Between Click and Consequence:
An Evaluation of Platform Reporting Procedures under the Digital Services Act
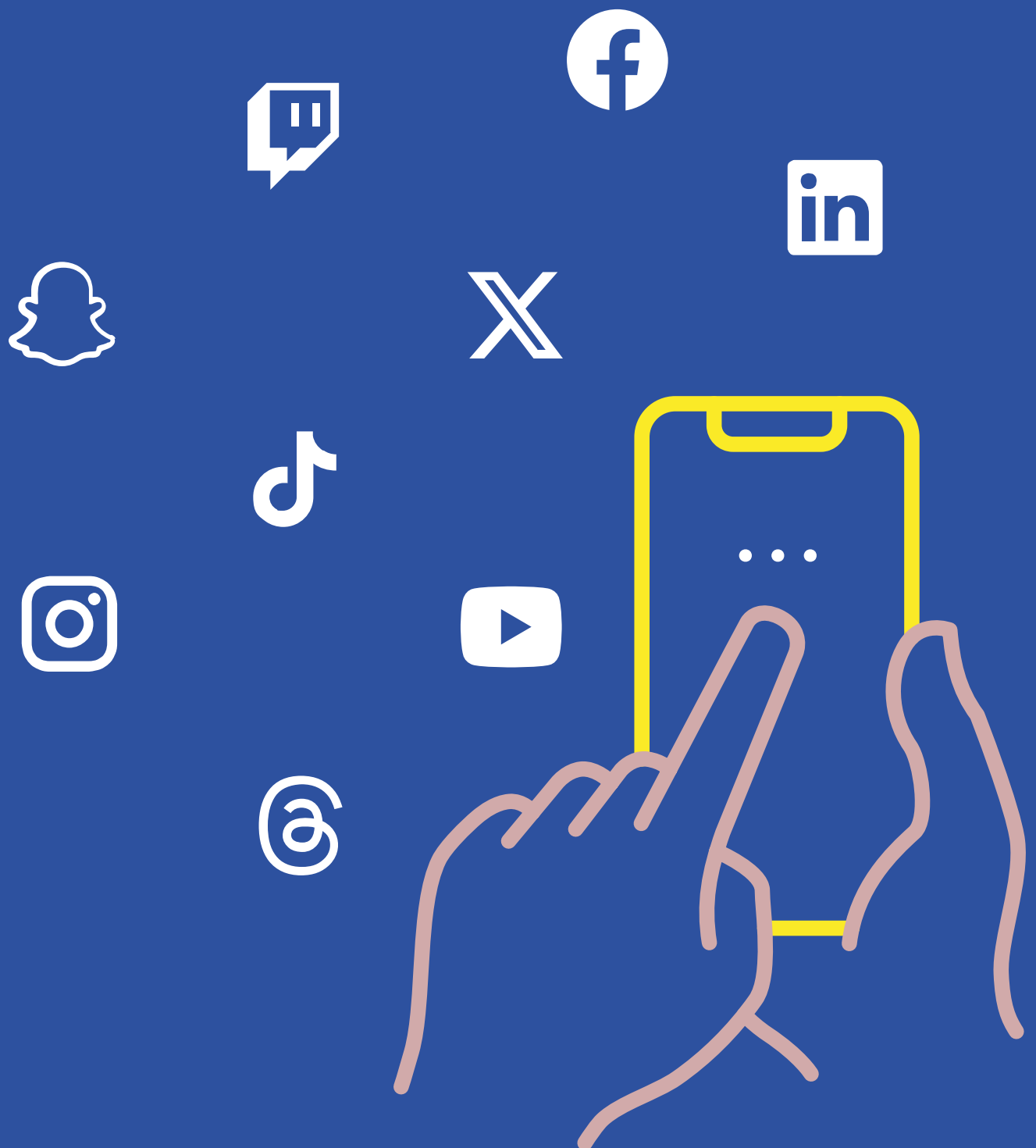
TABLE OF CONTENTS

1. Under Scrutiny: Reporting Channels and Platform Responsibility in the Digital Space

Users who encounter forbidden content online on online platforms should be able to report this to the platform operators in an easy and effective way. At least that is what the European Union's Digital Services Act (DSA) requires. The DSA sets out clear rules for handling illegal content. Content deemed illegal under national law must be removed from a platform without delay once it has been reported. Article 16 (1) of the DSA explicitly requires platforms to establish mechanisms that are "easy to access and user-friendly" and which offer users a simple and transparent tool for reporting illegal content. The legislation's aim is to ensure that illegal content on very large online platforms (VLOPs) and very large online search engines (VLOSEs)[1] can be reviewed and, if necessary, deleted.

Accordingly, the Digital Services Act stipulates that the channels through which users can report potentially illegal content or posts that violate community guidelines must be user-friendly and easily accessible. These reporting channels should be easy to find and well designed so that reporting content does not become a hurdle for users. To date, no platform has implemented uniform mechanisms for reporting illegal content or for demanding redress.

The present study, a collaboration between Das NETTZ together with Dr. Pablo Jost of Johannes Gutenberg University Mainz and Ursula Kristin Schmid at Ludwig Maximilian University of Munich, offers a comprehensive evaluation of the user-friendliness and effectiveness of reporting channels under the DSA in Germany and examines the motivations that drive platform users to report content. As researchers, we have deliberately placed the perspective of platform users at the center of our work in order to contribute to consistent oversight in the users' best interest.

The study focuses on four research questions:

1. What motivates users to submit a notice and what discourages them from doing so?

2. How clearly are reporting channels designed and what difficulties do users encounter?

3. How are the reporting channels of different platforms perceived and used?

4. Are there differences between reporting channels set up under the DSA and those set up under community guidelines?

Finally, we direct our recommendations to the European Commission, the national coordinating authorities, and the platforms themselves as regards implementation of Article 16 of the Digital Services Act.[2]

The findings of this study are crucial to holding platforms more accountable and ensuring that user safety and user rights remain at the core of how reporting channels are designed and developed.

---

1 European Commission (2025). "DSA: Very large online platforms and search engines." European Commission. https://digital-strategy.ec.europa.eu/en/policies/dsa-vlops.

2 While the European Commission has exclusive responsibility for enforcing and supervising compliance with the extended due diligence obligations imposed on VLOPs and VLOSEs by the Digital Services Act to address systemic risks, the Commission and national authorities share responsibility for all other obligations imposed on VLOPs and VLOSEs by the Digital Services Act.

2. Digital Due Diligence: The DSA in the Context of Political and Societal Expectations

After much anticipation, the DSA, a EU-wide regulation for digital services, entered into full force on 17 February 2024. Less than a month later, on 21 March 2024, the German Bundestag approved legislation implementing the Digital Services Act into national law as the Digitale-Dienste-Gesetz (DDG).[3] The law extends to internet providers, hosting services, cloud services, social networks, messaging services and online marketplaces. The European regulation has applied to VLOPs and VLOSEs – those with more than 45 million monthly users in the EU – already since 25 August 2023. This covers, for example, all major social networks and video-hosting platforms, Google Search, Microsoft Bing, but also online retailers (Zalando, Amazon), travel portals (Booking.com), and even porn sites (Pornhub).

The aim of the DSA is to introduce uniform rules across Europe for removing and restricting online content that violates the law or platform rules. In addition, platform operators must proactively mitigate the risks of such content arising and spreading, and they must provide transparent reasoning for how their content is moderated. As early as April 2024, the European Commission initiated formal proceedings against Meta, the parent company of Facebook and Instagram, regarding its reporting mechanisms.[4] The Commission suspects that Meta's notice-and-action mechanism violates obligations set out in the DSA, referring in particular to the requirement that such mechanisms be easily accessible and user-friendly. The proceedings are still ongoing. Ireland's supervisory and coordinating authority, Coimisiún na Meán, has also since sent formal requests for information to several platforms, seeking details on their reporting options.[5]

Article 16(1) Providers of hosting services shall put mechanisms in place to allow any individual or entity to notify them of the presence on their service of specific items of information that the individual or entity considers to be illegal content. Those mechanisms shall be easy to access and user-friendly, and shall allow for the submission of notices exclusively by electronic means.

The process for submitting notices of potentially illegal content and of violations of community guidelines is designed differently across all VLOPs and VLOSEs: while notices under the DSA are often tied to additional confirmation, the provision of personal data, and legal terminology, reports of community guideline violations can usually be submitted with just a few clicks. Reporting illegal content requires a certain degree of legal knowledge. Users must, for example, provide "a sufficiently substantiated explanation of the reasons why the individual or entity alleges the information in question to be illegal content" (Article 16(2)(a)), and this statement must confirm "the bona

fide belief of the individual or entity submitting the notice that the information and allegations contained therein are accurate and complete" (Article 16(2)(d)).

