



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

Warsaw, 3 July 2026

**DG Connect Draft Commission Communication *Commission Guidelines on the trusted flagger mechanism under Article 22 of Regulation (EU) 2022/2065* (version for targeted public consultation 29 May-10 July 2026) – opinion and remarks**



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

## Executive summary

- Draft Guidelines present a missed opportunity in relation to filling certain gaps in the DSA framework and providing interpretative guidelines allowing to safeguard the necessary balance between the need to moderate online content and preservation of the freedom of speech.
- Trusted Flaggers play a vital role in the DSA ecosystem, directly by their privileged status as the notifiers, and indirectly due to their role in identifying the systemic risks.
- Given the cursory character of the provisions related to the Trusted Flaggers in the DSA, the Guidelines should be viewed as *de facto* regulation of this institution.
- Thus, the importance of the Draft Guidelines for the DSA application lies at hand.
- Unfortunately Draft Guidelines perpetuate and, in fact, magnify the problems connected to the DSA, rather than alleviating them.
- Draft Guidelines seem to adopt an overly broad definition of illegality and encourage an overly broad treatment of this concept by the Very Large Online Platforms (“**VLOPs**”).
- Too little emphasis is put on the Trusted Flaggers being obliged to clearly demonstrate their legal expertise in particular Member States’ legal systems and providing the corresponding legal justifications.
- Draft Guidelines fail to account for the rights of the entities affected by the moderation (creators/recipients of the content).
- Draft Guidelines fail to introduce effective accountability mechanisms for the trusted flaggers.
- Draft Guidelines fail to stipulate cooperation mechanisms for the Digital Services Coordinators (“**DSCs**”) overseeing the Trusted Flaggers.



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

- Further, Draft Guidelines fail to introduce mechanism safeguarding the impartiality and objectivity of the Trusted Flaggers.
- This is especially disconcerting granted that this status may be awarded also to public-law entities.



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

## **Background: systemic problems regarding the DSC Framework**

1. The analysis of the DG Connect Draft Commission Communication *Commission Guidelines on the trusted flagger mechanism under Article 22 of Regulation (EU) 2022/2065* (version for targeted public consultation 29 May-10 July 2026) (“**Draft Guidelines**”) should be conducted in the light of the more general aspects of the Digital Services Act Framework. Only then can the provisions of the Draft Guidelines be fully appreciated. At this place, the analysis will concentrate on the aspects of the regulation related to the freedom of expression and the role of the Trusted Flaggers as the gatekeepers.
2. To begin with, the scope of regulation of the Digital Services Act is extremely broad, so as to cover the functioning of intermediary services that are to be provided on the territory of the EU (Articles 1-2). Consequently, it is applicable also to the social media platforms and search engines such as Google, Facebook or X. Its aim is to create **a safe, predictable and trusted online environment that facilitates innovation and in which fundamental rights enshrined in the Charter, including the principle of consumer protection, are effectively protected** (DSA Article 1.1). As elaborated further in the Preamble (para 9), its goal is to address **the dissemination of illegal content online and the societal risks that the dissemination of disinformation or other content may generate, and within which fundamental rights enshrined in the Charter are effectively protected and innovation is facilitated**.
3. In the pursuit of this goal, among others, the DSA foresees the creation of notice and takedown mechanisms (Article 16 DSA) and the sanctions should be accompanied by reasons (Article 17 DSA). Further, the online platforms providers should have also a redress system for the victims of content moderation. Furthermore, as per Article 23 DSA, the providers of online platforms should suspend accounts generating disproportionate amount of illegal content or manifestly unfounded complaints. The platforms are obliged to produce yearly reports on their activities (Article 24).
4. In the pursuit of this goal, among others, the DSA foresees the creation of notice and takedown mechanisms (Article 16 DSA) and the sanctions should be accompanied by reasons (Article 17 DSA). Further, the online platforms providers should have also a redress system for the victims of content moderation.



