

HateAid's priorities for the Digital Services Act Trilogue

In order to ensure effective protection for victims of online violence we urge negotiators to consider the following aspects in the Trilogue:

I. **Notice and Action Mechanism: Close the loophole that creates a zero-accountability regime for online platforms**

Notice and action mechanism should provide for a balanced approach to platform liability and exemptions from it – laying out the requirements for both - the users (substantiated notices when reporting illegal content) and the online platforms (assess and act upon a notification, when actual knowledge of illegal content is established). **However, the requirement that the European Parliament (EP) proposes under Article 14 paragraph 3 for establishing actual knowledge of illegal content is excessively high and nearly impossible to meet. It suggests that online service providers can only establish actual knowledge from the notices, if they do not need to conduct a legal or factual examination.**

It is not clear how actual knowledge could ever be established, when **online providers are granted a universal excuse** - most of the illegal content notified might require a factual or legal examination. In fact, this provision would defy the meaning of notice and action mechanism and intermediary liability regime as we know it. Furthermore, it could significantly reduce the quality of content moderation, because this loophole strips online service providers of any accountability – nearly all reported illegal content that is not manifestly illegal could be easily hosted by online providers even after being notified without any consequences. If actual knowledge cannot be established by using the official notification procedure the platform provides, it will hardly be relevant at all – relieving platforms of any meaningful accountability for specific items of content.

Furthermore, users should be able to **report illegal content anonymously**, and have their identity protected, not disclosing it to the recipient of service who provided the content. If any exceptions are established, like under the EP proposal for Article 15 paragraph 5a (new), personality rights violations should not be among them. Excluding cases of personality rights violations is not comprehensible, as this includes death and rape threats and all kinds of insults and could put the notifier in danger.

Moreover, **we warn against further restricting options to impose obligations on online platforms, especially by courts**, as this would also exclude them from ensuring a minimum level of protection for users. Especially when they are affected from mass attacks. In practice this means that illegal content could spread on online platforms, and even courts could not ask them to find and deal with a specific piece of content. It could be particularly relevant when court would want to order to find pieces of content that are already ruled illegal by courts, that could have spread further to closed groups and on private profiles. **A restriction suggested by the EP in the changes to Article 7 paragraph 1 would exclude courts from demanding further cooperation from platforms and leaving victims entirely helpless.** For example, this [landmark case](#) pursued by HateAid, would become impossible.

II. User redress: Empower **all users** to act against wrongful platform decisions

Make sure that **all users have access to redress mechanisms established under Article 17 and consequently - under Article 18**, as suggested in the Council General Approach. Limiting access to these crucial mechanisms only to content uploaders, would deprive users that unsuccessfully report illegal content of this right.

The only profiteers of limiting access to redress mechanisms are the very large online platforms, allowing to save money in the already disproportionally small and criticised in-size safety budgets. It could even provide an incentive to weaken the notice and action procedures and ignore notifications, as these decisions will remain unchallenged.

Allegedly feared misuse of complaint mechanisms by “bad actors” is a valid concern, that has to be addressed by looking at how automated decision-making works and leads to wrongful content removal and disabling of legit profiles. The misuse is possible and attractive because of automated decision-making that lacks human oversight. To be clear: misuse by “bad actors” is a result of bad content moderation by online platforms and should not be used to justify stripping users of their right to redress.

III. Points of contact: enabling an effective communication with users

We consider **points of contact that are accessible not only to authorities but also to recipients of the service to be particularly important**. Although the EP introduced such points of contact in its report (AM 180), the **Article 10 a (new) should be further strengthened to make contact points useful for victims of online violence**.

For the contact points to serve consumers, they need to be **meaningful** and **accessible**. Meaningful implies that users can communicate with online platforms, and this communication does not solely rely on chatbots and automated replies. Furthermore, for accessibility, considering the language diversity across the EU, VLOPs should make sure that users can communicate with their contact points in **all the official languages of the EU**. It can be especially important for contact point accessibility to younger and older users, as well as minorities.

Moreover, the contact points for users should also serve for delivery of official documents, e.g., evidence or formal requests that are made to initiate legal proceedings.

IV. Special rules for porn platforms: addressing image-based sexual abuse

We strongly support the position of the EP on establishing additional requirements for pornography platforms, introducing Article 24b (new). It provides additional safeguards for users of porn platforms and effectively protect people, especially women, from image-based sexual abuse. While there is no comprehensive data about the gravity of issue across the EU, the problem has only grown during the pandemic.

Researchers have [warned](#) that image-based sexual abuse is alarmingly common, and that victims report significant delays in removal of non-consensual material from porn platforms. In most cases police is unable to help, since it is difficult to establish the identity of the perpetrators, dismissing the case. Therefore, to act preventively and reduce the risk of image-

based abuse material being shared, it is necessary and proportionate to introduce verification obligation for content uploaders.

Moreover, the provision introduces requirement for porn platforms to guarantee reliable and skilled content moderation and fast removal of abuse material upon notification by victims. Researchers [report](#) that non-consensual content on porn-platforms is easily accessible, often even on the front pages of the most popular porn platforms.

V. Trusted Flaggers: cut the red tape for NGOs to be trusted flaggers

An effective system of trusted flagging heavily relies on the civil society - often publicly or donor funded NGOs, like HateAid, that have the best incentives to become a trusted flagger. These are organisations who have expertise in the respective field, justifying the trust that is put into these organisations and system as such.

It is important to **not overburden NGOs with a red tape, too strict requirements to application, expertise and obligations that may deter them** from becoming a trusted flagger and fulfilling these functions. Some of the requirements introduced under AM249 of EP's position are disproportionate and burden NGOs with extensive reporting obligations. This kind of obligations could be especially challenging to fulfil, considering scarcity of funding that many NGOs face or NGO grants that are designated for a specific purpose, leaving no flexibility for extra tasks. An NGO like HateAid would not have sufficient resources to meet all the requirements suggested in AM 249 of EP position and therefore most likely not even apply.

Instead, we suggest shifting the burden of reporting requirements concerning functioning of trusted flaggers from NGOs to online platforms, who could easily generate this information with help of a few clicks.

Furthermore, introduction of trusted flaggers system should not be a reason for limiting access to user redress mechanisms or treating notices flagged by users themselves with a lower priority.

VI. Removal orders: Give victims the option to seek help from the authorities

We urge you to support the proposal from the EP, empowering users to approach authorities to seek removal orders, in case their rights have been infringed by illegal content, and allowing authorities to issue such orders (AM159; Article 8 paragraph 4a). It is a safety net that users can turn to when personally faced with online violence. We consider it to be limited to rely solely on trusted flaggers to escalate a notification that has not received proper attention from the platforms. Trusted flaggers can be a particularly useful complementary tool to the reporting system; however, it is heavily relying on civil society to do the work that platforms should be doing. Sadly, this also means to rely on civil society to do this work voluntarily or using their own funds for it. We cannot expect the same type of support and accountability, as we should be expecting from more institutionalised authorities.

For this reason, we should give users additional options. We strongly support the EP's suggestion to amend article 8 paragraph 4a and give users an option to complain to authorities directly.



About HateAid

HateAid gGmbH was initiated in 2018. We are the first organization in Germany to offer protection from digital violence to those affected and at the same time to support effective sanctioning of the perpetrators. Moreover, we create social awareness of the destructive effects of digital hatred on our democracy. HateAid's aim is to relieve the burden of the victims of attacks, enforce their rights, deter the perpetrators, and overall strengthen our democracy and society. As part of the Landecker Digital Justice Movement, HateAid advocates for more platform responsibility on social media.