In May 2024, the nonprofit organization HateAid filed a formal complaint with the Federal Network Agency (Bundesnetzagentur, BNetzA), Germany's national coordinating authority for implementing the DSA, against TikTok's inadequate reporting mechanisms under Article 16.[6] That the Federal Network Agency views platform reporting mechanisms as a relevant instrument became clear in February 2025, when the agency tested platform reporting channels and mechanisms in its stress test ahead of the 2025 Bundestag election.[7]

According to HateAid, TikTok's complaint mechanisms are not sufficiently user-friendly. Working with legal scholar Daniel Holznagel, a judge at the Berlin Court of Appeal, they assessed the design and functionality of reporting mechanisms across major online platforms as a violation of Article 16(1).[8] The hypothesis: many very large online platforms and search engines (VLOPSEs) push users toward submitting weaker, largely unregulated notices under community guidelines. Users, so the argument, often do not think about the difference between notices submitted under platform guidelines and reporting under the DSA, since their focus is primarily on the act of reporting itself rather than its legal consequences. Nevertheless – in HateAid's view – platforms actively discourage users from submitting stronger notices that would fall under the DSA. The result is that such reports are unlikely to be included in the transparency reporting obligations for DSA evaluation (Articles 15 and 24) or to trigger the possibility of appealing moderation decisions through an internal complaint-handling system (Articles 20 and 21). Moreover, it is worth noting that the "statement of reasons" provided for under Article 17 – the obligation for platforms to inform users of the grounds for their moderation decisions – have so far played only a minor role in actual content moderation practice.

---

3 Deutscher Bundestag (2024). "Bundestag stimmt für die Umsetzung des Digital Services Act (DSA)." Deutscher Bundestag. https://www.bundestag.de/dokumente/textarchiv/2024/kw12-de-digitale-dienste-994508.

4 European Commission (2024). "Gesetz über digitale Dienste: Kommission leitet förmliches Verfahren gegen Facebook und Instagram ein." Representation of the European Commission in Germany. https://germany.representation.ec.europa.eu/news/gesetz-uber-digitale-dienste-kommission-leitet-formliches-verfahren-gegen-facebook-und-instagram-ein-2024-04-30_de.

5   Coimisiún na Meán (2024). "Coimisiún na Meán opens review of online platforms' compliance with EU Digital Services Act." Coimisiún na Meán. https://www.cnam.ie/coimisiun-na-mean-opens-review-of-online-platforms-compliance-with-eu-digital-services-act/.

6   Der Spiegel (2024). "HateAid reicht Beschwerde gegen TikTok ein." https://www.spiegel.de/netzwelt/hateaid-reicht-beschwerde-gegen-tiktok-ein-a-2db19652-4f70-46dd-ac02-4c999661a53c.

7   Bundesnetzagentur (2025). "Bundesnetzagentur testet Abläufe und Maßnahmen bei Verstößen gegen den Digital Services Act." Bundesnetzagentur (Digital Services Coordinator Germany). https://www.bundesnetzagentur.de/SharedDocs/Pressemitteilungen/DE/2025/20250131_DSC.html.

8   Daniel Holznagel (2024). "Follow Me to Unregulated Waters! Are Major Online Platforms Violating the DSA's Rules on Notice and Action?" Verfassungsblog. https://verfassungsblog.de/follow-me-to-unregulated-waters/.

3. State of Research
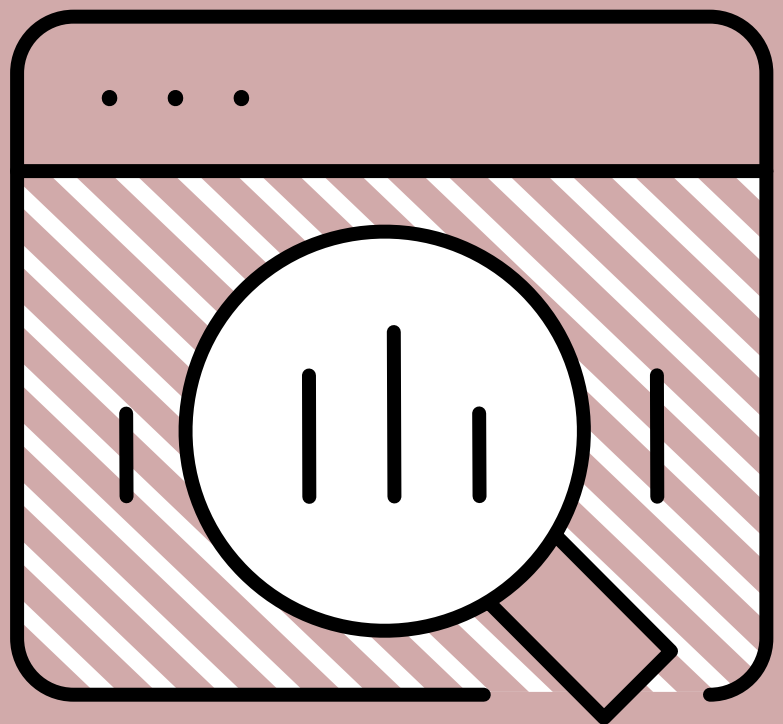
The question of how users handle the reporting channels available to them has long been the subject of academic and civil society debate. While the DSA has created new legal parameters for reporting rules, it remains unclear whether they will actually achieve the intended effect. Earlier regulations such as the German Network Enforcement Act (Netzwerkdurchsetzungsgesetz, NetzDG), which has since been superseded by the DSA, already showed that implementation can suffer from significant shortcomings. Still, there is a lack of empirical, systematic data on users' concrete experiences and perceptions.

As for what motivates users to submit a notice or prevents them from doing so, a 2024 study by the Bavarian Research Institute for Digital Transformation[9] on insults in social media provides figures: 54 % of respondents directly affected by online insults (N = 890) stated that they reported offensive posts either to the platform itself or to a government agency. Reasons cited by respondents for not reporting insults included their own lack of interest in reporting (66 %), the perceived futility of reporting on platforms (57 %), a lack of knowledge about existing reporting channels (nearly 44 %), and uncertainty as to whether the content was illegal or not (nearly 35 %). Wu (2024) categorized six factors influencing the decision to report illegal content or not: nature of the post, nature of the source, reactions from and impact on others, subject interest and knowledge, cultural norms, and consequences of reporting.[10] The last point is consistent with the findings of our qualitative survey, in which respondents said they would discontinue the reporting process if they feared legal consequences from false reporting. A study conducted by the Technical University of Munich in cooperation with the reporting office REspect! found that people who report content on platforms tend to distrust platforms, have already had negative online experiences, and report primarily out of a sense of civic duty.[11] Zhang, Montague and Jhaver (2018), in interviews about reporting motivation, found that content was reported when it ran counter to the respondents' own values, regardless of the platform's community guidelines.[12] More than half of the participants believed that every user is responsible for reporting rule violations. This conviction was closely linked to the idea of flagging content to protect others. What discouraged further flagging was, above all, the lack of or inadequate response from platforms to user notices. Content often remained online for weeks after being reported. The lack of action heightened frustration and gave reporters the feeling of being in the wrong.

The question of how understandable reporting channels are can be illustrated by looking at practice: as early as 2019, Germany fined Facebook because the platform failed to adequately forward to the relevant authorities complaints about illegal content submitted via

its internal reporting system, thereby violating NetzDG transparency requirements.[13] Despite the widely suspected shortcomings of reporting systems, however, empirical studies and data remain scarce. Earlier research on the design and enforcement of reporting mechanisms under NetzDG focused on how Facebook and X nudged users toward using the community guideline channel instead of the legal one by making the steps required for legal reporting more cumbersome.[14] Shim and Jhaver (2024), in their study on fair flagging, concluded that the legal complexity of flagging systems on social platforms often overrides user-friendliness. This leads to interfaces that are difficult for many users to navigate.[15]

Since early 2024, the nonprofit HateAid has been analyzing how user rights under Articles 16, 20 and 21 of the DSA are being implemented on Facebook, Instagram, TikTok, YouTube and X.[16] As with the present study, particular attention is paid to user rights from the perspective of users themselves. A quantitative evaluation of moderation responses revealed that only 44 % of the 151 items reported as illegal were removed; an internal appeal to the platform raised that figure only to 54 %. It can also be assumed – as in a study by Alliance4Europe on X[17] – that content moderation often happens automatically, given that some notices were processed in under two minutes. HateAid's final report also concludes that platforms do not clearly distinguish in their reporting channels between violations of community guidelines and illegal content. Further insights into reporting channels under Article 16 DSA come from semi-structured interviews conducted by Sekwenz et al. (2025).[18] Respondents noted that there are many overlaps between DSA reporting categories and community guideline categories, which complicates the reporting of illegal content, even when the legality of the content is unclear. A reporting structure oriented primarily toward compliance with statutory requirements may also reduce user motivation to report content, even though reporting channels are meant to protect user rights.