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

Furthermore, as per Article 23 DSA, the providers of online platforms should suspend accounts generating disproportionate amount of illegal content or manifestly unfounded complaints. The platforms are obliged to produce yearly reports on their activities (Article 24).

5. Operators of very large online search engines and platforms (e.g. Google, Youtube, Facebook, X) have additional obligations. Of them, of particular importance is the obligation to make a risk assessment. As per Article 34.1 DSA:

*This risk assessment shall be specific to their services and proportionate to the systemic risks, taking into consideration their severity and probability, and shall include the following systemic risks:*

*(a) the dissemination of illegal content through their services;*

*(b) any actual o.r foreseeable negative effects for the exercise of fundamental rights, in particular the fundamental rights to human dignity enshrined in Article 1 of the Charter, to respect for private and family life enshrined in Article 7 of the Charter, to the protection of personal data enshrined in Article 8 of the Charter, to freedom of expression and information, including the freedom and pluralism of the media, enshrined in Article 11 of the Charter, to non-discrimination enshrined in Article 21 of the Charter, to respect for the rights of the child enshrined in Article 24 of the Charter and to a high-level of consumer protection enshrined in Article 38 of the Charter;*

*(c) any actual or foreseeable negative effects on civic discourse and electoral processes, and public security;*

*(d) any actual or foreseeable negative effects in relation to gender-based violence, the protection of public health and minors and serious negative consequences to the person's physical and mental well-being.*

6. Although theoretically the concept of illegal content should encompass only content illegal under EU or specific Member States' law (Article 3 h) DSA), the whole DSA system seems to encourage a wide reading of this provision, as evidenced, e.g., by Motif 12 of the Preamble,



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

*In particular, the concept of **'illegal content'** should be defined broadly to cover information relating to illegal content, products, services and activities. In particular, that concept should be understood to refer to information, irrespective of its form, that under the applicable law is either itself illegal, **such as illegal hate speech or terrorist content and unlawful discriminatory content**, (...).*

7. For the avoidance of doubt, the systemic risks should be understood even broader and base on the wide reading of the illegality. As per para 84 of the Preamble:  
*When **assessing the systemic risks** identified in this Regulation, those providers should also focus **on the information which is not illegal, but contributes to the systemic risks** identified in this Regulation. Such providers should therefore **pay particular attention on how their services are used to disseminate or amplify misleading or deceptive content, including disinformation.***
  
8. Article 35.1 explains the ways to mitigate the risks as, among others:  
(...)  
*(g) **initiating or adjusting cooperation with trusted flaggers** in accordance with Article 22 and **the implementation of the decisions of out-of-court dispute settlement bodies** pursuant to Article 21;*
  
9. In light of the above it is thus clear, that the service providers invited to aggressively moderate the content of the users together with NGOs (chiefly Trusted Flaggers – see *infra*). The ideological character of these ambiguous concepts is evident in the light of the codes of conduct and other non-binding documents discussed below. The main legal basis for them is stipulated in Article 45 DSA. Its paragraph 2 stipulates that they may be created especially where there exists a systemic risk identified as per Article 34 DSA. What merits attention is that these codes should be drawn not only by the representatives of the industry, but also authorities and *civil society organisations and other relevant stakeholders*, Besides, the online service providers are expected to create and implement voluntary standards relating to technical aspects of their functioning (Article 44 DSA).



