Center for Democracy & Technology's Response to the European Commission's Consultation on the Digital Services Act Package Supporting Document

I. Safety and responsibilities - How to effectively keep users safer online?

1. Main issues and experiences

C. Activities that could cause harm but are not, in themselves, illegal

Question 5:
What good practices can you point to in tackling such harmful activities since the outbreak of COVID-19?

Digital services have played an essential role in helping European societies respond to the COVID-19 crisis. Online communications services have enabled people in lockdown to work from home, continue their education, and stay in touch with family and friends. Digital platforms have helped authorities provide essential healthcare information, and many have taken measures to counter COVID-19 related misinformation, scams, and fraud. In this sense, the pandemic has demonstrated how much modern societies rely on online services, which in turn shows the legitimate public interest in setting the right framework for how they operate – the objective of the Digital Services Act.

Examples of good practices can be found, for instance, in Wikipedia. With the aim to prevent the spread of dangerous misinformation, the organization has applied stricter editorial standards to articles regarding the pandemic, meaning that unregistered users are restricted from editing, and actual editors need to have acquired a certain amount of experience. In the area of online marketplaces, we have seen enhanced action against price gouging and increased support for small businesses. In March alone, Amazon suspended thousands of accounts for violating its pricing policies and removed hundreds of thousands of items from its website. Walmart took a similar action by automatically removing listings that were priced substantially higher than other listings. In addition to blocking or removing items that were making false health claims, eBay also launched an accelerator program to empower retailers without an e-commerce presence to transition to selling online. Social media services have also taken various steps to limit the spread of misinformation. For instance, Facebook has been directing its users to resources from the WHO and other reliable health authorities and expanded its fact-checking program for reviewing and rating content.

Many platforms, including Facebook and YouTube, have intensified the use of automated tools for content moderation, because human reviewers have been unable to work as usual. This reliance primarily on automation should be understood as a strictly emergency voluntary measure, not a blueprint for regulation, as automated content moderation often leads to overbroad restriction of content with negative implications for freedom of expression and access to information (see additional discussion in section 2, question 6). This demonstrates that human review continues to be essential for making the difficult judgments necessary for rights-respecting moderation. And it shows the need to ensure that responsible content moderation is conducted with transparency, accountability, and fairness, including the right to appeal. To better understand the impact of these measures, CDT together with other 75 organizations and researchers published an open letter to social media companies and content hosts (more information in question 19, section 2 of this module).

D. Experiences and data on erroneous removals

Question 1:
Are you aware of evidence on the scale and impact of erroneous removals of content, goods, services, or banning of accounts online? Are there particular experiences you could share?

Moderation on user-generated content (UGC) sites today often relies on a combination of user reporting ("flagging") and automation. Neither human nor automated moderation processes are infallible—both can make mistakes. There are also a variety of ways that a content moderation system can lead to an erroneous outcome. First, there may be error in consistently applying a service provider's content policy. Service providers need to adequately train their moderation staff, and to test their automated systems, to ensure that the decisions reached by the human and machine portions of their moderation system reach the provider's intended outcome in almost every case. Even with exemplary training and high consistency rates, however, error will still occur in content moderation because the inputs—the user-generated content—to that system are constantly changing. There also may be error in how the provider develops its policies. It may fail to account for particular scenarios, or to grasp the nuances of a specific culture, language, or context. In such cases, the letter of the policy may be applied correctly, but the outcome may still be contrary to the goal or intent of the policy.

Automated tools used for content moderation are prone to additional types of error. Examples range from erroneous content takedowns, mass account suspensions, misinterpretations of copyright infringement, wrong language translations and more. Some error in automated tools is due to the fact that even sophisticated machine-learning tools are not able to take the full context of a post into account, and thus essentially miss important information about the actual meaning of the post. (Human moderators, too, may fail to take context into account, especially if they are only provided with limited information about the account or surrounding content during the review process.) But machine learning tools can also “learn” a kind of error in the form of bias: tools that are trained on real-world data sets may pick up on inequities that exist in the world, and reproduce those biases in their own classifications. This is the kind of error that leads natural-language processing algorithms to conclude, for example, that “man is to computer programmer as woman is to homemaker”, or to identify positive, affirming speech by drag queens as “toxic”.

These various types of error in content moderation systems can be mitigated to a certain extent, by improving the quality of training for moderators and instituting regular processes for evaluating the results of the system. But error will never be entirely eliminated from content moderation—human communication is simply too complex and dynamic. Thus, it is imperative that any content moderation system includes a robust, transparent appeals process, including notifications to users about the reasons that their content has been removed or their accounts deactivated, and the opportunity to provide explanations or additional information. Any decisions on the legality of speech must remain the sole purview of the courts.

Question 8:
Does your organisation access any data or information from online platforms?

Yes, generally available transparency reports.

Question 9:
What data is shared and for what purpose, and are there any constraints that limit these initiatives?

CDT is a longtime advocate for transparency from technology companies about their treatment of user data and user-generated content. CDT sees a clear and compelling need for independent researchers to be able to access information held by these companies to conduct essential research into the dynamics that shape our online information environment. For example, during the COVID-19 pandemic, we have seen an unprecedented increase in the usage of automated tools compared to traditional human content moderation practices (the opportunities and risks of automated tools are further discussed in question 6, section 2 of this module; learnings from the COVID-19 pandemic can be viewed in question 5, section 1C - legal but harmful activities). This represents an opportunity to study how online information flows ultimately affect health outcomes, and to evaluate the macro- and micro-level consequences of relying on automation to moderate content in a complex and evolving information environment. For this reason, CDT and 75 other organizations and individual researchers have published an open letter that urges platforms to preserve this data so that it can be made available to researchers and journalists and included in the companies' transparency reports.

There has been a welcome increase in voluntary transparency reporting across the Internet industry. These reports take various forms and usually include a version of statistics on the enforcement of the company's Terms of Service and requests from governments to take action on specific pieces of content. Both metrics are important to show the amount of restricted content online, including erroneous removals, as well as the treatment of user data and privacy. For instance, Facebook provides statistics on its content moderation practices in 5 categories, including the enforcement of its Community Standards and Government Requests for User Data and Content Restrictions, i.e. content that is alleged to violate local law but does not violate the company's own policies. Similarly, Twitter reports about the enforcement of its Twitter Rules and discloses statistics about the Removal and Information requests that include legal demands from governmental and non-governmental bodies to withhold and/or remove content and produce information. In a total of 9 different categories, Twitter also publishes separate data on Copyright and Trademark notices, as well as on Platform manipulation referring to the use of the service to mislead others and/or disrupt their experience by engaging in bulk, aggressive, or deceptive activity. Analogical transparency reports are also being produced by Google, Microsoft and a number of other large digital service providers.

A significant constraint limiting companies' transparency initiatives is linked to privacy concerns around the retention of data, whether it's made available to third-party researchers or not (for further reasoning please refer to question 19, section 2).

2. Clarifying responsibilities for online platforms and other digital services

Question 1:
What responsibilities (i.e. legal obligations) should be imposed on online platforms and under what conditions? Should such measures be taken, in your view, by all online platforms, or only by specific ones (e.g. depending on their size, capability, extent of risks of exposure to illegal activities conducted by their users)? If you consider that some measures should only be taken by large online platforms, please identify which would these measures be.

	Yes, by all online platforms, based on the activities they intermediate (e.g. content hosting, selling goods or services)	Yes, only by larger online platforms	Yes, only platforms at particular risk of exposure to illegal activities by their users	Such measures should not be required by law

Maintain an effective 'notice and action' system for reporting illegal goods or content	X			
Maintain a system for assessing the risk of exposure to illegal goods or content				X
Have content moderation teams, appropriately trained and resourced				X
Systematically respond to requests from law enforcement authorities				X
Cooperate with national authorities and law enforcement, in accordance with clear procedures	X			
Cooperate with trusted organisations with proven expertise that can report illegal activities for fast analysis ('trusted flaggers')				X
Detect illegal content, goods or services				X
In particular where they intermediate sales of goods or services, inform their professional users about their obligations under EU law				
Request professional users to identify themselves clearly ('know your customer' policy)				X
Provide technical means allowing professional users to comply with their obligations (e.g. enable them to publish on the platform the pre-contractual information consumers need to receive in accordance with applicable consumer law)				X

Inform consumers when they become aware of product recalls or sales of illegal goods				
Cooperate with other online platforms for exchanging best practices, sharing information or tools to tackle illegal activities				X
Be transparent about their content policies, measures and their effects	X			
Maintain an effective 'counter-notice' system for users whose goods or content is removed to dispute erroneous decisions	X			
Other. Please specify				

Question 2:
Please elaborate, if you wish to further explain your choices.

While many of the concepts mentioned in the list in the preceding question have merit as potential good practice for content hosts, the Commission must also recognize that instituting them as legal mandates could create confusion or competing incentives with the N&A framework. For example, requiring systematic responses to law enforcement requests, or requiring cooperation with "trusted flaggers", could create a de facto notice-and-takedown regime if providers are concerned that they will be penalized for failing to comply with such requests. Further, CDT supports information-sharing and best-practice development across online services, but we caution that mandated "cooperation" with other service providers around content removal could effectively create a centralized censorship regime. And, in general, no liability regime can mandate that intermediaries take action against lawful content. It is also essential that any legal requirements for intermediaries to cooperate with or respond to requests from law enforcement meet the highest substantive and procedural protections for individuals' fundamental rights.

A core priority for the Commission should be to clarify the current liability regime under the E-Commerce Directive (ECD). A clear and stable liability framework is essential to promote freedom of expression online; absent such a framework, intermediaries of all kinds face strong incentives to overblock users' speech and to limit access to information. Any proposal by the

Commission should preserve the liability protections already present in the ECD and provide additional clarity to divergent interpretations of the ECD that have emerged in court opinions over the past twenty years. (See our response in module II for more.)

In general, CDT recommends that the Commission follow these principles in considering legal obligations for online platforms:

1. Preserve a Strong Baseline Liability Framework
   The ECD's current approach establishing immunity from liability for infrastructure intermediaries—including those providing "mere conduit" and "neutral hosting" services—should be maintained. It must also be clear that intermediaries of all types do not have an obligation to actively monitor and identify illegal content, and that a failure to proactively identify illegal content does not make them become liable (see additional discussion in module II, question 6).

2. Create Clarity and Include Safeguards in Notice & Action Systems
   A harmonised, transparent and rights-protective notice-and-action framework should be a key part of the DSA. It should enable users to flag potentially illegal content and set requirements for intermediaries to have processes in place to deal with such notifications with due regard for users' free expression rights. The N&A framework should include the opportunity for counter-notice by the speaker to rebut claims against their speech, and include penalties for notices sent in bad faith. Crucially, intermediaries should not face liability for failing to remove illegal content unless the notice is supported by a court order or similarly independent adjudication (see answer to module II, question 3). Any other approach would require intermediaries to make their own assessment of the illegality of third-party content, which would privatize an essential function of the courts.

