

Ten Principles for Future EU Policymaking on Intermediary Liability

Introduction

Many European policymakers and governments have concerns about the impact of several types of online content and user behaviour. These concerns are outlined in the [UK Government White Paper on Online Harms](#). Policymakers worry about content that may be illegal, such as some forms of hate speech, and content that is posted with the intent to incite violence for ideological and/or religious reasons. They are also concerned about content and online behaviours which are not illegal, but which they fear may cause harm to some users. These types of content include promotion of suicide or self-harm, cyberbullying, harassment, and disinformation.

Leading social media and content-sharing platforms have stepped up efforts and dedicated more resources to restricting the availability of both illegal content and legal content that, for reasons such as the aforementioned, is considered undesirable. They have done so in part because of public pressure, and in part to improve the services and user experience they provide.

Policymakers are considering policy and legislation that will ‘hold platforms accountable’ and make them ‘take more responsibility’ for the content they host. In European countries, there is a growing sentiment that existing legislation, notably the European [E-Commerce Directive](#) (ECD), should be updated. The ECD establishes the principle that content hosts are not liable for user-uploaded content, unless they have been notified of illegality. The Directive’s provisions are general, and have been implemented differently in different Member States. The Court of Justice of the European Union (CJEU) has issued a number of rulings that clarify certain questions, but guidance as to the expectations content hosts must meet to maintain safe harbour protection remains vague. The Center for Democracy & Technology (CDT) has argued that the Directive should be supplemented with additional notice-and-action guidelines or legislation, but the Commission decided not to move forward with this type of initiative.

Now, however, the Commission is understood to be preparing policy options for new rules for content hosting; a [Digital Services Act](#). The Act would add to several pieces of EU legislation adopted or proposed in the past few years, and to several Member State legislative initiatives focused on illegal and or harmful content, and overall regulatory supervision of content hosts.

Below, CDT proposes some [fundamental principles](#) that should inform future EU policymaking. This input is guided by CDT’s mission to protect the fundamental rights of internet users. While the concerns behind several new policy initiatives are legitimate, CDT emphasises that policy initiatives must be very carefully crafted so as not to harm free expression, access to information, and innovation and entrepreneurship on the internet.

Principles

- 1. Policy and legislation must respect human rights principles on freedom of expression.**
Legislators are obliged to abide by the principles laid down in human rights instruments, notably [Article 19](#) of the [International Covenant on Civil and Political Rights](#) and [Article 10](#) of the [European Convention on Human Rights](#). This means that any restriction on free expression must meet the three-part test: it must be provided in law, pursue a legitimate aim, and be necessary and proportionate for achieving that aim. Independent courts must remain the arbiters of what is and is not permissible speech, under clearly articulated laws. Some of the most problematic types of content are proscribed in European law, notably illegal hate speech and terrorist content. However, legal assessment is not straightforward, and even experts and courts differ when evaluating the legality of content. It should not be the case that *de facto* legal standards are set by company reviewers or automated content moderation tools, or delegated to administrative authorities. Moreover, if policymakers consider some types of content unacceptable and harmful, and it is not illegal, it is their job to legislate for it (respecting the human rights and rule-of-law principles referred to above). But it is inconsistent with these principles for governments to leverage private companies to limit speech that authorities cannot not directly restrict.
- 2. Policy should be based on the principle that content creators are responsible, under the law, for their online speech and behaviour.** Policy should empower users to post, share, and find content using platforms of their choice. It should also make it possible to hold users accountable for content they post, and how they otherwise behave. It should be clear to individuals that they are ultimately responsible under the law for what they choose to post online. People should be aware that if they post content that constitutes, e.g., illegal hate speech or defamation, they can be prosecuted for it. While online platforms can and should moderate content they host, enforce their terms of service, and restrict illegal content, intermediaries should not be held legally responsible for content authored by third parties.
- 3. Policy should be based on solid evidence, and targeted at well-defined and well-substantiated public interest concerns.** Policymakers should recall that several pieces of EU legislation have already been adopted (or are in the process of being adopted) that impose new obligations and responsibilities for platforms. These measures include the [DSM Copyright Directive](#), which obliges a broad range of content hosts to take particular measures, such as filtering, to prevent unlicensed copyrighted content from being uploaded. The [Audiovisual Media Services Directive](#) calls for the setting up of codes of conduct in order to ensure that minors are not exposed to types of content that may be considered harmful to them. The current draft [Terrorist Content Online Regulation](#) would impose duties of care as well as a requirement on content hosts to suppress content that is deemed illegal under the regulation, within one hour of notification. These pieces of legislation are not yet in force, and their effects are as yet unknown. New measures, such as the possibly forthcoming Digital Services Act,

should be carefully calibrated to focus on clearly defined problems that are not addressed in other legislation.