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

10. The provisions of the DSA should be enforced, in principle, by the Member States' organs (Article 49 ff. DSA), who should be ensured adequate resources to monitor their tasks. The DSA calls them Digital Service Coordinators, in practice this role is ascribed to digital market regulators (Article 49.1). They should have investigative and enforcement powers which, however, should be stipulated in detail only in the national regulations. In case of the very large online platforms and search engines, such as Youtube Google or Facebook, it is the Commission that should conduct the investigation and impose the penalties.
11. And the sanctions are indeed substantial. Theoretically, available remedies rely, in the first line, on requesting the online service providers to end infringements and provide compliance with the DSA. In practice this would mean that the targeted providers would have to adapt their business lines to the regulator's requests, submit action plans etc. While the detailed procedures and specific penalties are to be stipulated in the Member States' national regulations, the DSA sets certain minimum requirements. In particular, the maximal penalties for failing to remediate a violation of the Regulation should equal **6 % of the annual worldwide turnover of the provider of intermediary services concerned in the preceding financial year**, while the maximum penalty for supplying *incorrect, incomplete or misleading information, failure to reply or rectify incorrect, incomplete or misleading information and failure to submit to an inspection shall be 1 % of the annual income or worldwide turnover* (Article 52.3 DSA). In addition, there should be also a possibility of claiming money by the individuals (private enforcement), as stipulated in Article 54 DSA. The same 6% of the total worldwide annual turnover is foreseen for the penalties imposed by the Commission (Article 74.1. DSA). Consequently, it is clear that the VLOPs are strongly incentivised to take any input from the designated entities, such as Trusted Flaggers, seriously, in order to avoid escalations and the risks related thereto.
12. The role of the Trusted Flaggers in this ecosystem is defined primarily in DSA Article 22, according to which:
  1. *Providers of online platforms shall take the necessary technical and organisational measures to ensure that notices submitted by **trusted flaggers**, acting within their designated area of expertise, through the mechanisms referred to in Article 16, **are given priority** and are processed and decided upon without undue delay.*



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

**2. The status of ‘trusted flagger’ under this Regulation shall be awarded, upon application by any entity, by the Digital Services Coordinator of the Member State in which the applicant is established, to an applicant that has demonstrated that it meets **all of the following conditions:****

(a) it has particular **expertise and competence** for the purposes of detecting, identifying and notifying illegal content;

(b) it is **independent** from any provider of **online platforms**;

(c) it carries out its activities for the purposes of submitting notices diligently, accurately and **objectively**.

**3. Trusted flaggers shall publish, at least once a year easily comprehensible and detailed reports on notices submitted in accordance with Article 16 during the relevant period. The report shall list at least the number of notices categorised by:**

(a) the identity of the provider of hosting services,

(b) the type of allegedly illegal content notified,

(c) the action taken by the provider.

Those reports shall include an explanation of the procedures in place to ensure that the trusted flagger retains its independence.

Trusted flaggers shall send those reports to the awarding Digital Services Coordinator, and shall make them publicly available. The information in those reports shall not contain personal data.

Digital Services Coordinators shall communicate to the Commission and the Board the names, addresses and email addresses of the entities to which they have awarded the status of the trusted flagger in accordance with paragraph 2 or whose trusted flagger status they have suspended in accordance with paragraph 6 or revoked in accordance with paragraph 7.

**5. The Commission shall publish the information referred to in paragraph 4 in a publicly available database, in an easily accessible and machine-readable format, and shall keep the database up to date.**



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

6. Where a provider of online platforms has information indicating that a trusted flagger has submitted a significant number of insufficiently precise, inaccurate or inadequately substantiated notices through the mechanisms referred to in Article 16, including information gathered in connection to the processing of complaints through the internal complaint-handling systems referred to in Article 20(4), it shall communicate that information to the Digital Services Coordinator that awarded the status of trusted flagger to the entity concerned, providing the necessary explanations and supporting documents. Upon receiving the information from the provider of online platforms, and if the Digital Services Coordinator considers that there are legitimate reasons to open an investigation, the status of trusted flagger shall be suspended during the period of the investigation. That investigation shall be carried out without undue delay.

7. The Digital Services Coordinator that awarded the status of trusted flagger to an entity shall revoke that status if it determines, following an investigation either on its own initiative or on the basis information received from third parties, including the information provided by a provider of online platforms pursuant to paragraph 6, that the entity no longer meets the conditions set out in paragraph 2. Before revoking that status, the Digital Services Coordinator shall afford the entity an opportunity to react to the findings of its investigation and its intention to revoke the entity's status as trusted flagger.

