EU or national laws (e.g., regarding some type of *illicit* disinformation) by ensuring they can benefit from the liability safe harbors despite becoming active intermediaries; more importantly maybe, the whole chapter III of the DSA creates extensive due diligence obligations, mainly transparency requirements (which we examine in part 3 below). More action from the platforms is thus not only expected, but imposed under the DSA. With the DSA, we move from a liability-focused framework (defined early by the eCommerce directive and interpreted by the CJEU case law) to a due diligence regime; under the DSA, compliance is now key, not liability.[38] This also means that the important role of the judiciary will now be complemented (or superseded potentially) by the role of "agencies/regulators" (i. e., the Digital Services Coordinators, the Board for Digital Services and/or the Commission as the three main enforcers under the DSA).

22 **Platforms and the economy of attention.** The rise of online platforms since the 2000s has radically changed the situation for which the eCommerce framework was designed. Several companies such as Meta and Google follow an advertising-based business model that requires the collection of vast amounts of data from their users to serve targeted ads.[39] The social media companies have developed strategies aiming to maximize the engagement of their users. The more and longer attention they give to the platform, the more advertising revenues the platforms generate.[40] In order to attract visitors, platforms rank and organize the presentation of the content to make it addictive. Whether it is the search results from Google Search or a Facebook newsfeed, algorithms form an essential component of the ranking and moderating mechanisms used by platforms to determine the nature and the order of content shown to a specific user.[41] In addition, studies have shown that disinformation and polarizing content attracts more attention on online

32 Bennett and Livingston (n 10)., p. 159.

33 ibid.

34 United States: Congress: House of Representatives: Office of the Law Revision Counsel, 'Protection for Private Blocking and Screening of Offensive Material. Sec. 230', *TELEGRAPHS, TELEPHONES, AND RADIOTELEGRAPHS. Title 47* (2011th edn, US Government Publishing Office 2011) <https://www.govinfo.gov/app/details/USCODE-2011-title47/USCODE-2011-title47-chap5-subchapII-partI-sec230> accessed 16 August 2022.

35 Bennett and Livingston (n 10).

36 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 ('Directive on electronic commerce') 2000.

37 Folkert Wilman, Between preservation and clarification, The evolution of the DSA's liability rules in light of the CJEU's case law, 2 Nov. 2022, available at <https://verfassungsblog.de/dsa-preservation-clarification/>).

38 See also Miriam C. Buiten, The Digital Services Act: From Intermediary Liability to Platform Regulation, 12 (2022) JIPITEC p. 361.

39 Yongrui Duan, Yao Ge and Yixuan Feng, 'Pricing and Personal Data Collection Strategies of Online Platforms in the Face of Privacy Concerns' (2022) 22 Electronic Commerce Research 539.

40 Romain Badouard, *Les nouvelles lois du web: modération et censure* (Seuil 2020).

41 Jean-Gabriel Ganascia, *Servitudes virtuelles* (Seuil 2022).

platforms, encouraging the engagement of users and advertising revenues.[42]

23 **Platforms initiatives against disinformation.** In 2018, the revelations of a Canadian whistle-blower uncovered the Cambridge Analytica scandal[43]: this data analysis company had relied on the processing of massive amounts of personal data in order to influence electors during the 2016 US elections in favor of Donald Trump and the UK Brexit referendum.[44] The use of social media platforms by the Russian Internet Research Agency, which was able to disseminate a large amount of disinformation through online platforms during the 2016 elections, also raised suspicion against the platforms' ranking algorithms.[45] Similarly, obscure websites and bloggers are using fakes to develop a narrative above the weakness of the Taiwanese democracy and the alleged desire of Taiwanese people to join China, what might be called "cognitive warfare".[46] The COVID-19 pandemic that started in 2020 and the Russian invasion of Ukraine in 2022 were also accompanied with large campaigns of disinformation[47], putting even more pressure on the social media platforms.

24 Online platforms have responded to those criticisms by putting mechanisms in place to fight disinformation.[48] For example, online platforms work together with journalistic associations to develop fact-checking initiatives[49], but it appears that such attempts to "educate" people are not well-received and could even be counterproductive.[50] We do not review those interesting, although not fully convincing, initiatives here, but it is worth mentioning another self-regulatory scheme that applies in the EU and has been promoted by the European Commission.

25 **The EU Code of Practice on Disinformation.** More serious self-regulation measures have been adopted by platforms, such as Google or Meta, having subscribed to the EU Code of Practice on Disinformation, a strengthened version of which was issued in 2022.[51] The Code contains commitments as well as specific measures, focusing on the following areas:

- Demonetization of purveyors of disinformation;
- Transparency of political advertising;
- Ensuring the integrity of services, notably by preventing the manipulation of services for spreading disinformation;
- Empowering users, researchers and the fact-checking community;
- Strengthening the monitoring, notably by the establishment of a transparency center accessible to citizens.[52]

42 'Misinformation, Disinformation, and Online Propaganda (Chapter 2) - Social Media and Democracy' <https://www.cambridge.org/core/books/social-media-and-democracy/misinformation-disinformation-and-online-propaganda/D14406A631AA181839ED896916598500> accessed 16 August 2022.

43 In the US, this lead to the Dec. 23, 2022 settlement with the FTC, Meta having agreed to pay USD 725 million to settle a longstanding class action lawsuit accusing it of allowing Cambridge Analytica and other third parties to access private user data (see <https://edition.cnn.com/2022/12/23/tech/meta-cambridge-analytica-settlement/index.html>).

44 Christopher Wylie, *Mindf*ck: Cambridge Analytica and the Plot to Break America* (First edition, Random House 2019).

45 Renee DiResta and others, 'The Tactics & Tropes of the Internet Research Agency' [2019] U.S. Senate Documents <https://digitalcommons.unl.edu/senatedocs/2>.

