Pandemic, War and Half-Truths: Can The European Union's Digital Services Act Tackle Disinformation on Social Media?

By

Marie-Therese Burkard

Abstract

With the rise of social media platforms and the resulting decentralization of media channels, accessing accurate and well-researched news has become more challenging. The media world today requires that we criticize and evaluate each piece of news that we consume in passing on digital platforms to prevent ourselves from being misled by half-truths or outright incorrect statements. Aiming to protect consumer rights online, the European Union's newly implemented Digital Services Act has the potential to counteract such disinformation on social media platforms by setting out an expansive list of obligations. With COVID-19 vaccine hesitancy and Russian propaganda regarding its war with Ukraine, disinformation is flooding social media platforms. As such, especially the last two years have shown the inherent need to prevent the spread of disinformation in the face of a crisis. As such, the following paper will analyse the DSA's effectiveness in tackling disinformation to prevent further political and social polarisation and protect our democracies.

Table of Contents

Introduction

Since its launch in November 2022, ChatGPT, a text-based dialogue AI tool, has taken the world by storm, constituting the fastest growing digital service in history.[1] Whilst these so-called "real-time chatbots"[2] can be used for research, reasoning, or even light-hearted interactions, it has sparked growing concerns regarding future job opportunities, academics and even the development of disinformation in online spaces. This is because their ability to produce elaborate, nuanced, and emotive texts from a simple input can allow these tools to produce disinformation. Even Sam Altman, CEO of OpenAI, the company behind ChatGPT, expressed his concerns in an exclusive interview with ABC News' Rebecca Jarvis: "One thing I'm particularly worried about is that these models could be used for large-scale disinformation."[3] As such, the virality of ChatGPT is one recent example that highlights the danger of disinformation itself.[4]

It is drastically clear that the dissemination of false information is not a problem that is limited to the US, nor one that solely involves the accidental spread of misleading information. Within the European Union (EU), disinformation, misleading information that is created and disseminated with malicious intent,[5] has been repeatedly named as a threat to political stability, society, and democracy overall. With the emergence of social media, a digital environment

[1] Krystal Hu, *ChatGPT set record for fastest ´-growing user base - analyst note,* REUTERS (February 2, 2023), https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/.
[2] Tiffany Hsu and Stuart A. Thompson, *Disinformation Researchers Raise Alarms About A.I. Chatbots,* N.Y. TIMES (February 8, 2023), https://www.nytimes.com/2023/02/08/technology/ai-chatbots-disinformation.html.
[3] ABC News, *Open AI CEO, CTO on risks and how AI will reshape society,* YOUTUBE (March 18, 2023), 1:46-1:52, https://www.youtube.com/watch?v=540vzMlf-54.
[4] For more examples on how ChatGPT can produce disinformation: Tiffany Hsu and Stuart A. Thompson, *Disinformation Researchers Raise Alarms About A.I. Chatbots,* N.Y. TIMES (February 8, 2023), https://www.nytimes.com/2023/02/08/technology/ai-chatbots-disinformation.html.
[5] Don Fallis, *What is disinformation?,* 63(3) LIBRARY TRENDS, 401, 413 (2015).

exists that is increasingly exploited for large-scale disinformation campaigns.[6] The COVID-19 pandemic[7] and the ongoing Russo-Ukrainian war[8] has drastically brought the danger of disinformation campaigns conducted by China and Russia to the EU's attention.

Having entered into force in 2022, the Digital Services Act (DSA), regulating providers of digital intermediary services including social media platforms, imposes an extensive set of obligations with the aim to foster "a safe, predictable and trusted online environment".[9] Although the DSA does not solely focus on disinformation, it can potentially have significant implications on the behaviour of social media platforms and the dissemination of disinformation. Given its revolutionary nature and its recent entry into force, it is of interest to evaluate the relationship between the DSA, disinformation, and social media platforms, posing the question: Can the EU's DSA tackle disinformation on social media?

To evaluate this question, normative analysis of academic research and positive law, that being the DSA, will be conducted. As such, the question will be answered at the hands of three chapters. The first one on the danger of disinformation defines the term "disinformation", explaining how the Digital Revolution has facilitated its dissemination and how this poses a threat to democracy. The second chapter entails a descriptive and analytical section on the DSA, focusing on those sections which are relevant to social media platforms and the combat of disinformation. Lastly, the third chapter will analyze the

[6] *EEAS Special Report Update: Short Assessment of Narratives and Disinformation Around The COVID-19 Pandemic*, EUVSDISINFO (2020), https://euvsdisinfo.eu/uploads/2020/05/EEAS-Special-Report-May-1.pdf.
[7] *EEAS Special Report Update: Short Assessment of Narratives and Disinformation Around The COVID-19 Pandemic*, EUVSDISINFO (2020), https://euvsdisinfo.eu/uploads/2020/05/EEAS-Special-Report-May-1.pdf.
[8] *Disinformation about Russia's invasion of Ukraine - Debkunking Seven Myths spread by Russia,* EEAS (March 18, 2022), https://www.eeas.europa.eu/delegations/china/disinformation-about-russias-invasion-ukraine-debunking-seven-myths-spread-russia_en?s=166.
[9] 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) [2022] OJ L277/1, art 1.

effectiveness of the DSA in reducing online disinformation, elaborating on general limitations of the DSA and the overarching difficulty in combatting such a complex enemy as disinformation.

The limitation of this analysis resides in the nature of disinformation itself. It is a highly intricate phenomenon with numerous facets to disinformation, which cannot be summarized solely in one chapter. Nonetheless, the first chapter acts as a base for the understanding of the danger of disinformation and how it can be affected by the DSA.

The following argues that the DSA undoubtedly imposes a set of drastically needed duties for providers of online platforms, acting as legal and financial encouragement for social media platforms to monitor and moderate content, as well as coordinate with the EU's authorities. However, given the complex, ever-changing and highly adaptable nature of disinformation, the DSA may only have a limited effectiveness in tackling disinformation.

Chapter 1: The Danger of Disinformation

Defining disinformation

Since 2016 the term "fake news" has gained traction in news, politics and social media and has thus become ingrained in our day-to-day vocabulary. Academics specialized in journalism and media, such as Bennett and Livingston, however, explain that fake news cannot solely be considered as "isolated incidents of falsehood and confusion",[10] because it is more than just inaccurate information. This signals that behind the term "fake news" is not an issue solely concentrated in the realm of US politics, but a wider phenomenon: the circulation of misinformation and disinformation online.

Depending on the example in question, "fake news" can be misinformative or consist of disinformation.[11] As cited by Fallis, Floridi explains that misinformation is "well-formed and meaningful data (i.e. semantic content) that is false".[12] This often occurs during breaking news, for example, when people spread rumors to provide updates; their objective is not to circulate false information, instead they believe that this information is true.[13] Although misinformation does have negative effects on the consumption of media and influences people's outlooks,[14] disinformation could be considered as more dangerous. This is because "the activity of disinformation" can be defined as

[10] W. Lance Bennett and Steven Livington, *The disinformation order: Disruptive communication and the decline of democratic institutions,* 33(2) EUROPEAN JOURNAL OF COMMUNICATION, 122, 124 (2018).
[11] David M. J. Lazer, Matthew Baum, Yochai Benkler et al., *The science of fake news,* 359(6380) SCIENCE, 1094, 1094 (2018).
[12] Don Fallis, *What is disinformation?,* 63(3) LIBRARY TRENDS, 401, 409 (2015).
[13] Claire Wardle and Houssein Derakshan, *Journalism, "Fake News & Disinformation: Handbook for Journalism Education and Training,* UNESCO (2018), 7, 47, https://unesdoc.unesco.org/ark:/48223/pf0000265552.
[14] Alice Marwick and Rebecca Lewis, *Media Manipulation and Disinformation Online,* DATA & SOCIETY RESEARCH INSTITUTE (2017), 44, https://www.posiel.com/wp-content/uploads/2016/08/Media-Manipulation-and-Disinformation-Online-1.pdf.

"creating and distributing intentionally deceptive content,"[15] as is shown by large-scale disinformation campaigns, for example. According to Fallis, disinformation has three distinct qualities: it is a type of information, which is likely to create false beliefs, and has the intention of evoking exactly these false beliefs.[16] Therefore, the defining difference between misinformation and disinformation is the question of malicious intent. Given its stark impact on democratic systems, disinformation will be at the heart of this analysis.

