DETECT THEN ACT

TECHNICAL REPORT 1

ISSN 2736-6391

NEW HORIZONTAL RULES FOR ONLINE PLATFORMS ACROSS EUROPE

A COMMENT ON THE COMMISSION'S PROPOSAL FOR A DIGITAL SERVICES ACT FOR DTCT PARTNERS AND UPSTANDERS

WOLF J. SCHÜNEMANN

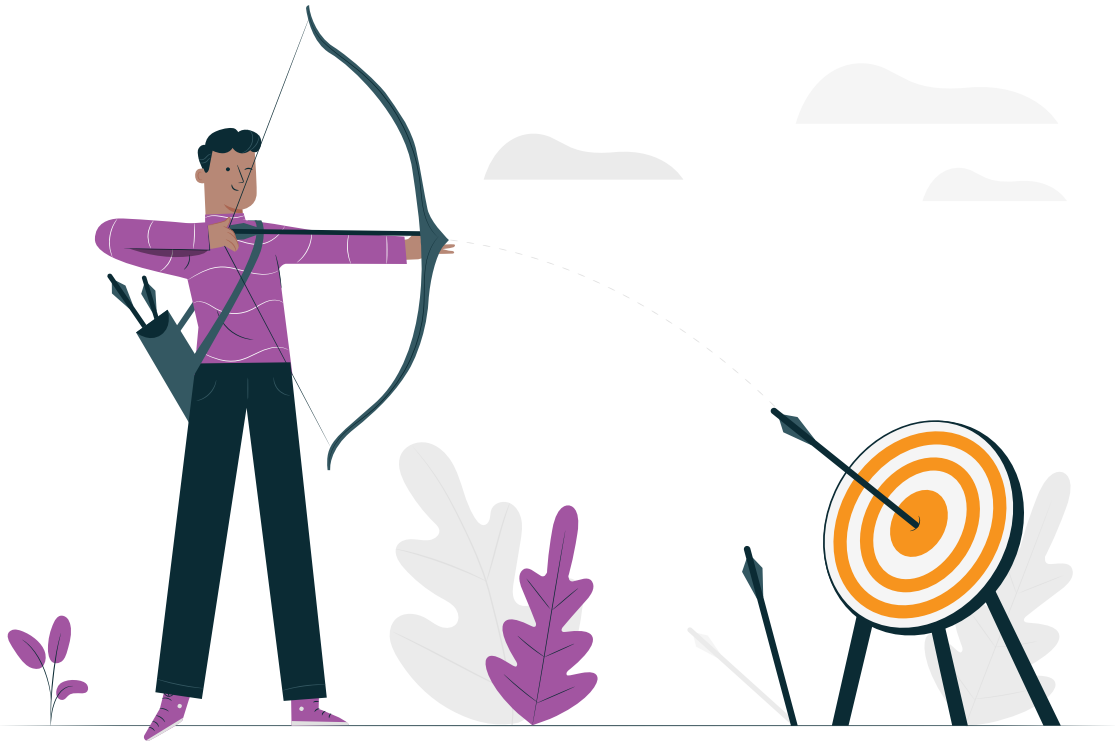
Mid-December 2020, the EU Commission presented a legislative package for the comprehensive regulation of the Digital Single Market. It includes two concrete proposals for EU Regulations: the Digital Markets Act (DMA)[1] and the Digital Services Act (DSA)[2]. While this development is of general interest as the EU continues on its way to updating and defining the basic rules for the digital economy and society in the 21st century with potential standard-setting effects beyond the scope of the Union, the DSA in particular will have important implications for the fight against hate speech. Thus, the *DTCT - Detect then Act* consortium must be aware of the process and should closely follow EU policy-making in the years to come. In order to prepare for its future work, DTCT partners should know about the essential provisions of the proposal with respect to their mission. The likely effects for their future work should be discussed with partners and upstanders in order to define a proper outlook and develop follow-up activities. This document is meant to serve this purpose by informing about the legislative initiative, by highlighting and commenting on the essential provisions of the DSA. Given the complexity of the issue, the author does not guarantee for a full coverage of all relevant aspects. Not just for that reason, all partners in the network are encouraged to take this document as a call for activity and make additional contributions to the discussion of current and future EU policy-making in the field.

[1] Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on contestable and fair markets in the digital sector (Digital Markets Act), COM(2020) 842 final.

[2] Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC, COM(2020) 825 final.