46 See Anne Applebaum, China's War Against Taiwan Has Already Started. How Bejing tries to make a democracy submit without putting up a fight, The Atlantic, Dec. 14, 2022, at <https://www.theatlantic.com/ideas/archive/2022/12/taiwan-china-disinformation-propaganda-russian-influence/672453/>.

47 'Disinformation: Online Platforms Continue the Code of Practice Revision in Light of the War in Ukraine and Report on First 2022 Actions to Fight COVID-19 Disinformation | Shaping Europe's Digital Future' <https://digital-strategy.ec.europa.eu/en/news/disinformation-online-platforms-continue-code-practice-revision-light-war-ukraine-and-report-first> accessed 16 August 2022.

48 Dawn Carla Nunziato, 'Misinformation Mayhem: Social Media Platforms' Efforts to Combat Medical and Political Misinformation' (2020) 19 First Amendment Law Review 32.

49 'L'AFP monte une opération mondiale de vérification des informations' (*L'AFP monte une opération mondiale de vérification des informations*) <https://www.facebook.com/journalismproject/afp-fighting-false-news-facebook?locale=fr_FR> accessed 16 August 2022. For example, platforms put specific stamps on certain content to inform their users that it does not conform to the scientific consensus (for ex. an anti-vaccination content) or that the user who posted it is related to a certain country. See also Government and State-Affiliated Media Account Labels' <https://help.twitter.com/en/rules-and-policies/state-affiliated> accessed 17 May 2022.

50 Wen-Ying Sylvia Chou, Anna Gaysynsky and Robin C Vanderpool, 'The COVID-19 Misinfodemic: Moving Beyond Fact-Checking' (2021) 48 Health Education & Behavior 9.

51 '2022 Strengthened Code of Practice on Disinformation | Shaping Europe's Digital Future' <https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation> accessed 16 August 2022.

52 ibid..

Critics have emerged regarding the Code, for example regarding the lack of details provided by the signatories in the annual reports they have to provide under the Code's commitments.[53] The strengthened version of the Code tries to further detail how platforms should implement the measures it contains. Other critics suggest that, while the Code is an appropriate tool to make online platforms more responsible regarding disinformation, it risks giving them too much power regarding the fine-tuning of the speech controls.[54] In any case, the self-regulatory nature of the Code means that there is a lack of oversight from public authorities as well as no compliance and enforcement mechanisms. Some have suggested to reinforce the Code through co-regulative measures that could allow for a better oversight and enforcement.[55] Despite their lack of teeth, the Code's provisions have become more persuasive in practice as the Commission threatens to adopt mandatory rules of hard law.

26 Lack of transparency of online platforms. Currently, platforms have to play a quasi-regulatory role as they are the one choosing which content will or will not stay on the platform and to whom it will be distributed.[56] Their decisions still lack the required transparency as they do not motivate their decisions, leaving users in the shadow. Even the initiatives proposed by the platforms to solve that issue, such as the creation of an Oversight Board by Facebook[57], raise questions of transparency and legitimacy.

27 The European Union, with the Digital Services Act, aims to better regulate online platforms, notably through the application of several transparency obligations helping regulators and researchers altogether to better understand the architecture of online platforms. We further develop those obligations in the next section.

D. The due diligence and transparency obligations of the DSA

I. Main DSA features and place of disinformation within the DSA

28 Legislative process. The European Commission unveiled the Digital Services Act ("DSA") proposal on 15 December 2020.[58] After a rather swift negotiation period, the final version of the text was voted by the European Parliament on 5 July 2022.[59] The DSA was published on 19 October 2022[60] and shall apply from 17 February 2024.[61]

29 Objective: a safer Internet. The goal of the legislation is to ensure a safe and accountable online environment.[62] The DSA aims to "fully harmonizes the rules applicable to intermediary services in the internal market with the objective to ensure a safe, predictable and trusted online environment, addressing the dissemination of illegal content online and the societal risks that the dissemination of disinformation or other content may generate, where fundamental rights enshrined in the

53 DG for Communications Networks, Content and Technology 'Study for The "Assessment of the Implementation of the Code of Practice on Disinformation" - Final Report' <https://imap-migration.org> accessed 9 January 2023.

54 The Eu Code of Practice on Disinformation and the Risk of the Privatisation of Censorship (Routledge 2020) <https://www.taylorfrancis.com/chapters/oa-edit/10.4324/9781003037385-20/eu-code-practice-disinformation-risk-privatisation-censorship-matteo-monti> accessed 9 January 2023.

55 DG for Communications Networks, Content and Technology (n 53).

56 Rotem Medzini, 'Enhanced Self-Regulation: The Case of Facebook's Content Governance' [2021] New Media & Society 1461444821989352.

57 'Oversight Board' (*Meta*) <https://about.fb.com/news/tag/oversight-board/> accessed 16 August 2022.

58 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 2020.

59 'Digital Services: Landmark Rules Adopted for a Safer, Open Online Environment | News | European Parliament' (5 July 2022) <https://www.europarl.europa.eu/news/en/press-room/20220701IPR34364/digital-services-landmark-rules-adopted-for-a-safer-open-online-environment> accessed 26 July 2022.

60 Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (Text with EEA relevance).

61 DSA, Article 93, 2.

62 'The Digital Services Act: Ensuring a Safe and Accountable Online Environment' (*European Commission - European Commission*) <https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en> accessed 26 July 2022.