Against this backdrop, difficulties that users face when reporting under the DSA versus community guidelines can also be seen in the results of an investigative inquiry by the Spanish NGO Maldita. It found that platforms' inadequate responses to flagging understandably lead to frustration among users. Of the 58 illegal posts Maldita reported to Facebook through the official DSA reporting channel, 93 % remained online a week later. Notices submitted under community guidelines fared differently: Meta removed only 6.8 % of posts reported as illegal through the DSA channel but deleted 17 % of posts flagged as guideline violations.[19] In other words, Meta was more attentive in enforcing its internal rules than in complying with its EU legal obligations. According to an analysis by Suzor et al. (2019), users facing the

decision whether to report content often feel deep frustration at how platforms enforce reporting mechanisms.[20]

It also becomes clear that users' perception and navigation of reporting channels depends largely on how understandable and effective they are. In addition, motivation to report content at all plays a central role in our study. Current figures are provided by a study from the Technical University of Munich and HateAid, which found that more than half of politically active respondents had already reported content on platforms.[21] According to the representative survey "Lauter Hass, leiser Rückzug" (Loud Hate, Silent Withdrawal) conducted by the Kompetenznetzwerks gegen Hass im Netz (Competence Network Against Hate Online) in 2024, only 34 % of all respondents had reported hate online directly to a platform, but 74 % of those personally affected by online hate had done so.[22] Those affected were also asked about their perception of reporting channels on the platforms they used: a majority of Facebook, Instagram and TikTok users (51–56 %) said the reporting channels on those platforms were fairly or very easy to use. Users of YouTube, X and dating platforms rated their reporting channels as easy with similar frequency (44–46 %).

At the heart of the Digital Services Act is the obligation for platforms to provide clear structures for dealing with problematic content and to assume responsibility for moderation and transparency. A comparative study by Riedl et al. (2021) on the handling of problematic user comments in Germany and the US illustrates the importance of how users themselves assign responsibility. The study shows that users place responsibility both on platforms and on the state, but also on themselves. In Germany, responsibility was assigned to all actors – including users themselves – to a greater degree than in the US, where assignments of responsibility were generally more restrained. At the same time, a strong commitment to freedom of expression reduces expectations of state intervention without lessening the responsibility of platforms, media or users. These findings underscore that users develop an integrated understanding of regulation in which different actors complement one another – a finding that supports the multilevel logic of the DSA.[23]

9   Steliyana Doseva, Fay Carathanassis, Hannah Schmid-Petri and Dirk Heckmann (2024). "Beleidigungen auf Social Media: Wie betroffen sind Nutzerinnen und Nutzer auf Social Media und wie wehren sie sich?" Bavarian Research Institute for Digital Transformation. https://www.bidt.digital/publikation/beleidigungen-auf-social-media-wie-betroffen-sind-nutzerinnen-und-nutzer-auf-social-media-und-wie-wehren-sie-sich/.

10 Shangyuan Wu (2024). "What Motivates Audiences to Report Fake News? Uncovering a Framework of Factors That Drive the Community Reporting of Fake News on Social Media." Digital Journalism 12, 6 (2023), 790–807. https://doi.org/10.1080/21670811.2023.2243489.

11 See Friederike Quint, Yannis Theocharis, Spyros Kosmidis and Margaret E. Roberts (2025). "From Bystanders to Reporters: Who Acts Against Illegal Online Content?" European Political Science Association (EPSA) Conference.

12 Alice Q. Zhang, Kaitlin Montague and Shagun Jhaver (2024). "Cleaning Up the Streets: Understanding Motivations, Mental Models, and Concerns of Users Flagging Social Media Content." arXiv. https://arxiv.org/pdf/2309.06688.

13 Thomas Escritt (2019). "Germany fines Facebook for under-reporting complaints." Reuters. https://www.reuters.com/article/us-facebook-germany-fine-idUSKCN1TX1IC/.

14 Ben Wagner, Krisztina Rozgonyi, Marie-Therese Sekwenz, Jennifer Cobbe and Jatinder Singh (2020). "Regulating transparency? Facebook, Twitter and the German Network Enforcement Act." FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 261–271. https://doi.org/10.1145/3351095.3372856.

15 Yunhee Shim and Shagun Jhaver (2024). "Incorporating Procedural Fairness in Flag Submissions on Social Media Platforms." arXiv. https://doi.org/10.48550/arXiv.2409.08498.

16 Anna-Lena von Hodenberg and Josephine Ballon (2025). "1 Jahr DSA: Sicherheit im digitalen Raum: eine Bilanz – und was jetzt nötig ist." HateAid. https://hateaid.org/wp-content/uploads/2025/07/HateAid_Policy-Brief_1-Jahr-DSA_barrierefrei-2.pdf.

17 Charles Terroille, Saman Nazari and Ewan Casandjian (2025). "Flagged and Ignored: Testing X's Response to EU Sanction Violations." Alliance4Europe. https://alliance4europe.eu/flagged-and-ignored.

18 Marie-Therese Sekwenz, Ben Wagner and Simon Parkin (2025). "'It is unfair, and it would be unwise to expect the user to know the law!' Evaluating reporting mechanisms under the Digital Services Act." FAccT '25: Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency, 532–546. https://doi.org/10.1145/3715275.3732036.

19 Maldita (2025). "Denunciamos a Meta 58 publicaciones fraudulentas en Facebook utilizando los mecanismos legales de la DSA y, una semana después, el 93 % siguen activas." Maldita.es. https://maldita.es/investigaciones/20250619/denunciamos-meta-publicaciones-fraudulentas-facebook-dsa/.

20 Nicolas P. Suzor, Sarah Myers West, Andrew Quodling and Jillian York (2019). "What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation." International Journal of Communication 13 (2019), 1526–1543. https://ijoc.org/index.php/ijoc/article/view/9736.

21 Luise Koch, Angelina Voggenreiter and Janina Steinert (2025). "Angegriffen und alleingelassen – Wie sich digitale Gewalt auf politisches Engagement auswirkt. Ein Lagebild." Technical University of Munich in cooperation with HateAid.

22 Das NETTZ, Gesellschaft für Medienpädagogik und Kommunikationskultur, HateAid and Neue deutsche Medienmacher*innen (eds.) (2024). "Lauter Hass – leiser Rückzug. Wie Hass im Netz den demokratischen Diskurs bedroht. Ergebnisse einer repräsentativen Befragung." Das NETTZ, Gesellschaft für Medienpädagogik und Kommunikationskultur, HateAid and Neue deutsche Medienmacher*innen. https://toneshift.org/wp-content/uploads/2024/02/Studie_Lauter-Hass-leiser-Rueckzug.pdf.

23 Martin J. Riedl, Teresa K. Naab, Gina M. Masullo, Pablo Jost and Marc Ziegele (2021). "Who is responsible for interventions against problematic comments? Comparing user attitudes in Germany and the United States." Policy & Internet 13, 3 (2021), 433–451. https://doi.org/10.1002/poi3.257.
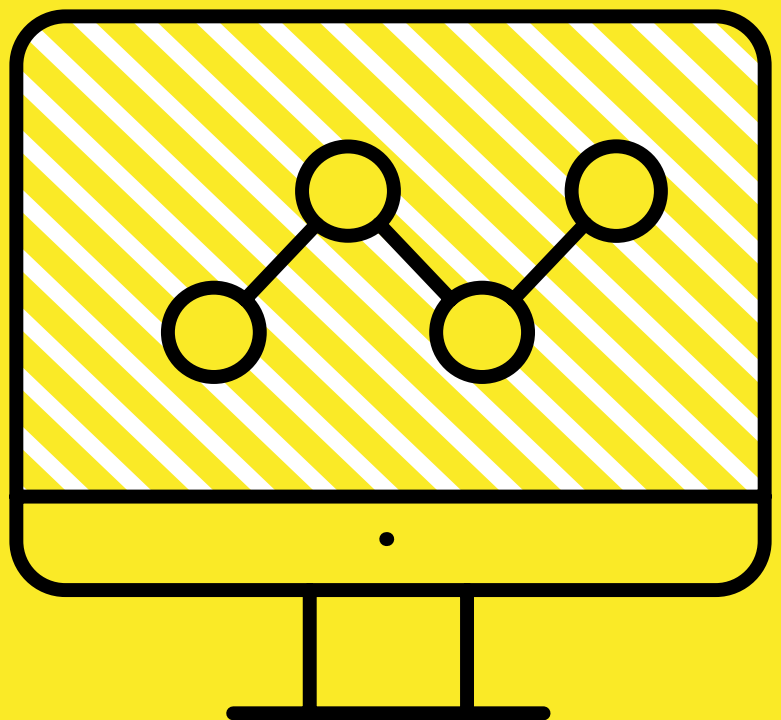
4. Method

To gain a more comprehensive understanding of the user-friendliness and effectiveness of in-platform reporting channels, we applied a mixed-methods approach for this study. To capture both motivation and willingness to report, as well as the clarity and effectiveness of reporting channels, we conducted two surveys: a qualitative exploratory survey from 13 May to 21 May 2025, with respondents recruited by S&O Market Research Germany, and a representative population survey from 10 July to 24 July 2025, covering all German-speaking users of social media aged 18 and above, carried out by Bilendi GmbH.

Methodologically, the qualitative survey followed the model of remote self-confrontation interviews.[24] Twenty users were first asked about their platform usage behavior and past reporting practices. They were then tasked with reporting a preselected comment on Facebook, Instagram or TikTok as an example. During this process, participants verbalized their thoughts using the "think-aloud" method, sharing impressions and reflections spontaneously and without filtering. Finally, their reporting experiences were discussed collectively, along with expectations for future reporting behavior.