8. The Commission, after consulting the Board, shall, where necessary, issue guidelines to assist providers of online platforms and Digital Services Coordinators in the application of paragraphs 2, 6 and 7.

13. This role was further elaborated upon in motifs 61-62 of the DSA Preamble:  
*Action against illegal content can be taken more quickly and reliably where providers of online platforms take the necessary measures to ensure that **notices submitted by trusted flaggers**, acting **within their designated area of expertise**, through the notice and action mechanisms required by this Regulation are treated with priority, without prejudice to the requirement to process and decide upon all notices submitted under those mechanisms in a timely, diligent and non-arbitrary manner. Such trusted flagger status*



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

*should be awarded by the Digital Services Coordinator of the Member State in which the applicant is established and should be recognised by all providers of online platforms within the scope of this Regulation. Such trusted flagger status should only be awarded to entities, and not individuals, that have demonstrated, among other things, that they have particular expertise and competence in tackling illegal content and that they work in a diligent, accurate and objective manner. Such entities can be public in nature, such as, for terrorist content, internet referral units of national law enforcement authorities or of the European Union Agency for Law Enforcement Cooperation ('Europol') or they can be non-governmental organisations and private or semi-public bodies such as the organisations part of the INHOPE network of hotlines for reporting child sexual abuse material and organisations committed to notifying illegal racist and xenophobic expressions online. To avoid diminishing the added value of such mechanism, the overall number of trusted flaggers awarded in accordance with this Regulation should be limited. In particular, industry associations representing their members' interests are encouraged to apply for the status of trusted flaggers, without prejudice to the right of private entities or individuals to enter into bilateral agreements with the providers of online platforms.*

14. What is more, the Trusted Flaggers are also referred to in other soft law acts regulating modalities of the DSA implementation. For example, they are explicitly mentioned in The revised Code of conduct on countering illegal hate speech online<sup>1</sup> (“**Hate Speech CoC**”) In addition to describing the role of the Trusted Flaggers in the notice and takedown procedures, the Hate Speech CoC expressly makes them eligible to act also as Monitoring Reporters (Definition in Annex I, p. 5). Monitoring Reporters are entitled not only to make privileged reports (section 2; Annex I section 3), but also play a prominent role in the standard setting exercises conducted by the industry and the relevant stakeholders (section 4). Consequently, the Trusted Flaggers in the context of the hate speech act not only as the enforcers of the existing regulations, but equally as the standard-setters. Importantly, they are neither bound by the Member States' law, nor accountable by any measure for their activities in this respect. Thus, in their role of both, enforcers and standard-setters,

---

<sup>1</sup> <https://digital-strategy.ec.europa.eu/en/library/code-conduct-countering-illegal-hate-speech-online> , accessed on 1 July 2026.



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

- they may effectively bypass the national organs in matters as sensible as balancing the freedom of expression and other freedoms. Other acts, such as, e.g., Commission Guidelines of 26 April 2024 for providers of Very Large Online Platforms and Very Large Online Search Engines on the mitigation of systemic risks for electoral processes pursuant to Article 35(3) of Regulation (EU) 2022/2065, C/2024/3014, while not stipulating any specific competences, underline the role of the Trusted Flaggers for their enforcement.
15. Even act that, such as, e.g., the The Revised Code of Conduct on Disinformation of 16 June 2022, integrated in the Digital Services Act legal framework on 13 February 2025, do not expressly refer to the category of the Trusted Flaggers while establishing extensive competencies of the “fact checking” organisations. Granted the institutional role provided for such, a clarification of the possible entanglements between the trusted-flagger and the Fact Checkers (i.e. empowered NGOs) is direly needed. In any case, the Draft Guidelines would be the right place to effectively explain the relationship between the Trusted Flaggers and “fact checkers”.
  16. Drawing from the above it may be said that since the DSA ecosystem poses tangible challenges to the freedom of expression on the internet, it is of particular importance to carefully consider the role to played in it by the gatekeepers, such as the trusted flaggers.
  17. The first issue of concern is the scope of moderation obligations: the social media platforms would be obliged to filter not only directly harmful content, such as incitation to or graphic depictions of violence; pornography etc., but also content falling under the umbrella of ambiguous concepts, such as “illegality” or “hate speech”, that may be used also to supress political opposition. Consequently, it is clear that it would be the political speech challenging the “consensus” would be particularly vulnerable to the threat of the over-moderation.
  18. The layered structure of the moderation framework also poses certain challenges. To begin with, the DSA introduces a robust notice and takedown framework, with little, if any regard being paid to the rights of the authors of the incriminating speech and their potential listeners. Within the framework of such proceedings, a preferential treatment should be provided for the complaints, in particular