3. Structural or Systemic Oversight Must Not Disincentivize Good Samaritan Moderation
   (see question 23 below for further discussion).

4. Commitments to Transparency and Accountability in Content Moderation
   The Commission should establish minimum requirements for meaningful and robust transparency mechanisms for both Member States and online platforms concerning the removal of user-generated content and the overall impact of their content moderation systems. Intermediaries should have in place adequate, accessible, and easy-to-use mechanisms to report illegal content and to flag content as violating the service's own policies.

5. Maintain Flexibility for Different Approaches to Content Moderation
   Effective content moderation will consist of different policies and practices for different types of services and different user-bases and communities. According to the standards

set by the Council of Europe, “States should take into account the substantial differences in size, nature, function and organisational structure of intermediaries when devising, interpreting and applying the legislative framework in order to prevent possible discriminatory effects.” Enabling effective 'Good Samaritan' moderation of harmful, but not illegal, content, requires recognizing that there is no one-size-fits-all approach, and ensuring that the legislative framework is not overly prescriptive as to the substance or method of content moderation, and does not create legal risk or onerous regulatory obligations that will discourage or constrain intermediaries’ content moderation efforts.

Question 3:
What information would be, in your view, necessary and sufficient for users and third parties to send to an online platform in order to notify an illegal activity (sales of illegal goods, offering of services or sharing illegal content) conducted by a user of the service?

- Precise location: e.g. URL - **Yes**
- Precise reason why the activity is considered illegal **- Yes**
- Description of the activity **- Yes**
- Identity of the person or organisation sending the notification. **- Yes**
  Please explain under what conditions such information is necessary:
- Other, please specify

Question 4:
Please explain

Ad 'Identity of the person or organisation sending the notification': The question is unclear, but if it is referring to notifications that create actual knowledge on the part of the intermediary that content is illegal, such notices can only come from courts or other independent arbiters who are accountable for making the determination of illegality. The identity of such arbiter (and its authority to make the determination) must be clearly communicated to the intermediary.

A notice-and-action framework should clearly specify the components of a valid notice — that is, a notice that can create actual knowledge on the part of the intermediary. This should include concrete elements such as those listed above, as well as a citation of the specific law and legal authority that authorizes the issuing of the notice. Notices must include precise URLs for the specified content, rather than a general description of the illegal content, as the latter effectively creates a monitoring obligation on the part of the intermediary. For video or audio content, inclusion of a precise timestamp of the offending content will also help the host identify the illegal content.

Formalistic requirements such as those listed above provide important safeguards for human rights in the N&A regime. They enable intermediaries to confidently reject improperly formed notices without risking liability. The sheer scale of user-generated content, with a recent study estimating that nearly half of the global population uses social media, means that intermediaries cannot carefully evaluate every notice they receive. But clear provisions delineating the components of a valid notice enable service providers to quickly identify which notices are inadequate, and are an important protection against fraudulent or malicious notices.

There may also be specific categories of content for which it is appropriate to consider an intermediary to have "actual knowledge" about the illegal nature of content in case it is beyond any dispute (the concept of "manifest illegality"). For instance, child sexual abuse material (CSAM) is generally considered illegal in all contexts and is typically apparent on its face. But the concept of manifest illegality has grown quite broad in case law interpreting the ECD (see discussion in module II, question 6); the Commission should clarify a set of criteria that define “manifestly illegal” content.

Otherwise, notifications from non-court sources should not be treated as providing the intermediary with "actual knowledge" of illegal content, and should not give rise to a liability risk.

Question 5:
How should the reappearance of illegal content, goods or services be addressed, in your view? What approaches are effective and proportionate?

Various online services have developed tools to enable them to identify and block re-uploads of content that they have previously decided to remove (whether because it violates their policies or has been subject to a legal notice). These tools, including PhotoDNA, Content ID, and the shared-hash database administered by the Global Internet Forum to Counter Terrorism, can identify repeated uploads of content in order to block it, flag it for human review, or (in the case of copyrighted material) to allow the original creator to monetize it.

CDT and a coalition of 16 human rights organizations have raised a variety of concerns about the lack of transparency and the risk of collateral censorship stemming from the GIFCT hash database. Similarly, CDT and many other NGOs, academics, journalists, Internet pioneers, and EU citizens opposed the inclusion of “re-upload” filters in the recent Copyright Directive, based in part over concerns about such filters’ ability to account for fair use and to respect freedom of expression.

The Commission must recognize that an obligation to “prevent the reappearance” of illegal content on an online service is another way of imposing a general monitoring obligation or filtering mandate. A “staydown” requirement for illegal content would threaten fundamental rights to freedom of expression and privacy by enacting prior restraints on speech; determining the presence of previously identified illegal content requires scanning and evaluating all content that is uploaded. As we discuss in module II, question 6, the prohibition against general monitoring obligations is an essential part of the liability framework and must be preserved.

As CDT and over 30 scholars and digital rights advocacy organizations advised the European Parliament, in the debates around the Terrorist Content Regulation, we would likewise urge the Commission to "reject proactive filtering obligations (...); and refrain from enacting laws that will drive internet platforms to adopt untested and poorly understood technologies to restrict online expression."

Question 6:
Where automated tools are used to detect illegal content, goods or services, what opportunities and risks does their use present as regards different types of illegal activities and the particularities of the different types of tools?

"Automated tools" can refer to a variety of automated processes which may be as simple as keyword filters or as complex as sophisticated machine learning tools. Automation can be used for proactive detection, evaluation and enforcement of a decision to remove, label, demonetize, or prioritize content. As such, using different types of tools can present different opportunities as well as risks.

An overarching technical limitation of automated tools is their failure to grapple with context. Current technologies struggle to parse historical, political, cultural, and other circumstances surrounding a piece of content. Yet it is the context that often determines whether a particular post violates the law or content policy of a site.

Some limitations of automated tools derive from the process of developing the tool. For instance, natural language processing (NLP) tools perform best in environments that closely match the data they were trained on, but lose reliability when used across a variety of sites, languages, or cultures. Algorithmic systems also have the potential to be biased against underrepresented groups, including racial and ethnic minorities and speakers of non-dominant languages, due both to the lack of training data and to the possibility of biased datasets. See question 1 of section 1D (erroneous removals) for further discussion of error and disparities caused by automated systems.

Algorithmic systems used for detecting particular types of content will always have so-called false positives (something wrongly classified as objectionable) and false negatives (something actually objectionable is missed). False positives pose a clear threat to individuals' right to freedom of expression. False negatives, on the other hand, can result in a failure to address hate speech, harassment, and other objectionable content. Moreover, bias in algorithmic systems risks discrimination of communities and individuals, including illegitimate silencing of their expression and failure to address harms to their communities. All these factors may then create a chilling effect on individuals' and groups' willingness to participate online.

Thus, CDT strongly believes that the element of human review remains a key component of any content moderation system. While humans are also prone to our own biases and errors, human involvement in moderation is an essential safeguard to mitigate the worst effects of filtering. Humans are able to bring cultural, linguistic, and historical context to their analysis of other people's speech in a way that machines cannot replicate. Human review is also a key feature of companies' appeals processes, which are an important procedural safeguard against the errors that both humans and machines can make. Policymakers should also realize that illegality is rarely manifest and the determination might already be difficult for an expert lawyer, let alone an automated tool. The use of these tools thus should not be mandated by law.

Question 7:
How should the spread of illegal goods, services or content across multiple platforms and services be addressed? Are there specific provisions necessary for addressing risks brought by:
a. Digital services established outside of the Union?
b. Sellers established outside of the Union, who reach EU consumers through online platforms?

CDT recognizes the need for strong legal frameworks that enable the sharing of relevant information across jurisdictions while respecting individuals' fundamental rights. While addressing these questions, the Commission should make sure not to impair privacy and security of users that are located outside the EU's territory and are not subject to its laws. In the initial e-Evidence proposals we have seen that each EU Member State would be given access for law enforcement purposes to the data of internet users worldwide. This is because each provider in the scope of the proposals can be compelled to disclose its users' data no matter where the user is located and no matter the country of citizenship of the user. This can create an enormous risk to privacy worldwide, and the DSA should make sure not to follow similar logic.

Question 8:
What would be appropriate and proportionate measures for digital services acting as online intermediaries, other than online platforms, to take – e.g. other types of hosting services, such as web hosts, or services deeper in the internet stack, like cloud infrastructure services, content distribution services, DNS services, etc.?

Liability for third-party content should not be imposed on intermediaries other than the content host. Infrastructure service providers, DNS services, cloud services, cybersecurity providers, and others should not be held responsible for content their customers host. These companies lack both the information to effectively make decisions about whether speakers have violated content policies and risk over-censoring in order to avoid liability risks. Others simply lack the technical ability/access to moderate content. Only the intermediary with a direct relationship with the uploader, and the possibility to take decisions on discrete pieces of content, should potentially be liable for it.

Question 9:
What should be the rights and responsibilities of other entities, such as authorities, or interested third-parties such as civil society organisations or equality bodies in contributing to tackle illegal activities online?

Year by year, digital platforms are recording an increasing number of government requests for user information or content restriction (more about company reporting can be found in question 9 of section 1D (erroneous removals). For instance, Facebook recorded nearly 270,000 requests for user data last year, an increase of around 500% compared to 2013 when the company started reporting this metric. While 9/10 of these requests were accompanied by a search warrant or a similar legal instrument, the rest were issued under an emergency regime. In nearly 3/4 cases, the company complied (in case of both legal and emergency requests). Similarly, Twitter recorded an increase of almost 650% in government requests for account information between 2018 to 2012. More than 20% of these were issued as emergency requests. The platform complied with around half of the total amount of requests in 2018. Please see also our response in Section IV on transparency requirements in relation to online ads.

While large companies disclose various types of data voluntarily, the same level of transparency should be required on the side of the governments. Government bodies, including law enforcement, are significantly behind the tech industry in providing regular transparency reporting about their demands for user data and content restriction. Without regular reporting on law enforcement efforts by public authorities themselves, the public, and policymakers, only ever have half of the picture about how illegal online activity is being addressed. The public has a legitimate interest in understanding government activity and holding governments accountable for any potential violations of human rights.

Question 10:
What would be, in your view, appropriate and proportionate measures for online platforms to take in relation to activities or content which might cause harm but are not necessarily illegal?