GENERAL OBJECTIVES

The Digital Services Act, if approved, will amend the e-Commerce Directive from 2000[3]. It will thus be the first EU regulation that explicitly addresses the more recent developments of digital platforms and communication spheres, in particular the widely discussed threats for society and democracy such as online hate speech. At the EU level, the member states' intensified activities of content regulation and legal requirements imposed on digital platforms are observed with growing concerns regarding the assumed negative effects for the Digital Single Market such as fragmentation. Acknowledging their crucial role for the current state of digitalisation and thus the development of societal, economic and democratic life, the proposal sets harmonised horizontal rules for providers of intermediary services, thus new responsibilities and obligations for all sorts of Internet Service Providers (ISPs) and Content Service Providers (CSPs). These intermediaries allow for or facilitate the use of the internet and digital services and thus stand between the citizens as users or customers on the one hand and the producers and providers of goods, services or information on the other hand. Given the wide range of providers of intermediary services and the varying risks their conduct bears for societies, the proposal follows a risk-based approach with graduated provisions depending on the type of intermediary service but also its size or reach.



[3] Directive 2000/31/EC.

One of the main goals of the proposed Regulation that affects the mission and practices of DTCT directly is to harmonise guidelines for the platforms' content moderation. This includes both the obligations that providers have with respect to illegal content on their platforms and obligations with respect to the content moderation that platforms exert on the basis of their own terms and conditions. As to the definition of illegal content, which naturally goes beyond hate speech, the Regulation does not prejudice nor overrule any definitions of illegal content as set out by other pieces of Union or Member State Law. Hence, the DSA does not include any kind of harmonised catalog of what constitutes illegal content that would be applicable across the Union. As to the question of what would be illegal in which jurisdiction, trainers and upstanders would thus have to turn to the known resources including the online hate speech manual prepared by *DTCT - Detect then Act*.[4] However, the proposed Regulation designs a mandatory framework on how content moderation has to be exerted including its documentation, its transparency and standardised ways of action and redress mechanisms. It is explicitly stated that no removal obligations are formulated nor envisaged for content that is considered harmful but is legal. For fighting or countering other harmful content, the DSA rules will be complemented by the existent co-regulatory and self-regulatory measures or new efforts to be taken under the *European Democracy Action Plan*.[5]

Moreover, in addition to the harmonised rules for content moderation, the proposal includes innovative provisions that would oblige platforms to guarantee greater transparency and accountability with regard to filter and recommendation algorithms and online advertisement. As the applied techniques of a digital attention economy are widely perceived as potential catalysts for detrimental phenomena (trolling, social bots, targeted advertising), the defined rules could have positive implications for the fight against hate speech and extremism too as they shall help to illuminate hidden and opaque practices of attentional steering. Moreover, empirical research is likely to benefit from the greater transparency required from providers. Data access for research is explicitly mentioned as a goal that should come with practical advantages for researchers, as it might improve data access also to some of the key platforms that do not grant sufficient transparency today. Thus, researchers might be enabled to play their essential role in providing qualified information and transparency on digital communication processes and the appropriate conduct of platforms.

[4] DTCT (2020). Online Hate Speech: Introduction into motivational causes, effects and regulatory con-texts, https://dtct.eu/wp-content/uploads/2020/11/DTCT_Online-Hate-Speech.pdf

[5] COM(2020) 790 final.

BASIC PRINCIPLES

Which are the basic principles of the DSA?

The first of the crucial principles of the DSA is the so-called marketplace principle that underlies the definition of scope for the Regulation. This kind of meta principle is already known as an innovative element from the GDPR for which it has allowed for an effective protection of fundamental rights and extraterritorial effects beyond the original (i.e. territorial) scope of the Union. The marketplace principle supersedes the former territoriality principle which – given transnational data flows or content and services provisions made by providers in foreign jurisdictions – would simply not live up to the constitutive features of digitalisation. The obligations as set in the proposal should thus include all actors that provide goods, information or services in the Union, regardless of their place of establishment. In contrast to the GDPR, however, for which infringements affect the personal rights of an individual, the marketplace cannot be defined in a manner as straightforward here because the practices of providers would be affected by the DSA in a much more general way. Nevertheless, the proposal lists a number of criteria that would qualify a service provider as addressing the EU market and thus as falling under this Regulation: significant reach within the Union, the use of language or currency, the use of a Top Level Domain or the availability in App Stores.

Secondly, the proposed Regulation takes a risk-based approach according to which it sets asymmetric rules for intermediary services depending on their type and size or reach. The latter differentiation is reasoned with the necessity to protect small and medium-sized companies from bureaucratic burdens in order not to hamper innovation and competition. Consequently, in addition to the general provisions proposed for all intermediary services, the proposal includes specific rules for all providers of hosting services including online platforms; the latter again will have more specified and in some ways stricter obligations than other hosting services. And finally there are additional obligations envisaged for very large online platforms, defined as those platforms that have an average monthly reach of 45 million users or higher within the EU.[6] This condition would be fulfilled only by some of the leading intermediaries like Google, Facebook, Twitter and the like.

