

Annex 2: policy options on managing online illegal or harmful content

Our policy options are articulated around two pillars - in the case the European Commission may consider the e-Commerce Directive needs to be reopened:

- **High-level obligations to apply to all “hosting service providers”** (HSPs - i.e. platform enabling users to share content publicly).
- **More targeted measures to apply to Data-monetizing platforms (DMPs)**, acknowledging the “incentive gap”.

General measures for all HSPs:

Our main overall suggestion would be to develop a **strengthened “duty of care” principle** for HSPs (similar to the UK White Paper approach, recently endorsed by France as well), whereby all HSPs should take measures to ensure online safety, suppress hate speech and more generally illegal content. New specific obligations would ensure the practical application of that principle:

- A notice and action framework, which includes the obligation to **set up an easy-to-use, effective reporting tool**, allowing for the “expeditious” removal of illegal content. Points 5 to 8 of the Commission’s [Recommendation on tackling illegal content](#) could serve as inspiration.
- A binding **“Good Samaritan”** provision, which would incentivize pro-active measures, including the use of **automated tools** to prevent the re-upload of illegal material, or to effectively ban/restrict users sharing illegal material/behaving illegally.
- A high-level provision stating that **HSPs should not “over-remove” content** and protect free speech, such as through an obligation to set up an effective redress mechanism for contents providers, as already foreseen in the draft terrorist content regulation.
- At the same time, an acknowledgement that platforms should **be free to act as per their terms and conditions**, and remove content or block users which contravene to these T&C.
- Incentives to **drawing up industry-specific codes of conducts** on tackling illegal content and ensuring user safety, the compliance to which would be deemed compliance to the law. These codes of conducts would be validated by the EU regulatory authority mentioned below.

A European **regulatory authority** or regulatory body would be created, overseeing the implementation of these rules. This authority would be the prime surveillance body and recipient of the various obligations aimed at DMPs (below). It could also serve as an effective industry-government cooperation forum, including in its mission to support and oversee Codes of Conducts. The authority would also have the power to request **data** from all HSPs (under conditions to be defined), in order to verify compliance with relevant regulations.

Specific additional measures for DMPs

On top of the above-listed provisions, measures strengthening oversight, transparency and sanctions would specifically apply to DMPs, thereby filling the “incentive gap” with direct-pay platforms.

- **Transparency:** DMPs would need to publish annual / quarterly transparency reports on online safety and content management, which would include:
 - o A description of its content management policies and tools;
 - o Figures on content flagged and removed as well as on complaints received and processed / reinstated content;
 - o These report should be drafted according to guidelines, to ensure consistency of reporting formulas and provide comparable numbers.
- DMPs should give the independent authority and third-parties (eg NGOs, under conditions to be defined) access to the corresponding raw data, in order to allow for the independent review of their content moderation policy.
- DMPs would need to be transparent on the main parameters of their automated filtering technologies.
- **Reporting tools:** DMPs would need to abide by quality standards developed to ensure user-friendly reporting, which could include a specific logo. There should be an “obligation of means”, controlled by the regulator, to ensure tech companies deploy the right resources (financial resources, appropriate staff including in terms of language – see France’s Mission Facebook report).
- **Independent audit:** DMPs could be obliged to include in their annual/quarterly report the outcome of an independent auditing of their content moderation policies.
- **Content removal timeframe:** there could be a 24 hours removal timeframe for illegal content, down to one hour for highly problematic content, such as terrorist content (as suggested in the French Avia report and the draft EU terrorist content regulation).
- **A specific sanction regime** for DMPs, to be defined (possibly aligned with GDPR, i.e. up 4% of the annual global turnover as suggested in the UK White Paper).
- **Automated content filtering tools:** the authority would be empowered to force DMPs to apply automated content-filtering measures if it notes that it has systemic failure to deal with illegal content. After the first six months/year of application, DMPs would be able to have the authority review the necessity of this measure. It should come as a last-resort measure, after prior warnings have been issued, and should be subject to appropriate proportionality safeguards.