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

submitted by the privileged entities, such as the Trusted Flaggers. In this context it has to be remembered that this compliance would be ensured also by the fact, that the same entities would also participate in the assessment of the existing risks; the VLOPs' reactions to their notifications will also be examined by the Commission. Furthermore, the very same privileged entities are invited to co-create the very standards they are going to implement within different frameworks, networks etc., as well as through educational actions aimed at the general public.

19. The above challenges are further amplified by the the harsh penalties for the online operators showing too much leniency towards the speech notified by the Trusted Flaggers.
20. All the above creates a structural imbalance in favour of over- rather than under-moderation, i.e. against the freedom of speech. This is particularly so, granted the striking imbalance between the rights of the content-creators (or recipients) and the Trusted Flaggers. One cannot lose out of sight that the notifications related to allegedly illegal content should always be seen not as a binary (Trusted Flagger-VLOP), but at least a triangular situation (Trusted Flagger – VLOP – content creator/consumer). Unfortunately, the DSA seems to put too little emphasis on this aspect.
21. Furthermore, granted that the “illegality” should be assessed by reference to the law of a given Member State, a proper weight should be given to the way this Member State’s legal system balances between the freedom of speech and the need to moderate the online content. Incentivising the VLOPs to over-moderate may result in the platforms violating the balance struck within the particular legal systems.
22. Seen from this angle Draft Guidelines would provide a perfect opportunity to mitigate at least some of the risks signalled above and providing more clarity as to the proper way of handling these discrepancies. Unfortunately, it is not the case.



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

## **Analysis of the Draft Guidelines**

23. Unfortunately, contrary to the expectations, the Draft Guidelines in their current form rather than alleviating the aforesaid imbalances and deficiencies seem to perpetuate, if not further strengthen them. In particular, despite their stated goal of striking balance between creating rules to tackle illegal content and safeguarding fundamental rights by elimination of the over- and under-removal of the content from the internet (para 2 of the Draft Guidelines), the Guidelines seem to be a rather one-sided instrument.

24. To begin with, the Draft Guidelines do little to remove the ambiguities surrounding the notion of illegality. On the one hand they remind that the concept is derivative and refers to the EU law and the laws of the Member States. As stated in para 11 of the Draft Guidelines:

*The Commission recalls that Regulation (EU) 2022/2065 itself does not define what ‘illegal content’ is. This is determined by national law in compliance with Union law or in other applicable Union law. Regulation (EU) 2022/2065 does not provide a legal basis to order the removal of illegal content either, nor does it empower the Commission, DSCs, or other competent national authorities to do so.*

25. At the same time, however, it contains many provisions suggesting that it should be understood more widely, e.g. by referring to EU taxonomies of illegal behaviours from the Regulation 2024/2835 or encouraging the VLOP operators not to conduct an in-depth illegality analysis of illegality of a given content under laws of different Member States. More importantly, this need to refer to the Member States’ legal orders is hardly emphasised throughout the document. On the other hand, at many instances the document emphasises the proactive role of the Trusted Flaggers as the content-moderators, while giving little if any weight to the need to preserve the freedom of speech.