In the entirety of this consultation and the forthcoming DSA, it is essential that the Commission maintains clear distinction between legal and illegal content and activities. By definition, harmful but lawful content falls outside the scope of legal prohibition, and intermediaries (of any sort) cannot be held liable for speech that is lawful. This does not mean that these same intermediaries cannot take action, themselves, against such content. Indeed, CDT recommends that the Commission consider how to more fully enshrine “Good Samaritan” or positive-intent protections for content moderation in the EU’s intermediary liability framework (for more in-depth discussion, see our recent paper 'Positive Intent Protections: Incorporating a Good Samaritan principle in the EU Digital Services Act').

CDT has long called for intermediaries to be clear to their users about what their rules are, and to provide transparency about their enforcement. Companies engaged in content moderation should provide meaningful due process to impacted speakers and better ensure that the enforcement of their content guidelines is fair, unbiased, proportional, and respectful of users’ rights. Their users should be provided detailed guidance about what content is prohibited, including examples of permissible and impermissible content and the guidelines followed by reviewers. Companies should also provide an explanation of how automated detection is used across each category of content.

CDT and other digital rights groups articulated a set of transparency and accountability best practices in the Santa Clara Principles in 2018, which recommend:

- Companies should publish the **numbers** of posts removed and accounts permanently or temporarily suspended due to violations of their content guidelines. This data should be provided in a regular report, ideally quarterly, in an openly licensed and machine-readable format, and be available to independent researchers to assess the effects of the content moderation practices.
- Companies should provide **notice** to each user whose content is taken down or account is suspended about the reason for the removal or suspension. Such notice should include a minimum level of detail including the means to identify the content at issue (e.g., URL), the rule assumed to be violated, the method of detection (e.g., automated detection, user or government flagging), and the explanation of future possible steps available to the user.
- Companies should provide a meaningful opportunity for timely **appeal** of any content removal or account suspension. Minimum standards for a meaningful appeal include a human review element, an opportunity to present additional information by the content creator, and a notification of the results including the reasoning used to make the final decision.

Better transparency and accountability of the companies' systems can also be facilitated by the policymakers. Future legislation should be aimed at providing legal certainty to intermediaries about their ability to moderate their users’ lawful speech. The current legal framework established with the e-Commerce Directive does not adequately promote the adoption of voluntary and proactive content moderation policies by private intermediaries, but rather the opposite. The more that intermediaries play an active role in monitoring the content they host, the more likely it becomes that they will find a potentially illegal piece of content. In this context the chances of overlooking a particular illegality, and therefore the risk of liability, grow significantly. The DSA should grapple with the distinction between active and passive hosting and provide significant additional clarity about the scope and requirements of notice-and-action procedures.

Question 14:
In special cases, where crises emerge and involve systemic threats to society, such as a health pandemic, and fast-spread of illegal and harmful activities online, what are, in your view, the appropriate cooperation mechanisms between digital services and authorities?

Even in a public emergency, such as a pandemic, any emergency measures by State authorities must be based on the rule of law and within the parameters provided by international human rights law. Any additional powers should be time-bound and only exercised on a temporary basis with the aim to restore a state of normalcy as soon as possible. Even without formally declaring states of emergency, States can adopt exceptional measures to protect public health that may restrict certain human rights. These restrictions must meet the requirements of legality, necessity and proportionality, and be non-discriminatory. The suspension or derogation of certain civil and political rights is only allowed under specific situations of emergency that "threaten the life of the nation". Some fundamental rights cannot be suspended under any circumstances i.e. obligations associated with the core content of the rights to food, health, housing, social protection, water and sanitation, education and an adequate standard of living remain in effect even during situations of emergency.

Cooperation between state authorities and digital services in the context of such a crisis must start from that basis. In the context of the coronavirus pandemic, CDT has convened a multistakeholder taskforce to examine issues of data requests by government and other uses of individuals' data for combatting the spread of COVID-19 and to provide observations and recommendations about how best to ensure the twin goals of protecting individual privacy and combatting the virus.

A particular area of concern during the current crisis is takedown of content at the request of governments or content that describes governments' responses to COVID-19. Russia has reportedly requested social media companies to censor media outlets that publish what the authorities deem to be "false information that is socially significant" about the coronavirus. China has ordered Zoom to shut down accounts of political activists. As governments exercise emergency powers to control people's movements, there have been rising reports of police brutality and abuse of power, including in Paraguay, the Philippines, India, Nigeria, and Kenya. Social media are a crucial tool for people to report on and document human rights abuses, and it is essential that this speech does not get blocked by companies' automated moderation systems or censored at the request of governments.

As CDT and many other human rights advocacy organizations have noted, states' emergency powers "must be time-bound, and only continue for as long as necessary to address the current pandemic."

Question 18:
In your view, what information should online platforms make available in relation to their policy and measures taken with regard to content and goods offered by their users? Please elaborate, with regard to the identification of illegal content and goods, removal, blocking or demotion of content or goods offered, complaints mechanisms and reinstatement, the format and frequency of such information, and who can access the information.

Based on CDT's years of experience researching and advocating for increased transparency from Internet companies, we offer the following recommendations:

- **Transparency for a purpose, not just transparency's sake.** Transparency is not an end goal in itself; rather, information from online service providers should enable concrete policy goals such as accountability of companies and governments over actions they take against user content, and increased user control over the information they share and receive online. Any effort around transparency should have a clearly identified set of goals that the transparency measures are directly designed to advance.

- **Transparency efforts need to be tailored to specific audiences.** The umbrella concept of "transparency" can encompass many things, from detailed data about actions taken against user content and accounts, to information about policies and practices, to independent evaluations of a provider's systems. Different audiences will benefit from different types of information:

    - **For users**, CDT recommends that online services provide clear and detailed information about their policies, illustrated with examples to help users understand where the service draws lines between permissible and prohibited speech. This should include clear information about how content is algorithmically targeted and promoted on the site. Services should provide notifications to users when their content is restricted, which should include specific information about the reasons for content restriction. Services should provide clear information about the ability to report content, the opportunity to appeal actions taken against content, and the tools available to users to control the use of their personal data and the targeting or recommendation of information that they see on the service. The over-arching goal of transparency aimed at users should be to empower users to make choices and exert control over their interaction with the service.

  - **For independent research**, services should make data available in structured, machine-readable formats. There are genuine privacy concerns associated with, for example, making detailed data about individuals' social media usage available to researchers; privacy and security controls over data made available to third-parties should align with the sensitivity of the information disclosed. For example, some data should be available in open-access formats, such as databases of advertising content and targeting information; generally, information that has been publicly available on a service should be made accessible in formats that enable independent research. More sensitive data, including information about non-public activity on a service, should only be provided to vetted researchers.

  - **For oversight,** whether by regulators, human rights watchdogs, or the press, services should provide regular reporting about the functioning of their systems, including information about the number of requests for content restriction received by government actors and other parties, and the rates at which the service's decisions were appealed by users. (For a full listing of the kinds of information that support accountability, please see the Santa Clara Principles.)

- **Transparency reporting will look different across different services.** Content moderation necessarily represents a series of trade-offs, and different services will (and should) experiment with different approaches to responding to the specific types of abuse that are most prevalent or problematic on their services. Extremely prescriptive requirements for the content and format of transparency reports could have the unintended consequence of constraining the ability of services to respond effectively to abusive content. For example, requiring services to report the length of time it takes to respond to notifications will exert a strong pressure on services to shorten that time, which will likely decrease the quality of the review that they conduct. Any framework for transparency reporting needs to be flexible and to account for necessary variation in content moderation across services.

- **Governments must provide complementary transparency.** As we discuss in question 9, section 2 of this module, there is little in the way of regular transparency reporting from government agencies about the actions they take to restrict content online. Reports from government authorities would provide an important public accountability mechanism for the exercise of state control over online content and would help the public identify issues of both over- and under-enforcement of the law. Such reports would also provide a useful cross-check to the information reported by technology companies about the orders and other requests they receive from governmental sources; currently, the public only has a view to half of the story.

Question 19:
What type of information should be shared with users and/or competent authorities and other third parties such as trusted researchers with regard to the use of automated systems used by online platforms to detect, remove and/or block illegal content, goods, or user accounts?

CDT believes it is crucial that technology companies provide transparency about their treatment of user data and user-generated content. There is a clear and compelling need for independent actors to be able to access information held by these companies to conduct essential research into the dynamics that shape our online information environment. Many large platforms already voluntarily participate in transparency reporting by disclosing statistics on the enforcement of their company policies and/or government requests for content restriction and user data (see question 9, section 1D (erroneous removals)).

The COVID-19 pandemic, and a related shift towards increased automation in content moderation, has demonstrated the substantial need for more information about the use of automation. This was driven by the important need for companies to send their moderator staff home for safety reasons. There are many questions about the ongoing consequences this shift to automation is having for people's access to information and ability to report on developments during this global public health crisis. As CDT and many other human rights advocacy organizations have noted, states' emergency powers "must be time-bound, and only continue for as long as necessary to address the current pandemic." The automation-reliant version of content moderation must also be seen as an "emergency power"—these measures cannot become the new status quo.

In this context, CDT together with 75 organizations and researchers published an open letter to social media and content-sharing platforms, urging them to enable future research and analysis about the "infodemic" side of COVID-19 by preserving information about what their systems are automatically blocking and taking down. The ways that social media companies design their algorithms to promote and demote content shapes what information reaches people, and therefore the nature of democratic discourse. These design decisions can translate into real world consequences for public health and our democracy. To assess the efficacy of efforts to share vital public health information while combating the spread of coronavirus scams, it is crucial to understand content moderation in practice. (Further reasoning can be found in question 6, section 2 of this module).

The COVID-19 crisis highlights how difficult it is to conduct solid empirical research on our online information environment. The data necessary for this research is held by multiple private companies, and some important information, such as the amount and type of content automatically blocked at upload, may not be recorded at all. Moreover, there are genuine and significant privacy concerns with companies retaining data, whether it's made available to third-party researchers or not. When companies retain data, they increase the risk that it gets exposed through a data breach or is unjustly demanded by government officials.

In CDT's view, it is thus crucial that legal frameworks establish a clear baseline and necessary safeguards to enable independent research and reporting. Safeguards should increase with the sensitivity of the data: open access to anonymized data sets is important to facilitate wholly independent research, but access to potentially re-identifiable information or content should be carefully mediated. And there must likewise be legal safeguards against exploitation by law enforcement or intelligence agencies of data that services may preserve for research purposes.

Service providers should include specific information about their use of automation in their regular transparency reports, including the rates at which content is surfaced for review by automated tools, what proportion of decisions to remove content are made through automated processes, and the rate of appeals of enforcement decisions involving automation. They should regularly evaluate their automated systems for bias or other unintended outcomes and should provide information to the public about the results of those evaluations. They should also provide information directly to users, as part of the standard notification about a content removal decision, about the role of automation in the flagging and removal of the user's content. Without this information, a user has no idea whether a mistaken takedown was the result of human error, disproportionate automated flagging of content, or erroneous automated decision making. With this information, users would be more empowered to appeal specific actions by service providers and hold providers accountable for their use of automation. It would also provide online services with important information about overbroad impacts or unintended consequences of their automated systems, that they may not perceive in their analysis of the operation of their systems at scale.