The quantitative survey was conducted online. A total of 3,683 respondents were recruited. Only people who reported using a digital platform "at least several times a month" were surveyed. After additional quality controls (processing time, response patterns), 2,040 valid cases remained for data analysis. Respondents were between 18 and 69 years old, with an average age of 44.5. Of these, 53 % were women and 47 % men, while a very small fraction (0.1 %) identified with another gender category. Nearly 53 % held higher educational qualifications, meaning at least a *Fachabitur* (secondary school diploma qualifying for study at universities of applied sciences). The most meaningful type of study would involve surveying users at the precise moment they encounter harmful content and are motivated to report it, thus allowing researchers to trace the full process through to moderation by the respective platform. As this is hardly feasible due to practical and ethical restrictions in research, we simulated the structure and flow of the reporting processes on Meta platforms, TikTok and X. This preserved the functional core aspects of the different reporting processes, although the user interface differed from the platforms' original reporting systems. For data protection reasons, we also did not require respondents to provide personal information normally requested for a DSA notice.

The questionnaire was developed by Das NETTZ in collaboration with Dr. Pablo Jost and Ursula Kristin Schmid. Feedback was obtained from the Federal Network Agency (BNetzA), the Federal Agency for Child and Youth Protection in the Media (BzKJ), as well as from civil society

and academic stakeholders (Institute for Strategic Dialogue, HateAid, Agora Digital Transformation, Technical University of Munich, Delft University of Technology).

---

24 See Ursula K. Schmid, Anna S. Kümpel and Diana Rieger (2023). "The Remote Self-Confrontation Interview Method: Gaining Insights into Users' Perceptions of Digital Media Content from a Distance." American Behavioral Scientist 69, 10 (2023), 1221–1237. https://doi.org/10.1177/00027642231205768.

5. Findings

5.1 Perception, use, impact: what the data reveals about reporting processes

The backbone of the findings presented here is the quantitative survey. It provides insight into how frequently certain phenomena occur in the population and allows for statements to be made about statistical correlations between them. We supplement these findings with results from the qualitative study in order to illustrate insights from the quantitative survey with concrete statements. This also enables us to generate and substantiate assumptions about the underlying processes.

5.2 Reasons for reporting problematic content

In the quantitative survey, participants were asked which types of content they perceive as problematic on platforms. It became particularly clear that fraud – such as financial fraud or identity theft – was almost unanimously rated as problematic: 93 % considered such content (somewhat) problematic. Very high levels of agreement were also found for threats against other users (88 %) as well as racist or antisemitic content (83 %). A large majority also rated threats against politicians (80 %) and insults against individual users (80 %) as problematic. Pornography (65 %) and insults against politicians (60 %) were rated as somewhat less problematic, though still by a majority.
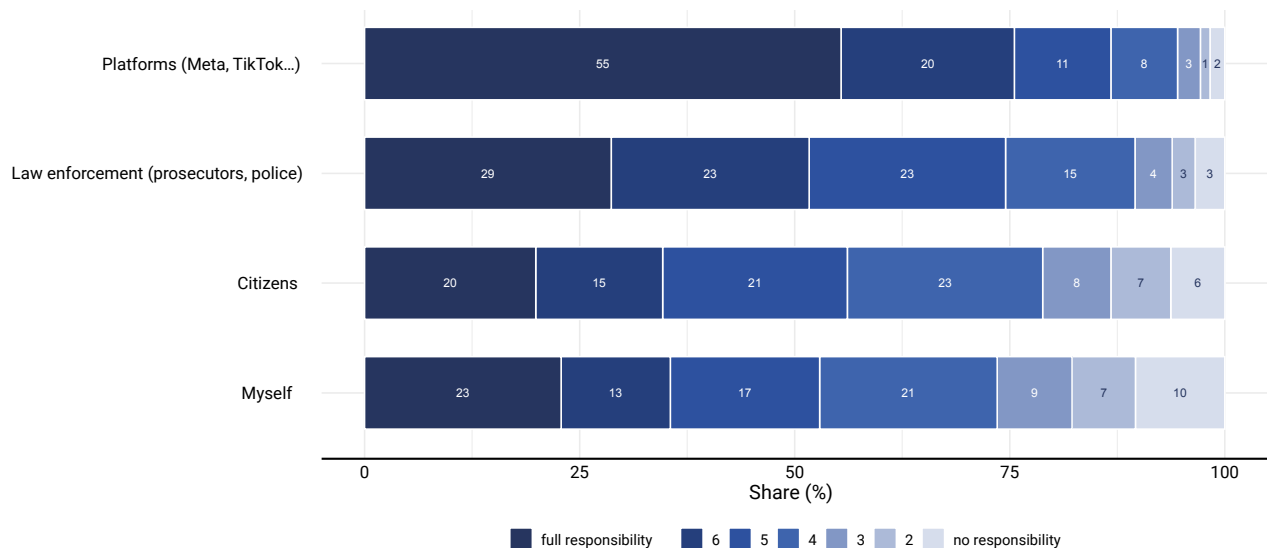
❶ Types of problematic content
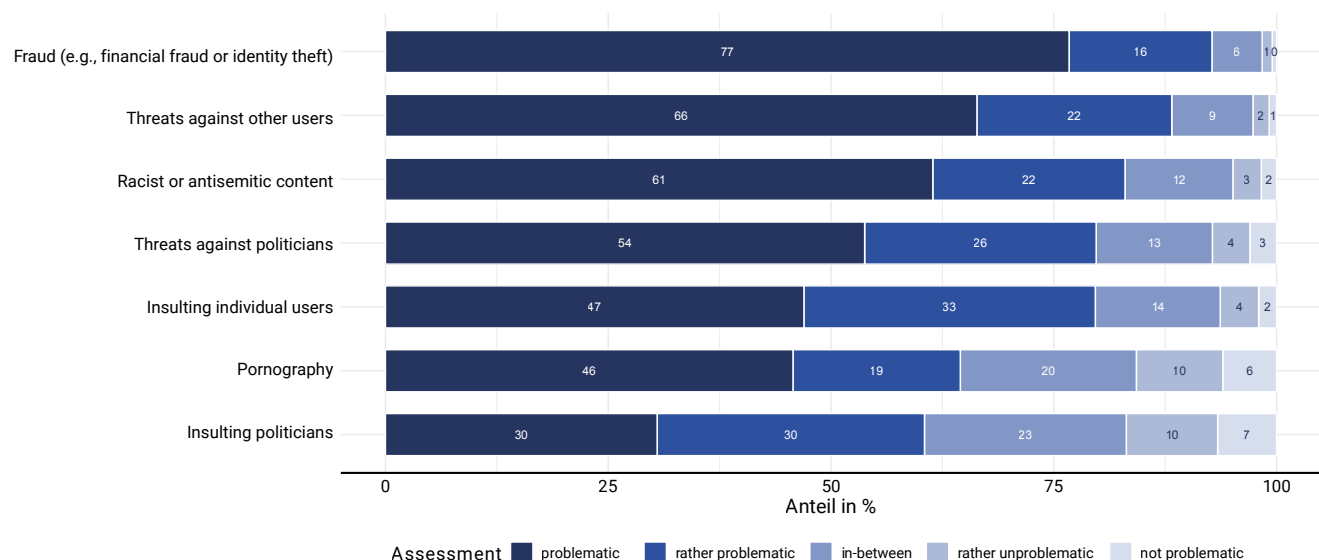
In the qualitative survey, participants described content showing and glorifying violence, or calling for violence, as particularly problematic: "So everything that has to do with violence, I personally think should definitely be reported." (Marek, 28).

Suicide announcements and child abuse or child pornography were also considered top priorities for reporting. Whether respondents actually reported insults, however, often depended on how personally affected they felt; users said they were more willing to go through the reporting process when they or someone they knew had been directly attacked.

Participants were also asked who they hold responsible for removing problematic content from platforms. Here the vast majority clearly saw the platforms themselves as bearing responsibility: 86 % of respondents stated that Meta, TikTok and other providers carry most or all of the responsibility for the removal of problematic content from their sites. Law enforcement agencies were also considered strongly responsible in this regard, with 75 % assigning them a central role in the removal process. Responsibility was also attributed, though less often, to private citizens in general (56 %) and to the users themselves (53 %). This clearly shows that, from the respondents' point of view, platforms are primarily responsible for removing problematic content, followed by government agencies, while individual responsibility is seen more as complementary.

② Perceived responsibility for dealing with content

Asked how they rate the possibilities for reporting problematic content on platforms, participants had a largely positive picture of their own competence: about three quarters (76 %) said they (somewhat) trusted themselves to report content. More than half (57 %) also stated that they knew and understood the reporting channels well. The assessment of platforms themselves, however, was more critical: 39 % agreed with the statement that platforms deliberately complicate the reporting process, and 37 % felt that reporting did not make any difference anyway. This shows that while users mostly evaluate their

own ability to submit notices as positive, significant doubts remain about the effectiveness and transparency of platform mechanisms.