Disinformation as a threat to democracy

Although the 2016 presidential election brought viral attention to disinformation, it is by no means the first or the last time that disinformation has seriously threatened access to reliable information, trust in media and the stability of democracies. Disinformation can be traced back to the Romans,[17] however its circulation initially soared with the invention of the printing press in 1436, allowing the efficient reproduction of printed information.[18] Similarly, the increase in the spread and significance of disinformation in recent years can be attributed to the development of the internet. The emergence of the internet and social media not only enables information to reach millions through a digital click, but also gives users access to platforms where they can publish any information, such as on Twitter, Facebook, or other online forums.[19] Any user can post anonymously or behind fake usernames and repost or share arbitrarily, decreasing the ability to maintain individual accountability for the creation or distribution of information. As such, the provision of information is no longer limited to journalists that are constrained to duties of honesty, transparency, and

[15] Peaks M. Krafft and Joan Donovan, *Disinformation by Design: The Use of Evidence Collages and Platform Filtering in a Media Manipulation Campaign,* 37(2) POLITICAL COMMUNICATION, 194, 195 (2020).
[16] Don Fallis, *What is disinformation?,* 63(3) LIBRARY TRENDS, 401, 404-407 (2015).
[17] Julie Posetti and Alice Matthews, *A short guide to the history of "fake news" and disinformation,* INTERNATIONAL CENTER FOR JOURNALISTS (July, 2018), 1, https://milunesco.unaoc.org/mil-resources/a-short-guide-to-the-history-of-fake-news-and-disinformation/.
[18] Ibid.
[19] Edson C. Tandoc, Zheng Wei Lim and Richard Link, *Defining "Fake News": A typology of scholarly definitions,* 6(3) DIGITAL JOURNALISM, 3 (2017).

reliability. Instead, journalists as "gatekeepers" of information have been exchanged with intermediaries, such as social media platforms.[20]

Furthermore, disinformation campaigns exploit the system of the internet by using bots to disseminate false information on a mass scale.[21] They mimic human behavior by "[posting] content, [interacting] with each other, as well as real people, and … [targeting] people that are more likely to believe disinformation."[22] Similarly to how the identity of many users is unknown or non-existent, algorithms construct users' feeds using personal data without the user knowing how these recommendations are given. This allows for the targeting of specific user groups which are considered vulnerable to being interested in or believing the presented disinformation.[23] The result is "source blindness", a phenomenon described by Pearson as a situation "whereby individuals fail to consider source information when processing news content."[24] This creates a digital environment that is prone to exploitation by malicious actors. Therefore, Freelon and Wells argue that disinformation is "*the* defining political communication topic of our time."[25] The intricate and complex tactics used by actors of disinformation have the power to influence election outcomes, create social and political instability and amplify polarizations amongst a population.

Moreover, disinformation threatens democracy by undermining society's right to free and reliable information. Before the Digital Revolution, such was guaranteed by reading a newspaper or watching the news on television, even if

[20] Lucas Graves and C.W. Anderson, *Discipline and promote: Building infrastructure and managing algorithms in a "structured journalism" project by professional fact-checking groups,* 22(2) NEW MEDIA & SOCIETY, 342, 344-345 (2020).
[21] Samantha Bradshaw and Philip N. Howard, *Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation*, UNIVERSITY OF OXFORD (2017), 11, https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2017/07/Troops-Trolls-and-Troublemakers.pdf.
[22] Kai Shu, Amrita Bhattacharjee, Faisal Alatawi et al., *Combating disinformation in a social media age,* 10(165) WILEY INTERDISCIPLINARY REVIEWS, 1, 8 (2020).
[23] Samantha Bradshaw and Philip N. Howard, *Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation*, UNIVERSITY OF OXFORD (2017), 10, https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2017/07/Troops-Trolls-and-Troublemakers.pdf.
[24] George Pearson, *Sources on social media: Information context collapse and volume of content as predictors of source blindness,* 23(5) NEW MEDIA & SOCIETY, 1181, 1183 (2021).
[25] Dean Freelon and Chris Wells, *Disinformation as a Political Communication,* 37(2) POLITICAL COMMUNICATION, 145, 145 (2020).

different media outlets are also biased politically. However, today's digital world blurs our knowledge of where accurate information is found and who publishes it. A process of fact-checking and questioning content is an important part of consuming news on social media to ensure the information is truthful.

Due to the difficulty in detecting disinformation, anyone can fall into its trap; as the common saying goes, *the best lies contain a bit of the truth*. As digital tools and systems now exist that enable the fast dissemination of such lies or half-truths, democracy has found a new enemy in disinformation. Simply said, democracy is based on the foundations of freedom, equality, and representation of the population in government. However, the ability of a population to have an individual, independent, and informed voice is severely undermined by the oblivious confrontation with manipulative information. Due to the effectiveness of algorithms, disinformation is recommended to targeted groups, cementing belief in political extremism, anti-government sentiment, and hence aggravating polarization. It is unavoidable that a society has divisions, but disinformation further cements and increases these divisions.

The relationship between news, social media, and disinformation

In order to analyze the increasing threat that disinformation poses, the relationship between news and social media, and its implications, need to be analyzed. It will be illustrated how the structure of social media has challenged the position of journalists and enabled foreign powers to infiltrate media, fostering an environment for disinformation to grow.

Firstly, the emergence of social media uprooted traditional means of accessing news, resulting in a change of news distribution.[26] Before the Digital Revolution, news sources were characterized by a certain degree of centralization; information was accessed via public or private newspapers, television, or radio

[26] Edson C. Tandoc, Zheng Wei Lim and Richard Link, *Defining "Fake News": A typology of scholarly definitions,* 6(3) DIGITAL JOURNALISM, 3 (2017).

stations. Today, an additional layer of the digital world exists, where these traditional media outlets additionally produce digital content whether this is for their own specific application or for social media. According to the EU's annual Media and News Survey of 2022, 45% of respondents across all age groups stated that they used social media for information purposes, i.e., to follow the news and stay up-to-date with current affairs.[27] It is therefore clear that social media acts as a facilitator for information, or disinformation, distribution by "[blurring] the conception of information source".[28] The development of news from a system of centralization to one of fragmentation is only amplified by social media's characteristically low barriers to entry: anyone can make an account from anywhere in the world and can connect and reach anyone. Due to social media now being a hub of anonymity, it is more difficult to control the accuracy of published information, trace its origins and assign responsibility. This new system of fragmentation and anonymity can therefore easily be exploited by agents willing to create and distribute disinformation.

Secondly, due to the development away from traditional, centralised news sources, how we consume news and value it has changed. Apart from news now also being consumed through social media, what information we are recommended has been altered by algorithms of social media platforms. Posts that have a higher number of likes, comments, or shares (e.g., reTweets) are more likely to be viewed by even more people.[29] Such algorithmic amplification therefore results in more attention being generated for posts that people interact with more, which may lead to internet virality. A 2018 study by the Massachusetts Institute of Technology on news stories on Twitter, found that false news often spreads more extensively than true ones, which may be linked to the bigger emotional reaction to such false stories.[30] This illustrates how

[27] *Media & News Report 2022,* EUROPEAN PARLIAMENT (2022), 29, https://europa.eu/eurobarometer/surveys/detail/2832.
[28] Edson C. Tandoc, Zheng Wei Lim and Richard Link, *Defining "Fake News": A typology of scholarly definitions,* 6(3) DIGITAL JOURNALISM, 3 (2017).
[29] Ibid.
[30] Soroush Vosoughi, Deb Roy and Sinan Aral, *The spread of true and false news online,* 359(6380) SCIENCE, 1146, 1146 (2018).

rapidly disinformation can spread online. Furthermore, social media algorithms create echo chambers, those being virtual "environments in which the opinion, political leaning, or belief of users about a topic gets reinforced due to repeated interactions with peers or sources having similar tendencies and attitudes."[31] In short, this means that online we see what we want to see. Due to echo chambers further activating our confirmation bias,[32] disinformation that a user is confronted with which fits into their political, ideological, or moral beliefs, is more likely to be considered as true. Our news consumption has therefore changed in platform and is subject to algorithms, which can allow disinformation to flourish.

Thirdly, the emergence of social media has changed the journalistic profession, triggering the so-called third wave of journalism.[33] Whilst many positive results have also emerged due to this development, social media has posed significant challenges. Journalists have come under pressure to compete for attention, views, virality and hence advertising revenues with social media.[34] Due to such an attention economy, "low-quality but high-performing posts over high-quality journalism" are being incentivized.[35] Moreover, where it was solely journalists' role to provide reliable news, social media has now blurred the responsibility for news production and distribution. In turn, this is making the process of fact-checking and verification of information more difficult. Whilst social media has not changed the fundamental nature of journalism,[36] it has undermined journalism's ability to fulfil its core journalistic aims, including executing its role as the "watchdog" or "fourth power" and ensuring the population is accurately informed and fostering healthy democratic debate.

[31] Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi et al, *The echo chamber effect on social media*, 118(9) PROCEEDINGS OF THE NATIONAL ACADEMY FOR SCIENCE, 1, 1 (2021).
[32] *Confirmation Bias*, THE DECISION LAB, https://thedecisionlab.com/biases/confirmation-bias.
[33] Emily Bell and Taylor Owenthe, Peter Brown et al., The Platform Press: How Silicon Valley reengineered journalism, COLUMBIA UNIVERSITY (2017), 16.
[34] Alice Marwick and Rebecca Lewis, *Media Manipulation and Disinformation Online,* DATA & SOCIETY RESEARCH INSTITUTE (2017), 42, https://www.posiel.com/wp-content/uploads/2016/08/Media-Manipulation-and-Disinformation-Online-1.pdf.
[35] Ibid.
[36] Emily Bell and Taylor Owenthe, Peter Brown et al., The Platform Press: How Silicon Valley reengineered journalism, COLUMBIA UNIVERSITY (2017), 15.