Charter are effectively protected and innovation is facilitated".[63]

30 **Tiered structure.** The DSA embraces a tiered structure: the more important the role of an online intermediary is, the more obligations it is subject to.[64] Four classes are defined in the digital services act: providers of online intermediary services[65], providers of hosting services[66], online platforms[67] and very large online platforms (VLOPs).[68] The large social networks on which disinformation circulates with potential systemic effects, such as the erosion of the trust in democracy and in the institutions, are to be considered as VLOPs (see below). With regard to VLOPs (and very large online search engines or VLOSEs[69]), the DSA will enter into force four months after their designation as such by the European Commission. On 25 April 2023, the Commission designated 17 VLOPs and 2 VLOSEs. [70]

31 **Online platforms and VLOPs.** Online platforms are defined as "a provider of a hosting service which, at the request of a recipient of the service, stores and disseminates to the public information"[71], while VLOPs are "online platforms which reach a number of average monthly active recipients of the service in the Union equal to or higher than 45 million".[72] Due to the higher systemic and societal risks VLOPs pose, the DSA imposes higher transparency obligations on VLOPs as well as specific obligations related to risk management.[73]

32 **Illegal content and disinformation under the DSA.** The DSA does not bring any modification to the liability exception granted to online intermediaries by the eCommerce directive[74], but imposes strengthened due diligences obligations on intermediaries and an obligation to delete illegal content when requested by the relevant authorities.[75] Illegal content is now defined as: "any information, which, in itself or in relation to an activity, including the sale of products or the provision of services, is not in compliance with Union law or the law of any Member State, irrespective of the precise subject matter or nature of that law".[76]

33 The DSA provisions on illegal content could directly affect the fight against disinformation if the content shared would be considered illegal speech according to the Member States' legislation. While some disinformation during an election campaign is considered illicit and banned under strict conditions in France[77], most EU Member States have not legislated on this delicate issue. (In principle, free speech is highly protected during an election period, and many excessive and inaccurate allegations made by candidates and their supporters thus pass the proportionality test). Against disinformation which remains licit, despite being wrong, the content removals' obligations of the DSA do not provide for a solution.

34 Disinformation is not completely absent from the DSA. While it lacks a definition of the term, the DSA targets the disinformation phenomenon in several recitals[78] and identifies the fight against the spread

63 DSA, recital 9.

64 Alain Strowel and Laura Somaini, 'Towards a Robust Framework for Algorithmic Transparency to Tackle the Dissemination of Illegal and Harmful Content on Online Platforms' [2021] CRIDES Working Paper <https://cdn.uclouvain.be/groups/cms-editors-crides/droit-intellectuel/CRIDES_WP_2_2021_Alain%20Strowel%20and%20Laura%20Somaini.pdf> accessed 12 April 2022.

65 DSA, Article 2 (g).

66 ibid.

67 ibid., Article 2(i).

68 ibid., Article 33.

69 ibid., Article 33. When dealing with the reinforced transparency provisions, we will refer only to VLOPS, although VLOSES are also concerned – for the present contribution, the very large social platforms (one example of VLOPS) are indeed the main propagators of disinformation (and at least more than the VLOSES).

70 ibid., Article 92, see: <https://ec.europa.eu/commission/presscorner/detail/en/IP_23_2413>.

71 ibid., art 2(i).

72 ibid., art 33.

73 Strowel and Somaini (n 64).

74 Miriam Buiten, 'The Digital Services Act: From Intermediary Liability to Platform Regulation' (Social Science Research Network 2021) SSRN Scholarly Paper 3876328 <https://papers.ssrn.com/abstract=3876328> accessed 25 April 2022.

75 DSA, Article 9.

76 DSA, Article 2 (h).

77 In 2018, one year after the presidential election, France adopted a law regulating online disinformation during elections. The actual effects on this law during the 2022 French presidential campaign are yet to be studied.

78 See DSA Recital 2, Recital 9, Recital 69, etc.: the recitals mostly consider disinformation as one of the societal risk online platforms should be aware of.

of disinformation as an objective of the regulation.[79] Notably, several transparency obligations imposed on online platforms could help understand and correct the design of online platforms in a way that could curb the dissemination of disinformation. For example, the DSA requests VLOPs to "also focus on the information which is not illegal, but contributes to the systemic risks identified in this Regulation. Providers should therefore pay particular attention on how their services are used to disseminate or amplify misleading or deceptive content, including disinformation".[80]

35 We have already established that the problem of online disinformation is reinforced by the importance of online platforms in the public debate. Therefore, we will now concentrate on the transparency obligations which are specific for VLOPs, as their impact on the public conversation is considerable. As VLOPs are also bound to the obligations imposed on other online providers, we will first briefly describe those requirements.

II. Transparency and due diligence requirements applicable to all online providers

36 **Point of contact or legal representative.** Intermediaries will have to designate a single point of contact for communication with users and Member States.[81] Intermediaries not based in the EU also have to appoint a legal representative inside the Union.[82]

37 **Terms and conditions.** Article 14 of the DSA defines specific obligations regarding the terms and conditions of online intermediaries. These should include "information on any policies, procedures, measures and tools used for the purpose of content moderation, including algorithmic decision-making, and human review as well as rules of procedure of their internal complaint handling system".[83] Those should be "set out in clear, plain, intelligible, user friendly and unambiguous language, and shall be publicly available in an easily accessible and machine-readable format".[84] Furthermore, their application should respect fundamental rights of users, including freedom of expression.[85]

38 **Reporting obligations.** Finally, Article 15 of the DSA imposes reporting obligations for intermediary services providers regarding the following information:

- Orders regarding the removal of illegal content based on Article 8 of the DSA[86];
- Information regarding their moderation practices[87], provided that they engage in such activities;
- The number of complaints received through the internal complaint-handling system[88];
- Any use of AI for the purpose of content moderation.[89]