### ③ Perceived (self-)efficacy in reporting
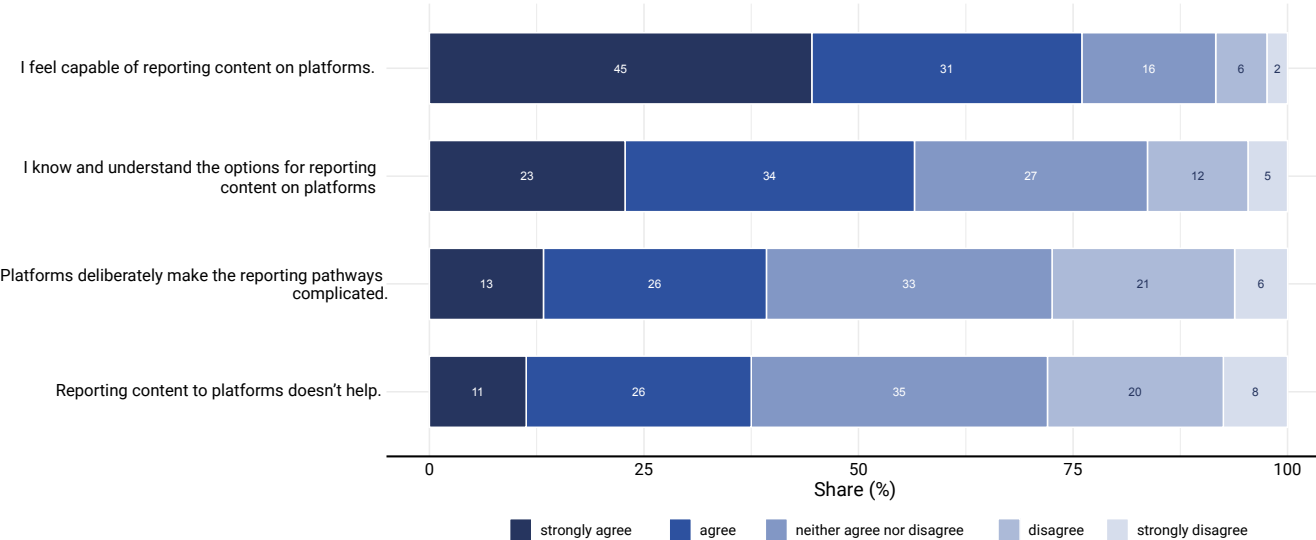
In the qualitative survey, participants likewise considered moderation and regulation to be fundamentally important – especially the option of reporting content. Most then saw platforms as responsible, though many admitted they "don't really trust the whole process one hundred percent" (Marek, 28). At the same time, they emphasized that using such functions should depend on context and severity. Some respondents considered counter-speech more effective than immediate deletion or reporting, especially in less serious cases. One participant summarized that "[there are] certain comments where you're just better off confronting the people, addressing them, and trying to get into a discussion on the subject. Everything that [counter-speech] can cover, you just have to endure first. That too is a sign of democracy." (Lucas, 45). Others criticized that reporting systems are sometimes too easy to use and thus open to abuse or overly strict application.

5.3 Experience with reporting content

In both the quantitative survey (45 %) and the qualitative study, about half of the participants stated that they had reported problematic content in the past. We examined the factors that encouraged or discouraged past reporting behavior using logistic regression. The results show that younger people and those with higher levels of education were more likely to have reported content. People who identified as politically left-leaning and who feared negative consequences of problematic content for society also reported content more often. Gender or migration background, by contrast, showed no link to reporting behavior.

④ Factors encouraging and discouraging past reporting behavior
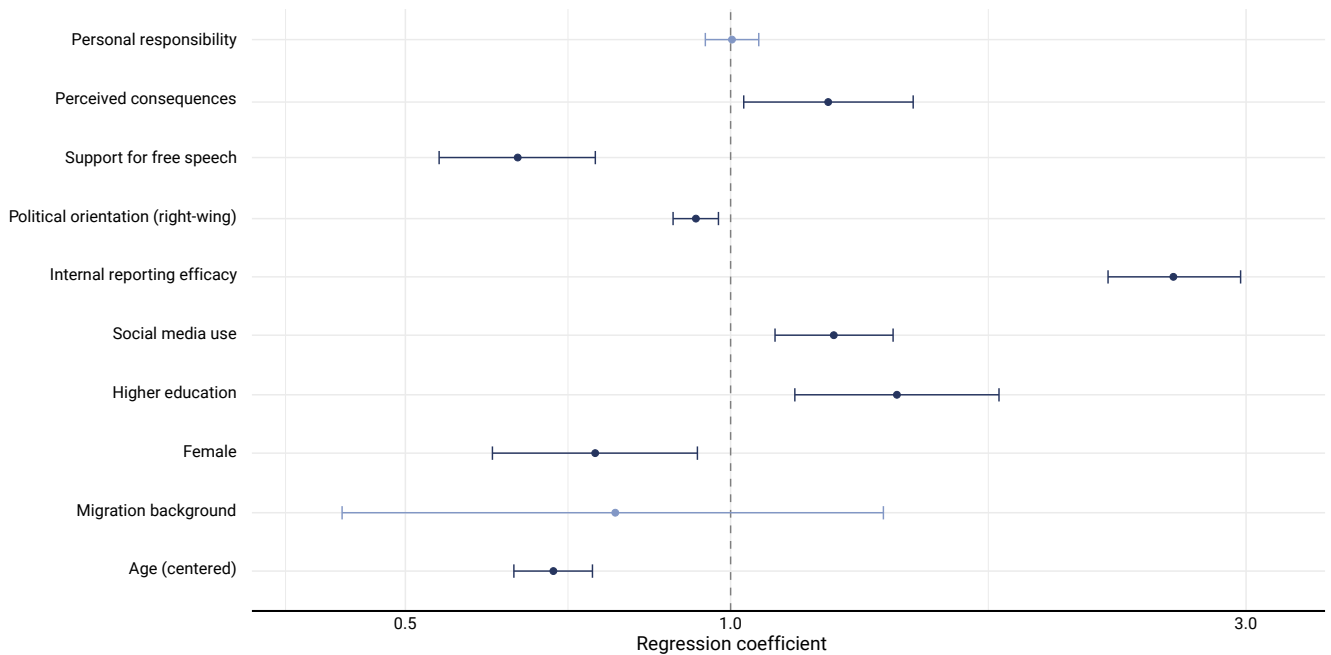
A regression coefficient > 1 indicates a positive relationship with past reporting behavior, while a value < 1 indicates a negative correlation.

Among the participants in the qualitative study, younger politically engaged users with high levels of online activity were more likely to encounter problematic content and report it. In the past, reporting was usually done under community guidelines, primarily because reporting under the DSA was not known. The participants in the qualitative survey reported frequent experiences with common insults and hate against minorities. They also often encountered fake profiles and fraud attempts, which were frequently tied to negative personal experiences. The most common reasons for reporting such content were the perceived severity of the content and the general desire to improve their (own) online environment and to achieve that "less hate is posted as a result" (Lina, 26).

As for reasons not to have made reports in the past, several patterns emerged. More than a third said that reporting content to the platform was pointless (36 %). Another common reason was simply a lack of personal interest in the problematic content (32 %). Nearly 28 % said they had never come across problematic content, and 26 % reported not finding a suitable reporting channel. Other reasons included lack of time (19 %), negative experiences with reporting systems (19 %), and uncertainty about the consequences of reporting (18 %). Less commonly cited was the fear of emotional strain from reporting (13 %).

Similar comments were also made by the participants in the qualitative survey without stating reasons. The lack of trust in the platforms again had a particularly negative effect on their reporting behavior. This was especially the case when they had already experienced that

reporting did not lead to content being removed, as one participant explained: "When I see, hey, that doesn't do anything, because TikTok says they'll review it and then nothing happens... then at some point you think, hey, why should I bother?" (Lina, 26). Besides the lack of necessity, the perceived high effort involved, and the avoidance of taking responsibility themselves, some users simply found it more comfortable to ignore problematic content, pay no further attention to it, or otherwise tell themselves: "Let them do it! It's not my problem." (Marek, 28).

5.4   Simulation of reporting channels

Before the simulated reporting process, respondents were told they would see content on the basis of which they should then act out the process of reporting the content. The simulations involved illegal content (insults, threats of violence, racist comments or hate speech) that had already been classified as criminal by the courts and that could all have been reported under DSA provisions. Respondents could abort the process at any time.

First, respondents were asked to rate the content along several dimensions. This revealed that, overall, the stimuli were seen as negative: on average they were considered aggressive and unfriendly, as well as problematic and in part dangerous. The ratings on the respective scales show that the content was not perceived as neutral by the participants but that it was clearly classified as negative and problematic communication.

