



UvA-DARE (Digital Academic Repository)

The EU's regulatory push against disinformation

What happens if platforms refuse to cooperate?

Fahy, R.; Appelman, N.; Helberger, N.

DOI

[10.17176/20220805-182037-0](https://doi.org/10.17176/20220805-182037-0)

Publication date

2022

Document Version

Other version

License

CC BY-SA

[Link to publication](#)

Citation for published version (APA):

Fahy, R. (Author), Appelman, N. (Author), & Helberger, N. (Author). (2022). The EU's regulatory push against disinformation: What happens if platforms refuse to cooperate?. Web publication or website, Verfassungsblog. <https://doi.org/10.17176/20220805-182037-0>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

The EU's regulatory push against disinformation

Ronan Fahy

2022-08-05T10:42:36

Tech billionaire Elon Musk's [surprise bid to buy Twitter](#) on 14 April 2022 questions the wisdom of the current EU efforts to combat the spread of disinformation, which has relied to a large extent on platforms' voluntary cooperation. After a whirlwind few weeks, Twitter's board [accepted his offer](#) on 25 April 2022. However, Musk [pulled out of the deal suddenly](#) at the beginning of July, with [Twitter suing](#) him in response. Even though what will happen with the takeover is uncertain, these tumultuous developments shed light on the potential weaknesses of the EU's voluntary approach to disinformation.

One motivation behind the takeover, according to Musk, was to [protect free speech](#), with Musk calling Twitter the digital equivalent of a "town square" where people should be able to speak freely. Musk, who calls himself a "[free speech absolutist](#)", has often repeated his [views on content moderation](#) which he feels [should not go further than](#) the law. Essentially, Musk would abolish most existing efforts to fight disinformation, as most EU measures on disinformation rely on voluntary actions on the side of the platform, given that disinformation is in most cases only considered "harmful" but not unlawful by the law. This abrupt change of policy, were it to happen, would take Twitter's content moderation policies back more than a decade to when it [described itself](#) as "the free speech wing of the free speech party". Only moderating illegal content would have far-reaching consequences for the spread of disinformation, which is in most cases [considered legal](#) yet potentially very harmful. Even though Musk has partially changed course, stating in an [awkwardly staged video](#) with the European Commission's Thierry Breton that he "agrees with everything in the DSA", the potential abrupt change in leadership and content moderation policies raises serious questions on EU disinformation policy's reliance on platforms' discretion to moderate this category of speech. It is likely to put pressure on the carefully constructed web of self- and co-regulatory measures and legislation the European Commission has spun to counter the spread of disinformation. For example, in June 2022, the Commission [presented](#) the Strengthened Code of Practice on Disinformation, which Twitter co-signed, while the EU's landmark Digital Services Act ("[DSA](#)") has only been [agreed](#) to in April 2022.

In this context, this piece aims to map out the possible consequences and measures available to member states and the EU, should a platform like Twitter decide to *not* intervene on disinformation. Most of the legal debate has focused on the space to regulate disinformation in light of freedom of expression, with, for example, national court [rulings](#) that platforms have the freedom to ban disinformation if they [choose](#). Given the growing complexity of the EU regulatory framework, it is an open question to what extent [Very Large Online Platforms](#) (known as VLOPs) such as Twitter also have the freedom to allow disinformation, should they want to.

1. EU intervention

The first question to consider is whether a platform like Twitter will actually have the freedom to not intervene against disinformation under the DSA, and what tools the DSA offers to address such a platform policy. Crucially, the question of how much leeway Musk hypothetically has, at least for the territory of the EU, will largely depend on the interpretation of Articles 26 and 27 DSA.¹⁾ According to Article 26 DSA, providers of VLOPs shall diligently identify and mitigate any systemic risks stemming from the “design, including algorithmic systems, functioning and use made of their services in the Union”. The spread of disinformation could constitute such a systemic risk in three situations:

