Information Law and Policy Lab

Protecting fundamental rights online with a “User Protection Duty”

Towards an intermediary liability framework which respects fundamental rights

Naomi Appelman, Thijs Bakker & Paula Hooyman

Amsterdam, August 2, 2019

This position paper is written by Naomi Appelman, Thijs Bakker and Paula Hooyman of the Glushko & Samuelson Information Law and Policy Lab (ILP Lab) of the Institute for Information Law (IViR) of the University of Amsterdam. The ILP Lab is a student-run, IViR-led institution which develops and promotes research-based policy solutions that protect fundamental rights and freedoms in the field of European information law.

It has been written in partnership with the Dutch digital rights organisation Bits of Freedom. This position paper reflects the recommendations and conclusions of the authors of the ILP Lab. As part of the research conducted, we consulted the following stakeholders: Buma/Stemra, EDRi, Facebook, Marktplaats, Medium, Microsoft, the Dutch Ministry of Justice and Security and Nederland ICT. We are also very grateful for the input received from Egbert Dommering, João Quintais, Joris van Hoboken, Ot van Daalen, Paul Keller and Stef van Gompel.

Executive summary

The current legal framework for intermediary liability is focused on the removal of unlawful information and neglects the protection of lawful information. To remedy this, we propose a "User Protection Duty": the duty for intermediaries such as hosting providers to protect the fundamental rights of their subscribers and others affected by the information they host, when interfering with this information. This duty is necessary to prevent excessive removal of lawful information online.

The User Protection Duty applies to the *action* providers take when they have *become aware* that information is (1) in violation of their Terms of Service or (2) unlawful. It does not apply to *acquiring* this knowledge. Thus, the User Protection Duty should be seen as an addition to the safe harbour framework.

As part of the User Protection Duty, hosting providers have to, when responding to information on their platform:

1. differentiate their responses in relation to different types of information;
2. adhere to a due process-requirement and protect user privacy;
3. take the nature of the reporter into account;
4. undertake action within a reasonable timeframe;
5. require a minimum level of proof of unlawfulness before taking action;
6. publish a transparency report and adhere to an accountability requirement.

Further guidance should be given in the form of EU guidelines. An independent regulatory body should oversee enforcement and implementation of the User Protection Duty.

1. Introduction

1.1 Problem: lack of protection for user shared information

Both fundamental freedoms and safety online are important values. However, the current legal framework for intermediary liability is focused on the removal of unlawful information and neglects the **protection of legal information**.

Firstly, while hosting providers in principle are free to remove **User Shared Information ("USI")** in **violation of their Terms of Service ("ToS")**, also in view of the freedom to conduct a business, these terms, currently often do not sufficiently protect the interests of users. As a result, providers do not weigh different fundamental rights in response to such a violation. This creates a risk of excessive removal of (lawful) USI without a viable legal recourse.

Secondly, under the current e-Commerce Directive, hosting providers can only be held liable for unlawful information when they have knowledge of the unlawful nature and do not act upon that knowledge[1]. This safe harbour framework stimulates hosting providers to unhesitatingly remove (lawful) USI when obtaining knowledge of alleged unlawfulness, leading to excessive censorship.

It is necessary to protect the communication freedom and the rights of uploaders through additional regulation when hosting providers handle responses to USI. Even though hosting providers vow they have their users interests at heart, their users rights are currently not sufficiently protected.

1.2 Solution: User Protection Duty

In order to combat excessive censorship of (lawful) USI and to promote legal certainty, we propose to create the following duty:

> **"Hosting providers, when interfering with the information they host, have the duty to protect the fundamental rights of their users and others affected by this information."**

The User Protection Duty only applies when a hosting provider plans to remove, suspend or otherwise interfere with the USI they host. Below we will elaborate further by setting out the minimum set of duties of hosting providers when acting upon a violation of their ToS or knowledge of alleged unlawful information (hereafter referred to as "responsive USI"). The User Protection Duty is a duty of care and therefore should be considered a **best efforts obligation**.[2] By implementing this duty of care, hosting providers are obliged to be as good as their word when they promise to protect their users interests.

[1] See article 14 of the e-Commerce Directive.
[2] Such a duty of care is explicitly compatible with the e-Commerce Directive as stated in in recital (48).

2. User Protection Duty: Scope

2.1 The User Protection Duty applies to hosting providers

Hosting providers are service providers that host information provided by their users. The term hosting provider covers a broad range of companies, from the more passive companies that can host your website to the more actively involved social media platforms, such as Wordpress, Facebook and YouTube.

2.2 What actions fall under the User Protection Duty?

The User Protection Duty applies when a hosting provider plans to remove, suspend or otherwise interfere with the USI they host.

