



LEGAL & POLICY BRIEFING · REF. EFE/DSA/2026-01

Who Flags the Flaggers?

Closing the Independence Gap in Article 22
of the Digital Services Act

*Prepared for: DG CONNECT, European Board for Digital Services,
and national Digital Services Coordinators*

Goran Kolouch

Independent Researcher · Founder, EqualFaith Europe

August 2026

EQUALITY · DIGNITY · JUSTICE

Executive Summary

Article 22 of the Digital Services Act (“DSA”) created the trusted flagger mechanism to give expert, high-confidence notices of illegal content priority processing on online platforms. The mechanism rests on a single, load-bearing safeguard: that a trusted flagger “represents collective interests and is independent of any provider of online platforms.” This briefing argues that the safeguard, as drafted, is too narrow. It secures independence from platforms but says nothing about independence from governments, political movements, or funders — the very actors most likely to want privileged, fast-tracked access to content removal.

Drawing on documented cases in Germany and Ireland, and on the European Commission’s own transparency database of designated trusted flaggers, this briefing shows that the current framework cannot detect or prevent political or state capture of the flagging function until after harm has occurred, if at all. It further shows that this gap falls unevenly across protected groups: while organisations monitoring antisemitism have an established, documented pathway to trusted flagger status, no equivalent Muslim-led or Islamophobia-focused civil society organisation currently appears in the Commission’s public trusted flagger register — despite EU Agency for Fundamental Rights (FRA) data showing that four in ten Muslims in the EU report experiencing discrimination, and despite online “toxicity” research showing Islam among the most targeted topics in EU online hate speech.

EqualFaith Europe proposes three concrete amendments to Article 22 and its implementing guidance: (1) mandatory disclosure of funding sources and institutional or political affiliations above a materiality threshold, both at application and in annual transparency reports; (2) an independent, cross-community advisory review of contested designations, with standing for affected civil society groups; and (3) proactive outreach by Digital Services Coordinators to under-represented communities — including Muslim civil society — to correct the current asymmetry in who has meaningful access to the fast lane the DSA created.

1. The Promise of Article 22

Article 22 DSA requires providers of online platforms to give priority, undue-delay processing to notices submitted by entities certified as “trusted flaggers” by a national Digital Services Coordinator (DSC). The underlying policy rationale is sound: platforms receive an overwhelming volume of low-quality user reports, and expert organisations — hotlines for child sexual abuse material, consumer protection bodies, hate speech monitors — can cut through that noise with more reliable notices. Trusted flagger status is awarded EU-wide once granted by any Member State’s DSC, and is meant to be earned through demonstrated expertise, accuracy, and diligence.

Properly designed and supervised, this is a genuine public good. The concern raised in this briefing is not with the existence of the mechanism, but with a specific and correctable weakness in how “independence” is defined and policed.

2. The Legal Gap: What Article 22 Requires — and What It Does Not

Under Article 22(2), a DSC may award trusted flagger status only where the applicant demonstrates that it has particular expertise and competence in detecting, identifying and notifying illegal content; that it represents collective interests and is independent of any provider of online platforms; and that it carries out its activities in a timely, diligent and objective manner.

The independence criterion is deliberately, and narrowly, platform-facing. Nothing in Article 22(2), nor in Recitals 61–62, requires an applicant to disclose who funds it, whether it is controlled by or answerable to a government ministry, or whether its leadership holds public political positions relevant to the content it will be flagging. Once designated, Article 22(3) requires only an annual transparency report explaining “the procedures in place to ensure that the trusted flagger retains its independence” — self-reported, and again framed around platform independence rather than funder or political independence.

Revocation under Article 22(6)–(7) is possible where a DSC determines, following investigation, that an entity no longer meets the Article 22(2) conditions. But because those conditions never asked about funding or political affiliation in the first place, an investigation triggered by exactly those concerns has no statutory hook to act on. The result is a structural blind spot: the DSA can discipline a trusted flagger for being too close to a platform, but has no textual basis to discipline one for being an instrument of a government ministry or a partisan campaign, however clearly documented.