III. Transparency and due diligence requirements applicable to hosting services (including online platforms)

39 **Notice-and-action mechanisms.** Article 16 of the DSA imposes the hosting providers to put in place notice-and-action mechanisms, allowing users to notify host of illegal content. Hosting services have to allow users to easily communicate a series of information about the content, and those notices "shall be considered to give rise to actual knowledge or awareness for the purposes of Article 6 in respect of the specific item of information concerned where they allow a diligent provider of hosting services to identify the illegality of the relevant activity or information without a detailed legal examination".[90] Hosting services also have to inform the person submitting the good reception of the notice[91] and of their decision[92] and specify if this decision was made through the use of AI.[93] Article 15 also requires hosting services to issue information on those notices as part of their reporting obligation. The fact that an obligation specific to hosting services is

79 DSA, Recital 9.

80 DSA, Recital 84.

81 DSA, Article 11 and 12.

82 DSA, Article 13.

83 DSA, Article 14.

84 ibid.

85 ibid.

86 DSA, Article 15 (a).

87 DSA, Article 15 (b).

88 DSA, Article 15 (d).

89 DSA, Article 15 (e).

90 DSA, Article 16, 3.

91 DSA, Article 16, 4.

92 DSA, Article 16, 5.

93 DSA, Article 16, 6.

contained in an article that is normally relevant to all intermediaries is regretful, as including a specific article for reporting obligation specific to hosting services would have improved clarity (see below for the same comment regarding the obligations for trusted flaggers).

40 Statement of reasons. Article 17 requires the hosting services to communicate a statement of reasons to the recipients affected by the measures restricting their usage of the service, whether it is restrictions on the visibility of the content, demonetization or suspension of the services or of the user.[94] This statement of reasons shall include information related to the impact of the decision on the relevant information as well as the facts and circumstances leading to the decision.[95] Such a statement of reasons is not necessary when the removal of content stems from the order of an official authority pursuant to Article 9 of the DSA.[96]

41 Suspicion of criminal offences. Article 18 requires hosting services who are aware of any information related to a criminal offence involving a threat to the life or safety of individuals to notify the appropriate law enforcement or judicial authorities.[97]

IV. Additional transparency and due diligence requirements applicable to online platforms

42 Internal complaint-handling systems. Article 20 requires online platforms to put in place an internal complaint-handling system against the measures taken by the platform to restrict their usage of the service, whether it is restrictions on the visibility of the content, demonetization or suspension of the services or of the user.[98] Furthermore, Article 21 allows online platforms users to rely on out-of-court settlement body which have been certified by the appropriate Digital Services Coordinator.[99]

43 Trusted flaggers. Article 22 introduces the notion of trusted flaggers, a status awarded by a Digital Services Coordinator to individuals with sufficient expertise and independence from online platforms.[100] Notices sent out by those trusted flaggers within their area of expertise should be prioritized by online platforms.[101] Trusted flaggers shall issue specific reports[102], and online platforms have to include information on trusted flaggers as part of their reporting obligation under Article 15 of the DSA.[103]

44 Reporting obligations. Article 24 imposes specific reporting obligation for online platforms. Online platforms have to report on the following information:

- Disputes submitted to out-of-court dispute settlement bodies[104];
- The number of suspension of users pursuant to Article 20 of the DSA, which requires platforms to take measures against the misuse of their services[105];
- Information on the average monthly active recipients of the service in the Union[106]; and
- Decisions and statements of reasons pursuant to Article 17, while preserving their users' privacy.[107]

45 Clear marking of advertising. Article 26 requires online platforms to provide recipients with sufficient information regarding advertising, including the clear marking of commercial communication.[108] To do so, online platforms should provide recipients with the possibility to declare whether the content they provide contains commercial communication or not.[109]

46 Recommender system transparency. Finally, Article 27 requires platforms relying on a recommender system to "set out in their terms and conditions, in plain and intelligible language, the main parameters used in their recommender systems, as well as any options for the recipients of the service to modify or influence those main

94 DSA, Article 17, 1.

95 DSA, Article 17, 2.

96 DSA, Article 17, 5.

97 DSA, Article 18.

98 DSA, Article 20, 1.

99 DSA, Article 21.

100 DSA, Article 22, 2.

101 DSA, Article 22, 1.

102 DSA, Article 22, 3.

103 DSA, Article 15, 2.

104 DSA, Article 24, 1. (a)

105 DSA, Article 24, 1. (b)

106 DSA, Article 24, 2.

107 DSA, Article 24, 5.

108 DSA, Article 26, 1.

109 DSA, Article 26, 2.

parameters".[110] Where several options are available, users should have the ability to modify their preferences at any time.[111]

V. Additional transparency and due diligence requirements applicable to VLOPs

47 **Risk management.** Articles 34 and 35 impose risk management obligations on VLOPs. The DSA, considering the social impact and means of VLOPs, mandates VLOPs to assess, manage and mitigate systemic risks. Those risks stem from the very design of platforms, based on "behavioral insight and advertising-driven business models".[112] Yearly risk assessments should address risks related to online safety, the shaping of public opinion and discourse and online trade. VLOPs should also assess the impact of their content moderation, recommender and advertising systems on systemic risks including "the potentially rapid and wide dissemination of illegal content and of information contrary to their terms and conditions".[113] VLOPs should put mitigating measures in place in order to correct the risks they have assessed and some of these measures, such as discontinuing advertising revenue for specific types of content or enhancing the visibility of authoritative information sources, could benefit the fight against disinformation.[114]

48 The final version of the text further specifies the different risks that need to be considered and adds some categories, such as negative effects related to gender-based violence or the protection of public health and imposes better accountability for risk assessments as they have to be kept by VLOPs for at least 3 years.[115] Risk mitigations measures for some situations that directly relate to information disorders, such as the circulation of deep fakes, have also been included in the regulation.[116] Deepfakes (or "manipulated image, audio or video" falsely appearing authentic or truthful) should be flagged through "prominent markings" on the platforms' interfaces, and recipients should be provided with an easy tool to communicate their inauthentic character. The obligations to adapt the content moderation processes to reduce illegal (hate) speech or cyber violence could as well contribute to reduce some verbal excesses associated with disinformation.[117] These moderation measures against unlawful expressions will prompt a reduction in awful content.