Hosting providers take these actions when the USI is either in **violation of their ToS** (1) or allegedly **unlawful** (2). Knowledge of a ToS violation or unlawful information can be obtained either via notices by users, governed by notice and action ("NA") policies or proactive measures implemented by the hosting provider itself, such as automatic monitoring/filtering.

The User Protection Duty only applies to the **action** that hosting providers take *when* this knowledge is obtained: it does not impose a duty in relation to how such knowledge is **acquired**. In particular, the User Protection Duty does not imply a filtering or monitoring obligation of hosting providers.

2.3 Proportionality of the User Protection Duty: risk based assessment

The User Protection Duty allows for flexibility in its application. The extent of the duty can vary, depending on the **societal impact**, **capabilities** and **business model** of the hosting providers. This should lower the burden for non-commercial or low impact hosting providers, enabling new hosting providers to enter the market and promoting a diverse online debate.

The larger the possible societal impact of responsive USI, the larger the responsibilities under the User Protection Duty. The societal impact and associated risks can be assessed based on factors such as the breadth of coverage, number of users and the business model of the hosting provider. A lighter duty is also imposed on hosting providers with limited capabilities, often non-profit providers, such as Wikipedia. With regard to capabilities, factors such as the time zone where a provider operates, and its capacity, working days and hours and turnover are relevant. The business model is also relevant: where a hosting provider is a for-profit organisation, the User Protection Duty should already apply more stringently, and where this provider makes its money with interfering with hosted information, for example when selecting and organising USI, it should be applied even more diligently.

The possible societal impact, associated risks and actual capabilities of the hosting provider should be substantiated in the transparency reports, enabling enforcement of the User Protection Duty (see 'Transparency and accountability' §3.6).

2.4 Violation of the User Protection Duty

The violation of the User Protection Duty creates the possibility of **civil liability** and **administrative sanctions** (see §4). This is a form of primary liability. Possible secondary liability of the hosting provider for the underlying responsive USI (such as IP infringement) remains unchanged when the User Protection Duty is not adhered to. However, when the User Protection Duty is properly adhered to, hosting providers are exempted from both this primary and secondary liability. Consequently, even when the final action taken by the provider turns out to be incorrect, complying with the User Protection Duty exempts the hosting provider from the secondary liability. For example, a provider receives a notification of a possible IP infringement. In correctly following its User Protection Duty, the provider decides the responsive information falls under the parody exception, and should not be removed. When in a possible later court ruling this decision turns out to be incorrect (the responsive information does not fall under the parody exception) the provider cannot be held liable for not removing the information because it reached its decision in accordance with the User Protection Duty.

3. Six requirements of the User Protection Duty

The general obligation that the User Protection duty creates can be divided into six specific duties.

3.1 Duty 1: differentiate between different types of information

Various types of responsive USI can be identified. Examples are hate speech, terrorist information, images of sexual abuse of children, revenge porn, defamation and intellectual property infringements. The current legal framework for responding is the same for all these types of information. This 'one-size-fits-all' approach is unsuitable and hosting providers should implement different action policies when interfering with different types of USI. These different action policies should also take into account the violations of ToS.[3]

1. Knowledge-and-notification
   As the least invasive tool, this policy simply follows the minimum requirements of the due process norm (see §3.2-I). The hosting provider is merely obligated to notify the uploader of the knowledge. The uploader can remove the content or respond to the notification within a limited period of time. This approach is best suited for alleged **IP infringements**, because this is the only type of information that inflicts monetary harm (as opposed to the more grave personal harm). For **copyright** infringements, and exception should be made. For these type of infringements, the new Copyright Directive and its obligations will apply.

2. Knowledge-wait-and-takedown
   Hosting providers are required to wait (for example a week) after giving the obligatory notification, before engaging in removal or blocking. If the uploader consents, does not respond or fails to convince the hosting provider, it may proceed with the takedown. This approach introduces a moment of judicious and/or judicial reflection, making room for a viable decision. It is best suited for **defamation** claims and **violations of ToS**. Defamation claims are often difficult to determine, especially when hosting providers do not have in-house legal expertise. When assessing the violation of the ToS, the fundamental rights involved should be respected. The ToS should not constitute a breach of the relevant fundamental rights.

3. Knowledge-and-suspension
   Upon obtaining the knowledge, the hosting provider temporarily suspends the hosted information, while waiting for the uploader to respond to the allegations. The information can be easily reuploaded when it has been ascertained that the USI is lawful. Depending on the intensity of the harm inflicted, the hosting provider can also choose to temporarily suspend the user *account*, preventing recidivism. This consideration should be substantiated in the transparency report. Given the greater harm involved and the fundamental rights at stake, **hate speech**, **revenge porn** and **terrorist information** should be the object of this more powerful action. For terrorist information, this tool is additional to the proposed Terrorist Content Regulation.