To summarize, the relationship between news, social media and disinformation is double-pronged. On one hand, social media acts as an amplifier for disinformation:[37] it has increased the possibility of disinformation to be viewed, disseminated and most importantly to be believed as the truth. On the other hand, journalism as a profession has undergone changes, reducing its control of the information that is broadcast to the public and adding additional pressures of competition on social media. These two inputs create an environment where disinformation can flourish.

Disinformation campaigns in practice

Social media has placed journalism in "a system built for scale, speed and revenue,"[38] a system that does not reward true news and accurate reporting but sensationalism and anonymity. This has been exploited by various actors in the past, and will continue to be exploited in the future, to execute disinformation campaigns.

In the 2020 Global Inventory of Organized Social Media Manipulation of the University of Oxford, evidence was found which illustrated that 81 countries use social media to spread disinformation and hence shape public opinion, representing an increase of 11 countries from the previous year's report.[39] Strategies for disinformation campaigns include the use of so-called cyber troops, "government, military or political party teams … to [manipulate] public opinion over social media".[40] Apart from actual people being behind the screen, countries

[37] W. Lance Bennett and Steven Livington, *The disinformation order: Disruptive communication and the decline of democratic institutions,* 33(2) EUROPEAN JOURNAL OF COMMUNICATION, 122, 124 (2018).
[38] Emily Bell and Taylor Owenthe, Peter Brown et al., The Platform Press: How Silicon Valley reengineered journalism, COLUMBIA UNIVERSITY (2017), 15.
[39] Samantha Bradshaw, Hannah Bailey and Philip N. Howard, *Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manipulation,* UNIVERSITY OF OXFORD (2021), 1,
https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2021/01/CyberTroop-Report-2020-v.2.pdf.
[40] Samantha Bradshaw and Philip N. Howard, *Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation*, UNIVERSITY OF OXFORD (2017), 4,

also employ bots, accounts that are coded to interact with and imitate human users, which further virtually bombard social media with disinformation.[41] By posting automatic messages or posting keywords to trigger algorithmic amplification and control what content is trending[42] such bots support the dissemination of disinformation.[43] Given that journalists now also use social media in their look out for report-worthy news, such disinformation campaigns can also indirectly manipulate the news agenda.[44]

One of the most recent disinformation campaigns in the EU and the USA conducted by the Russian and Chinese governments occurred during the beginning of the COVID-19 pandemic.[45] In 2020, Guy Berger, Director for Policies and Strategies regarding Communication and Information at UNESCO stated that "there seems to be barely an area left untouched by disinformation in relation to the COVID-19 crisis."[46] Nine key recurring topics at the heart of disinformation campaigns were identified, including origins of the coronavirus, medical science and the discrediting of journalists,[47] some of which were produced and disseminated by Russian state media and pro-Kremlin outlets, and further spread through social media.[48] Some stories of Russian disinformation

https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2017/07/Troops-Trolls-and-Troublemakers.pdf.

[41] *Id*, 11.

[42] Samantha Bradshaw and Philip N. Howard, *Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation,* UNIVERSITY OF OXFORD (2018), 6, https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2018/07/ct2018.pdf.

[43] It is important to note here that the emergence of chatbots, such as ChatGPT, have posed another strategy for disinformation campaigns. For further information: Tiffany Hsu and Stuart A. Thompson, *Disinformation Researchers Raise Alarms About A.I. Chatbots,* N.Y. TIMES (February 8, 2023), https://www.nytimes.com/2023/02/08/technology/ai-chatbots-disinformation.html.

[44] Franziska Keller, David Schoch, Sebastian Stier and JungHwan Yang, *Political Astroturfing on Twitter: How to Coordinate a Disinformation Campaign,* 37(2) POLITICAL COMMUNICATION, 256, 258 (2023).

[45] *EEAS Special Report Update: Short Assessment of Narratives and Disinformation Around The COVID-19 Pandemic*, EUVSDISINFO (2020), https://euvsdisinfo.eu/uploads/2020/05/EEAS-Special-Report-May-1.pdf.

[46] UN News, *During this coronavirus pandemic, "fake news" is putting lives at risk,* UNITED NATIONS (April 13, 2020), https://news.un.org/en/story/2020/04/1061592.

[47] Julie Posetti and Kalina Bontcheva, *Disinfodemic: Deciphering COVID-19 disinformation,* UNESCO (2020), 6, https://unesdoc.unesco.org/ark:/48223/pf0000374416.

[48] Robin Emmott, *Russia deploying coronavirus disinformation to sow panic in West, EU document says,* REUTERS (March 18, 2020),

campaigns that supported conspiracy theories, for example, illustrated a link between COVID-19 and 5G networks or COVID-19 has an origin in biological warfare, as well as denounced the effectiveness of vaccines, especially those produced in the West.[49] The aim was to increase the disruption and confusion caused by the public health crisis and undermine European powers.[50] One recurring story that was disseminated by Russia and China regarded alleged biological labs where the USA produced the coronavirus.[51] It is clear that in the midst of a public health crisis, a time of worry and loss, such false stories add to the overarching level of confusion within a population and pose a significant danger in that they may prevent citizens from taking action to ensure their health.

To illustrate how disinformation campaigns are used for political gains, the Russian disinformation campaign in Russia must also be mentioned. Given Russian political interests in Ukraine as a former Soviet country and border state, Russian disinformation campaigns have been prevalent in Ukraine, especially regarding the Russian illegal annexation of Crimea in 2014 and the current ongoing war between the two nations.[52] As tracked and recorded by the EEAS, pro-Kremlin media outlets have claimed that Ukraine has committed genocides towards the Russian-speaking populations in the East for example.[53] This has since been proven as untrue in reports by the Council of Europe and the

https://www.reuters.com/article/us-health-coronavirus-disinformation/russia-deploying-coronavirus-disinformation-to-sow-panic-in-west-eu-document-says-idUSKBN21518F.

[49] *EEAS Special Report Update: Short Assessment of Narratives and Disinformation Around The COVID-19 Pandemic*, EUVSDISINFO (2020), https://euvsdisinfo.eu/uploads/2020/05/EEAS-Special-Report-May-1.pdf.

[50] Robin Emmott, *Russia deploying coronavirus disinformation to sow panic in West, EU document says,* REUTERS (March 18, 2020), https://www.reuters.com/article/us-health-coronavirus-disinformation/russia-deploying-coronavirus-disinformation-to-sow-panic-in-west-eu-document-says-idUSKBN21518F.

[51] *EEAS Special Report Update: Short Assessment of Narratives and Disinformation Around The COVID-19 Pandemic*, EUVSDISINFO (2020), https://euvsdisinfo.eu/uploads/2020/05/EEAS-Special-Report-May-1.pdf.

[52] *Disinformation about Russia's invasion of Ukraine - Debkunking Seven Myths spread by Russia,* EEAS (March 18, 2022), https://www.eeas.europa.eu/delegations/china/disinformation-about-russias-invasion-ukraine-debunking-seven-myths-spread-russia_en?s=166.

[53] Ibid.

OSCE.[54] Apart from the disinformation spread by state supported news within Russia, disinformation campaigns have reached social media and across Europe.[55] Deep fake videos of Ukrainian President Zelensky, for example, were spread across social media, alleging that he had fled the country and encouraged his army to surrender.[56] Similarly, videos that impersonated those of traditional media outlets, such as the BBC and Al Jazeera, were created and disseminated to illustrate reliability and undermine popular trust in traditional media.[57] This illustrates how Russia is using disinformation to erode foreign and domestic support for Ukraine and its leadership. Therefore, Russian disinformation campaigns are exploiting the already existing horror and confusion of war to further promulgate their political goals.

It is therefore clear that this new form of warfare, i.e., information warfare, poses a decisive threat to democracy.[58] Amongst many others, the distinguishing factors between democracies and authoritarian regimes are based on the involvement in free and fair elections, as well as the ability to independently produce and access reliable information that is not censored by the government. Apart from undermining a population's ability to be well informed, disinformation increases political polarization, lowers trust in traditional media outlets and subverts the honesty and accuracy of electoral processes, all of which contradict core democratic principles.[59] It is therefore of increasing importance that disinformation is tackled to maintain political stability, unity, and a healthy societal environment. The DSA has been named as one of such measures that may be able to reduce disinformation and hence ensure the maintenance of a

[54] Ibid.
[55] Roman Osadchuck et al., *Undermining Ukraine: How the Kremlin Employs Information Operations to Erode Global Confidence in Ukraine,* ATLANTIC COUNCIL (2023), 1, https://www.atlanticcouncil.org/wp-content/uploads/2023/02/Undermining-Ukraine-Final.pdf.
[56] Id., 12.
[57] Id., 1.
[58] Samantha Bradshaw, Hannah Bailey and Philip N. Howard, *Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manipulation,* UNIVERSITY OF OXFORD (2021), 21, https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2021/01/CyberTroop-Report-2020-v.2.pdf.
[59] Sophie L. Vériter, Coreneliu Bjola and Joachim A. Koops, *Tackling COVID-19 Disinformation: Internal and External Challenges for the European Union,* 15(4) THE HAGUE JOURNAL OF DIPLOMACY, 569, 572 (2020).

stable and strong democratic system. How this may exactly be possible will be outlined in the following chapter.