49 The Commission may issue guidelines recommending best practices and possible measures, which could shed further light on the risk assessment process as well as on the mitigating measures that could be taken by platforms.[118]

50 **Crisis response mechanism.** Article 36 of the DSA gives the possibility to the Commission to impose specific measures on VLOPs at a time of crisis.[119] "A crisis shall be deemed to have occurred where extraordinary circumstances lead to a serious threat to public security or public health in the Union or in significant parts thereof".[120] In this situation VLOPs have to assess how their services can solve the crisis as well as apply specific measures to reduce such impact and to report on those to the Commission.[121]

51 **Independent audits.** Article 37 of the DSA requires platforms to perform independent audits of their services. The audit should give sufficient information and help inform, and, if necessary, suggest improvements, regarding compliance. Audits shall assess compliance with due diligence obligations imposed on the provider as well as the respect of any relevant code of conduct.[122] VLOPs may be forced to adopt mitigating measures in case the audit report is not satisfactory.[123]

52 In the adopted version of the DSA, the transparency obligations related to the audit of VLOPs have been reinforced. Limitations to the audit have been strongly limited: confidentiality should not be an obstacle to the audit itself.[124] Article 28 also details the various circumstances under which the audit

110 DSA, Article 27, 1.

111 DSA, Article 27, 3.

112 Strowel and Somaini (n 64).

113 DSA, Article 34, 2.

114 DSA, Article 35.

115 DSA, Article 34.

116 DSA, Article 35, 1, k.

117 DSA, Article 35, 1, c.

118 DSA, Article 35, 3.

119 DSA, Article 36, 1.

120 DSA, Article 36, 2.

121 DSA, Article 36, 1.

122 DSA, Article 37, 1.

123 DSA, Article 37, 6.

124 DSA, Article 37, 2.

should be performed, and the Commission also receives additional powers to further determine how such an audit should be realized.[125]

53 **Recommender systems.** Article 38 of the DSA reinforces Article 27 in relation to recommender systems.[126] In the original proposal, most of the requirements imposed on online platforms were contained in Article 38 and therefore limited to VLOPs. On top of the requirements set out in Article 27, VLOPs have to allow their users to provide at least one option not based on profiling for their recommender systems.[127]

54 **Additional online advertising transparency.** Article 39 of the DSA reinforces Article 26 on the advertising requirements for online platforms, by requiring VLOPs to put in place a public repository containing information related to the advertisement present on the platforms for at least one year after the last diffusion of the commercial communication.[128]

55 **Data access and scrutiny.** Article 40 imposes data access and scrutiny obligations on VLOPs. It requires platforms to "make data available for regulatory scrutiny and research through access rights".[129] Access to data for externals actors such as the Commission, the Digital Services Coordinators or the vetted researchers allows for a better monitoring of compliance as well as "to assess the risks and possible harms of the platforms' systems".[130] Data related to the risk assessment made by the VLOP may be shared with vetted researchers under certain conditions. The original DSA proposal contained several limitations to the sharing of their data by VLOPs, notably in relation to data privacy and the protection of trade secrets, limiting the efficiency of the scrutiny imposed on platforms—despite the fact that the EU Commission or vetted researchers could be bound by confidentiality agreements.[131]

125 DSA, Article 37, 3. and 7.

126 See supra, chapter 3, section d.

127 DSA, Article 38.

128 DSA, Article 39, 1.

129 Strowel and Somaini (n 64).

130 ibid.

131 ibid; on 25 April 2023, the Commission opened a consultation to obtain additional evidence from interested parties on the framework for vetted researchers's access to data from VLOPs/VLOSEs. see: <https://algorithmic-transparency.ec.europa.eu/news/call-evidence-delegated-regulation-data-access-provided-digital-services-act-2023-04-25_en>.

56 This section was reinforced in the adopted version of the DSA, including with Article 40(3) imposing on VLOPs to "explain the design, logic the functioning and the testing of their algorithmic systems, including their recommender systems"[132] upon request from the Digital Service Coordinator or from the Commission. Access to information for vetted researchers has been broadened, it now covers not only the identification of systemic risks, but also the measures taken to mitigate those risks.[133] VLOPs still have the power to request an amendment of the access requests to the Digital Services Coordinator under article 40(5)—it remains to be seen whether this could undermine the impact of this obligation.[134]

57 **Reporting obligations.** Article 42 imposes specific reporting obligations for VLOPs, in addition to those already contained in Articles 24 and 15. VLOPs will have to issue those reports every 6 months, instead of once year for other intermediaries.[135] Furthermore, VLOPs have to report on the following information:

- Information specific to their human resources involved in moderation, including their qualification and linguistic expertise[136];
- Their number of active users in each Member State[137];
- Information related to the risk assessments and mitigation measures pursuant to Articles 34 and 35[138];
- Information related to the independent audit pursuant to Article 37(4).[139]
- VLOPs have the possibility to publish versions of those reports redacted of certain confidential information. In that case, however, VLOPs have to transmit the complete report to the relevant Digital Services Coordinator and the European Commission.[140]

VI. Enforcement mechanisms in the DSA

58 The enforcement roles in the DSA have been divided between the newly created Digital Services

132 DSA, Article 40, 3.

133 DSA, Article 40.

134 DSA, Article 40, 5.

135 DSA, Article 42, 1.

136 DSA, Article 42, 2.

137 DSA, Article 42, 3.

138 DSA, Article 42, 4.

139 ibid.