⑤ Evaluation of the content shown

Vignette No.   ○ 1   ● 2   ● 3   ● 4   ● 5   ● 6
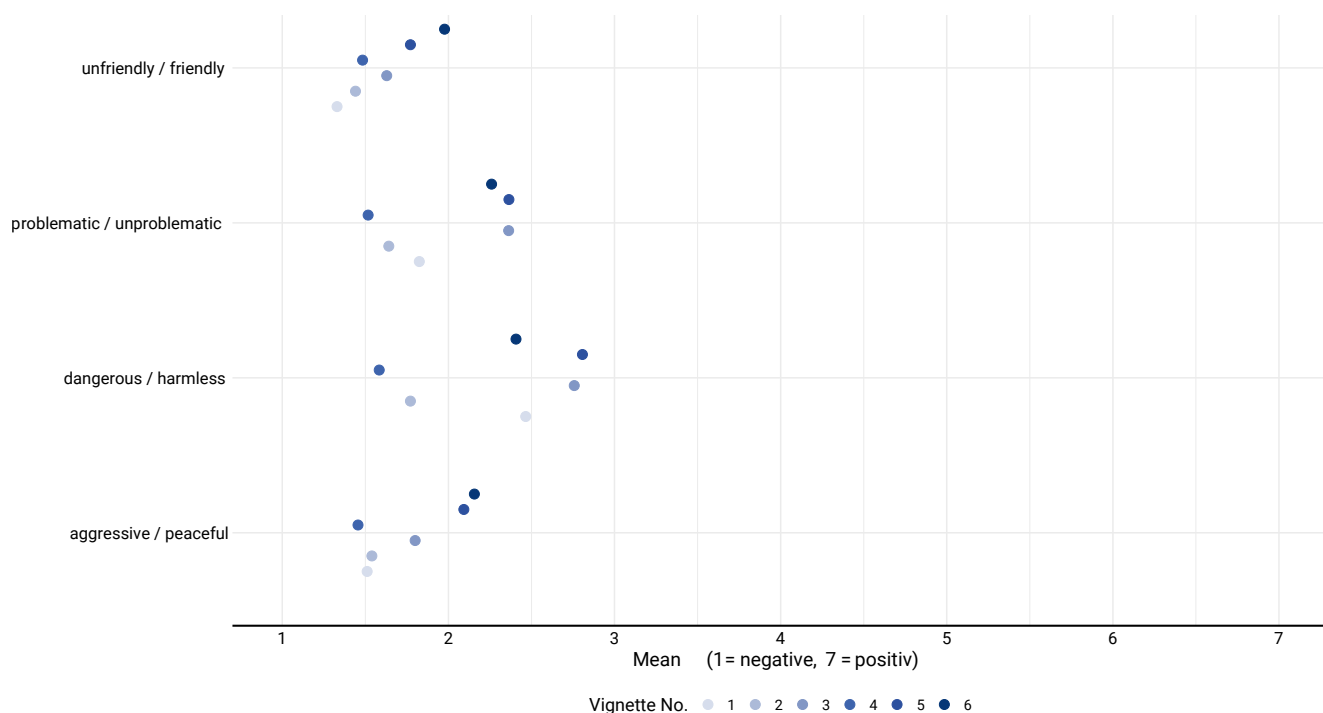
Respondents were then randomly assigned to one of the reporting channels of Meta (Facebook or Instagram), TikTok or X. We selected these four platforms to cover different age groups and because the DSA reporting interfaces differ significantly among them. While Facebook and Instagram list the relevant sections of the criminal code, TikTok requires users to name the law and provide their own statement. X, meanwhile, offers two separate channels in the first step of reporting – one for illegal content and one for content that violates community guidelines. The following statements about how the reporting channels work can therefore only be made for the platforms discussed.

We first assessed whether the participants categorized content correctly. The categorization of content was deemed correct if it was assigned to the appropriate categories in the first two steps, either under the platforms' community guidelines (e.g. glorification of violence, hate speech, insults) or under the DSA.[25] If the participants chose unrelated categories that deviated significantly from the main content of the post (e.g. fraud), the reporting process was terminated to test whether the categorization itself was an obstacle.

Overall, nearly three quarters (73 %) categorized content correctly. Results differed by platform: correct categorization rates were 61 % on the Meta platforms, 75 % on TikTok, and 82 % on X. This may be connected to the fact that DSA reporting was also chosen with different frequency by the respondents: on average, around 9 % of participants used the DSA reporting channel on the Meta platforms, compared to just 2 % on TikTok, while a comparatively high proportion of respondents (11 %) used this channel on X.

⑥ Use of DSA reporting

Share of reports under DSA in %
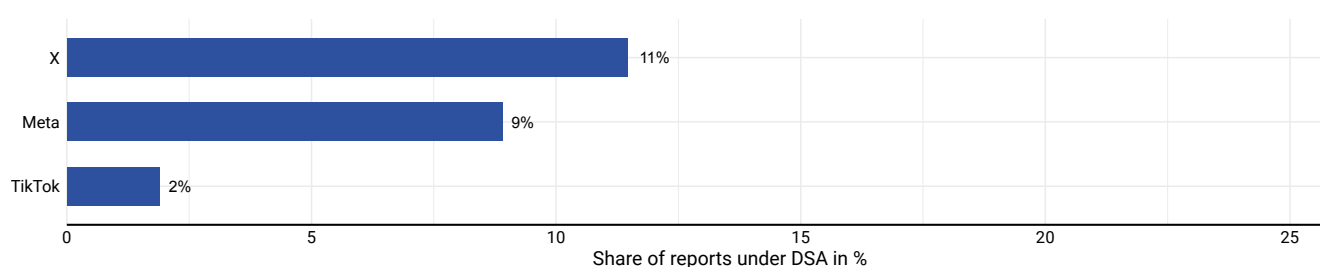
One possible explanation is that X requires a clear decision between community guideline reporting and illegal content (DSA) right at the start of the reporting process. On other platforms, the DSA reporting channel appears only as one of many subcategories, making it less visible and less likely to be used.
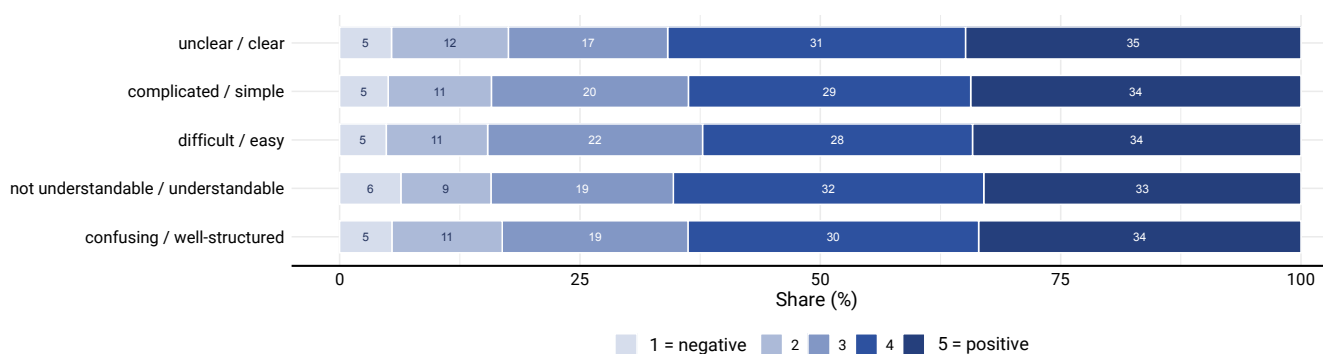
This was also reflected in the qualitative study. Most respondents intuitively chose one of the first community guideline categories without ever reaching the "illegal content" option – often because the categories already seemed to cover the case in question. One participant described his reasons for reporting a hateful comment to the community guidelines as follows:

> *"Because if there was something like hate speech in the comment, then I would have already reported it before, and I wouldn't have gone here [DSA reporting]. I would only have chosen that point if I hadn't found myself in any of these [community guidelines] categories." (Salim, 35)*

Another barrier to using the DSA channel was uncertainty about what counts as illegal. Many opted against DSA reporting because that meant less uncertainty and a quick decision was required: "I wouldn't know explicitly what the difference is, because for me the question is, when is something actually illegal? I couldn't judge that ad hoc." (Laura, 43).

After completing the process, respondents rated the reporting channel on various scales. Overall, the assessment was (somewhat) positive: 66 % of respondents considered the reporting channels to be (fairly) clear, 65 % found them comprehensible, 64 % described them as (fairly) straightforward, 63 % said the process was (fairly) simple, and 62 % regarded it as (fairly) easy. A majority therefore considered the processes understandable and manageable, even if a considerable share was neutral or critical. Differences between platforms were minor, though the process on X was perceived more critically – surprising given the higher likelihood of DSA reporting there, but possibly reflecting lower trust in X among users. Reasons for evaluating the reporting process were not measured.

7 Evaluation of the reporting process

|	1 = negative	2	3	4	5 = positive
unclear / clear	5	12	17	31	35
complicated / simple	5	11	20	29	34
difficult / easy	5	11	22	28	34
not understandable / understandable	6	9	19	32	33
confusing / well-structured	5	11	19	30	34

Share (%)

Clearer differences emerged between DSA and community guide-line reporting. The process for DSA reporting was consistently rated as more difficult, more confusing and more complicated, as well as somewhat less clear and less comprehensible than reporting via community guidelines.