Question 20:
In your view, what measures are necessary with regard to algorithmic recommender systems used by online platforms?

In general, users need more information about how algorithmic amplification and recommendation of content is shaping the information they have access to and affecting the reach of their own speech. Algorithmic ranking and promotion of content is not inherently problematic; indeed, due to the overwhelming volume of content available online, algorithmic tools are essential to enabling people to sort and identify content that is relevant and interesting to them.

Measures to provide more transparency into the operation of ranking algorithms and recommender systems should start from this basis, of providing better understanding and control for users over the content they see. Key questions include: what are the values that motivate these systems, and how well do the systems actually express those values? For example, a search engine could describe that it intends to provide links to “relevant” information, but users need to understand more about how the search engine determines relevance, to understand what may and may not be shown to them.

To evaluate the impact of an algorithmic system, it is necessary to identify specific values or features the system is intended to express, and to determine effective ways to measure those values. For example, to evaluate a content-promotion algorithm on a social media service for its impact on media pluralism or exposure to a diversity of viewpoints, it would be necessary to identify metrics to assess this value (- Number of different sources of information shown to a user per day? - Proportion of news stories from dominant news media versus independent journalists?) and then to obtain the relevant data to conduct that assessment.

Along with transparency, services should provide enhanced user control over the criteria and values that inform what a recommender system displays to them. Providing user control requires recommender systems to be more transparent and explainable. This can encourage users to look beyond their known interests and generally improve user satisfaction and trust. For instance, users could opt to receive recommendations outside their ordinary consumption habits and/or view content in chronological order rather than curated. Some would warn that increasing user control can also enable users to deliberately view extremist or contentious content. Much depends on how the tool is implemented and designed, and more empirical research (and access to data) is needed to study its effects in practice.

One area of evaluation of recommender systems that deserves particular attention is the use of “downranking” or “shadowbanning” as a part of content moderation on a service. Online content hosts are increasingly turning to measures beyond a simple “take down/leave up” paradigm for content moderation, to include actions against content that limit the incentives for users to post such content (e.g., demonetization, removing comment features) and that limit the content’s reach (e.g., downranking and deprioritizing content). Such responses can be beneficial to free expression, because they avoid a total silencing of speech that does not actually violate the service’s content policy, while also being effective at mitigating abuse. But when the operation of algorithmic systems is generally opaque, the potential use of downranking creates an environment ripe for confusion and conspiracy theories about exactly how a service is or is not manipulating content. Without transparency into policies around downranking and the general operation of amplification algorithms on a site, it is easy for bad actors and genuinely confused users alike to claim that a service is suppressing their speech. Without general transparency into how content promotion is handled on a service, it is difficult to hold providers accountable for the decisions they actually are making.

Finally, we note that there is a tension inherent in transparency of recommender algorithms, in that bad actors (or merely commercially motivated actors) will also use information about how the algorithmic system works in order to game it. This can be mitigated to a certain extent by providing higher-level information about how the system works and withholding very granular information that could be more easily exploited. But it is important to acknowledge that transparency and mitigation of abuse of algorithmic systems will likely need to unfold in a perpetual, iterative cycle.

Question 21:
In your view, is there a need for enhanced data sharing between online platforms and authorities, within the boundaries set by the General Data Protection Regulation? Please select the appropriate situations, in your view *(multiple choice)***:**

- Specific request of law enforcement authority or the judiciary

Question 22:
Please explain. What would be the benefits? What would be concerns for companies, consumers or other third parties?

Any enhanced data sharing between companies and authorities need to be weighed against risks to individual privacy and security. For instance, in the case of ride-sharing services, location information can reveal a person's sensitive activities and interests, including those being pursued in real time. For that reason, authorities should aim to rely on aggregated data disclosed by the mobility service provider, rather than individualized data. To the extent authorities compel disclosure of disaggregated data, they should protect user privacy by limiting their collection to data which is necessary to achieve a clear and narrowly stated purpose, controlling the flow of such data to law enforcement entities and third-party aggregators, deleting unneeded data and securing the data that is needed, and by being transparent about the data they are collecting and the uses to which it is being put. Equally, there are significant legal risks that authorities incur when collecting user data. These risks stem from the potential of shared mobility companies and their users challenging compelled collection of location information under privacy laws, as well as the security risks and danger of public disapproval of overly broad data collection.

Similarly, data sharing for the purposes of supervising platforms' content moderation procedures needs to be justified with a specific and narrowly defined regulatory objective. For example, sharing of information for regulatory purposes should never become a backdoor for information collection as part of a law enforcement investigation of a specific individual. Also, platforms should never be required to disclose users' personal data without a due legal process. A generalized oversight regime would create increased pressure on companies from the state and raise serious privacy, surveillance and censorship risks. In return, people could feel a chilling effect discouraging them from exercising their free speech rights, knowing that their content comes under increased and inappropriate governmental scrutiny. It should be reiterated that governments already do send requests for various user data to online platforms, which are partly being issued under an emergency regime without a proper legal procedure. In this context, CDT calls for authorities to provide regular transparency reporting about their demands for obtaining user data and restricting content (see question 9 in section 2 for more information).

Question 23:
What types of sanctions would be effective, dissuasive and proportionate for online platforms which systematically fail to comply with their obligations (See also the last module of the consultation)?

The Commission should ensure that the DSA creates a clear and predictable liability regime for illegal content and facilitates companies' voluntary efforts to moderate harmful, but not illegal, content without the risk of liability discouraging them from such efforts (see module II, question 4 for further discussion). In principle, sanctions around the failure to moderate harmful-but-lawful content would violate the rule of law principles and should not be imposed.

If intermediaries are subject to certain structural or systemic duties aimed at tackling the spread of illegal content, these duties need to be commercially reasonable, transparent, proportionate, and generally flexible. Such obligations should not focus on the outcomes of content moderation processes, i.e. intermediaries should not be evaluated on whether they have removed “enough” illegal content, as this creates a strong incentive towards over-removal of lawful speech. Intermediaries should not face penalties, for example, for failing to “consistently” or “comprehensively” enforce their policies against illegal content, as this creates a disincentive towards having specific and nuanced policies aimed at combating abuse of their platforms, and effectively creates a mandate for general monitoring/filtering. Legal regimes must make a clear difference between administrative responsibility related to failure to fulfil regulatory obligations and loss of immunity regarding hosted content upon receiving an external notice. Sanctions should only be applied in cases of demonstrated systemic failure to respond to valid notifications of illegal content.

Additionally, the DSA could consider facilitating exchanges of data about companies' content moderation and curation systems with outside researchers by providing a processing ground under data protection law (or clarifying existing grounds) (see question 20, section 2 of this module for further information on algorithmic systems). In that case, sanctions could serve as an appropriate measure against breaches of confidentiality.

II. Liability regime - Reviewing the liability regime of digital services acting as intermediaries?

Question 2:
The liability regime for online intermediaries is primarily established in the E-Commerce Directive, which distinguishes between different types of services: so called 'mere conduits', 'caching services', and 'hosting services'. In your understanding, are these categories sufficiently clear and complete for characterising and regulating today's digital intermediary services? Please explain.

While the technical complexity of providing online services that host, store, and transmit user content may have increased since the E-Commerce Directive was enacted, the ECD still embodies important principles for any liability framework. It is important to maintain a clear distinction, that liability for user-generated content can only ever attach to the service that hosts it directly (and then only under certain conditions; see question 3 below); liability for specific user content should not attach to providers of mere conduit, caching, or similar services. And no category of intermediary should face a general obligation to monitor or a requirement to employ "proactive measures" to block illegal content (see discussion in question 6 below).

Question 3:
For hosting services, the liability exemption for third parties' content or activities is conditioned by a knowledge standard (i.e. when they get 'actual knowledge' of the illegal activities, they must 'act expeditiously' to remove it, otherwise they could be found liable). Are there aspects that require further legal clarification?

There are three components of the "actual knowledge" standard in the current framework that CDT believes could benefit from clarification given divergent interpretations by member states.

First, in most countries the process for notice is not defined; instead, "actual knowledge is sufficient however acquired." CG v. Facebook. Although some countries have statutory frameworks for notice, the statutory framework is not the exclusive means by which intermediaries can obtain notice. (See Overview of the legal framework of notice-and-action procedures in Member States SMART 2016/0039.) This also disincentivizes proactive monitoring measures (see question 4 below): if intermediaries can stumble upon notice by doing the right thing, they will be incentivized to bury their heads.

Second, there is no consensus on the standard of specificity required for valid notice. That is, when a rights-holder has identified the content or user with enough specificity for the intermediary to identify it. In L'Oreal SA v. eBay, the CJEU ruled that notice must not be "insufficiently precise or inadequately substantiated." However, there is little consensus on member states on what suffices; for example, to consider service providers generally notified about "identical" or "equivalent" content, as in the Glawischnig-Piesczek v. Facebook case, undermines the prohibition of mandated filtering and monitoring in Article 15.

As we discuss in module I, section 2, questions 3 & 4, the notice-and-action regime should provide a clear statutory framework detailing the requirements for notice of illegal content online and failure to comply with those requirements is equivalent to a failure to provide notice. Such specificity provides important certainty to service providers and to those who seek to have content removed: if plaintiffs comply with the requirements, they can be certain that they have provided notice, and if they do not, the defendants can be certain that notice has not been provided.

Third, there is no consensus on the standard of reviewing the illegality of content; there is confusion as to whether providers need to receive specific knowledge that allegedly illegal content is on their service, or true knowledge, in the form or a court order or similar independent adjudication, that specific content is in fact illegal. In CDT's view, actual knowledge as to the illegality of content can only come from court orders. To allow otherwise would force intermediaries to determine illegality on their own, which they are not well-equipped to do and which will inevitably cause them to over-censor speech. Putting private companies in the position of making binding determinations about the lawfulness of people's speech also undermines the rule of law and limits people's ability to hold either governments or companies accountable for limitations on their freedom of expression.

CDT recognizes that holding a full adversarial hearing on the merits of each allegedly unlawful post is impractical; as much as intermediaries struggle with content regulation at scale, national court systems would be entirely overwhelmed if they made a serious attempt to adjudicate every case concerning allegedly illegal content. But the answer cannot be to put the full burden--or power--of interpreting the law entirely in the hands of online service providers. Instead, we note with interest discussions about "e-courts" and other dispute resolution mechanisms that could provide the ability to address a much larger volume of cases while still upholding (and potentially improving upon) essential features of the arbitration done by courts, namely: fairness, accountability, independence, transparency, and effectiveness. Any "e-court" innovation would need to preserve fundamental rights to due process, including the right to remedy and appeal, and to enshrine fair trial safeguards.