140 DSA, Article 42, 5.

Coordinators, the European Board for Digital Services as well as the European Commission.

59 Digital Services Coordinators (or DSCs). Digital Services Coordinators are designed by Member States. Even though more than one authority could be responsible for the enforcement of the Digital Services Act, the DSC should be responsible for ensuring coordination at national level of all authorities in charge of enforcing the DSA.[141] Mechanisms are put in place in order to allow for cooperation between DSCs across borders[142], as well as with the Board and the Commission.[143] DSCs are assigned investigation[144] and enforcement[145] powers, which includes the power to require audits from online platforms, impose fines and require immediate actions or commitments in order to remedy harmful situations.[146]

60 European Board for Digital Services. The European Board for Digital Services is an EU-level independent advisory group whose role is to ensure the consistency of the application of the DSA across Member States and to provide assistance and guidance on relevant emerging issues across the EU and regarding the supervision of VLOPs. It does not have investigating nor enforcement powers towards online platforms.[147]

61 European Commission. The European Commission, while not fully in charge of enforcing the DSA, still has a role to play in its enforcement. Its role is more subsidiary for online platforms under the 45 million users mark where it can assist DSCs in case of inconclusive investigation or repeated infringements. In the context of VLOPs, however, the Commission notably has the authority to launch an investigation[148], to issue fines to non-compliant VLOPs[149], to put interim measures in place in case of urgency[150], to require commitments for platforms to ensure compliance[151] and to effectively take actions to monitor the effective application of the DSA.[152]

62 Limits of enforcement by national authorities. The Member States-centered approach taken by the European legislator with the DSA is similar to the one proposed in the General Data Protection Regulation ("GDPR"), where Member States designate one (or more) Data Protection Authority in charge of data protection. Through a one-stop-shop mechanism, also similar to the one instituted in the GDPR, the DSA aims to better resolve cross-borders conflicts involving platforms.[153] This situation could lead to potential discrepancies between the Member States regarding the DSA, as some have already pointed regarding the GDPR.[154] Lack of uniformity between the means at the disposal of various data protection authorities has been highlighted as an issue regarding GDPR enforcement and the same could be true for the DSA.[155] Finally, the concentration of VLOPs' main establishment in a few Member States, notably Ireland, could put additional workload on specific DSCs as well as political pressure[156] in order not to see VLOPs move their establishment to a Member State which is less strict (or less staffed) in terms of enforcement.[157]

63 The enforcement by the Commission might be more effective. The European Commission has already issued large fines to corporations, such as the $2.4B fine imposed on Google for abusing its dominant position.[158] Letting the Commission enforce the DSA

141 DSA, Article 49.

142 DSA, Article 57.

143 DSA, Article 49, 2.

144 DSA, Article 51, 1.

145 DSA, Article 51, 2.

146 DSA, Article 52.

147 DSA, Article 61.

148 DSA, Article 66.

149 DSA, Article 74.

150 DSA, Article 70.

151 DSA, Article 71.

152 DSA, Article 72.

153 DSA, Article 58.

154 J. Ryan, 'Europe's Governments are failing the GDPR: Brave's 2020 Report on the enforcement capacity of data protection authorities', 2020, <https://brave.com/wp-content/uploads/2020/04/Brave-2020-DPAReport.pdf>.

155 'Has GDPR Delivered on Its Central Promise?' (*Law.com International*) <https://www.law.com/international-edition/2022/01/31/lawyers-say-gdpr-has-failed-to-deliver-on-its-central-promise/> accessed 16 August 2022.

156 See for instance the allegations concerning Facebook's investigations by the Irish Data Protection Authority, B. Goodwin, 'Max Schrems accuses Ireland of 'Kafkaesque' delay in Facebook GDPR investigation', Computer Weekly, 26 May 2020, https://www.computerweekly. com/news/252483668/Schrems-accuses-Ireland-of-Kafkaesque-delay-in-Facebook-GDPR-investigation.

157 Strowel and Somaini (n 64).

158 'Antitrust: Commission fines Google €2.42 billion for abusing

might create some uncertainty in case of a change of the political composition and/or inclination of the Commission (while some of today's Commissioners, for example Thierry Breton, are in favor of robust intervention).

64 **Means in the hands of the EU.** Another issue regarding enforcement that is common to both the DSCs and the EU Commission is the discrepancy between the means at the hands of public powers and the large pockets on which VLOPs can rely on. Effectively regulating platforms will require additional personnel and expertise. New funds should be allocated to this mission. Article 43 of the DSA will allow the Commission to charge VLOPs a supervisory fee that should, in theory, cover the expenses incurred for their supervision.[159]

65 Thierry Breton, in a press release following the final vote on the DSA by the EU parliament, gave a few insights of how the Commission will supervise the enforcement of the DSA for VLOPs. He insists on the cooperation within the Commission itself, but also on a reliance on "a network of trusted flaggers, such as NGOs, hotlines or rightsholders, to ensure that platforms react to the flagged illegal content as a priority".[160] During the same address, he also mentioned the creation of a high-profile European Centre for Algorithmic Transparency (ECAT). The ECAT, hosted by the Joint Research Center of the Commission, has been launched on 18 April 2023, it should closely cooperate with DG CONNECT and with industry representatives, academia and civil society, fostering the multi-stakeholder model of regulation that the DSA aims to promote.[161]

E. Conclusions on the DSA contribution to the fight against disinformation and some possible improvements

66 **A necessary yet only first step.** During the debates preceding the adoption of the DSA, most Members of the European Parliament welcomed the draft legislation. At the same time, they admitted that the DSA is only a first step towards an efficient regulatory framework for online platforms in the EU.[162] When the whole DSA will apply (as of 17 February 2024), we will see how the changes regarding liability for illicit content (see above under C) together with the new diligence obligations (see above under D) and the implementing measures will (or will not) make a difference in practice. The regulators have less than a year to prepare. We have some doubt about whether the new reporting and transparency obligations of the DSA will really make a difference, in particular for reducing online disinformation, but it is also clear that the DSA is part of a European and global trend towards platform regulation (even the US is now clearly considering to introduce such regulatory framework[163]), and the convergent regulatory initiatives might well prompt some platforms to partly revise their (ad-based) business model and their treatment of awful and illicit content.