⑧ Evaluation of the reporting process by reporting route

⑨ Predictors for the evaluation of the reporting process
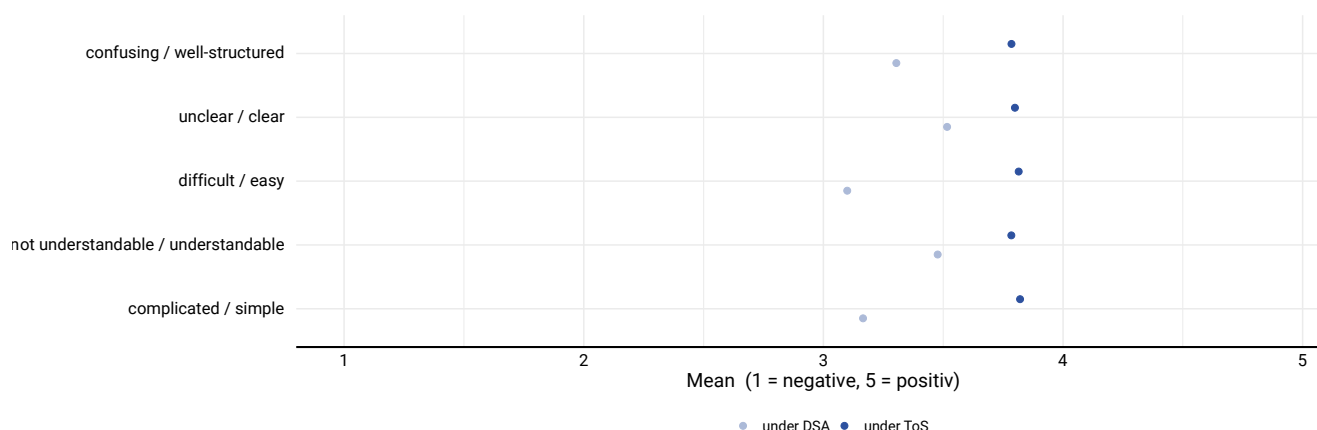
Regression analysis confirmed this statistically: DSA reporting was associated with more critical assessments. Younger people and those perceiving stronger consequences of problematic content for society rated the channels more positively; education, migration background, gender, social media use and political orientation showed no significant impact on the evaluation of the reporting process.
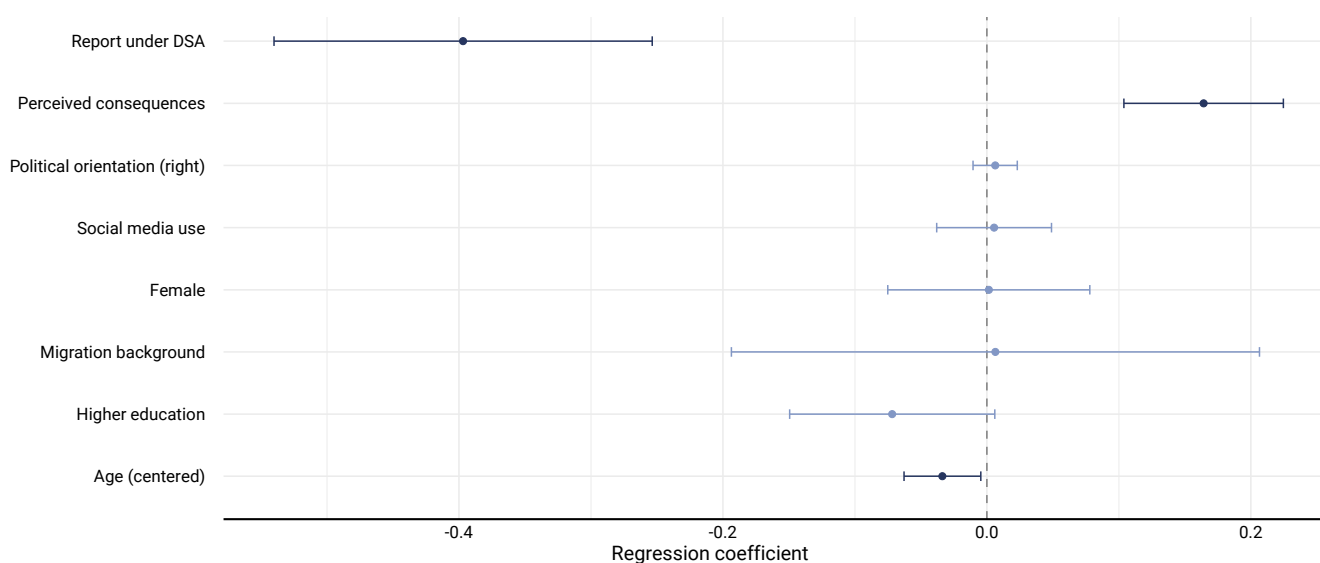
A regression coefficient > 1 indicates a positive relationship with the rating of a report, while a value < 1 indicates a negative correlation.

This is consistent with the qualitative findings. Looking back, many participants described the DSA reporting process as overall manageable and in some cases even less complicated than expected. In particular, persons who had already filed reports in the past more frequently found the steps comprehensible and easy to handle. They said it was "not that much in the end" (Elif, 29). Providing detailed input during reporting gave some participants the impression of submitting a more differentiated report, one that would be taken seriously and therefore more likely to actually succeed. One young participant summed it up by saying that he found the DSA reporting channel "easy enough for what it's meant for, since it really is about illegal stuff." At the same time, he added: "But still... for our generation, I think a lot of people wouldn't bother to take the extra step of writing something. […] So it's kind of in-between." (Jonas, 18).
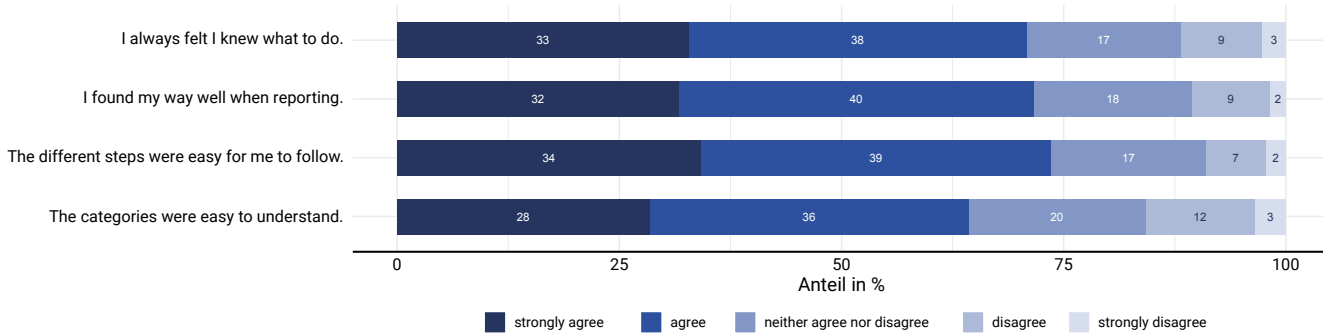
In contrast, community guideline reporting was seen as simpler and more straightforward. DSA reporting was more often viewed as time-consuming and discouraging, leading respondents to prefer community guideline reporting: "If I'm not entirely sure, I'd rather just report it the normal way." (Jonas, 18). Fear of personal legal consequences could also discourage reporting and result in too little illegal content being flagged. Another factor was that many users felt uncomfortable providing or storing personal data. One participant described being put off by having to sign a declaration: "I'd immediately be afraid of doing something wrong. That I'd end up doing something illegal myself." (Elif, 29).

This results in the paradox that the "juridification" of the user's right to report content, requiring legal expertise, may have made the process too complex and daunting. This has led, on the one hand, to users more frequently aborting the process, and, on the other hand, to a reduced willingness to report content in the future:

"I think it will tend to make people – or me personally – less likely to report here or not as often, because it takes a huge amount of time if you want to do it conscientiously." (Dennis, 31)

The evaluation of the reporting process was followed by concrete statements about participants' experiences. Here too the overall assessment was largely positive: 74 % of respondents found the individual steps easy to follow, 72 % said they had no trouble navigating the process, and 71 % reported always knowing what to do at each stage. Approval was somewhat lower regarding the clarity of categories, with 64 % rating them positively. Taken together, these results confirm that the majority of users perceived the reporting process as manageable and understandable.

	strongly agree	agree	neither agree nor disagree	disagree	strongly disagree
I always felt I knew what to do.	33	38	17	9	3
I found my way well when reporting.	32	40	18	9	2
The different steps were easy for me to follow.	34	39	17	7	2
The categories were easy to understand.	28	36	20	12	3

Anteil in %

In our study, 6 % of respondents aborted the reporting process. Clear differences emerged between platforms: drop-offs were rare on Meta (3 %), somewhat more common on TikTok (6 %), and highest on X (9 %). This suggests that the design and handling of reporting channels can lead to varying drop-off rates depending on the platform. Even more decisive than platform design, however, was the type of reporting: more than a quarter of DSA reports (27 %) were aborted, compared with only 4 % of community guideline reports.

As reasons for not continuing with the reporting process, about a quarter of the 117 respondents who had done so said they were "overwhelmed by the options". Another third reported no longer knowing "how to continue" (32 %).