Chapter 2: Defining The Digital Services Act

As outlined in the previous chapter, disinformation is a grave threat to democracy, one that is amplified through the mechanisms of social media and exploited by foreign powers. In February of this year, the EEAS published its first report on Foreign Information Manipulation and Interference Threats, analyzing disinformation campaigns and establishing a common framework for policy choices.[60] In the report, Josep Borrel, High Representative of the Union for Foreign Affairs and Security Policy and Vice President of the European Commission, wrote: "We need to work with democratic partners around the world to fight information manipulation by authoritarian regimes more actively. It is time to roll up our sleeves and defend democracy, both at home and around the world."[61] The EU is therefore clearly aware of the grave threat of disinformation.

The EU's fight against disinformation

Even before the COVID-19 pandemic and the Russian invasion of Ukraine, the EU has put significant efforts into establishing a framework that can prevent disinformation from thriving. These significant steps will be briefly outlined to illustrate the context in which the DSA was introduced. In 2015, the EU launched EEAS East StratCom Task Force, which focuses specifically on analyzing and monitoring disinformation in Eastern Europe,[62] by for example publishing and correcting disinformation on the EUvsDisinfo website.[63] Here, it

[60] *First EEAS Report on Foreign Information Manipulation and Interference Threats*, EEAS (2023), 7, https://euneighbourseast.eu/news/publications/first-eeas-report-on-foreign-information-manipulation-and-interference-threats/.
[61] *Id*, 6.
[62] *Questions and Answers about the East Stratcom Task Force*, EEAS, https://www.eeas.europa.eu/eeas/questions-and-answers-about-east-stratcom-task-force_en#11234.
[63] *About,* EUVSDISINFO, https://euvsdisinfo.eu/about/.

is also important to mention the 2019 EU Action Plan Against Disinformation, which aims to "to build up capabilities and strengthen cooperation between Member States and EU institutions to proactively address disinformation."[64] This plan outlines how the EU aims to further reduce disinformation through four pillars: improving detection and monitoring of disinformation, improving coordinated responses, encouraging the private sector to fight disinformation, and raising awareness within society.[65] Especially important for the mobilization of the private sector is the 2022 Strengthened Code of Practice on Disinformation, which sets out regulatory standards for combatting disinformation, including the transparency of political advertising and the amplification of awareness through content flagging and moderation.[66] Signatories include a wide range of private sector companies, including Google, Meta, TikTok and Twitter.[67]

What is the DSA?

Whilst the EU's Action Plan Against Disinformation has included measures of absolute importance, the efforts may be limited. The 2022 Strengthened Code of Practice on Disinformation, for example, relies on a self-regulatory approach[68] and as such may not provide the necessary encouragement to adhere to the measures.

Considered a landmark piece of legislation, the introduction of the DSA may therefore be able to fill this gap by imposing stronger and consistent legal

[64] *Action Plan Against Disinformation,* EEAS (2019), https://www.eeas.europa.eu/sites/default/files/disinformation_factsheet_march_2019_0.pdf.
[65] *Audit Preview: EU Action Plan Against Disinformation,* EUROPEAN COURT OF AUDITORS (2020), 8, https://www.eca.europa.eu/lists/ecadocuments/ap20_04/ap_disinformation_en.pdf.
[66] *2022 Strengthened Code of Practice on Disinformation,* EUROPEAN COMMISSION (June 16, 2022), https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation.
[67] *Signatories of the 2022 Strengthened Code of Practice on Disinformation,* EUROPEAN COMMISSION (June 16, 2022), https://digital-strategy.ec.europa.eu/en/library/signatories-2022-strengthened-code-practice-disinformation.
[68] Ibid.

obligations for online platforms.[69] The European Commission describes it as "a first-of-a-kind regulatory toolbox globally", which "sets an international benchmark for a regulatory approach to online intermediaries."[70] This chapter will outline exactly this new benchmark, i.e. the new obligations for social media platforms due to the implementation of the DSA, and will make first remarks regarding its implications for the spread of disinformation.

The DSA constitutes one half of the EU's Digital Services Package,[71] the other one being Digital Market Act (DMA), which focuses on the economic positions of digital platform companies and competitiveness within the digital services market.[72] Based on the e-Commerce Directive (ECD), which was a long-standing pillar for the regulation of digital services,[73] the main aim of the DSA however is to protect consumers of digital platforms and to ensure that their fundamental rights are upheld by regulating intermediary services.[74] Although the DSA aims to fill the gaps of the ECD and hence does not solely address disinformation, this landmark regulation for online platforms could nonetheless have significant implications for the EU's fight against harming information.[75]

[69] 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) [2022] OJ L277/1.
[70] *Digital Services Act: EU's landmark rules for online platforms enter into force,* EUROPEAN COMMISSION (November 16, 2022), https://ec.europa.eu/commission/presscorner/detail/en/ip_22_6906.
[71] *The Digital Services Act Package,* EUROPEAN COMMISSION, https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package.
[72] *The Digital Markets Act: ensuring fair and open digital markets,* EUROPEAN COMMISSION (October 12, 2022), https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets_en.
[73] 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] OJ L178/1.
[74] *The Digital Markets Act: ensuring fair and open digital markets,* EUROPEAN COMMISSION (October 12, 2022), https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets_en.
[75] *Digital Services Act: EU's landmark rules for online platforms enter into force,* EUROPEAN COMMISSION (November 16, 2022), https://ec.europa.eu/commission/presscorner/detail/en/ip_22_6906.

To ensure that the legal approach of the DSA is understandable, its categorization of different services must first be outlined. The DSA lays out four categories of service providers varying in their size and impact on the digital ecosystem, which are targeted by the Regulation, those being: intermediary services, hosting services, online platforms and very large online search engines (VLOSEs)[76] and very large online platforms (VLOPs).[77] This does not mean that these are completely separated entities, but rather that they are interlinked: a hosting service is a subcategory of an intermediary service, an online platform a subcategory of a hosting service, and a VLOP a subcategory of an online platform.[78] As such, more rules and obligations apply to VLOPs than a simple intermediary service. This categorization is of significance, as they result in large social media platforms, such as TikTok, Facebook, Instagram and Twitter, being categorized as a VLOP and hence held to the highest standard of due diligence under the DSA.[79] As it is estimated that these VLOPs reach more than 10% of the EU population,[80] it is understandable that the DSA introduced such an extensive list of obligations for social media platforms, facing them with an increased level of scrutiny.

With the rules of the DSA only having entered into force on the 16th of November 2022 and only fully applying from the 17th of February 2024,[81] the

[76] The obligations outlined under Chapter III Section 4 of the DSA apply to VLOPs and VLOSEs equally. However, given the nature of the discussion on disinformation and social media platforms, only "VLOPs" or "social media platforms" will be mentioned. Nonetheless, the same obligations apply to those online platforms that will be designated as VLOSEs.

[77] *The Digital Services Act: Ensuring a safe and accountable online environment,* EUROPEAN COMMISSION (October 27, 2022), https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en.

[78] 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) [2022] OJ L277/1, recital 41.

[79] Clothilde Goujard, *TikTok, Twitter, Facebook set to face EU crackdown on toxic content,* POLITICO (February 17, 2023), https://www.politico.eu/article/tiktok-confirms-it-faces-highest-content-moderation-obligations-under-eu-law/.

[80] *Digital Services Act: EU's landmark rules for online platforms enter into force,* EUROPEAN COMMISSION (November 16, 2022), https://ec.europa.eu/commission/presscorner/detail/en/ip_22_6906.

[81] 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) [2022] OJ L277/1, art 93.

DSA's effectiveness is not yet known. Nonetheless, given the lack of obligations for digital platforms in relation to their content prior to the DSA, the analysis of the DSA is necessary to understand the future of disinformation in Europe.

The objectives, definitions, and structure of the DSA

In order to outline how and to what extent the DSA may be able to reduce the threat of disinformation, its aim, scope and application must be further portrayed. Article 1(1) of the DSA stipulates that the subject matter of the regulation is "to contribute to the proper functioning of the internal market for intermediary services by setting out harmonized rules for a safe, predictable and trusted online environment that facilitates innovation and in which fundamental rights enshrined in the Charter, including the principle of consumer protection, are effectively protected."[82] Specifically, the DSA also intends to "[address] the dissemination of illegal content online and the societal risks that the dissemination of disinformation or other content may generate."[83]

In order to achieve these objectives, the DSA imposes new obligations on providers of digital services which should harmonize the laws of the Member States, foster increased transparency and responsibility of platforms and reduce confrontation with illegal content, including disinformation.[84] As stipulated by Article 2 of the DSA, these obligations apply to all intermediary services that provide such services within the Union.[85] Therefore, the DSA is not only a milestone for the legal regulation of digital services within the EU but creates

[82] 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) [2022] OJ L277/1, art 1.
[83] Id., recital 9.
[84] *The Digital Services Act: Ensuring a safe and accountable online environment,* EUROPEAN COMMISSION (October 27, 2022), https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en.
[85] 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) [2022] OJ L277/1, art 2.

this "international benchmark"[86] for ensuring protection online, which is likely to affect regulation outside of the EU as well.