Academic commentary on Article 22 has independently identified this as one of the framework’s central weaknesses, alongside risks of content over-removal and insufficient transparency, and has specifically flagged that state entities acting as trusted flaggers create a heightened fundamental-rights tension precisely because states carry particular responsibility for protecting freedom of expression — a responsibility that is difficult to reconcile with a state-funded body simultaneously acting as the platform’s trusted gatekeeper of what counts as illegal.

3. Illustrative Cases: Where the Gap Becomes Visible

3.1 Germany: REspect! and ministry funding

Germany’s Digital Services Coordinator certified the organisation REspect! as a trusted flagger in 2024. Documents obtained by a Member of the European Parliament have reportedly shown that the organisation receives the large majority of its funding — cited at approximately 95% — from the federal family affairs ministry. If accurate, this does not on its own establish misconduct; it does establish that an entity whose funding is almost entirely dependent on a single government ministry has been granted a fast lane to flag content across all major platforms operating in the EU, without any DSA requirement that this funding relationship be disclosed as part of the certification record made public to researchers, journalists, or affected communities.

3.2 Ireland: contested designation and unanswered transparency requests

Ireland’s Coimisiún na Meán granted trusted flagger status to the Jewish Representative Council of Ireland in June 2026 for illegal antisemitic

material online — a legitimate and important category of harm. Irish investigative outlet The Ditch has separately reported that the Council's chair holds and has publicly expressed a political position opposing Ireland's recognition of the State of Palestine, and reported that he arranged a meeting at the Taoiseach's office between a then junior minister and a former Israeli intelligence officer at which the Council's trusted flagger application was discussed. The same reporting states that the regulator declined two Freedom of Information requests concerning this contact and would not confirm whether records exist.

EqualFaith Europe takes no position on the accuracy of the underlying reporting, which we have not independently verified, and it should not be read as a challenge to the legitimacy of combating antisemitism online — a goal this organisation shares without reservation. The case is cited here for a narrower, procedural reason: under the current Article 22 framework, a citizen, journalist, or rival civil society organisation with concerns of exactly this kind has no statutory disclosure mechanism to test them. The only tools available — national freedom of information law — were reportedly refused. That is precisely the accountability gap this briefing asks the Commission to close, for every community's trusted flaggers alike.

3.3 The United Kingdom: a cautionary comparator outside the DSA

Although outside the DSA's scope, the United Kingdom's experience under the Online Safety Act is instructive as to what happens when a moderation-adjacent enforcement power is exercised without a transparent, independent gatekeeping standard. In 2023, 37 police forces made 12,183 arrests under older communications offences — an average of 33 a day — of which only 1,119 resulted in conviction, a conversion rate below 10%. The lesson for Article 22 is not that flagging or enforcement is illegitimate, but that any mechanism which fast-tracks content decisions needs proportionate, independently auditable standards, or it will be used far beyond its intended target.

4. A Structural Asymmetry: Who Currently Benefits from the Fast Lane?

FRA's EU-wide research finds that four in ten Muslims in the EU report having experienced discrimination, with the proportion higher still among Muslim women, much of it linked to religious dress. Separately, FRA's European Observatory of Online Hate recorded a more than 30% rise in online toxicity in a single nine-month period, with Islam ranking among the most targeted topics alongside Judaism. The European Commission's own coordination process on combating anti-Muslim hatred records a marked deterioration since October 2023, with rising threats against Muslim communities across the Union.

Despite this documented and rising harm, the Commission's public trusted flagger transparency database does not, at the time of writing, list a Muslim-led or Islamophobia-focused civil society organisation among designated trusted flaggers in any Member State — in contrast to antisemitism-focused organisations, which are cited in the Commission's own explanatory materials as a paradigm example of the kind of specialist NGO the mechanism was designed for. This is not evidence of bad faith by any DSC; it more plausibly reflects that Muslim civil society organisations in Europe are frequently smaller, less institutionally funded, and less established within national regulatory relationships — the very capacity gap EqualFaith Europe exists to help close.