1. An EU member state has declared certain forms of disinformation as unlawful (Art. 26 (1)(a) DSA). As we explain below, this is indeed the case in a number of EU member states;
2. The dissemination of disinformation has “actual or foreseeable” negative effects for the exercise of fundamental rights (Art. 26(1)(b) DSA). One possible example could be forms of disinformation coupled with illegal hate speech that incite hatred toward certain (groups of) people, and therefore conflict with the right to human dignity. A complicating factor here is that many forms of disinformation will in themselves be protected by freedom of expression, and that the operationalisation of this provision can be extremely difficult in practice. Especially, seeing that fundamental rights have so far been applied between states and individuals. Accordingly, little experience or effective procedures exist yet to identify situations in which the conditions of Art. 26 (1)(b) DSA are fulfilled;
3. Situations in which disinformation has an actual or foreseeable negative effect on the civic discourse and electoral processes, or public security (Art. 26 (1) (c)). Given the potentially disruptive effect of disinformation on public discourse, alongside the increasing framing of disinformation as a matter of public security, many if not most forms of disinformation could potentially fall under Art. 26 (1)(c) DSA.

According to Art. 26 DSA, it is up to VLOPs to carry out risk assessments, and adopt effective risk mitigation measures according to Art. 27 DSA. If a platform is led by someone, like Musk, who believes that disinformation does not have a negative effect on the public sphere, the DSA does seem to leave Musk considerable leeway to lead the platform by his convictions. The difficult task of proving that disinformation does have a negative effect on the civic discourse will then fall to the independent auditors (Art. 28 DSA) and the European Board for Digital Services (made up of national Digital Services Coordinators), in cooperation with the Commission when assessing the risk assessment (Art. 27 (2) DSA). Seeing the difficulty of defining adequate metrics to identify and measure such effects reliably, the devil with Art. 26 DSA probably sits in enforcement.

Further measures

A more potent tool for the European Commission to intervene could be the new crisis response mechanism in Art. 27a DSA. It grants the Commission considerable

leeway in (temporarily) intervening in the content moderation decisions of a platform such as Twitter. The ability to do so is limited in time (not exceeding three months) and reserved to crisis situations. “Crisis” is defined rather broadly as “extraordinary circumstances” that can lead “to a serious threat to public security or public health” (Art. 27a(2) DSA). The definitional broadness is one reason why Art. 27a DSA has been heavily [criticised](#), for example, by civil society organisations.

The second issue to consider is that on 16 June 2022, Twitter officially signed the [2022 Strengthened Code of Practice on Disinformation](#) (“2022 Code”), committing itself to implement a whole slew of new measures. The 2022 Code follows the [2018 Code of Practice on Disinformation](#) and seeks to implement the objectives the Commission set out in its [2021 Guidance](#) to strengthen the Code, even though the Commission is, again, at pains to emphasize that it is a self-regulatory effort. The 2022 Code creates an extensive framework with measures ranging from demonetization, transparency, user and researcher empowerment, fact-checking, and integrity of services. For example, signatories have committed to adjust their recommender systems “to improve the prominence of authoritative information and reduce the prominence of Disinformation”. Out of the Code’s 44 commitments and 128 measures, signatories could choose whether to commit to each of them individually, with [Twitter committing](#) to 109.

Although the 2022 Code is considered a self-regulatory measure, it is intimately connected to the DSA. The preamble explicitly states that it “aims to become a Code of Conduct under Article 35 of the DSA, after entry into force, regarding [VLOPs]”, and that “signing up to all Commitments relevant and pertinent to their services should be considered as a possible risk mitigation measure under article 27 of the DSA”. This intertwinement questions the extent to which a platform could abandon the commitments it has voluntarily made. Preamble 68 of the DSA states that “the refusal without proper explanations by a provider of an online platform [...] to participate in the application of such a code of conduct could be taken into account, where relevant, when determining whether [it] has infringed the obligations laid down by this Regulation.” The 2022 Code as an industry standard, and the clear key performance indicators it provides, could mean that it will play a role in the Commission’s judgement of a platform’s compliance with the DSA. However, the Commission and independent auditors’ enforcement powers are very limited. Clearly, the self-regulatory nature of the codes means platforms could withdraw at any moment. The [agreed text](#) of Article 35 DSA explicitly provides that the codes of conduct are “voluntary”. However, the Commission’s explicit endorsement of the 2022 Code, as well as its setup to function as a possible risk mitigation measure for complying with Article 27a DSA, could mean that it is, in practice, difficult for VLOPs to completely abandon taking any voluntary measures on disinformation.