To fully understand the aim of the DSA, and let alone its application, certain definitions need to be clarified, which are laid out in Article 3 of the Regulation. A "hosting service" is the type of intermediary service that is significant in the analysis of social media platforms and disinformation, as these "[consist] of the storage of information provided by, and at the request of, a recipient of the service".[87] Hence, services which allow the sharing of information and content online fall under this category,[88] which is the purpose of social media platforms. Whilst the DSA does not explicitly define social media platforms, it does provide a description of an "online platform", explaining it as "a hosting service that, at the request of a recipient of the service, stores and disseminates information to the public, unless that activity is a minor and purely ancillary feature of another service or a minor functionality of the principal service …".[89] As the storing and dissemination of information to the public is not an ancillary feature to social media platforms, as indicated by recital 13 of the DSA,[90] social media platforms are considered as those of a hosting service and, more specifically, an online platform in the context of the DSA.

Furthermore, the biggest social media platforms would be considered as VLOPs. Article 33(1) of the DSA stipulates that to be designated as a VLOP, an online platform must have 45 million average monthly users or higher within the EU. Until the 17th of February 2023, online platforms had to publish their number of average monthly users,[91] which confirmed that TikTok, Facebook, Instagram and

[86] *Digital Services Act: EU's landmark rules for online platforms enter into force,* EUROPEAN COMMISSION (November 16, 2022), https://ec.europa.eu/commission/presscorner/detail/en/ip_22_6906.
[87] 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) [2022] OJ L277/1, art 3(g)(iii).
[88] Id., recital 29.
[89] Id., art 3(i).
[90] Id., recital 13.
[91] 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) [2022] OJ L277/1, art 24(2).

Twitter would be considered VLOPs.[92] To ensure transparency and legal clarity, the Commission has published its first list of platforms that have been designated as VLOPs, including those mentioned above.[93] Therefore, these large social media platforms will be subject to the additional obligations imposed on VLOPs.

Apart from showing under which category social media platforms fall, the DSA places the concept of disinformation under the term of "illegal content." Under Article 3(h) DSA, this is defined as "any information that, 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 which is in compliance with Union law…". This definition alone seems to refer solely to stark examples such as images that show child sexual abuse or content that promotes the sale of illicit goods. However, recital 12 of the DSA clarifies that the phrase should be considered as an umbrella term, including "information relating to illegal content, products, services and activities."[94] Although disinformation is not generally illegal within the EU,[95] disinformation campaigns can be considered as an illegal activity due to the disruptive nature of disinformation in terms of democracy, freedom of speech, public policy, and public health.

Enforcement mechanisms

Regulations imposing new obligations can only be as effective as their enforcement mechanisms. Therefore, before analyzing any new obligations that

[92] Clothilde Goujard, *TikTok, Twitter, Facebook set to face EU crackdown on toxic content,* POLITICO (February 17, 2023), https://www.politico.eu/article/tiktok-confirms-it-faces-highest-content-moderation-obligations-under-eu-law/.

[93] *Digital Services Act: Commission designates first set of Very Large Online Platforms and Search Engines,* EUROPEAN COMMISSION (April 25, 2023), https://ec.europa.eu/commission/presscorner/detail/en/ip_23_2413.

[94] 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) [2022] OJ L277/1, recital 12.

[95] Meyers, Z. (2022, April 21). *Will the Digital Services Act save Europe from disinformation?* Centre for European Reform. Retrieved April 15, 2023, from https://www.cer.eu/insights/will-digital-services-act-save-europe-disinformation.

came to fruition through the DSA, it is important to understand the general structure and methods of monitoring and enforcing the commitment to this Regulation, where specific focus needs to be given to its two key players: the Digital Service Coordinators (DSCs) and the Commission.

Apart from imploring each Member State to "designate one or more competent authorities to be responsible for the supervision of providers of intermediary services and the enforcement of this Regulation",[96] the DSA's created the new supervisory position of the DSCs, which will consist of one of the competent authorities.[97] As such, each Member State will have a DSC that will be responsible for establishing nation-wide coordination of the matters of the DSA and coordinating with the DSC of the other Member States,[98] whilst maintaining full independence from external influences.[99]

In order for the DSCs to achieve their objective of monitoring and ensuring compliance with the responsibilities of the Regulation, they will be appointed certain powers. These include the power to demand information from providers of services, to perform themselves or seek a judicial authority of a Member State to inspect a possible infringement of the Regulation.[100] Furthermore, the DSCs may take actions to stop an infringement of the DSA. The supervisory body can adhere to this responsibility by "[requiring] the management body of those providers … to examine the situation, adopt and submit an action plan setting out the necessary measures to terminate the infringement", ensuring that the action plan is realized and reporting on it.[101] If this is not sufficient, the DSC may also "request that the competent judicial authority of its Member State order the temporary restriction of access of recipients to the service concerned by

[96] 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) [2022] OJ L277/1, art 49(1).
[97] Id., art 49(2).
[98] Ibid.
[99] Id, art 50(2).
[100] Ibid.
[101] Id, art 51(3)(a).

the infringement."[102] Therefore, it is clear that the DSCs are given significant powers that aid the supervision of compliance with the Regulation.

The other key player in the regulation of VLOPs is the Commission. Whilst the DSC is responsible for all intermediary services, it is the Commission that has "exclusive powers to supervise and enforce Section 5 of Chapter III", that which is solely applicable to VLOPs.[103] As such, the Commission may initiate proceedings against VLOPs,[104] request information,[105] take interviews and statements,[106] and even conduct inspections,[107] representing the new supervisory powers also gained by the Commission.[108]

The national DSCs and Commission will be supervised, advised and supported by the creation of another supervisory body: the European Board for Digital Services ("the Board").[109] It shall be tasked with aiding the coordination of investigations, supporting the analysis of reports and audits of VLOPs, as well as providing advice and opinions on matters of the DSA.[110] As such, the Board will be the player that overarchingly ensures that the DSA is efficiently applied and that the relevant parties fulfil its responsibilities under it.

Furthermore, it is the threat of a significant penalty that will play a role in encouraging compliance with the DSA. Under Article 52 of the Regulation,

[102] 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) [2022] OJ L277/1, art 51(3)(b).
[103] *Id*, art 56(3).
[104] *Id*, art 66(1).
[105] *Id*, art 67(1).
[106] *Id*, art 68(1).
[107] *Id*, art 69(1).
[108] *Digital Services Act: EU's landmark rules for online platforms enter into force,* EUROPEAN COMMISSION (November 16, 2022), https://ec.europa.eu/commission/presscorner/detail/en/ip_22_6906.
[109] 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) [2022] OJ L277/1, art 61(1).
[110] *Id*, art 63(1).

providers of services may be fined up to 6% of their annual worldwide turnover in the preceding financial year if they fail to fulfil an obligation.[111]

How can the new obligations for social media platforms reduce disinformation?

General obligations

Before discussing obligations that are solely specific to VLOPs, some general requirements for all intermediary service providers that are relevant to reducing disinformation within the EU must be outlined.

Firstly, it is important to note that as is the case under the ECD,[112] social media platforms, as a hosting service, are not liable for the content on their platforms if the providers do not have "actual knowledge of illegal activity or illegal content".[113] This means that social media platforms need to remove content once they are aware of its illegality, however that no general obligation for proactive monitoring and possible removal exists.

All hosting services, including VLOPs, need to provide notice and action mechanisms to "to allow any individual or entity to notify them of the presence on their service of specific items of information that the individual or entity considers to be illegal content."[114] Such notices must contain various requirements, including the reasons why the individual believes the content to be illegal, as well as where this information can be found online, for example by providing a URL.[115] This shall allow for the efficient submission of notices and

[111] 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) [2022] OJ L277/1, art 52(3).
[112] 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] OJ L178/1, art 15.
[113] 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) [2022] OJ L277/1, art 6.
[114] *Id*, art 16.
[115] *Id*, art 16(2).

hence facilitate a quick response by providers where needed. When reacting to such a submission of notice, the provider of the online platform can remove the content in question, as well as suspend or terminate the provision of services to the recipient who published the content.[116] If such is the case, the service provider must provide a statement of reasons to the affected user, establishing what decision has been taken and reasons for this decision.[117] Furthermore, the provider of the service must inform the law enforcement or judicial authorities of the Member State if this process makes the provider aware of criminal offences, which involve "a threat to the life or safety of a person or persons."[118] Therefore, when a user becomes confronted with information they consider to be dangerous, such as disinformation, it is easier for the user to bring it to the attention of the platform and for the provider of the platform to know where to react. The intention is that the reaction speed of social media platforms towards removing or reducing access to content that contains disinformation, for example, can be increased, as the users themselves direct the platforms to where the threat is.