The consequence is a double exposure. Muslim communities face rising, well-documented online hate with no dedicated fast-lane flagger of their own, while simultaneously having no meaningful transparency tool to scrutinise whether other designated flaggers operate with the neutrality Article 22 assumes but does not verify. Closing the independence gap and building genuine capacity for under-represented civil society are, in practice, the same reform.

5. Proposed Amendments to Article 22 and Implementing Guidance

- **Mandatory funding and affiliation disclosure.** Amend Article 22(2)–(3) to require applicants and designated trusted flaggers to disclose, at application and annually, any funding source exceeding a materiality threshold (e.g. 20% of annual income) and any formal institutional or political affiliations of governing officers relevant to the content categories flagged. This mirrors long-standing lobbying-transparency logic already familiar in EU law and imposes no burden beyond what accountable, publicly funded bodies already provide in other contexts.
- **Independent, cross-community review of contested designations.** Establish a standing advisory mechanism at European Board for Digital Services level, with participation from affected civil society sectors (including religious minority organisations), empowered to review and issue non-binding findings on contested designations or revocation decisions — addressing the current gap in which only the DSC and the platform have standing to trigger review.
- **Proactive DSC outreach to under-represented communities.** Require DSCs to publish, alongside their annual trusted flagger register, an assessment of which protected characteristics and communities are and are not represented among designated flaggers in their jurisdiction, and to support capacity-building applications from qualified organisations serving under-represented groups — extending the logic already applied to INHOPE-network child-safety hotlines to religious and ethnic minority civil society.
- **Extend Article 40(12) data-access logic to flagger-pattern research.** Permit vetted researchers to study aggregate trusted flagger notice patterns under the same public-interest research access framework already used for platform systemic-risk data, enabling independent empirical testing of over- or under-flagging by protected category — rather than relying, as at present, on investigative journalism and parliamentary document requests.

None of these proposals weaken Article 22's core function. On the contrary: a trusted flagger regime that can demonstrate its independence, rather than merely assert it, is the version of Article 22 most likely to retain public and judicial confidence as enforcement scales up across the Union.

6. Conclusion

The trusted flagger mechanism is one of the DSA's more promising tools for making online spaces safer — and one of the more fragile, precisely because it depends on public trust that is easily lost if capture is ever demonstrated rather than merely suspected. EqualFaith Europe supports the mechanism's aims and wants to see it strengthened, not dismantled. The reforms proposed here — transparency of funding and affiliation, independent review, proactive outreach to under-represented communities including Muslim civil society, and research access — would close a gap that currently leaves the DSA unable to distinguish a genuinely independent expert flagger from a state or political proxy wearing the same

certification. We welcome the opportunity to discuss this briefing further with DG CONNECT, the European Board for Digital Services, and interested Digital Services Coordinators.

Respectfully submitted,

Goran Kolouch *Independent Researcher; Founder, EqualFaith Europe*

Sources

European Union. Regulation (EU) 2022/2065 (Digital Services Act), Article 22 and Recitals 61–62.

European Commission, Digital Strategy. “Trusted flaggers under the Digital Services Act (DSA).” digital-strategy.ec.europa.eu

van de Kerkhof, J. “Article 22 Digital Services Act: Building trust with trusted flaggers.” Internet Policy Review / Utrecht University research portal, 2025.

European Union Agency for Fundamental Rights (FRA). “Being Muslim in the EU,” 2024; European Observatory of Online Hate, “Online Hate Speech in 2023,” periodical report.

European Commission. “Combating anti-Muslim hatred.” commission.europa.eu

Collectif contre l’Islamophobie en Europe (CCIE). Annual Report on Islamophobia in Europe, 2025.

William Fry. “Trusted Flaggers under the DSA — what you need to know,” 2024.

TechPolicy.Press. “Building Trust in the Digital Service Act’s Trusted Flaggers,” 2024; “More Questions Than Flags: Reality Check on DSA’s Trusted Flaggers,” 2024.

Reporting on national trusted flagger designations (Germany, Ireland) and related transparency requests as circulated in public commentary, August 2026; underlying primary-source documents not independently verified by EqualFaith Europe and cited here for illustrative purposes only.