



Comparative Report

Introduction

Under the Digital Services Act (DSA), Trusted Flaggers play an important role in identifying and reporting potentially illegal online content. Their annual transparency reports provide valuable insight into how this role is being implemented in practice, including the types of content reported, the platforms concerned, and the outcomes of their notices.

This comparative report examines the annual reports published by selected Trusted Flaggers working on online hate speech and related harms. It identifies common trends, differences in reporting practices and recurring challenges in cooperation with online platforms. The comparison also considers whether the information currently published provides sufficient transparency to assess the effectiveness of the Trusted Flagger mechanism across the European Union.

Most Interesting Findings

The figures presented in the annual reports, particularly removal rates, are not directly comparable. The organisations submitted very different numbers of notices, worked on different types of content and used different reporting categories and methodologies. Nevertheless, their experiences reveal several recurring patterns in the practical implementation of the Trusted Flagger mechanism.

First, onboarding remains unnecessarily fragmented. Although Trusted Flagger status applies across the EU, organisations must often establish separate arrangements with each platform. ZARA sometimes waited weeks or months to be onboarded and, in some cases, could not identify an appropriate contact person. HateAid encountered different accounts, forms and procedures for each platform, as well as considerable variation in the space available to explain the legal basis of a notice. RESpect! was certified in October 2024 but could only begin submitting notices in January 2025. In practice, therefore, the mechanism still resembles a collection of separate platform programmes rather than a coherent EU-wide system.

The reporting tools themselves can also make it more difficult to submit legally substantiated notices. Reported problems include strict character limits, an inability to submit multiple URLs, mandatory CAPTCHAs after only a few reports, the absence of case or reference numbers, and no option to export reporting histories. Organisations also experienced difficulties reporting individual comments, requests for repeated or disproportionate documentation, and situations in which

they had to use ordinary user-reporting channels because no dedicated Trusted Flagger channel was available.

Another important finding is that “removal” does not necessarily mean that content is no longer accessible. LICRA and HateAid distinguish between worldwide removal and territorial restrictions or geoblocking. LICRA reports that much of the content it successfully challenges is blocked only within the EU. HateAid similarly found that some content was restricted only in Germany, whereas other material was removed worldwide under the platform’s terms and conditions. Future transparency reports should therefore distinguish systematically between worldwide removal, EU-wide restriction, restriction within a single Member State, reduced visibility, account-level measures and no action.

A lack of feedback from platforms also remains a significant problem. HateAid received no outcome decision for approximately one quarter of its notices, while another group of notices received an acknowledgement but no final result. ZARA encountered platforms that provided neither a decision nor confirmation of receipt. RESpect! received no response from Telegram, and Point de Contact observed substantial differences in the quality and speed of providers’ responses. This indicates that the priority treatment required under Article 22 of the DSA is not yet supported by a consistently functioning communication and feedback infrastructure.

Finally, cases involving individual victims create particular difficulties. ZARA, HateAid, SafeLine and Rat auf Draht frequently support people affected by cyberstalking, doxing, impersonation, non-consensual images or threats. Platforms may require information or evidence that only the affected person can provide, even though one purpose of involving a Trusted Flagger is to prevent victims from having to navigate complex reporting processes alone. SafeLine encountered cases in which platforms requested identity documents or required victims to submit complaints personally. ZARA found that repeated requests for evidence sometimes caused already vulnerable individuals to abandon the process. Private groups and inaccessible content can also prevent Trusted Flaggers from verifying potentially illegal material. Priority treatment is therefore insufficient when the underlying reporting procedures are neither accessible nor trauma-informed.

These reporting experiences closely reflect INACH’s own monitoring findings. Our data similarly reveal considerable variation in what platforms classify as “removal”: content is less frequently removed altogether than geoblocked or subjected to visibility restrictions. The frequent absence of a clear outcome decision is also familiar to INACH and its members, as are many of the technical difficulties encountered when submitting and tracking notices.

Conclusion

Taken together, these first reports demonstrate the value of the Trusted Flagger mechanism while revealing that it does not yet operate as a coherent EU-wide system. Its effectiveness remains highly dependent on the platform receiving the notice, the category of illegal content involved and the reporting organisation’s resources. Trusted Flaggers must navigate different onboarding procedures, interfaces, evidentiary requirements and feedback systems for each platform,

meaning that certification may precede effective operation by several months. Moreover, a rapid response does not necessarily result in meaningful action: platforms may provide only automated acknowledgements or inadequate explanations, or restrict content in a single country while leaving it accessible elsewhere. Common reporting standards, dedicated contact points, interoperable reporting channels and clearer requirements concerning response times, statements of reasons and the territorial scope of measures would make the mechanism more consistent. They would also allow Trusted Flagger reports to serve as a stronger source of evidence about systemic problems in platforms' implementation of the DSA.



Bringing the Online In Line with Human Rights



Contact Us

For further information, please reach out to us at **adinde@inach.net**.

INACH

Website: www.inach.net