Moreover, we emphasize that court orders are far from the only type of notification that intermediaries regularly receive and respond to--it is practically guaranteed that more actually illegal content will be removed in the course of a content host enforcing its Terms of Service, than in enforcing a court order. This is why CDT champions both Good Samaritan protections (see question 4 below) and transparency reporting from governments and companies (see module I, section 2, question 18) as crucial elements of accountable content moderation by intermediaries that addresses abusive content. But, for the reasons discussed above, it is essential that these informal, non-adjudicated notices do not create liability for intermediaries over specific content.

Question 4:
Does the current legal framework dis-incentivize service providers to take proactive measures against illegal activities? If yes, please provide your view on how disincentives could be corrected.

Yes. The lack of clarity around the meaning of “actual knowledge” and the question of whether proactive review of content could create liability risk for intermediaries can serve as strong disincentives to engage in proactive content moderation. In our recent paper, Positive Intent Protections: Incorporating a Good Samaritan Principle in the EU Digital Services Act, CDT describes the important role that so-called “Good Samaritan” moderation plays in addressing abusive and “lawful-but-awful” content.

The “Good Samaritan” principle ensures that online intermediaries are not penalized for good faith measures against illegal or other forms of inappropriate content. This rule applies with particular relevance to intermediaries that provide hosting services. When intermediaries are granted immunity for the content they handle, this principle in fact incentivizes the adoption and implementation of private policies regarding illegal and other types of lawful but offensive or undesirable content.

At EU level, the E-Commerce Directive contains the general intermediary liability regime applicable to hosting services and establishes a series of provisions regarding the imposition of possible monitoring obligations to intermediaries. Intermediaries enjoy liability immunities inasmuch as they perform a role of a mere technical, automatic and passive nature. This requirement of “passivity” is compatible with certain activities identified by the case law of the CJEU. However, intermediaries become liable in cases where they fail to act expeditiously to remove or to disable access to the illegal content upon obtaining knowledge or awareness, or they are simply proven to have overlooked a particular illegality when implementing voluntary and proactive monitoring measures in such a way as to create actual or constructive knowledge that strips them of immunity.

This legal framework, however, does not adequately promote the adoption of voluntary and proactive content moderation policies by private intermediaries, but rather the opposite. The more that intermediaries play an active role in monitoring the content they host, the more likely it becomes that they will find a potentially illegal piece of content. In this context the chances of overlooking a particular illegality, and therefore the risk of liability, grow significantly.

In order to incentivize content moderation under the Good Samaritan principle, and thereby enable intermediaries to address problematic but lawful content on their services, CDT endorses a number of recommendations for the Digital Services Act. Given the importance of a strong liability framework to promoting freedom of expression, access to information, and innovation online, the future DSA needs to keep the liability protections already present in the ECD. At the same time, it also needs to create additional clarity about the scope and requirements in notice-and-action systems. Intermediaries should not be required to make determinations of illegality of third-party content; that is the function of courts. Uploaders of content should have the right to issue a counter-notice and the framework should include penalties for notices sent in bad faith, among others. Exceptions to these general rules should be limited and narrowly defined.

Moreover, intermediaries should be transparent regarding the impact of their content moderation systems and develop mechanisms to evaluate their effectiveness. Reporting mechanisms for content that is illegal and content as violating the service's own policies should be kept distinct so that it is clear whether there is an allegation of illegality. Liability penalties should not arise from notifications of violations of content policies or Terms of Service.

Question 5:
Do you think that the concept characterising intermediary service providers as playing a role of a 'mere technical, automatic and passive nature' in the transmission of information (recital 42 of the E-Commerce Directive) is sufficiently clear and still valid? Please explain.

The existing law that articulates the relationship between certain activities by a service provider and their risk of liability needs to be clarified. The current CJEU case law has developed a confusing standard to determine the application of liability immunities to "active" and "passive" hosting intermediaries. This approach is based on the wording of Recital 42 of the ECD, which provides that the liability exemptions are applicable when the role of the intermediary "is of a mere technical, automatic and passive nature, which implies that the information society service provider has neither knowledge of nor control over the information which is transmitted or stored". Despite the fact that a consistent reading of this Recital would suggest that this neutrality requirement would only be applicable vis-à-vis "mere conduit" and "caching" activities, and thus to the immunities established in articles 12 and 13 of the Directive, the CJEU has also considered it applicable to hosting activities.

In the Google France ruling, and regarding the web search and advertising services provided by this company, the CJEU states, as per its "technical, automatic and passive" nature, that "the mere facts that the referencing service is subject to payment, that Google sets the payment terms or that it provides general information to its clients cannot have the effect of depriving

Google of the exemptions from liability". Equally, the decision also affirms that "concordance between the keyword selected and the search term entered by an internet user is not sufficient of itself to justify the view that Google has knowledge of, or control over, the data entered into its system by advertisers and stored in memory on its server".

In the L'Oréal case, the Court limits liability to cases where the intermediary "plays an active role of such a kind as to give it knowledge of, or control" over the hosted content. This active role would not occur when "the operator of an online marketplace stores offers for sale on its server, sets the terms of its service, is remunerated for that service and provides general information to its customers". However, it does qualify as an active role to provide "assistance which entails, in particular, optimising the presentation of the offers for sale in question or promoting those offers". There are also pending cases before the CJEU where the Court will have the opportunity to provide some additional clarifications. This is an important question to clarify, as thus far it is difficult to determine the general principles according to which intermediaries' interventions can clearly be classified as active or passive (with the corresponding consequences in terms of liability), as we only have a few specific examples derived from individual court cases.

CDT cautions that proposals to create specific liabilities for "active" hosts that employ automation to promote and recommend content should not be premised on the mere fact of these services' use of automation. As discussed in module I, section 1D (erroneous removals), question 1, and section 2, questions 6 & 20, such systems are opaque and prone to error, and do not engage in in-depth qualitative assessment of the content that they promote; promotion or recommendation of UGC does not equate to actual knowledge of the nature of that content.

Question 6:
The E-commerce Directive also prohibits Member States from imposing on intermediary service providers general monitoring obligations or obligations to seek facts or circumstances of illegal activities conducted on their service by their users. In your view, is this approach, balancing risks to different rights and policy objectives, still appropriate today? Is there further clarity needed as to the parameters for 'general monitoring obligations'? Please explain.

Any attempt to overturn Article 15's prohibition on imposing general monitoring obligations would be a serious mistake. Any duty to monitor will functionally mean, for large sites through which most expressive activity on the internet is conducted, mandated automated filters. CDT opposes mandated filters because they fundamentally change the relationship between people, their speech, and the law: filtering is a kind of prior restraint on speech, and poses the same risks to human rights as other forms of prior restraint:

- Filtering exposes more speech to evaluation and pre-approval before expression can happen, inverting the offline norm in which most people's speech is rarely evaluated under the law.
- Filters remove procedural hurdles to censorship, such as requiring independent adjudication and the opportunity for the speaker to defend herself. Absent such procedural safeguards, the scale of decisions is enormous, which increase the impact of errors in the filters. Filters (of widely varying degrees of sophistication) are notoriously both under- and over-inclusive, often in ways that harm vulnerable communities.
- Filtering reduces the ability of people to understand the systems of censorship that apply to their speech; not only can filters exacerbate the problems of vague rules or standards by attempting to apply them comprehensively across all speech, but they are also typically opaque in their operation, leaving users unaware that an automated system has evaluated their speech.

Furthermore, a general monitoring requirement might violate Article 10 of the European Charter. It is clear that general obligations to filter all illegal speech violate Article 10, and to the extent monitoring obligations are functional filtering obligations, they too unduly burden online free expression. And although this argument has been underdeveloped in the case law, mandated filtering—or mandating monitoring of any kind—raises serious privacy issues, including possible human rights violations of the European Charter, Articles 7 and 8.

The Commission should articulate clear parameters for the prohibition against "general monitoring obligations". Injunctions against specific content identified by a specific URL relatively uncontroversial. These are clearly not "general" injunctions and do not raise the free expression concerns of other injunctions. However, after Glawischnig-Piesczek v. Facebook it is unclear to both intermediaries and member states what other injunctions might be compatible with Article 15. In both Scarlet Extended SA v. SABAM and SABAM v. Netlog, the CJEU made it clear that past violations of copyright laws by users did not justify an injunction requiring a general filter of all user activity. But under the Glawischnig-Piesczek case, prospective injunctions blocking all "equivalent" or "identical" violations can be ordered, so long as the blocking is not done by a human.

There is a lack of clarity over what either of these terms mean. For content to be truly "identical", it must not only be the precise content, but it must appear in the same context as the original violation. Filters are generally not capable of assessing the relevant context to make such a determination, so mandates for filtering of "identical" content are likely to be overbroad.
"Equivalent" content is even less clear, and identifying it requires content hosts to engage in precisely the kind of general inspection and evaluation of all user content that Article 15 is intended to prevent. The only way to comply with this type of injunction would be to closely monitor a huge universe of information, or to broadly ban all information related to the violation.

III. Gatekeeper platforms - What issues derive from the gatekeeper power of digital platforms?

1. Main features of gatekeeper online platform companies and the main criteria for assessing their economic power

Question 1:
Which characteristics are relevant in determining the gatekeeper role of large online platform companies?
Please rate each criterion identified below from 1 (not relevant) to 5 (very relevant):

- Large user base -- **2**
- Wide geographic coverage in the EU -- **2**
- They capture a large share of total revenue of the market you are active/of a sector -- **4**
- Impact on a certain sector -- **5**
- They build on and exploit strong network effects -- **3**
- They leverage their assets for entering new areas of activity -- **5**
- They raise barriers to entry for competitors -- **5**
- They accumulate valuable and diverse data and information -- **2**
- There are very few, if any, alternative services available on the market -- **4**
- Lock-in of users/consumers -- **5**
- Other

Question 3:
Please explain your answer. How could different criteria be combined to accurately identify large online platform companies with gatekeeper role?

As an initial matter, we applaud the Commission's work to ensure that consumers continue to receive the benefits of fair and open competition. CDT has a 25-year history of putting the digital user first, and one of the greatest challenges of the next decade will be to develop competition policies that ensure that consumers benefit from the great promise of the digital economy. The digital economy, and especially online platforms with their inherent network effects and winner-take-all characteristics, raise competition concerns warranting scrutiny.

In that vein, we urge the Commission to develop its approaches here in a manner that is informed by data, that emphasizes the primacy of the user, and that gives fair notice to platforms about the rules. We also hope that, while political attention is on various platforms, the Commission continues to be vigilant about other competition issues in the digital economy that do not involve platforms, including wired and wireless network providers, device OEMs, chipset manufacturers, and the like.