Risk management and crisis response for VLOPs

Apart from the provisions applicable to intermediary services, hosting services and online platforms, Section 5 of Chapter III of the DSA provides additional obligations for providers of VLOPs and VLOSEs to manage systemic risks. As such, Section 5 lays out specific responsibilities that pertain to social media platforms as VLOPs, the most important of which will be discussed in the following.

Firstly, in regard to risk assessments, VLOPs are obliged to "diligently identify, analyze and assess any systemic risks in the Union stemming from the design or functioning of their service and its related systems, including algorithmic

[116] 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) [2022] OJ L277/1, art17(1).
[117] *Id*, art17(2).
[118] *Id*, art18(1).

systems."[119] Such risk assessments should be conducted at least once a year.[120] The DSA then identifies four categories of systemic risks that should be taken into account in the annual report: (a) the dissemination of illegal content; (b) negative impacts on the exercise of fundamental rights, explicitly mentioning the freedom of expression and information, and freedom and pluralism of the media as guaranteed by Article 11 of the Charter; (c) negative impacts on civic discourse, electoral processes and public security; and (d) negative impacts to the protection of public health, and person's physical and mental wellbeing amongst others.[121] As disinformation campaigns interfere with access to reliable information, free and fair elections, public security and public health, it is clear that disinformation counts as a systemic risk that needs to be fought by social media platforms.

Following the assessment of systemic risks, VLOPs are tasked to implement "reasonable, proportionate and effective mitigation measures"[122] that tackle those risks. The DSA provides a non-exhaustive list of measures, however, those applicable to weakening the threat of disinformation include adapting content moderation processes and algorithmic systems, or taking measures to raise awareness.[123] In collaboration between the Board and the Commission, comprehensive reports shall be published once a year which illustrate recurrent systemic risks and the best practices to mitigate them.[124] Therefore, through the duty to perform regular risk assessments, social media platforms are obliged to stay alert to the threat of disinformation, providing proper identification of the content of disinformation, data on its spread and its results. As this formal research and recording illustrates the issues that need to be faced, it may allow social media platforms to respond in a way that is more finetuned to the risk and hence more efficient at reducing disinformation.

[119] 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) [2022] OJ L277/1, art 34(1).
[120] Ibid.
[121] *Id*, art 34(2).
[122] *Id*, art 35(1).
[123] Ibid.
[124] *Id*, art 35(2).

Furthermore, the DSA establishes a crisis response mechanism under Article 36, which gives the VLOPs guidance in situations that require acute risk identification and mitigation. Given a "crisis" is defined as a situation where "extraordinary circumstances lead to a serious threat to public security or public health in the Union",[125] the provision incorporates the threat of disinformation campaigns that occurred during the COVID-19 pandemic and are still occurring during the ongoing Russo-Ukrainian war. Firstly, following a decision of the Commission on the recommendation of the Board, VLOPs may be asked one or more of the following: to assess "whether, and if so to what extent and how, the functioning and the use of their services significantly contribute to a serious threat", to determine and apply mitigation measures of such risk, and to report to the Commission on the taken steps and their impact.[126] As such, the Commission will become an active participant in the crisis response by discussing the measures' effectiveness and proportionality with the provider,[127] and by monitoring the application of the measures.[128] Furthermore, if the Commission deems it necessary, it may initiate a decision obligating the provider to review the identified measures and/or their application.[129] These decisions will be taken with and reported to the Board,[130] allowing the Board to retain oversight of the developments.

Apart from the responsibility of crisis management being transferred to the VLOPs themselves, the Board may also recommend the Commission to develop voluntary crisis protocols to mitigate crisis situations.[131] These protocols shall include the role of each participant, the measures that shall be implemented and used to safeguard fundamental rights guaranteed by the Charter, and a process

[125] 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) [2022] OJ L277/1, art 36(2).
[126] *Id*, art 36(1).
[127] *Id*, art 36(6).
[128] *Id*, art 36(7).
[129] Ibid.
[130] *Id*, art 36.
[131] *Id*, art 48(1).

allowing for the regular reporting of the measures and their effects.[132] Crisis management and responses are therefore conceived by the social media platforms, where disinformation is found, and the Commission, the supervisory figure. It is hoped that through the adoption of such measures the spread of disinformation campaigns in situations of crisis can be effectively hindered by the collaboration between social media platforms, the Commission, and the Board.

Recommender system, transparency, and compliance requirements for VLOPs

Although the DSA already sets out obligations regarding the recommender system, transparency and compliance for all intermediary services, Section 5 includes additional responsibilities that solely apply to VLOPs.

As established in the previous chapter, the algorithms of social media platforms act as amplifiers of disinformation, adding to the targeting of disinformation campaigns and its spread. Therefore, it is necessary to discuss new regulations for the recommender systems of social media platforms and how they may combat this problem. In their terms and conditions, all online platforms that use recommender systems should describe the "main parameters used in their recommender system" and the options provided to the users "to modify or influence those parameters".[133] These parameters should explain the factors that determine which information is presented to them[134] and shall be explained "in plan and intelligible language"[135] to foster understanding among the recipients of such services. Following the obligations set out all online platforms, further requirements for the recommender systems of VLOPs are imposed under Article 38. It demands that if a VLOP uses a recommender system, it "shall provide at least one option for each of their recommender systems which is not based on

[132] 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) [2022] OJ L277/1, art 48(4).
[133] *Id*, art 27(1).
[134] *Id*, art 17(2).
[135] *Id*, art 27(1).

"profiling".[136] The EU's General Data Protection Regulation (GDPR) defines the term "profiling" as "any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person,"[137] including location or movements, personal preferences or the health or economic situation of an individual user. As such, it obliges VLOPs to provide at least one option where recommendations are given without the need of automatically processing personal data, allowing for transparency and control of user's data.

Lastly, all providers of intermediary services must publish regular reports to ensure transparency in reporting to the DSC of their Member State and to the Commission. Given their status as VLOPs, social media platforms need to provide such a report every six months[138] instead of at least once a year, as is foreseen for other providers of intermediary services,[139] and must include a wider set of information that encompasses requirements of Article 15 (for providers of intermediary services), Article 24 (for providers of online platforms) and Article 42 (specifically for VLOPs and VLOSEs). Firstly, all providers of intermediary services shall include specific information regarding content moderation in their annual reports, including what orders for content moderation were received by Member State authorities pursuant to Article 9 and Article 10, what notices were submitted pursuant to Article 16, what content moderation was actually undertaken, as well as the median time required for such response.[140] Furthermore, upon request online platforms must communicate to the DSC and the Commission the average monthly active recipients of their services within

[136] 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) [2022] OJ L277/1, art 38.
[137] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119/1, art 4(4).
[138] 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) [2022] OJ L277/1, art 42(1).
[139] *Id*, art 15(1).
[140] *Id*, art 24(3).

the EU.[141] Solely applicable to VLOPs, social media platforms must also illustrate the human resources used in content moderation and compliance with other previously mentioned obligations.[142] Here, Article 42 DSA also reiterates the transparency reporting duties found in its requirements for an independent audit, and risk mitigation efforts.[143] Therefore, it is clear that the EU legislators are strongly ensuring that information on actions undertaken by social media platforms, for example, flow freely and consistently between the providers, the DSCs and the Commission. This will allow the supervisory figures to maintain oversight about the content that is spread throughout the social media platforms, as well as the actions they are taking to limit its negative impacts.

In order to ensure that the VLOPs adhere to these new obligations set out, Section 5 of the DSA includes provisions that direct VLOPs themselves to ensure compliance with the act. Firstly, VLOPs are instructed to conduct an independent audit at least once a year at their own expenses.[144] This independent entity shall assess whether the VLOP complies with the provisions of Chapter III of the DSA, as well as the codes of conduct established in Article 45, 46 and 48,[145] and make "operational recommendations" where compliance is lacking.[146] To illustrate that these recommendations are taken seriously, the VLOP shall adopt an "audit implementation report" that either shows that they have implemented the recommendations or gives reasons as to why this has not been done.[147]

Furthermore, the VLOPs must "establish a compliance function, which is independent from their operational functions and composed of one or more compliance officers" in order to monitor compliance with the Regulation.[148] The

[141] 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) [2022] OJ L277/1, art 15.
[142] *Id*, art 42(4).
[143] *Id*, art 42(2).
[144] *Id*, art 37(1).
[145] Ibid.
[146] *Id*, art 37(6).
[147] Ibid.
[148] *Id*, art 41(1).

head of such a compliance function must cooperate with the DSC and the Commission, ensuring, amongst others, that risk reporting and mitigation, as well as independent audits are conducted appropriately.[149]

Apart from implementing methods to ensure and monitor compliance with the Regulation internally, VLOPs must also secure the ability of the DSC and the Commission to monitor conformity. Under Article 40 DSA, VLOPs must provide the DSC or the Commission access to data that is necessary in monitoring the fulfilment of the obligations when required by these supervisory bodies.[150]

Lastly, given that the DSA establishes new tasks for Commission in terms of monitoring, assessing, and proposing recommendations in the field of activities on and of online platforms,[151] VLOPs are instructed to pay a yearly supervisory fee.[152] This is also intended to cover the costs for the hiring of almost 200 new staff members that the Commission expects to hire to enforce the Regulation.[153]

To summarize, it is visible that VLOPs, and hence social media platforms, are subject to a larger number of new duties than other providers of intermediary services in order to ensure that the online environment is a safe and fair place for users. It imposes responsibilities of transparency reporting, identifying, monitoring, and reacting to possible risks on social media platforms, and ensures their compliance and cooperation with the DSC and the Commission. As such, a framework is established that allows the institutions of the European Union to maintain an overview of risks and issues of the online sphere. Whilst disinformation may not be the primary subject matter of the DSA, the imposed

[149] 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) [2022] OJ L277/1, art 41(3).
[150] *Id*, art 40(1).
[151] *Id*, art 43(2).
[152] *Id*, art 43(1).
[153] The Associated Press, *EU law targets Big Tech over hate speech, disinformation,* NPR (April 23, 2022), https://www.npr.org/2022/04/23/1094485542/eu-law-big-tech-hate-speech-disinformation.

obligations may nonetheless significantly support the EU's fight against disinformation and hence reduce this threat against democracy.