26. Consequently, the references to the need of understanding the illegality differently for different states seem to serve as a sort of lip-service to the Member States’ competencies rather than a good-faith attempt at giving due to their inherent powers. And this has further, far-reaching consequences: Since there is no uniform concept of illegality across the EU, there is also no universal measure for the



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

balancing between the need to protect the internet users and the rights of content-creators, in particular their freedom of speech.

27. To begin with, it creates a situation where the balancing between the competing fundamental rights is conducted completely outside of the jurisdictional reach of the “native” Member States. To be more specific, actions of the Trusted Flaggers would not be controlled, even indirectly, by the national courts of this state. Secondly, the DSC overseeing the Trusted Flagger would not have the necessary qualifications to control the Trusted Flagger’s actions (Polish DSC has no knowledge of German law and vice-versa, etc.). Thirdly, structural factors favour the VLOPs being deferential to even more “aggressive” Trusted Flaggers. It is foremostly due to the VLOPs incurring substantial risks only in case of dismissing the Trusted Flaggers’ reports. In particular, should such decisions not to remove specific content significantly impact the statistics, the VLOP could attract the attention of the local DSC or the Commission, in particular in the context of the systemic risk assessment. As explained in para 13 of the Draft Guidelines:

*Accordingly, the role of trusted flaggers may also be relevant in identifying systemic risks stemming from the design or functioning of their service and its related systems or from the use made of very large online platforms’ and of very large online search engines’ services.*

In fact, in para 14 it is further elaborated that:

*The involvement of independent experts such as trusted flaggers may be useful for the purposes of identifying and mitigating such systemic risks. For example, measures to mitigate the systemic risk of dissemination of illegal content in the European Union may include, pursuant to Article 35(1)(g) of Regulation (EU) 2022/2065, initiating or adjusting cooperation with trusted flaggers.*

28. On the other hand, an aggressive approach to the removal of the reported content would, at most, involve a threat of isolated court proceedings at the local fora. This structural imbalance is further amplified by the fact that the actions of the Trusted Flaggers may be controlled only by the DSC of the Country of Registration, i.e. a DSC not only having no practical knowledge of foreign legal systems, but also having



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

little if any incentives to adopt a foreign definition of illegal content, narrower than its own. To remove any doubts, Draft Guidelines themselves underline that the role of the Trusted Flaggers is not to assist in finding a new balance, but rather *shedding light on potential gaps* in the providers' moderation systems (Draft Guidelines, paras 6, 13).

29. In light of the above the statement in para 11 of the Draft Guidelines, that

*It therefore remains the responsibility of each provider concerned to assess the notices and undertake the relevant content moderation action, if any, in a diligent, non-arbitrary, and objective manner.*

fails to convince.

30. The above threats are even more evident if to take into account that little was done to ensure the objectivity of the Trusted Flaggers. In **Chapter 3** dedicated to the Trusted Flaggers' requirements little was done to address this issue.

31. To begin with the Guidelines failed to take into account that the objectivity of the Trusted Flaggers required by Article 22.2.b) DSA could be negatively influenced not only by their ties to VLOPs, but also to other political and economic actors. In fact, the issue of the Trusted Flaggers' potential political involvement or even activism, that could negatively affect their impartiality, was simply not thematized in the document. What is even more striking, The document expressly allows that the Trusted Flaggers may not only use public funding, but may be public law entities. The only requirement set for the public law bodies is that they are organisationally distinct from the DSCs. Other than that, they should operate in the same regime as the extra-governmental. Further, while referring to the need to maintain the diversity and integrate different areas of expertise, the draft Guidelines at no place do refer to the need to reflect the plurality of views and opinions in the larger society. The above institutional deficits are accompanied by lack of emphasis on the need to safeguard a necessary level of proficiency in the different Member States' legal systems and the need for the DSC's to cooperate in this respect, in particular with regard to Trusted Flaggers active in more than one Member States.