[3] We have gladly elaborated on the PhD thesis of C.J. Angelopoulos, 'European intermediary liability in copyright: A tort-based analysis' (2016).

4. Knowledge-and-stay-down
   In this approach, the hosting provider will, in addition to removing the information, ensure that that same information cannot be reposted on the platform. Since this is the most drastic measure, it is best suited for information which requires no context to assess, in particular **images of sexual abuse of children**.

5. Knowledge-and-judicial-takedown
   The reporter must obtain a court order for the removal or blocking of information before the hosting provider can be obliged to take action. As a part of general civil law, this option remains applicable to **all types** of information. Caution is advised for removal obligations for equivalent content[4]. These obligations raise concerns given the prohibition of a general obligation to monitor information.

   The circumstances of the case and the nature of the hosting provider can necessitate a different appropriate response than illustrated above. Relevant factors for deviation can be the intensity of the harm inflicted (to an individual or society as a whole), the newsworthiness, the possibility of automating the process and the nature and context of the information. Deviations should be explained in the transparency report.

3.2 Duty 2: adhere to a due process-requirement and protect user privacy

The User Protection Duty requires hosting providers to adhere to a due process-requirement when acting upon responsive USI. Five core requirements are given below. Further norms should be established in EU guidelines (see "Implementation & Enforcement" §4). The core requirements are:

1. A **notification** should always be given to the uploader that the hosting provider suspects his USI to be unlawful or in violation of the ToS. The grounds for this suspicion, and how the uploader can contest this allegation should be included. When a reporter is involved, he/she must be informed that the alleged infringer has been contacted.

2. The uploader and the reporter should always have the **right to appeal** a decision of the hosting provider internally.

3. The reporter should be required to attest under legal penalty to a good-faith belief of the **truth of the facts** stated. This helps to limit bad faith restriction requests, and can provide the basis for sanctions against those who send false notices.

4. Reporting **anonymously** should at all times remain possible for natural persons. Legal entities should provide their identity. An exception should be made for reporting IP infringements and defamation: these notices should be made by the person/entity that is actually harmed or their legal representative. Their identity should be made clear in their notices.

[4] See the AG's opinion on the case of Glawischnig-Piesczek v Facebook.

5. Throughout the entire process, hosting providers have to respect the **privacy** of both the reporter and the uploader when they are natural persons. As a general rule, hosting providers should not be allowed to provide the identity of the reporter to the uploader and vice versa, except when court ordered.

3.3 Duty 3: take the nature of the reporter into account

The third duty relates to the nature of the reporter when knowledge is obtained via notices. Firstly, when the reporter is a **government authority** (like the police or a ministry of justice), it may not report information to the hosting provider if it has specific legal competences to order a takedown. This safeguards the legal guarantees that such legal means offer for the fundamental rights of uploaders. When such authorities do not have specific legal competences for a takedown, there is no objection to allow them to file a regular notice.

Secondly, **trusted flaggers** can potentially play a role for hosting providers to detect information that should be removed. Trusted flaggers are individuals or organisations who have been given a special status as reporter, as a result of their particular knowledge of a specific type of information. The choice whether to use trusted flaggers should be up to the hosting providers themselves.[5] An important addition is that government (appointed) institutions should not be able to be trusted flaggers. Otherwise, the privileges of a trusted flagger would form a circumvention of the limited competences of a governmental institution. Trusted flaggers should adhere to some minimum safeguards concerning transparency and accountability. This can be further developed in EU Guidelines (see §4).

3.4 Duty 4: undertake action within a reasonable time frame

The fifth duty relates to the time frame in which a hosting provider should take action after obtaining knowledge of the problematic USI. The time frame should be 'reasonable'. This will be determined based on the **type of information** involved and the **potential harm** (see §3.1) . A too limited time frame will lead to disproportionate restriction to the freedom to conduct a business, and could also lead to automation in the processing of removal orders, with a further negative impact on the freedom of expression. For example, IP infringements warrant a less swift response than revenge porn, given the potential harm. Other relevant factors to establish a reasonable time frame can be further developed in EU Guidelines (see "Implementation & Enforcement" §4). Hosting providers should, as part of the transparency obligation, demonstrate why their policies contain a certain time frame for response.

[5] Using trusted flaggers puts less strain on the resources of the hosting providers to assess whether information is considered unlawful or not. However, caution is advised. There are risks in using trusted flaggers, as they might get priority over other reporters which disrupts the fair balance of fundamental rights. Also, they might be biased towards the interests of one party (e.g. when the trusted flagger is an advocacy group) and they will not make a well-balanced assessment of the information. Therefore, if trusted flaggers are used, they should only be used for information which is easy to assess, not requiring any context (like images of sexual abuse of children).