But we appreciate that this consultation focuses on powerful online platforms. We agree that market power in this arena can be abused, and the Commission is right to develop policies about how to detect and remedy any such abuses. We respectfully submit that defining platforms with gatekeeper power is inherently difficult, and rules about special competition issues related to gatekeeper power must be clearly defined, explained, and documented so companies may comply. If the Commission's approach varies between online and offline platforms, we encourage it to explain why.

In our view, gatekeeper power in a platform may exist (1) when a service is of particular importance to users (e.g., important to daily life versus one that adds ephemeral value), (2) where there are significant barriers to entry, (3) where switching costs are high, (4) where consumers do not multi-home, and (5) where monopoly leveraging is common.

Where there are several competing providers, where switching rates are high, where consumer relationships are short-term, where consumers regularly use multiple services for similar services (e.g., Uber Eats, GrubHub, Deliveroo, and Just Eat for food delivery, or Bumble, Happn, Her, and Tinder for dating), and where new competitors succeed regularly, it is unlikely that an online platform will possess gatekeeper power.

In our view, the "digital gatekeeper" concept should not constitute any new status independent of the traditional criteria set out in Articles 101 and 102 to allow competent competition authorities to intervene. It should be defined according to clear criteria and concrete examples, so as not to be subject to different interpretations and consequent uncertainty. That being said, we encourage the Commission to articulate the ways in which powerful digital platforms may raise issues under Articles 101 and 102 and to be vigilant in protecting users from potential abuses.

2. Emerging issues

Question 9:
Are there specific issues and unfair practices you perceive on large online platform companies?

Large platform companies provide a wide array of meaningful services to users around the world. They give us the ability to connect with friends and family, listen to music, travel nearly anywhere, and connect with others in a variety of other helpful ways. The increased growth of many platform companies serves a helpful purpose, as their widespread use leads to "network effects" that enhance the utility of their product. However, as some large platform companies grow, they may abuse their dominant positions by blocking competitors from fair competition. Among the main examples of anticompetitive practices we can list: distorting information availability, lock-in measures, and practices against multi-homing.

If digital platforms manipulate the availability of truthful information or promote disinformation as a means of achieving competitive advantages, that has the potential to distort competition. Consumers need truthful information to make informed purchasing decisions. Markets are more competitive when consumers have access to competitively-relevant information such as what is available, at what price, and for what quality. For example, a targeted campaign to post fake negative reviews for competitors could affect competition. Relatedly, suppressing truthful but potentially negative information on digital platforms distorts what information is available for consumers. Antitrust enforcers should be attentive to the incentives that companies may have to suppress competitively relevant information, assess their competitive effects, and consider taking action against platforms that engage in such practices if competition is harmed. One challenge here is that such deals are often confidential, so antitrust enforcers should consider **requiring disclosure** – ideally, publicly – **of any payments from companies to restrict consumer access to information that could distort competition. (**Such conduct should be distinguished from paid-for/advertising content, which does not raise the same competition concerns provided its status as an advertisement is disclosed. We address transparency recommendations for advertising content in module IV.)

Relatedly, a lack of transparency on other dimensions can distort competition. Pricing on digital platforms is often opaque, especially related to advertising. The Commission should consider whether there are ways to promote greater transparency about pricing to both consumers and advertisers to help identify predatory or otherwise troubling conduct. Further transparency on how user data is used in this manner would also help people to better enforce their rights under GDPR Art. 9, on the processing of special categories of personal data.

Importantly, transparency is likely insufficient when it comes to digital platforms' use of consumer information. There are valid concerns that some platforms' large size allow them to set conditions for how they use and share consumers' data without consumers having any real choice but to "consent" to these practices. To address this issue, CDT supports reasonable regulatory limitations on the collection and use of users' data.

The implementation of lock-in measures is another unfair practice that may occur, and evidence of such conduct should raise competition concerns. Lock-in measures are features designed by tech platforms to prevent consumers from switching to alternate providers. For example, a platform might try to lock in customers by requiring multi-year contracts with heavy termination fees. (Such conduct is also suspect in the non-platform context.) That lock-in can cause competition concerns because it insulates the platform from the fear that the customer will switch to a competitor's offering. Another example would be companies charging fees to consumers to switch providers, or erecting barriers to data portability. The Commission should clearly delineate the differences between organic network effects, which usually benefit users, and steps to lock in users, which are often harmful to them.

Additionally, many users engage in multi-homing. Common examples of this practice include users who drive for both FREE NOW and Uber, and vacationers who check both AirBnB and Booking.com. Multi-homing is an important practice that increases the choice available to consumers. Despite this, if large online platforms were to actively encourage or require users to use their platform exclusively, that may warrant investigation.

Question 10:
In your view, what practices related to the use and sharing of data in the platforms' environment are raising particular challenges?

Many companies use data to improve their services. For example, companies can use data about when consumers shop to make sure they have enough customer service representatives available to help at peak demand times. Companies use data to identify bottlenecks in manufacturing processes. Data analysis is often essential to a firm's continuous improvement efforts. Competition authorities should be wary of disrupting such consumer-benefiting activities. But data could also be leveraged in ways that harm competition. For example, if the ability to leverage data from one part of its business into another makes a firm so entrenched that it is very difficult for any new entrant to offer services in either business, those barriers to entry warrant examination. We urge the Commission to put users first and to encourage companies to act in creative ways that use data to enhance consumers' experiences while simultaneously guarding against anticompetitive uses of data. For example, in cases where it is shown that large databases of information are necessary to train a machine learning algorithm, the Commission should explore ways that such information can be made more broadly available in a privacy protective manner.

Data can also be used to reinforce entrenched social problems. Real life prejudices are often translated online, and vulnerable and at-risk groups tend to be disproportionately harmed by automated decision-making. When artificial intelligence, algorithms and other forms of automated decision-making is employed by platforms – for example, in content moderation, in ad selection, in product promotion or pricing – the European Commission should explore ways to make the code auditable. Innovative ways to pursue means for enabling external and independent auditing should be considered separately from regulatory requirements.

Question 11:
What impact would the identified unfair practices can have on innovation, competition and consumer choice in the single market?

Unfair practices harm innovation, competition, and consumer choice. When companies engage in lock-in behaviors, when they act against multi-homing, and when they attempt to leverage power in one market to another--those may be anticompetitive acts that harm users.

Multi-homing happens when users or service providers form ties with multiple platforms at the same time. This generally occurs when the cost of adopting an additional platform is low and lock-in measures are not in place. For instance, in the riding app space, many drivers and riders use both FREE NOW and Uber. This way, riders can compare prices and wait times, and drivers may reduce their inactive time and take advantage of higher payments. Multi-homing can also take place if close non-platform substitutes are available to users.

When multihoming is absent, platforms may face reduced incentives to compete fiercely on price and quality. They may be less competitively constrained by users easily moving to a rival platform. Lock-in measures hinder customers from changing suppliers in response to changes in a platform's efficiency. These measures, such as the use of proprietary data formats or contractual agreements with exclusivity clauses, are often opaque and may harm the interests of both users and sellers. Lowering multi-homing on both sides of the market may decrease the competitive intensity and allow platforms to increase a dominant position in a non-transparent way.

Furthermore, switching among platforms can be facilitated by some elements of consumer data rights, such as data portability. Data portability, ex. GDPR Art.20, *is a user's right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and the right to transmit those data to another controller without hindrance from the controller to which the personal data have been provided.* This right lets users bring their data to new services outside the control of the original platform and might be an effective tool to counteract the power of large online platforms. In 2016, the EC relied on the right to data portability to protect consumers against lock-in effects with respect to a Joint Venture between subsidiaries of Google and Sanofi to offer services for diabetes, including data collection, processing, and analysis.

However, data portability regimes can differ significantly depending on the purposes they intend to achieve. In light of this, policymakers should clearly articulate the specific goals of such a regime. For instance, if only historic data are provided to be transferred at one point in time, this may not facilitate multi-homing with complementary services that rely on continuous transfers of consumer data. Finally, policies relating to data portability and interoperability may need to balance trade-offs between stimulating competition and protecting incentives for investment.

Question 12:
Do startups or scaleups depend on large online platform companies to access or expand? Do you observe any trend as regards the level of dependency in the last five years (i.e. increases; remains the same; decreases)? Which difficulties in your view do start-ups or scale-ups face when they depend on large online platform companies to access or expand on the markets?

We encourage the Commission to explore these issues. Access to data or other essential aspects of competition is relatively uncertain in the digital economy. We are wary of the ability of dominant platforms to affect the competitive impact of smaller companies.

Question 13:
Which are possible positive and negative societal (e.g. on freedom of expression, consumer protection, media plurality) and economic (e.g. on market contestability, innovation) effects, if any, of the gatekeeper role that large online platform companies exercise over whole platform ecosystem?

As the Commission has noted, large online platforms have brought benefits to the digital economy and society in general. They expand consumers' access to a wide range of products, often with lower prices and enhanced quality. They can facilitate access to information, offering the potential to enhance citizens' participation in society and democracy.

In the digital environment consumers and/or businesses prefer to be on platforms populated by as many users as possible, to have an efficient one-stop-shop experience. The direct and indirect effects of large networks can benefit users. Indeed, travellers use Airbnb because homeowners are using the platform, and vice versa. The business model of many large digital service providers is based on significant initial capital costs, massive scale and very low marginal costs. Efforts aimed at increasing ecosystem contestability could ultimately result in less technological development and less growth for the digital economy.

Nevertheless, without the pressure of small platforms to challenge their power, large platforms can become competitively complacent. That competitive complacency can manifest itself in the form of sluggish innovation on the platform, higher prices, or direct quality degradation, such as failure to safeguard sensitive data. Platforms may also engage in conduct that is competitively suspect, such as designing its technology and business practices to make it hard for customers to switch platforms. Platform growth benefits users in a virtuous cycle at first, but may spiral into a vicious cycle that harms users because the switching costs to consumers who fear losing the benefit of the network is too high. The trick is to spot when the cycle changes from virtuous to vicious. For indicators of when that may happen please see the factors set out in question 8.

As we explain in our response to question 9, large platforms may also use data in ways that are unknown to users and that raise both privacy and equity concerns. In module I, section 2, question 20, we discuss how services can use this data to affect people's access to information and opportunities to speak in opaque ways. While algorithmic ranking and promotion of content is not inherently problematic, users need more information and control over how their information environment is being shaped.