Another fundamental, yet more specific principle regarding the objectives of this regulation would be the decision to hold on to general liability exemptions for platforms and hosting providers when hosting or spreading third party content. Thus, there are no ex-ante obligations made for providers for cases in which they are not co-responsible for the content available on their platforms.

[6] The number is supposed to correspond to 10% of the EU's population. In case of significant changes to the population size the threshold can be adapted by the Commission through a delegated act.

Fourthly, this entails a pertained prohibition of general monitoring obligations for intermediary services. There will be no general obligation for total monitoring by intermediary services of the content transmitted or stored through their platforms. Providers will not be obliged to actively seek for illegal activity. This affects the practices towards illegal content, e.g. illegal hate speech, including the application of upload filters or other automated detection tools.

Finally, in order to protect user rights and the freedom of expression, anonymity of users shall be guaranteed wherever possible in the online environment.

MEANINGFUL PROVISIONS FOR DTCT PARTNERS AND UPSTANDERS

The Regulation will set harmonised due diligence obligations for providers, as they in particular would have to take immediate action upon information on illegal activities on their platforms. Such notice-and-action mechanism shall be facilitated, streamlined and harmonised throughout the European Single Market. The obligation imposed on intermediary services to set up easily accessible and user-friendly notice-and-action mechanisms following harmonised guidelines would probably have the most substantial effect on DTCT partners and upstanders in their fight against illegal hate speech. For co-ordinated activities as exerted by DTCT and some of its partners it would be particularly helpful if the required possibility to hand in multiple notices at once made it into the final version of the Regulation.

The DSA also requires providers of intermediary services to set up accessible, user-friendly and transparent complaint-handling systems and redress mechanisms. Moreover, there shall be possibilities for out-of-court dispute settlement in case of ongoing controversies.

Furthermore, the legislation will facilitate the work of DTCT partners and upstanders with respect to compliance, reliability of providers and acknowledgement of their work because providers will be obliged to respond to their notices in due course. As providers of intermediary services under the DSA will have to establish a single point of contact to which governments and other entities can turn with regard to content moderation practices, at least the corporate actors among DTCT partners might have much better chance to contact providers and to get a response in return.

Moreover, the DSA takes up the idea and role of so-called "trusted flaggers" as known from the platforms' self-regulatory arrangements already. In future, entities, thus corporate actors including civil-society organisations,[7] that have shown expertise and competence in tackling illegal content online shall be able to apply for registration as trusted flaggers. Applications would then be submitted to a newly established official authority, the Digital Services Coordinators at the national level. Having their status officially acknowledged, notices of trusted flaggers would have to be given priority by providers, and they would have the privilege to directly turn to the single point of contact at a given platform.

As to the Digital Services Coordinators, these new institutions would be established or appointed under the DSA as independent public authorities at the national level. The model of Digital Services Coordinators with their role in monitoring and sanctioning resembles the role of Data Protection Authorities. Like the latter, the new national coordinators would gather in a European Board for Digital Services that would serve as an independent advisory and coordinating body at the Union level.

Online platforms will have to abide to new harmonised transparency rules which include the obligation to publish a report on their content moderation activities including information on submitted notices, the action that has been taken and on complaint management. Reports would have to cover both the treatment of illegal content and of content that is contrary to their own terms and conditions.

Finally, as highlighted above, researchers must be granted better access to data sources that would help illuminate the content moderation, filtering, recommendation and advertisement practices of the providers of intermediary services. Transparency reports would also have to include information with respect to online advertisement.

[7] The glossary for the DSA published on the Commission's website does also mention individuals as eligible for trusted flaggers status, but this contradicts the actual provision as made by the proposal, see https://ec.europa.eu/digital-single-market/en/glossary-digital-services-act.

RECOMMENDATIONS FOR DTCT PARTNERS AND UPSTANDERS

The Digital Services Act as part of a package proposal with the Digital Markets Act requires the subsequent approval of the European Parliament and the Council of the EU. Both legislative organs can still make substantial amendments to the document until they have to find a compromise and approve the final version. This process will take about one and a half years at least so that the actual Regulation is to be expected to enter into force by mid-2022. Following the example of the GDPR, the legislative chambers might agree on a longer transition period after its entry into force so that the rules will become binding only later.

However, it seems rational for DTCT partners and upstanders to prepare for the upcoming policy development. The following recommendations are based on the current proposal. DTCT partners and upstanders should:

- **…closely follow the policy-making process.**
- **…actively gather information on the proposed standard notice-and-action and complaint management mechanisms.**
- **…prepare for engagement in the implementation of the DSA.**
- **…closely observe the national institution-building in particular with respect to the Digital Services Coordinator.**
- **…in case of being entities, they should discuss their willingness and eligibility to serve as trusted flaggers according to the newly set rules and gather information on the registration process.**
- **…keep an eye on new opportunities for data access given to researchers.**