3.5 Duty 5: meet a minimum level of proof of unlawfulness before taking action

The fifth duty is primarily applicable to notices, and governs the level of proof a notice should meet in order to require a hosting provider to undertake action. The level of proof should be made public in the transparency report. Minimum requirements can be established in EU Guidelines (see "Implementation & Enforcement" §4). The specific level of proof can differ according to the **type of information** at hand. We propose that, at a minimum, notices must contain:

1. A **detailed description** of the specific information alleged to be unlawful.

2. An explanation of the **grounds for suspicion**. When possible, a reference to the law allegedly being violated and the country where that law applies, or the parts of the ToS that are violated should be included, but this is not mandatory.

3. A specification of the **exact location** of the material, such as a specific URL.

When determining what action to take with regard to alleged unlawful information detected by self-implemented **proactive measures**, a standard of proof must be established as well. Hosting providers should make explicit in their policies what standard they are employing, and publicize this in the transparency report.

3.6 Duty 6: publish a transparency report & adhere to an accountability requirement

Crucial for the functioning and enforcement of the User Protection Duty is transparency about its implementation by the hosting provider. In a yearly published transparency report, hosting providers should publish both the **qualitative** data on what their policies consist of, as well as the **quantitative** data pertaining to the actual functioning of these policies. Aggregation of the quantitative data is sufficient. This enables enforcement of the User Protection Duty, since the transparency report demonstrates whether a hosting provider has fulfilled its requirements (see "Implementation & Enforcement" §4).

Hosting providers should firstly be transparent about how they obtain knowledge of alleged unlawful information. When the hosting provider uses self-implemented **proactive measures**, this should be made clear, together with the functioning of these measures. The assessment of the (un)lawfulness of the information should be substantiated, making clear how the hosting provider has properly balanced the fundamental rights involved. Furthermore, hosting providers should be transparent about their NA policies and publish the quantitative data on how many and what type of notices are received and that types of actions followed.

Secondly, hosting providers should be transparent about the different **types of actors** that report unlawful USI (see duty 4, §3.4). Thus, when a hosting provider has received a government notice, it should be transparent about how these notices

were handled. The transparency report also covers any cooperation with possible trusted flaggers.

In addition, parallel to the GDPR, hosting providers should adhere to the principle of **accountability** and should be able to demonstrate their compliance with the User Protection Duty when requested. This way, hosting providers are responsible for their own compliance. At the same time, this approach creates and preserves evidence that can be used in, for example, court cases. Non-compliance with the transparency or accountability requirement is regarded as non-compliance with the User Protection Duty itself and leads, as stated in §2.4 and §4, to liability and possibly administrative sanctions.

4. Implementation & Enforcement

Two final crucial matters are how the User Protection Duty should be implemented and how, after implementation, this norm should be enforced. The User Protection Duty should be implemented in a new **EU Regulation,** as part of the expected reform of the e-Commerce Directive. The implementation of the proposed text itself is however not sufficient, especially from the perspective of legal certainty. Further guidance is needed. This should be governed in the form of **EU Guidelines** on a broad range of topics that are drawn up in consultation with relevant stakeholders. Guidelines should at the very least be created on the following subjects:

1. As the **extent** of the User Protection Duty **can vary** according to the possible impact, capabilities and business model, these factors have to be further defined.

2. The **transparency** requirement should be clarified. A guideline should elaborate in depth on what type of information, and in what form hosting providers are required to make their report public.

3. The **due process** norms hosting providers should adhere to when acting upon the different alleged unlawful types of information have to be clearly defined in a specific guideline.

4. A guideline should be created covering the **Notice and Action** policies themselves. This guideline should describe how hosting providers can concretely respond to different types of information and illustrate what time frame and standard of proof should be adopted.

5. The use of **trusted flaggers**. The minimum safeguards they should adhere to and their potential risks and advantages should be thoroughly analysed.

For the proper functioning of the User Protection Duty, it is imperative to create adequate enforcement measures. As stated earlier, the primary measure through which the User Protection Duty will be enforced is the **transparency report**. The information contained in the transparency report combined with the principle of accountability forms the basis of enforcing this norm.

As stated earlier, non-compliance with either the User Protection Duty or the underlying transparency and accountability requirement will lead to both civil liability and possibly administrative sanctions. It will also leave open any liability based on the responsive USI itself. Finally, we propose that an **independent regulatory body** should oversee the implementation and function of the User Protection Duty, and that this body can create the suggested guidelines. It would also be authorised to hand out **administrative sanctions** in the case of non-compliance by the hosting providers and sanctions against reporters who, on purpose, repeatedly send false notices. It should be considered to oblige larger hosting providers with high societal impact to appoint a **USI protection officer**, similar to a Data Protection Officer under the GDPR.