3. Regulation of large online platform companies acting as gatekeepers

Question 1-2:
Do you believe that in order to address any negative societal and economic effects of the gatekeeper role that large online platform companies exercise over whole platform ecosystems, there is a need to consider dedicated regulatory rules? Please Explain:

While we appreciate the Commission's desire to protect consumers, we find it difficult to respond to a question about a regulatory framework without more information about what the regulations might be and to whom they might apply. But the Commission could consider the following:

Data about performance. Digital platforms could report data on the topics listed below, to help identify areas of concern:

- How many new users the platform acquired?
- How many users has the platform lost?
- If known (or estimated), for those users lost, how many switched to each alternative platform?
- If known (or estimated), how many users single-home on the platform for its services, compared to how many are multi-homing?
- For platforms that offer more than one service, how many of its users for each service are also active users of the firm's other services? Does the firm offer incentives for those who use multiple services offered by the platform?

Data about requests to suppress truthful information about competitors. We know relatively little about whether companies are paying digital platforms to suppress information to distort competition. For example, a dominant firm might attempt to pay a restaurant reservations platform to suppress the reservations slots of its closest competitors, or of new entrants, without disclosure of the payment. To avoid the potential competition-distorting effects of restricting access to information, the Commission could encourage digital platforms to:

1. Disclose whether they accept requests to suppress truthful information that could be competitively relevant, and if so, whether they are paid to do so. Companies should also disclose the relative volume of such requests.

2. Disclose whether they accept requests to degrade a competitor's results or listings, including not showing results for competitors and distorting rankings for a given company's products and services. If so, is the platform paid to do so? Again, companies should disclose the relative volume of such requests.
3. Voluntarily commit that they will not accept requests to suppress competitively relevant information or degrade third-party listings. In addition, some companies may contractually bar suppliers, vendors, competitors, or employees from complaining to antitrust authorities about conduct that might be anticompetitive.

We also urge **digital platforms characterised by significant network effects** to develop and **implement interoperability standards.** Any intervention standard would need to strike the appropriate balance between interoperability to avoid consumer lock-in and sufficient flexibility such that platforms could continue to compete based on differentiation and innovation. Innovation might be more effectively delivered, for example, if competition occurs between networks or digital platforms, rather than between service providers operating over individual platforms. Moreover, the standards should be open, so other businesses may freely adopt them.

Question 17:
Specifically, what could be effective measures related to data held by very large online platform companies with a gatekeeper role beyond those laid down in the General Data Protection Regulation in order to promote competition and innovation as well as a high standard of personal data protection and consumer welfare?

Antitrust enforcers, politicians, reporters, and the public are all concerned about the increasing power of digital platforms. Some have suggested that access to data might increase competition among digital platforms. Our response here is focused on such data sharing issues, which raise novel questions.

First, mandatory data sharing may encourage entry. If data is a barrier to participating meaningfully in a market, access to others' data could mitigate entry issues. We note that some companies have voluntarily allowed users to access their data and port it to other companies and that, to date, those programs do not appear to have impacted competition meaningfully.

Second, any mandatory data sharing likely distorts incentives. For example, if a platform with gatekeeper power is told that it must share its data with competitors, it may stop collecting data, the result of which could be diminished service quality. Platforms might also try to avoid reaching widespread adoption to avoid triggering such an obligation, or they might create more niche offerings that would evade a gatekeeper power designation. We encourage the Commission to seek input from a variety of stakeholders to better understand how these incentives could affect users.

Third, mandatory data sharing could have privacy implications if the data is shared with entities that do not protect it carefully. The Commission should carefully consider what data security and encryption obligations would apply and how those costs would be borne.

Fourth, there are a variety of logistical questions. Which companies would be subject to such a requirement? Would start-ups, middle stage platforms, and other competitors with gatekeeper power all be able to access a platform's data? What are the implications for GDPR compliance?

Fifth, such data sharing would need to stem from some legal basis. As a general matter, successful companies are not obligated to share their assets with competitors or potential competitors. A deviation from that general practice would require sound legal authority and evidence of competitive harm that can only be mitigated by data sharing. The Commission could start, instead, with encouraging large platforms to consider voluntary data-sharing programs, such as the Data Transfer Project among Apple, Google, Microsoft, Twitter, and Facebook. Studying voluntary data sharing programs could inform future regulatory approaches.

Sixth, data sharing could raise collusion concerns, especially if it included pricing information. Even absent collusion, reverse-engineering of pricing algorithms from data sharing could reduce price competition in Europe.

Finally, we urge the Commission to consider whether access to other data sets, such as those held by government agencies, might provide significant opportunities. For example, agencies may have data that would enable start ups or middle-stage companies to train a machine learning algorithm designed to compete with a large digital platform.

Question 26:
Please explain which of the options, or combination of these, would be, in your view, suitable and sufficient to address the market issues arising in the online platforms ecosystems.

We are encouraged by the Commission's focused attention on these issues. We believe that action may well be warranted to protect competition for the benefit of consumers. We share the Commission's concerns that large platforms may have the ability to act in ways that foreclose competition and harm users.

We are willing and able to respond to any proposed regulations or other approaches that make the digital economy competitively free, open, and well-functioning. These markets operate best when there is equipoise among competition, access to information, human rights, opportunities for to lift marginalized voices, and innovation. In general, regulatory proposals should seek to enhance the user experience and should be measured against the yardsticks of quality, price, and the pace of innovation.

In particular, we look forward to responding to proposals about preventing lock-in, promoting multi-homing, reducing the risk that secret side payments restrict consumer choice, and determining whether access to some assets, like large data sets, is essential to new entry. In addition, if the Commission seeks to enact regulations that apply differently to technology companies, or to platforms with gatekeeper power, or in some other way that limits their general applicability to commerce, we hope the Commission will define carefully to which companies such rules might apply and will seek input on potential consequences of such an approach.

IV. Advertising and Smart Contracts - Other emerging issues and opportunities, including online advertising and smart contracts

1. Online advertising

Question 14:
Based on your experience, what actions and good practices can tackle the placement of ads next to illegal content or goods, and/or on websites that disseminate such illegal content or goods, and to remove such illegal content or goods when detected?

The fundamental challenge in controlling the placement of ads next to illegal content is the same as the fundamental challenge of content moderation overall: the intermediary controlling the placement of the ad would have to know that the content it is placing an ad next to is illegal. When intermediaries that host user-generated content receive actual knowledge that a post is illegal, they should remove it and disable advertising alongside it. They should also provide information about the removal of the illegal post to the advertising intermediary. This would allow the ad intermediary to provide information about the prevalence of ads appearing alongside illegal content in their ad transparency library or other transparency reporting (see question 19 below). This transparency would encourage greater accountability for both ad buyers and content-hosting platforms, to each other and to Internet users in general.

Advertising intermediaries need to accurately convey to advertisers the risk that their ad will appear next to illegal UGC, as well as the extent to which they can mitigate those risks.

Intermediaries that host UGC and those that place ads will also need to work together to determine whether there are categories or general types of UGC alongside which ads should not appear, due to the risk that this UGC may be illegal--but this is a challenging determination that is also vulnerable to overbroad application. Moreover, it is also the case that marginalized groups often end up bearing the burden of an ad-platform's demonetization policy; for example, YouTube's efforts to reduce the appearance of ads alongside sexual content has frequently disproportionately affected people in the LGBTQ community.

Question 15:
From your perspective, what measures would lead to meaningful transparency in the ad placement process?

Generally, online platforms should provide users with relevant information on the ad placement process in a concise, intelligible and easily accessible way, in compliance with the GDPR transparency principle and information obligations set out in Arts. 12, 13 and 14. Layered fair processing notices can be an effective way to reach this purpose. However, beyond a one-time informative pop-up, online platforms should guarantee transparency on an ongoing basis. For example, the "Protection Dashboard" of Firefox 78, the Ghostery app/web extension, and the new Ads Transparency Spotlight extension for Chrome provide information to users about the trackers and scripts on a webpage in a clear but not overwhelming manner. Beyond transparency, these tools also provide users with some control over the use of their data in the ad placement process (e.g. allowing users the ability to easily block trackers).

According to the UK ICO's report, the Real Time Bidding (RTB) process may involve the processing of special categories of data such as political or religious affiliation, ethnicity, mental or physical health. The GDPR expressly prohibits processing such information unless a condition within Art. 9 applies. The only applicable exception here could be that of explicit consent, as set out in Art. 9.2.a. This means that when special categories of data are involved in the ad placement process, an explicit consent request should be provided.

For behavioral advertising based on internet traffic. consumer control means, in line with GDPR Art.7, that even after consumers have opted in to the data collection, it must be as easy to revoke as to give consent. Upon revocation, behavioral advertising networks and their ISP partners should stop using any data collected while the consumers were opted in. Otherwise, processing previously collected Internet traffic content data should be based on legal grounds other than consent and the consumers should be informed at the time when they revoke their consent.

Finally, companies should be further encouraged to use open data archives to increase ad transparency. For some political ads, platforms like Facebook and Google have developed open political ad libraries**.** (See discussion in Q18/19 below about the difficulty of defining "political" ads.). Existing archives provide information about who paid for an ad, how it was targeted, the size of the audience that saw it, and other relevant information.

Question 16:
What information about online ads should be made publicly available?

Users should be able to easily determine whether online content has been placed or promoted by an individual or entity who is paying for that placement. Some intermediaries already insert labels to alert users that certain content is placed pursuant to a financial relationship, which helps users understand why they may see a particular ad or search result placement. For example, Google search, Facebook, and Amazon all label content placed as part of a sponsorship agreement. This information can help users evaluate the merits of search results relative to un-sponsored content placement and can help users distinguish between organic and paid content in social media interactions. Users would also benefit from additional information, such as criteria for ad targeting, to help them understand why they were served a particular ad. We suggest that the above information (identification of ads or sponsored content and targeting criteria) should be available to end users with minimal additional effort (no more than a click).

Other kinds of information could be useful to researchers, such as through the ad API described by Mozilla, but the Commission should be wary of the potential impacts of mandating disclosure of identities (see discussion in question 19 below).

Question 17:
Based on your expertise, which effective and proportionate auditing systems could bring meaningful accountability in the ad placement system?

In regard to the GDPR's compliance auditing of the ad placement process, CDT's preference would be to have an independent third party (such as the EDPB) conduct compliance reviews. These reviews should be conducted in response to user complaints, and potentially also at random. Having an independent third party auditor will lend credibility to the results of the reviews, and ensure that all members receive fair treatment in the review process. In addition, the documentation of compliance should be made public.

Real life prejudices are often translated online, and vulnerable and at-risk groups tend to be disproportionately harmed by automated decision-making and algorithms that push online advertising. When automated decision-making is employed by platforms – for example, in content moderation, in ad selection, in product promotion or pricing – the European Commission should explore ways to make the code auditable. Innovative ways to pursue means for enabling external and independent auditing should be considered separately from regulatory requirements.