4. **Policy should ensure clarity about requirements for responding to notifications of illegal speech.** In general, platforms should have adequate and transparent [notice-and-action processes](#) that include safeguards and sanctions against wrongful or malicious notification. A content host should not be sanctioned for refusing to remove or downgrade content solely because it has been labeled by a non-judicial actor as illegal. Any new legislation should also be flexible enough to enable platforms to remain passive hosts with regard to some content, and to be active curators and moderators with regard to other content. New legislation will need to grapple with the distinction between active and passive hosting, and determine what level of responsibility companies should take for content it engages with in different ways. It is essential that platforms' efforts to restrict illegal content does not lead to a presumption of knowledge of illegality. Any new legislation should introduce a version of the [Good Samaritan principle](#), in order to ensure that intermediaries are not penalized for good faith measures against illegal content. Sanctions should only be applied in cases of proven systemic failure to respond to valid notifications of illegal content.
5. **Content hosts should not be discouraged from, or limited in their capacity to moderate content.** As a principle, it is both legitimate and desirable that platforms restrict types of lawful content they do not consider, for whatever reason, appropriate for the service they provide. Different platforms serve different communities and purposes, and not all content is suitable for all platforms. It is important to note, however, that the legal status of such content moderation is currently not clear. European courts have in certain cases ruled that a content host may not restrict lawful content, while in other cases ordering hosts to restricting which the host had not considered in violation of either the law or its own terms of service. Any future regulation should be aimed at providing legal certainty to hosts of user-generated content about their ability to moderate their users' lawful speech. Policy should seek to incentivise [human rights-respecting content moderation](#), as recommended by the UN Special Rapporteur on Free Expression.
6. **Use of technological solutions for online content moderation should not be mandated by law.** Content moderation technologies are being used increasingly by a broad range of internet companies. These technologies evolve constantly and will continue to do so, but currently remain quite rudimentary. For example, tools for automating social media content analysis have limited ability to parse the nuanced meaning of human communication, or to detect the intent or motivation of the speaker. They are not able to understand the subtlety of context and meaning, which is necessary to determine whether a statement posted on social media may be considered to violate the law, or terms of service. Policymakers must [understand these limitations](#) and should not mandate the use of endorsing or adopting automated content analysis tools or impose time limits on responding to notifications of illegal content, which in

practice will necessitate the use of automated filters to comply with the law. The DSM Copyright Directive has already imposed a *de facto* requirement to use filtering technology. This approach should not be followed in future legislation.

7. **Responsibility for content should not be imposed on other companies than the content host.** Infrastructure service providers, payment providers, advertisers, 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. Only the company with direct relationships with uploaders, and ability to take decisions on discrete pieces of content, should be responsible for it.
8. **Policy should promote, not hinder, innovation and entrepreneurship.** One of the most important and successful features of the limited liability provisions in the ECD is their capacity to encourage innovation and entrepreneurship. Had it not been for these protections, the many thousands of online sites and services that have appeared in Europe and beyond would not have grown and prospered. If new legislation undermines these protections, and introduces new responsibilities and obligations calibrated for global internet companies across the board, it will have a disproportionate negative impact on small companies and start-ups, and could further shrink the diversity of platforms and hosts available to support a broad range of expression online. Compliance, implementation, and litigation costs would disadvantage small companies, while larger, incumbent technology companies would be much better positioned to absorb these costs. It will be essential to ensure that obligations that may be suitable for the largest global networks are not applied to smaller operators.
9. **Content hosts should not be forced to apply one Member State's restrictions on free expression outside that country's territory.** Cross-border enforcement of restrictions would lead to unacceptable infringement of free expression and access to information rights. EU Member State laws vary considerably in how they restrict free expression. For example, some countries criminalise content such as blasphemy, while others have abrogated blasphemy laws; many countries prohibit hate speech but apply those prohibitions differently based on the cultural and historical context of their particular state. If hosts are required to apply one country's speech restrictions broadly, they will inevitably encounter conflicts of law and the space for free expression and public debate would be severely curtailed.