The third point, following Russia’s invasion of Ukraine, is that the Council of the EU has recently been implementing measures targeting disinformation *far beyond* those under the Code of Practice. Namely, banning certain Russian media outlets, including on online platforms. This first occurred in March 2022, when the Council adopted a [decision and regulation](#) prohibiting Sputnik and Russia Today in the EU, with a further three media outlets [added to the list](#) in June 2022, and the

General Court [rejecting](#) Russia Today's challenge to the ban. These bans were implemented (without a court order) on the basis of the Council stating that the Russian government was engaged in "gravely distorting and manipulating facts" through a "systematic, international campaign of media manipulation" in order to "destabilis[e]" the EU and its member states, and had repeatedly targeted European political parties, civil society, and the "functioning of democratic institutions" in the EU. The broad scope of the ban included a prohibition on broadcasting "any content" by the banned media outlets, including "through transmission or distribution by any means such as cable, satellite, IP-TV, internet service providers, internet video-sharing platforms or applications". Online platforms, including YouTube, Facebook and TikTok, all responded by removing Sputnik and Russia Today channels from their platforms, and it was later [revealed](#) that the European Commission had [pressured](#) platforms to remove these channels, including search engine results. Further, Google [removed](#) RT and Sputnik's apps from the Play Store, while Apple also [removed](#) the apps from its App Store.

While these measures have been seriously questioned by [journalist organisations](#), [media law scholars](#), and [human rights organisations](#), the main point here is that the EU has used one of the most draconian tools available to target disinformation – an executive order banning an entire publication, without a court order. With the new Art. 27a DSA, there is a real risk the European Commission will retain considerable discretion to engage in similar interventions against an online platform in the future, especially, where the Council finds that it allowed "gravely distorting and manipulating facts" to circulate, which target EU populations, groups, or democratic institutions. The risk is also heightened where a platform may be connected to a foreign government. Indeed, during the Covid-19 pandemic, the European Commission adopted a [Joint Communication](#) specifically finding that the Chinese government engaged in "disinformation campaigns around COVID-19 in the EU", and "seeking to undermine democratic debate and exacerbate social polarisation". More recently, the governments of Poland, Lithuania, Latvia and Estonia [demanded](#) that YouTube, Facebook, and Twitter remove more Russian disinformation, which had been "tolerated" on these online platforms "for years". Platforms were now an "accessory to the criminal war of aggression the Russian government is conducting against Ukraine and the free world". In a future crisis, it's conceivable that a platform may be targeted for adopting a policy of not intervening on disinformation.

2. Member State intervention

A further question is whether EU member states could take measures against a platform over its policy on disinformation. The first point on this question, as highlighted in our recent [article](#) in [Internet Policy Review](#), is that numerous EU member states have national laws on disinformation, false information, and false news, including criminal laws. In 2020, the European Commission specifically singled out Hungary for introducing a "[new criminal offence for spreading of disinformation](#)". However, it is not just Hungary; numerous EU member states have laws on false information and false news. For example, Malta's [criminal code](#) makes it an offence to "maliciously spread false news which is likely to alarm public opinion or disturb public good order or the public peace". There are similar national

provisions in Croatia, France, Greece, Romania, the Slovak Republic, the Czech Republic, and Cyprus.

These laws have been subject to criticism. Romania, for example, was [criticised](#) by the Council of Europe's Commissioner for Human Rights, for its decree during the pandemic, allowing the authorities to block websites with "false information". Recently, Lithuania's media regulator issued [orders](#) for ISPs to block access to numerous websites, including for "disseminating disinformation". And ISPs in the Netherlands blocked access to the websites of Russia Today and Sputnik following the EU Council's ban, and the ISP trade association issued a [statement](#) that it had advised its members to block the websites "under protest", criticising the "extremely unclear" scope of the ban.