Finally, according to GDPR art.35.3.a, online platforms involved in the RTB process are obliged to publish a data protection impact assessment (DPIA) before beginning the related personal data processing operations and to consult the competent DPA if high risks remain following the DPIA.

Question 18:
What is, from your perspective, a functional definition of ‘political advertising’? Are you aware of any specific obligations attached to 'political advertising' at national level ?

Attempting to distinguish “political” from “non-political” ads could pose high risks for the fundamental rights of individuals and civil society organisations. While campaign ads from candidates may be clearly political, messages addressing issues such as abortion, education, climate change, and immigration can be difficult to categorize. Online ads are cost-effective ways for nonprofits and advocates to reach audiences and raise citizens’ awareness on critical issues for the public debate. An overbroad definition of “political” ads would chill the speech of many organizations lacking the resources to utilise traditional media**.** Democratic elections depend on citizens informing themselves within an environment characterised by pluralism and access to diverse viewpoints.

As an alternative, the Commission should consider a content-agnostic approach to ad transparency by seeking the same kind of disclosures from all online advertisers. (See Q. 19 for more.) Source and targeting information about ads helps users understand why they see the ads they see online, but requiring intermediaries to discern political from non-political ads will likely lead to both overbroad and underinclusive categorization. As we have seen in efforts to create political ad databases, attempts to draw these distinctions can have significant unintended consequences for news media, bookstores, civil society organizations, and other non-political speakers.

The Commission should exercise great caution if it decides to define “political” ads. A narrow definition could reduce the Commission’s ability to address certain aspects of political influence online, but it would also reduce the negative consequences for free expression. If further restrictions were considered, for example during an election period, such measures should apply only to content that an online business has been paid to host, that expressly advocates for the election or defeat of a candidate or political party for public office. It should not apply to content posted by individual users or other organic content, or to content voicing a position on policy issues, even if those issues are associated with a political platform or party. The Commission should be wary of creating a rule or definition specific to existing online content formats. It should strive to be agnostic as to delivery methods and should consider other exceptions, e.g. for media coverage or for paid ads below a minimum expenditure threshold.

If the Commission requires disclosures in any form, it should also establish a centralized, open access, machine-readable database for the disclosed information. Through this database, electoral commissions, researchers and civil society organizations could analyze and identify trends in advertising, such as targeting efforts, uses of sensitive criteria, or discriminatory outcomes. This research could encourage voluntary efforts to address problematic practices and inform further regulatory approaches.

Question 19:
What information disclosure would meaningfully inform consumers in relation to political advertising? Are there other transparency standards and actions needed, in your opinion, for an accountable use of political advertising and political messaging?

In Q18, we note the inherent challenge of defining “political” ads in a way that does not inappropriately sweep in diverse forms of issue-based speech. Instead, we recommend that disclosures for *all* ads should include information about the purchaser of the ad and the nature of any targeting criteria. This baseline transparency approach would give people access to information about the sponsors, techniques, and amounts spent on political (and non-political) advertising online.

If further transparency for election-related advertising is desired, it could be achieved by requiring advertising systems to collect and provide information that includes the candidate, party(ies) the candidate is related to (if any); the identity, nationality, and country of residence of the sponsor; the total amount spent on the ad campaign; the specific election the ad is referring to if any; and targeting information for the ad. As noted in Q18, if this information is created, the Commission should also establish a centralized, open access, machine-readable database for the disclosed information.

At national level across the EU, very few states have updated their electoral laws to include online campaigning. The Commission may wish to advise Member States to update national laws. Placing obligations on campaigning parties, and not intermediaries, is a more desirable intervention as it avoids the risk of intermediaries improperly chilling the speech of civil society advocates and individuals.

We also emphasize to the Commission that ad transparency comes with tradeoffs. “Ads”, or content that is provided by users for display by advertising services at the user’s expense, can include endorsement of political candidates or positions, promotion of news articles, and targeting of advocacy messages from NGOs aiming to provide services and information to specific populations. Regulation of online advertising potentially touches all of this speech.

In general, identification of the source of the funding behind an ad is an important piece of information for users, to understand how specific content is reaching them, and for accountability of the speaker and advertising system. But mandating the disclosure of individuals' identities can also harm individual privacy and undermine people's willingness to promote their speech online, whether it is a political opinion or an advocacy message that places the speaker at risk of reprisal. Anonymous and pseudonymous protects privacy and individuals' safety (e.g. persecuted minorities, political enemies of the state). Stringent transparency measures for all paid messages could interfere with individuals' political speech and could undermine the election law goals of equalizing political influence, improving the quality of electoral debate, and ensuring competitive elections. As it considers issues of ad transparency, the Commission must consider the serious privacy threat to individuals posed by identification requirements.

Question 21:
Are there other emerging issues in the space of online advertising you would like to flag?

According to the ICO's report "Update report into adtech and real time bidding", in 23 EU countries there are no rules for political parties to report on campaign spending on online platforms in a transparent way. The European Commission should call on Member States to introduce new rules for political parties, and candidates competing in elections, to report on campaign spending on online platforms in a transparent manner.

Moreover, we believe that:

- In line with the GDPR privacy by default principle, third party cookies should be blocked by default in every web browser (e.g. Tor, Brave, Safari, Firefox, DuckDuckgo, Chrome, Opera);
- In line with the GDPR storage limitation principle, non-essential cookie data should be dropped at the end of each session by default in every web browser (e.g. Firefox ETP 2.0);
- In line with the GDPR data minimisation principle, cookies should be designed to authenticate a user without using direct identifiers.(e.g. Trust Tokens).

We also called on the Commission to investigate and enforce against discriminatory advertising practices, to protect users from some of the harms of targeted ads. The Commission can use its authority to study and enforce against data-driven discrimination in the digital advertising ecosystem. The data flows that lead to targeted ads or offers are opaque to consumers and often involve hidden inferences or data from companies with which consumers have no direct relationship. This makes it nearly impossible for individuals to accurately assess or avoid harm. The Commission should assume a flexible, case-by-case approach to protect consumers while preserving innovation and growth in digital advertising. Moreover, when investigating predatory ad targeting, the Commission should consider whether the targeting involves collecting sensitive data and/or inferring sensitive information from data and using it in ways that are likely to exploit particularly vulnerable groups. Finally, creating a centralized, open-access, machine-readable database of advertising information would enable academic researchers to analyze the ad system for evidence of discrimination and other harms.

VI. Governance and Enforcement - What governance for reinforcing the Single Market for digital services?

Governance of digital services and aspects of enforcement

Question1:
Based on your experience, how would you assess the cooperation in the Single Market between authorities entrusted to supervise digital services?

The country of origin principle does provide for more predictable and clearer rules. At the same time, it also creates a situation, to take the example of data protection, whereby due to the high concentration of tech companies in Dublin, Ireland, the Irish data protection authority has a proportionately much higher burden in terms of enforcement of the related rules. Art. 52 (4) of GDPR provides that EU Member States should ensure 'adequate resources' to the competent authority to allow for the effective enforcement of the rules.

In order to ensure that relevant authorities are adequately resourced in practice, the effect of the country of origin rule should be considered and if necessary additional EU financial resources could be provided to the relevant national authority. It is particularly pertinent that data protection enforcement authorities be adequately resourced as many of the policy concerns highlighted in this survey could be resolved through better enforcement of those rules.

The Commission should anticipate that more aggressive enforcement of national laws against illegal content will bring to the fore the pre-existing conflicts between national legal standards for protected expression. There are already long delays in cases of notice-and-action disputes where two or more EU member state jurisdictions are involved. The Digital Services Act should not attempt to harmonize national legal standards but it should provide clear procedures for cross-border cooperation on such cases.

Question 2:
What governance arrangements would lead to an effective system for supervising and enforcing rules on online platforms in the EU in particular as regards the intermediation of third party goods, services and content (See also Chapter 1 of the consultation)? Please rate each of the following aspects, on a scale of 1 (not at all important) to 5 (very important).

- Clearly assigned competent national authorities or bodies as established by Member States for supervising the systems put in place by online platforms -- **4**
- Cooperation mechanism within Member States across different competent authorities responsible for the systematic supervision of online platforms and sectorial issues (e.g. consumer protection, market surveillance, data protection, media regulators, anti-discrimination agencies, equality bodies, law enforcement authorities etc.)-- **5**
- Cooperation mechanism with swift procedures and assistance across national competent authorities across Member States-- **5**
- Coordination and technical assistance at EU level -- **5**
- An EU-level authority-- [no answer]
- Cooperation schemes with third parties such as civil society organisations and academics for specific inquiries and oversight -- **5**
- Other: please specify in the text box below

Question 3:
Please explain.

This survey implies that further EU regulatory laws and instruments could be developed in response to the consultation. The area of possible regulatory areas is very broad, from consumer rights, content moderation to oversight of electoral laws and the sale of goods. It will be essential to examine each area and assess what the appropriate governance arrangements might be. In recent years, the European Commission has itself acknowledged and taken legal action in relation to concerns about the rule of law in some EU Member States. Considering that there is potential for the Digital Services Act to touch upon regulatory matters which speak to the cornerstones of democracy such as media pluralism and the safeguarding of free and fair elections, it will be crucial that any bodies which are created, or any existing bodies given new authorities, have key rule of law safeguards. Governance based on the rule of law implies that all persons, institutions and entities, public and private, including the State itself, are accountable to laws that are publicly promulgated, equally enforced and independently adjudicated, and which are consistent with international human rights norms and standards. Concretely, in the case of any regulatory body, this would imply such factors as the separation of powers, participation in decision-making, legal certainty, avoidance of arbitrariness, and procedural and legal transparency are built into the mandate by design.

Meaningful participation in decision-making on an operational level would mean adequate resources should be dedicated to ensure robust multi stakeholder consultation processes including with civil society actors.

Each area will need careful consideration on what type, if any, of regulation and governance bodies are needed. For example, the UN Special Rapporteur on Freedom of Expression has stated that states should refrain from adopting models of regulation where government agencies, rather than judicial authorities, become the arbiters of lawful expression. In the case of content moderation then, it would be more appropriate to envisage cross-border cooperation between EU Courts. (See our discussion of necessary safeguards in the notice-and-action framework in module I, section 2, questions 3 & 4.) An EU level regulatory body should not decide on the legality of speech.

For more information, please contact:

Emma Llansó, Director, Free Expression Project, ellanso@cdt.org

David Nosák, European Affairs Associate, dnosak@cdt.org

Pasquale Esposito, European Affairs Associate, pesposito@cdt.org

Avery Gardiner, General Counsel and Senior Fellow for Competition, Data, and Power, agardiner@cdt.org

Stan Adams, Deputy General Counsel & Open Internet Counsel, Center for Democracy and Technology, sadams@cdt.org