Chapter 3: Can The Digital Services Act Overcome The Threat of Disinformation?

From the previous chapter, there is no doubt that the DSA has set out significant new obligations for providers of VLOPs, including social media platforms, aiming to improve consumer protection and ensure the moderation of illegal content online, such as disinformation. Given the introduction of this extensive legislative system, the DSA has been described as the "constitution for the internet,"[154] which has the potential to limit the stark influence of Silicon Valley in Europe. The previous chapter has demonstrated the significant demands made for the VLOPs, illustrating the advantages the DSA can produce when combatting disinformation. The strengths of the DSA are derived from two aspects: its impact on the practice of social media platforms itself, and the benefits to the DSCs and the Commission. As such, the social media platforms are obliged to take actions, such as monitoring and moderating of disinformation, and provide the supervisory bodies information on this, allowing them to remain under oversight and deepen their understanding of the issues online. The hope is that this will allow the supervisory bodies in collaboration with the social media platforms to produce successful strategies that combat systemic risks, including disinformation online.

However, no matter how revolutionary a legislative piece may be, limitations will undoubtedly exist to its effectiveness. This may also be the case for the DSA, which does face significant criticisms regarding its interference with other rights, liability, and enforcement mechanisms amongst other factors. Furthermore, whilst the DSA provides a legal framework for content moderation, its effectiveness in reducing disinformation may remain limited. Therefore, the

[154] *Europe Calling "DSA Deal: A constitution for the internet!",* THE GREENS/EFA IN THE EUROPEAN PARLIAMENT (April 29, 2020), https://en.alexandrageese.eu/video/europe-calling-dsa-deal/.

following will analyze the DSA's practical effectiveness, focusing on the possible limitations to the DSA's ability in reducing disinformation and its threat to democracy.

Limitations to the DSA's enforcement mechanism

Enforcement mechanisms of a legislative act are of such significance, because "substantive rules are nothing but a "paper tiger" without effective enforcement."[155] It is therefore important to firstly note that the DSA has strengthened its enforcement mechanisms from the GDPR, having learnt from its "serious failures".[156] The DSA enjoys a more centralised enforcement scheme, providing significant powers to the newly introduced DSCs and the Commission.

However, whilst the DSCs and the Commission not only can but should monitor, investigate, and demand information of providers of social media platforms in all Member States, there exists the risk of a lack of harmonization in implementation. Whilst the DSA lays out specific requirements for processes and information that needs to be included in reports, such as those requirements of the transparency reports,[157] it may be possible that the enforcement mechanisms may be applied differently in each Member State. This is already derived from the fact that the DSA has elicited differing opinions amongst its stakeholders regarding the necessity and strictness.[158] It is likely that the Member States' opinions will be divided similarly. Given that Article 6 of the DSA only provides rules for when a provider of an intermediary service cannot be held liable, "when a provider can be held liable" is up to the "applicable rules of Union or national law to determine."[159] Such freedom may "limit the capacity of the DSA to create a

[155] Joris van Hoboke, Ilaria Buri, João Pedro Quintais et al., *The DSA has been published – now the difficult bit begins,* VERFASSUNGSBLOG (October 31, 2022), https://verfassungsblog.de/dsa-published/.
[156] Ibid.
[157] 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) [2022] OJ L277/1, art 42.
[158] Caroline Cauffmann and Catalina Goanta, *A New Order: The Digital Services Act and Consumer Protection,* 12(4) EUROPEAN JOURNAL OF RISK REGULATION, 758, 759 (2021).
[159] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October

level playing field throughout the EU."[160] This is only reinforced by DSC's omission of specific provisions on the allocation of resourcing for each DSC, which may create "uneven enforcement" across the EU.[161] Therefore, the DSA's enforcement mechanisms may result in a lack of harmonization, which may especially pose an issue in the topic of disinformation, a phenomenon which aims to disseminate as fast and as far as possible. Due to the nature of disinformation, it is important that Member States apply similarly harsh enforcement mechanisms in order to actively and efficiently combat disinformation.

Furthermore, concerns regarding the rule of law have been raised with the responsibilities of the Commission under the DSA. As "guardian of the Treaties", the Commission is the head of the executive branch of the EU. However, the DSA "effectively assigns an implementation role to the European Commission", which may reflect a conflict of interests when the enforcer of the DSA also brings the social media platforms to court if they do not fulfil their obligations.[162] Similarly, the role distribution between the Commission, DSCs and the Board is not as clear, which may lead to overlap and conflict amongst them.[163]

Apart from the structural difficulties of the enforcement mechanism, it is also important to note the possible lack of effectiveness in allocating the responsibility of VLOPs, and hence disinformation, to the Commission.[164] Given the ever evolving and highly complex phenomenon of disinformation, it is possible that the Commission, not being an expert agency of the digital world, may not have the knowledge and abilities required to effectively ensure the reduction of disinformation.

2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act) [2022] OJ L277/1, recital 17.

[160] Caroline Cauffmann and Catalina Goanta, *A New Order: The Digital Services Act and Consumer Protection,* 12(4) EUROPEAN JOURNAL OF RISK REGULATION, 758, 766 (2021).

[161] Asha Allen and Ophélie Stockhem, *A Series on the EU Digital Services Act: Ensuring Effective Enforcement,* CENTER FOR DEMOCRACY & TECHNOLOGY (August 18, 2021), https://cdt.org/insights/a-series-on-the-eu-digital-services-act-ensuring-effective-enforcement/.

[162] Ibid.

[163] Ibid.

[164] 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) [2022] OJ L277/1, art 56(3).

Limitations to the DSA's effectiveness in combatting disinformation

Although a key element of the DSA is the monitoring, evaluation and potential removal of illegal content, the DSA does not provide a clear definition of what constitutes illegal content. This is already highlighted by the lack of certainty regarding whether the DSA targets solely illegal or also harmful content.[165] Rather it uses an umbrella term which encompasses not only hate speech and videos portraying sexual abuse, but also disinformation.[166] As different categories of illegal content are not outlined, the DSA also does not provide different obligations or recommendations depending on the illegal content found on social media platforms.[167] This uniform approach to regulating such content provides room for ineffectiveness in the application of the DSA. Simply removing a Tweet that contains hate speech and barring the user from using the platform, for example, may be a successful content moderation act. However, doing the same for a Tweet that contains disinformation would not necessarily prevent the spread of a disinformation campaign. This difference in effectiveness is derived from the fundamental differences in the nature of types of illegal content: whilst hate speech may come from one individual, a post containing disinformation may be part of a larger campaign that is systematically organized from a governmental entity abroad. As such, the removal of a single Tweet cannot fight the dissemination of such disinformation, especially as this disinformation is spread by numerous bots and trolls. Therefore, the DSA may lack recommendations to tackle different types of illegal content, which may mean that no type of illegal content is regulated as effectively as could be possible.

[165] Aina Turillazzi, Mariarosaria Taddeo, Luciano Floridi and Federico Casolari, *The Digital Services Act: An Analysis of Its Ethical, Legal, and Social Implications,* 15(1) LAW, INNOVATION AND TECHNOLOGY, 83, 95 (2023).
[166] 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) [2022] OJ L277/1, recital 12.
[167] Joan Barata, *The Digital Services Act and its impact on the right to freedom of expression: special focus on risk mitigation obligations,* PLATAFORMA POR LA LIBERTAD DE INFORMACIÓN (2021), 17, https://libertadinformacion.cc/wp-content/uploads/2021/06/DSA-AND-ITS-IMPACT-ON-FREEDOM-OF-EXPRESSION-JOAN-BARATA-PDLI.pdf.