67 As the DSA alone will thus not be sufficient to curb the spread of disinformation and other nefarious content, we suggest three paths of additional improvements to regulate online disinformation in Europe:

- A broader principle of transparency for online platforms which would provide a positive right for some collectives (such as consumers organizations) to initiate actions;
- A co-regulation model with enhanced involvement of users and third parties in accessing and adjusting the parameters for content recommendation on the platforms; and
- An independent authority to regulate online platforms.

dominance as search engine by giving illegal advantage to own comparison shopping service' (*European Commission - European Commission*) <https://ec.europa.eu/commission/presscorner/detail/es/MEMO_17_1785> accessed 16 August 2022.

159 DSA, Article 43.

160 'Sneak Peek: How the Commission Will Enforce the DSA & DMA' (*European Commission - European Commission*) <https://ec.europa.eu/commission/presscorner/detail/en/STATEMENT_22_4327> accessed 16 August 2022.

161 ibid; see: <https://algorithmic-transparency.ec.europa.eu/index_en>.

162 'Sitting of 04-07-2022 | Plenary | European Parliament' <https://www.europarl.europa.eu/plenary/en/vod.html?mode=chapter&vodLanguage=EN&vodId=c53414f9-469d-6196-fa8d-be169c87c94e&date=20220704#> accessed 26 July 2022.

163 See J. Biden op ed in the Wall Street Journal, 11 Jan. 2023, "*Republicans and Democrats, Unite Against Big Tech Abuses.* Congress can find common ground on the protection of privacy, competition and American children".

I. Need for an additional transparency principle generating a right to get an explanation and a remedy

68 **A general transparency principle.** We believe that a general transparency principle and a related right to transparency for the users of platforms should be imposed. The current reporting obligations imposed on platforms by the DSA will allow for the opening of platforms' data and mechanisms, but we believe more could be done. The importance of social media platforms for our democratic societies justifies that transparency should be the norm, not the exception; platforms should offer their users and society, in general, an accurate picture of the way they operate and make decisions about prioritizing and spreading information.[164] We therefore propose to impose a general transparency principle on platforms, similar to the one applicable to public administration[165] (and to some extent to the transparency principle included in the GDPR).[166]

69 Transparency obligations related to the design of online platforms and their moderation policies could foster accountability and allow users and external actors to test the efficiency and effectiveness of the platforms' moderation tools, just as administrative transparency theoretically allows citizens to oversee the actions of the administration.[167] Furthermore, obligations analog to those of administrative transparency could reduce the secrecy around the operations of platforms, allowing for better oversight thereafter.

70 This transparency principle should be accompanied with an accountability principle. Platforms should not only comply with the various transparency obligations contained in the legislation but should also be able to demonstrate their compliance. Shifting the burden of proof of compliance on VLOPs makes sense given their role as gatekeepers online.[168] The extent of such change in the burden of proof should be further analyzed, and in any case well-targeted, as VLOPs' freedom to conduct their business cannot be disproportionately curtailed.[169]

II. Co-regulation with vetted researchers and other certified stakeholders involved in the process

71 **A more active role for users and third parties.** For years, the circulation of information and the mitigation of disinformation have been ordered by platforms through tech design twists and self-regulation. Content orientation and recommendations were thus only left to "private ordering". The DSA marks a step towards more intervention by public authorities. However, such an approach should remain minimal as freedom of expression rightly limits how far the State can interfere in the public debate and in the process leading to the collective construction of truth. To go further, we believe it is important to empower users and third parties such as academic scholars and NGOs so that they can play a more active role in the fight against disinformation. The DSA takes some steps in that direction (see above on data access and Article 40), but more could be done.

72 **Empowering users through middleware.** Middleware has been defined in this context as "software and services that would add an editorial layer between the dominant internet platforms and in-

164 Amélie Heldt and Stephan Dreyer, 'Competent Third Parties and Content Moderation on Platforms: Potentials of Independent Decision-Making Bodies From A Governance Structure Perspective' (2021) 11 Journal of Information Policy 266.

165 In the 1970s, a growing movement called for more transparency on the part of public administrations. The doctrine of administrative transparency developed itself in opposition to the culture of secret which had been prevalent in the public administration. Jacques Chevallier, « Le mythe de la transparence administrative », in Information et transparence administratives, PUF, 1988.

166 Élise Degrave and Yves Poullet, *L'e-Gouvernement et La Protection de La Vie Privée: Légalité, Transparence et Contrôle* (Larcier 2014) 314. Its aim would be to allow citizens to understand "how the governments operate on their behalf". See Christopher Hood and David Heald, *Transparency The Key to Better Governance?* (2012) 49.

167 ibid.

168 Philip M Napoli, 'Social Media and the Public Interest: Governance of News Platforms in the Realm of Individual and Algorithmic Gatekeepers' (2015) 39 Telecommunications Policy 751.