Notably, these laws can be operationalised against platforms under the DSA. Article 8 DSA will create an explicit legal mechanism for national judicial and administrative authorities to issue orders for online platforms to "act against" "illegal content". Crucially, illegal content is defined quite broadly under the DSA, including "any information" "not in compliance with Union law or the law of a Member State, irrespective of the precise subject matter or nature of that law". It thus captures all national provisions applicable to disinformation. However, a notable scenario might arise were a platform to publicly adopt a policy of not intervening on disinformation, though ordered to do so by an administrative authority in an EU member state which criminalises disinformation. Article 8 DSA includes language clarifying that the orders concern national law "in compliance with Union law", which leaves room for a platform to refuse to follow an Article 8 order because it considers that it violates EU fundamental rights law on freedom of expression. Even more interestingly, Twitter under Musk could adopt the argument that under [international and European human rights standards](#), laws on disinformation violate freedom of expression. A Twitter policy of not intervening on disinformation would actually be more consistent with these standards than certain EU member states' policy.

3. Implications

A number of implications arise. First, until now, there has been strong reliance on the voluntary cooperation of platforms in the tackling of disinformation. Elon Musk's bid for Twitter and rejection of the voluntary commitments that Twitter has made so far is a wake-up call, showing how fragile the current arrangement can be, and also how much it can depend on the political convictions of an individual person in charge. Second, another interesting implication is what Musk's pledge to uphold free speech means vis-à-vis national efforts to enforce (questionable) disinformation laws, challenging freedom of expression. Will Musk be a defender of free speech rights, if so, how, and who decides where to draw the line? Third, should Musk come in conflict with some national disinformation laws, then much will depend on enforcement, and whether the DSA enhances the power of member states to enforce their national disinformation rules. Depending on the type of disinformation, it will be more difficult to escape liability under Art. 26 DSA if disinformation is indeed treated as a systemic risk. While, as *ultima ratio*, the Commission has reserved

the right to initiate heavy-handed interventions under to the new Art. 27a “crisis” mechanism.

Further, if a platform decides to adopt a policy of non-intervention over disinformation, there may be a shift to pressurizing Apple and Google to remove this platform from their app stores, and a shift in focus on Apple and Google’s rules on apps being required to moderate harmful content. This happened, for example, when Apple and Google [removed the Parler app](#) from their app stores for failing to “[moderate or prevent the spread](#)” of harmful and illegal content. Both Apple and Google have long succumbed to pressure to remove apps, including from the [Chinese](#) and the [Russian](#) governments. What we could now see is EU governments pressuring Apple and Google to remove apps over non-moderation of specific types of content, including disinformation. There will also be an increased focus on the rules Apple and Google set for social media apps. For example, Apple [requires](#) social media apps have mechanisms for “filtering objectionable material from being posted” (including “false information”), and mechanism reporting content and “timely responses to concerns”. Finally, an interesting and still open question is how other platforms will respond if one decides to step out of the EU Code of Conduct. There is some evidence from the Parler case that the pressure and repercussions from other platforms could potentially be as or even more effective than the European Commission, and certainly member states.

4. Conclusion

In sum, it does seem that if Twitter under Elon Musk were to engage in a shift away from intervening on disinformation, the upcoming DSA provides enough wriggle room for such a policy shift to occur, while the revised EU Code of Practice on disinformation still remains largely voluntary. However, there remain some policy options, although drastic, for the European Commission and member states to take. But a Musk takeover of Twitter and possible shift in policy toward disinformation does raise a more fundamental point: when the political economy of online platforms does not align with EU regulatory goals, there is a genuine risk of the whole regulatory system coming under intense pressure.

References

- According to Art. 1a DSA, the Regulation does apply to Twitter because individuals in European are recipients of Twitter’s services.