32. **Chapter 5**, dealing with the submission of notices, and **Chapter 6**, pertaining to reporting requirements is similarly unsatisfactory as it presents a lost opportunity



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

of providing at least some balance between the need to moderate and the protection of freedom of speech. It does so chiefly by referring to the EU taxonomy of illegal activities, rather than the Member States' legal systems and by encouraging the VLOPs to aggressively moderate the content by resigning from conduct in an in-depth legal analysis whether content illegal under one Member State's law is illegal also under the law of the other. This lacuna should be assessed even more negatively as the Draft Guidelines do not require the notices to contain a legal reasoning indicating precisely the conditions for the purported illegality.

33. **Chapter 7**, concerning the integrity of trusted flaggers, is particularly disappointing, as it provides neither clear accountability criteria, nor empowers the parties affected by the Trusted Flagger's activity (content creators and recipients) with any substantial means of redress.

34. Most importantly, while acknowledging that the DSCs of the Trusted Flagger's seat may investigate their activities of their own motion or basing on the third-party information, Draft Guidelines provide not framework for reporting irregularities on the part of the Trusted flaggers by the affected parties (content creators and recipients), or by the DSCs of other Member States that were affected by the actions of a Trusted Flagger registered in another state. In addition, Draft Guidelines do not encourage the DSCs to cooperate in their investigations, even if a Trusted Flagger's activities covered more than a single Member State. Further, the Guidelines fail to address the problem of instrumental treatment of the notices by the Trusted Flaggers. Lastly, they do not list failure to provide legal justification for the notices as a separate assessment criterion. This, again, does not encourage them to provide effective legal justifications to their decisions.

35. Lastly, **Chapter 9** pertaining to the institutional support for the Trusted Flaggers fails to provide the guidance on how to organize the support scheme in a way safeguarding the Trusted Flaggers independence.

36. More specific remarks and proposals are listed in the following section.



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

### Specific recommendations:

- **Para 6**

The aim of the Trusted Flaggers involvement should not be limited to *closing gaps* in the moderation system, but should also encompass *providing assistance in striking the right balance between competing fundamental rights*

- *Proposed recommendation should contribute to striking the right balance between the need for moderation and the protection of fundamental rights.*

- **Para 20**

The circle of entities entitled to the initiate a control of the Trusted Flaggers' notifications should be widened so as to encompass, in addition to VLOPs, also the affected Members of the Public.

- *Proposed recommendation would give voice also to those affected by the Trusted Flaggers wrongful activities*

There should be also a specific procedural solution allowing to report imprecise or inadequately substantiated notices also to DSCs of the Member States other than the Trusted Flagger's seat (potentially) affected by the Trusted Flagger's notifications that would later cooperate with the DSC of the seat in the assessment of the Trusted Flagger's status.

- *Proposed recommendation would ensure that the DSC assessing the Trusted Flaggers' notices would have all the necessary expert knowledge.*

- **Para 26**

The eligibility criteria should expressly indicate that the Trusted Flagger's expertise should also indicate the proficiency in legal systems of particular Member States.

- *Proposed recommendation would ensure that the Trusted Flaggers would have the necessary expertise to assess the illegality of particular content in light of the applicable Member States' law and would facilitate the assessment of their expertise by the VLOPS and DSCs.*

Independence Criteria should reach beyond lack of fiscal ties to VLOP and encompass also impartiality and apolitical character.

- *Proposed recommendation would prevent misuse of the Trusted Flagger's position to further goals unrelated to the content moderation.*



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

- **Para 29**

The circle of entities capable of acting as trusted flaggers is inconsequential. Granted the Flaggers wide competencies and extraterritorial reach, one should opt either for public, or private entities.

- *As public and private entities operate on different principles, they should be governed by different rules. Mere organizational decoupling of the public trusted flaggers from organs with decision-making powers in paras 30-31 does not suffice.*

- **Para 32 b)**

The provision should also cover assessment of other factors capable of negatively affecting the impartiality of the Trusted Flaggers than representing collective interests, such as political involvement.