169 Article 16 of the EU Charter of Fundamental Rights only timidly recognises the freedom to engage in business activities, and the CJEU interpretation of this general principle of law, which predates its incorporation in the Charter, is not a bar to an increased burden of proving some level of compliance (still to be defined). See for ex. Thierry Leonard and Julie Salteur, Article 16 - Liberté d'entreprise, in Fabrice Picod, Cecilia Rizcallah et Sébastien Van Drooghenbroeck (eds.), *Charte des droits fondamentaux de l'Union européenne. Commentaire article par article*, Larcier, 2023, 3rd ed., p. 401 ff.

ternet users".[170] Middleware would allow users to better control the content they receive on social media. While several solutions for users do exist, those solutions are usually reserved to tech-handy users.[171] Articles 27 and 38 of the DSA, which encourage platforms to give users more choice regarding the way content is formatted, prioritized and proposed to them, is a first step towards a broader and user-friendlier introduction of middleware on online platforms.

73 **Broader access right beyond the vetted researcher status.** In order to better integrate additional actors in the regulatory process, a larger opening of the vetted researcher status should be considered. Under the DSA, the DSC are responsible for granting the status of vetted researchers allowing them to access the data related to risks assessments of VLOPs and their mitigation measures. We suggest to leave the certification process in the hands of the ethics committees of the research institutions (thus reducing the role of the DSCs). More could also be done to allow some NGOs to benefit from the vetted researchers' access rights.[172] A strong certification mechanism should therefore be put in place in order to safeguard online privacy as well as the commercial interests of platforms.

74 **Better compliance and enforcement for self-regulatory instruments.** The self-regulatory tools used by platforms are currently left mostly unchecked. Tools such as the Code of Practice against Disinformation (see above under C) have been criticized for the lack of enforcement and compliance mechanism. Article 45 of the DSA specifically addresses codes of conduct such as the Code of Practice and allows the Commission as well as the Board to take actions in case of systematic failure to comply with a code of conduct—providing the existing Code of Practice with co-regulatory features.[173] (An explicit reference to the Code of Practice is by the way included in recital 106 of the DSA.) However, the current wording of Article 45 only allows the Commission and the Board to "invite the signatories to the codes of conduct to take the necessary action". It does not seem to give the EU authorities the necessary power to go further.

III. An independent authority to regulate online platforms

75 **Potential issues with the DSA enforcement.** The current enforcement methods of the DSA, splitting responsibilities between DSCs at the Member State level and the Commission, might cause issues similar to what has already been observed with the enforcement of the GDPR[174]: domestic issues might hinder the efficiency of the DSC in some EU countries[175] while there might be some pushback from certain DSCs to adequately address pressing issues, justified for instance by a lack of resources.[176] This could open the way for a form of forum shopping between Member States.[177] However, the European Commission services (in particular, the division on platforms at DG CONNECT) will be directly involved for the DSA enforcement. The prominent role given to the Commission might make the regulation of online platforms dependent on the political willingness of the Commission to use its new regulatory powers. A shift of policy objectives could therefore undermine the long-term enforcement of the obligations contained in the DSA.

170 Francis Fukuyama et alii, Middleware for Dominant Digital Platforms: Technological Solution to a Threat to Democracy, Stanford Cyber Policy Center, available, but not dated, at: <https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/cpc-middleware_ff_v2.pdf>, accessed 8 Sept. 2022 .

171 Several browser extensions are available online to tweak the ranking mechanisms of social media. For example, Social Fixer for Facebook (<https://socialfixer.com/>) allows the user to disable certain features of the platform, such as infinite scrolling or advertised posts.

172 Strowel and Somaini (n 64).; T Marsden, I Brown and M Veale, 'Responding to Disinformation: Ten Recommendations for Regulatory Action and Forbearance' in M Moore and D Tambini (eds), *In: Moore, M and Tambini, D, (eds.) Regulating Big Tech: Policy Responses to Digital Dominance. (pp. 195-230). Oxford University Press: Oxford, UK. (2021)* (Oxford University Press 2021) <http://doi.org/10.1093/oso/9780197616093.003.0012> accessed 25 April 2022.

173 DSA, Article 45.

174 'Has GDPR Delivered on Its Central Promise?' (n 155).

175 See for example the numerous accusations of malfunctioning of the Belgian Data Protection Autority (APD/GBA), which almost led to an official procedure of the European Commission against Belgium in front of the ECJ – see https://www.lesoir.be/438557/article/2022-04-27/lapd-est-inoperante-un-et-demi-dalertes-de-ses-deux-codirectrices.

176 See for example the tensions between the European Commission and the Irish DPA regarding Meta - <https://iapp.org/news/a/what-the-dpc-meta-decision-tells-us-about-the-gdprs-dispute-resolution-mechanism/?mkt_tok=MTM4LUVaTS0wNDIAAAGJO9479tKXSTPebi5oJZaJ5y7hxaF3KMUwUiTwQamXWTXoNesognmhoyE5N2RKcskx-N27jhO14TlzjA_TzQK1xIWS9SMpQGcu7vvQ1a2pD3nY>.

177 Dan Jerker B Svantesson, 'EDPB's Opinion 8/2019 on the Competence of a Supervisory Authority in Case of Establishment Changes Reports: European Union' (2020) 6 European Data Protection Law Review (EDPL) 98.

76 **An independent EU authority to regulate platform.** We therefore suggest the creation of an independent, European-wide entity solely in charge of the regulation of online platforms. The creation of the European Board for Digital Services is a positive first step. Such an independent authority would be responsible for the enforcement of the transparency principle described above and to organize the relations between platforms, their users and the different stakeholders involved in the production and regulation of content online. This authority should fulfill the standards imposed on any regulator, such as independence and accountability towards the public, and be well-equipped (sufficient funding and staffing with data and algorithms experts).