Due to this lack of differentiation between illegal content, the DSA creates concerns regarding the balancing of two rights: that of consumer and user protection against freedom of expression. The latter is guaranteed by Article 10 of the European Convention on Human Rights (ECHR),[168] as well as by Article 11 of the EU Charter of Fundamental Rights.[169] Whilst both rights contain limitations, these rights can only be restricted according to principles of legality, necessity and proportionality.[170] Whilst many provisions of the DSA express that action should be taken in a "reasonable, proportionate and effective" manner, as for example stipulated in Article 35 on the mitigation of risks,[171] it is not clarified what is "illegal enough" to demand action. Although there is no longer a general liability of social media platforms for the content produced by their users,[172] the lack of specificity in definitions and guidelines of the DSA may trigger the over-removal of content in an attempt to adhere to other obligations of the Regulation.[173] Such disproportionate infringement of users' rights as well as the disproportionate mitigation of risks,[174] is only amplified by the different interpretations of Member States in terms of what constitutes illegal expression and the lack of guidelines within the DSA.[175]

The issues of infringing on the right to freedom of expression is also present when trying to reduce disinformation. It is difficult to identify whether content

[168] Convention for the Protection of Human Rights and Fundamental Freedoms, Council of Europe (1950), art 10.
[169] Charter of Fundamental Rights of the European Union [2000] OJ C364/1, art 11.
[170] Joan Barata, *The Digital Services Act and its impact on the right to freedom of expression: special focus on risk mitigation obligations,* PLATAFORMA POR LA LIBERTAD DE INFORMACIÓN (2021), 16, https://libertadinformacion.cc/wp-content/uploads/2021/06/DSA-AND-ITS-IMPACT-ON-FREEDOM-OF-EXPRESSION-JOAN-BARATA-PDLI.pdf.
[171] 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) [2022] OJ L277/1, art 35(1).
[172] Id., art 6.
[173] Amélie P. Heldt, CHAPTER 4: EU DIGITAL SERVICES ACT: THE WHITE HOPE OF INTERMEDIARY REGULATION in Terry Flew and Fiona R. Martin, DIGITAL PLATFORM REGULATION: GLOBAL PERSPECTIVES ON INTERNET GOVERNANCE 79 (2022).
[174] Joan Barata, *The Digital Services Act and its impact on the right to freedom of expression: special focus on risk mitigation obligations,* PLATAFORMA POR LA LIBERTAD DE INFORMACIÓN (2021), 19, https://libertadinformacion.cc/wp-content/uploads/2021/06/DSA-AND-ITS-IMPACT-ON-FREEDOM-OF-EXPRESSION-JOAN-BARATA-PDLI.pdf.
[175] Ibid.

qualifies as misinformation, an unintentional expression of false information, or disinformation, a purposeful and systematic dissemination of false information. Content that may just be an expression of unconventional or unaccepted views, may be flagged or removed, which could infringe on a user's right to freedom of expression and may undermine public discourse online. Whilst regulating misinformation is also important to society and democracy, the real threat is derived from actors aiming to undermine political and social stability through the intentional dissemination of false information. As such, the lack of specificity regarding types of illegal content in the DSA may not only lead to a difficulty in balancing rights but may also waste the efforts and resources of social media platforms.

Given the complex nature of disinformation due to the variety of actors and digital mechanisms involved and the speed of dissemination, social media platforms may not have the knowledge or resources available to remove disinformation from their platforms. In order to distinguish misinformation from disinformation and to follow the leads to other accounts involved in the disinformation campaign, professionals with extensive knowledge in the nature of disinformation are needed to investigate, monitor and deal with disinformation. Without enough resources or dedication to fighting disinformation, it easily falls under the radar of content moderators. Whilst such teams do exist, for example at Facebook, they are not a priority in the business model and their effectiveness remains questionable.[176] Given that under the DSA platforms are not ordered to take actions specifically to disinformation, no obligation to proactively investigate illegal content exists,[177] and no liability for platforms exists for content published by users,[178] the DSA also does not encourage or oblige platforms to properly tackle disinformation.

[176] Steven Lee Myers and Nico Grant, *Combating Disinformation Wanes at Social Media Giants,* N.Y. TIMES (February 14, 2023), https://www.nytimes.com/2023/02/14/technology/disinformation-moderation-social-media.html.
[177] 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) [2022] OJ L277/1, art 6.
[178] Id., art 7.

Disinformation: can legislation alone fight such a complex enemy?

Even if the DSA included obligations specific to fighting disinformation, it is unlikely that a legal framework can truly halt disinformation campaigns. This is attributable to the nature of disinformation, the speed at which it disseminates and the response of the audience.

The first barrier to the ability of legislative approaches to effectively reduce disinformation is that disinformation flexibly adapts to the restrictions put in place. Firstly, disinformation is not reliant on certain platforms to spread. If one social media platform removes certain content, the users can migrate to a different platform, which may not have tracked this disinformation campaign yet.[179] And when that social media platform removes the disinformation as well, users can even migrate to niche platforms, portals, or blogs. As such, "disinformation-related risks are endemic to the whole ecosystem" and not just to specific platforms.[180] Secondly, even if disinformation is tracked effectively, it can be hidden to avoid the detection of AI moderation tools.[181] Instead of using hashtags that are directly linked to the topic of disinformation, communities can use different ones to evade simple monitoring.[182] This means that AI tools used to detect disinformation must be updated regularly to adhere to the fast adaptation of dissemination. Thirdly, the use of bots in disinformation campaigns and the existence of AI which speeds up the production of disinformation means that content moderators cannot catch up to disinformation; it is simply produced, published, and disseminated too fast.

[179] Paolo Cesarini, *Countering disinformation: Is the DSA Punching Below its Weight?,* EURACTIV (February 19, 2021), https://www.euractiv.com/section/digital/opinion/countering-disinformation-is-the-dsa-punching-below-its-weight/.
[180] Ibid.
[181] Iris Malone, *Will the EU's Digital Services Act Reduce Online Extremism?,* JUST SECURITY (May 16, 2022), https://www.justsecurity.org/81534/will-the-eus-digital-service-act-reduce-online-extremism/.
[182] Ibid.

Moreover, a root cause of the success of disinformation campaigns is that it divides already existing social and political divisions and exploits a population's distrust in conventional media and the government. To put it simply, if the population were united in their opinions, and would continuously believe traditional news sources, then disinformation could not grow into a threat to democracy. However, especially since the COVID-19 pandemic, trust in the government[183]and traditional media outlets has been challenged,[184] making the population vulnerable to intentionally false information. Such individuals can effectively be targeted by disinformation campaigns.[185] These sociological factors are only amplified by the structures of social media, as discussed in a previous chapter, and the overload of accessible and recommended information, which require constant questioning. Therefore, a legislative act may change the behaviour of the intermediaries but cannot address social and political beliefs.

As such, the DSA can tackle the response of social media and hence the speed at which disinformation is disseminated. Given the wide scope of obligations imposed on platforms and the extensive regime of enforcement, it may be possible that disinformation is reduced, even if limitations exist. However, it cannot dive into the root issue of disinformation: its production and a society's willingness to believe it. As such, any attempts at reducing the impact of disinformation can only show effective results if society is united internally and can accept governmental institutions and those within them.

[183] Cécile Jacob, Pierre Hausemer, Adam Zagoni-Bogsch, Audra Diers-Lawson, *The effect of communication and disinformation during the COVID-19 pandemic,* EUROPEAN PARLIAMENT (2023), 20, https://www.europarl.europa.eu/RegData/etudes/STUD/2023/740063/IPOL_STU(2023)740063_EN.pdf.

[184] Alice Marwick and Rebecca Lewis, *Media Manipulation and Disinformation Online,* DATA & SOCIETY RESEARCH INSTITUTE (2017), 40, https://www.posiel.com/wp-content/uploads/2016/08/Media-Manipulation-and-Disinformation-Online-1.pdf.

[185] Samantha Bradshaw and Philip N. Howard, *Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation*, UNIVERSITY OF OXFORD (2017), 10, https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2017/07/Troops-Trolls-and-Troublemakers.pdf.

Conclusion

There is no denying that the Digital Revolution has forever changed our perceptions of and our interactions with the world around us. Whilst this has brought significant advantages, it has also created new threats, as well as amplified old ones. Such is the case with disinformation. Social media has provided a digital environment where disinformation can foster, exploiting deep-rooted fears, contradicting beliefs and political polarization, and resulting in a serious threat to democracy and its institutions.

There is no doubt that the DSA serves as a landmark Regulation to protect consumers of the digital world in a manner and to an extent that has not existed prior. However, it is likely that the DSA's effectiveness in providing protection from disinformation is only limited. There are structural criticisms that can be made of the DSA, however, its content cannot solely be criticized for this lack of impact. Disinformation is an inherently complex phenomenon whose threats cannot wholly be solved by a Regulation that does not even target it directly. To reduce the threat of disinformation, a wide range of actions need to be taken by the EU and further obligations must be imposed on social media platforms to target the dissemination of information that is threatening to democracy. As such, the DSA is a revolutionary piece of legislation that has the potential to influence legislative systems abroad, however, it may only slow down the speed of dissemination, not prevent it.