- *Proposed recommendation would prevent misuse of the Trusted Flagger's position to further goals unrelated to the content moderation.*

- **Para 34 (in particular let. a)**

The provision should also expressly introduce the possibility of the DSCs being entitled to consult their peers in matters related to the verification of the relevant qualifications, in particular related to given national legal systems.

- *Proposed recommendation would ensure that the assessing DSC would have all the necessary expert knowledge, in particular in case of multinational Trusted Flaggers.*

- **Para 36, 39**

The provision should also cover assessment of other factors capable of negatively affecting the impartiality of the Trusted Flaggers than having ties to VLOPs, such as prior involvement in societally contentious issues, membership in political parties, external financing, etc.

- *Proposed recommendation would prevent misuse of the Trusted Flagger's position to further goals unrelated to the content moderation.*

- **Para 42**

The provision should be more specific as to the impartiality requirement, as it provides no substantive guidelines to the stakeholders.



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

- **Para 60**

The provision should encourage not only to safeguard a diversity and balance of expertise, but also a pluralism and representativeness.

- *Proposed recommendation would enhance the legitimacy and representativeness of the Trusted Flaggers.*

- **Para 82, corresponding sections of Annex II**

The provision should not refer to the taxonomy of illegal content in Regulation 2024/2835, but rather correspond with the national regulations on the illegality of particular content.

- *Proposed recommendation will ensure that the Trusted Flaggers' declared experience will correspond with the legal categories characteristic of the legal systems of their seat. To give some examples, concepts such as "Category 4d Gendered Disinformation" or "8a Misinformation disinformation, foreign information, manipulation and interference" are strange to the Polish legal system.*

- **Para 83**

The provision should be changed so as to underline that the identification of the illegality of a specific provision of also in different Member States necessarily must be preceded by a detailed legal examination.

- *Proposed recommendation will ensure the declared goal of the Guidelines, i.e. helping to strike the right balance between the under- and overmoderation. Furthermore, it would pay respect to the need for respecting the differences between the Member States' legal systems. Last but not least, the proposed change would reflect the obligation to respect the Trusted Flaggers' areas of expertise.*

- **Para 99 b)**

The provision should be changed so as to underline that the legal justification should form a mandatory part of the Trusted Flaggers' submission.

- *Proposed recommendation will ensure that the Trusted Flaggers' notices will fully take into account the relevant legal systems of the Member States and facilitate for the VLOPs the assessment of the submission.*



*Razem w obronie  
życia, wolności, rodziny  
i suwerenności!*

- **Para 100**

The provision should also list the notifications from outside the Trusted Flagger's field of expertise as a separate criterion.

- *Proposed recommendation will help to streamline the activities of the Trusted Flaggers so as to help them to act within the remit of their competences.*

- **Para 101**

The provision should also directly indicate that instrumental treatment of the Trusted Flagger's status should be recognized as acting in bad faith.

- *Proposed recommendation will enhance the impartiality and thus, the legitimacy of the Trusted Flaggers.*

- **Para 104**

The list of information submitted to the DSC should correspond with the modified paras 99-101.

- **Paras 109-111**

The provision should be more explicit in acknowledging the DSCs' competence to initiate investigation against the Trusted Flaggers on the basis of third parties' information and the modalities of processing the third parties' motions.

- *Proposed recommendation would enhance the legitimacy of the DSC's oversight over the Trusted Flaggers.*

One should also introduce a provision directly empowering and encouraging the DSCs to consult their peers in matters requiring their expertise, e.g. in cases of the notifications concerning misusing other Member States' laws by the Trusted Flaggers.

- **Para 133-135**

The provisions should be reformulated so as to fully appreciate the problems connected with maintaining the independence and impartiality of the Trusted Flaggers vis-à-vis the institutional actors providing them with the necessary support.