This also came through in the qualitative findings. The selection of categories, which appears only later in the reporting process, was often seen as too complex and as a potential reason for aborting the submission of a report. Many participants found the categorization to be "overwhelming, to be honest. [...] it's kind of overloaded. [...] and then you think, oh whatever. Or in case of doubt you just click 'Other'." (Leon, 26). The complexity of the process was deemed to be particularly negative by participants at this point. Above all, they questioned and criticized the required level of detail: "They can't be serious. [...] what difference does that make now? Why does it have to be broken down in such detail, making it so hard for people to report this stuff? I just don't get it." (Florian, 38).

Instead of aborting the process at that stage, some respondents simply settled for choosing the category that seemed to fit best – the focus being on getting the content reported: "So I just always think, maybe it doesn't matter which one you click. As long as you click something in case of doubt." (Leon, 26).

We also used a regression model to examine the factors influencing process drop-offs. Women were more likely to discontinue the reporting process prematurely, while intense social media use and the

perception of negative consequences of problematic content made it more likely that participants would stick with it. Most striking was that drop-offs became significantly more likely when the notice was submitted under the DSA. Age, education, migration background, and political orientation played no significant role.

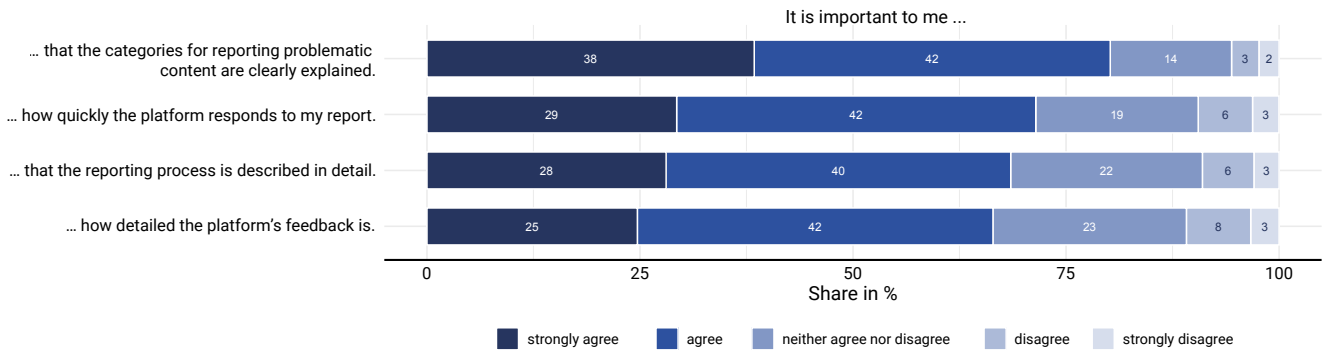**⑪ Factors influencing the decision to abort the reporting process**
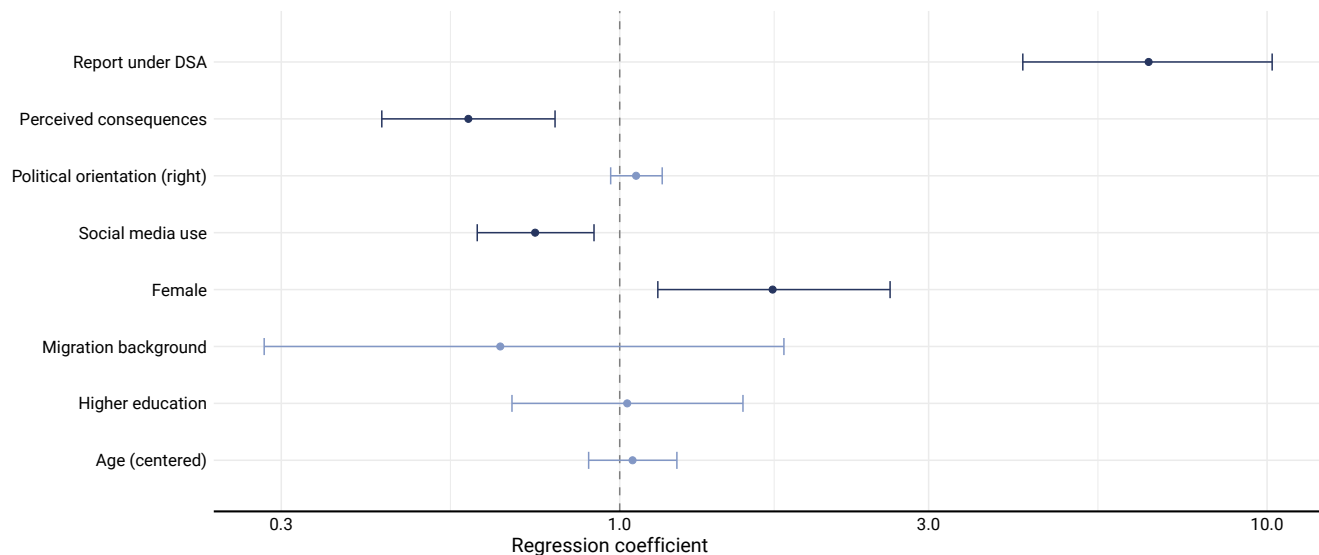
A regression coefficient > 1 indicates a positive relationship with the abort of a reporting process, while a value < 1 indicates a negative correlation.

5.5 Future reporting behavior

Finally, we asked which factors would be decisive for participants in reporting problematic content in the future. In principle, respondents continue to see themselves as responsible for reporting problematic content – though mainly as informants. They expect platforms to handle the further review, evaluation and classification, stressing that the question as to what the actual offense is "[should] happen at the next level up. People like me, the average user, should only be required to provide the indication that something isn't right." (Lucas, 43).

The findings from the quantitative survey highlight factors that could encourage future engagement. Transparency and comprehensibility of the process play central roles here. The most important thing was that reporting categories be clearly explained – 80 % considered this decisive. Many respondents also emphasized the speed of platform response (71 %). Two other factors were also rated as important by a majority of respondents: that the reporting process be described in detail (68 %) and how complete the platform's feedback is (67 %). In sum, users are most likely to report content in the future when the process is clear, transparent and provides tangible feedback.

It is important to me ...

	strongly agree	agree	neither agree nor disagree	disagree	strongly disagree
… that the categories for reporting problematic content are clearly explained.	38	42	14	3	2
… how quickly the platform responds to my report.	29	42	19	6	3
… that the reporting process is described in detail.	28	40	22	6	3
… how detailed the platform's feedback is.	25	42	23	8	3

Share in %

legend: ■ strongly agree ■ agree ■ neither agree nor disagree ■ disagree ■ strongly disagree

The design of the reporting process also appears improvable, with some respondents describing it as unintuitive, overloaded or half-hearted: Besides the presentation, the structure also gives the impression that the platforms are not putting much effort into it:

"The whole thing looks a bit old school. Like it was last updated ten years ago. And it's also kind of half-hearted, because up top it's in English with some kind of links, and below I'm supposed to decide what the exact law is that was violated. I don't know if I, if that even makes sense." (Markus, 28)

Most participants rejected a detailed categorization or legal assessment by the user, saying they felt neither responsible nor had the required expertise. Too much responsibility placed on users was seen as discouraging:

"I'm just a regular user. [...] I don't even really understand why this has to be evaluated in such detail. [...] I think the platform itself should want to make sure there's no illegal content." (Markus, 28)

---

25 All content shown to the participants was classified as illegal. According to the DSA, illegal content should ideally be reported via the DSA reporting channel rather than under community guidelines. Most participants, presented with illegal content, opted to use the community guidelines, as was also the case in the qualitative survey. The reasons for this were not measured.

6. Limitations

To examine how reporting channels are perceived and what hurdles they present, it is essential to create controlled conditions. Studying reporting behavior under controlled conditions, however, inevitably entails certain limitations that must be taken into account when interpreting the findings. Our reconstruction of the reporting structures on TikTok, Facebook/Instagram and X was designed to be as realistic as possible. Nevertheless, some legal requirements in the reporting process – such as the step of users entering their real name and giving their consent before submission – had to be omitted in our simulation to ensure the anonymization of the survey in line with the General Data Protection Regulation (GDPR). This reduces the external validity of the study, as it removes one hurdle identified in the qualitative investigation of DSA reporting – the requirement to provide personal data – from the reporting process. In addition, only the reporting channels of four VLOPSEs were reconstructed. This allows for specific statements to be made about how the reporting mechanisms function for a selected platform, but not for others. At best, design recommendations can be derived for platforms with similar setups. Furthermore, our focus was limited to reporting channels in Germany, whereas procedures for reporting illegal content vary across EU member states. In some countries, for example, there is no categorization of content, but rather an open-text field from the outset.[26] Finally, this study only considers reports of content, not of advertising placed on platforms. The examples shown to respondents were based primarily on user comments rather than posts published by individual profiles – although on almost all platforms the process of reporting content differs slightly from that of reporting comments.

---

26 See Marie-Therese Sekwenz, Ben Wagner and Simon Parkin (2025). "'It is unfair, and it would be unwise to expect the user to know the law!' Evaluating reporting mechanisms under the Digital Services Act." FAccT '25: Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency, 532–546. https://doi.org/10.1145/3715275.3732036.

7. Policy Recommendations: Strengthening the DSA's Impact, Clarifying Reporting Channels
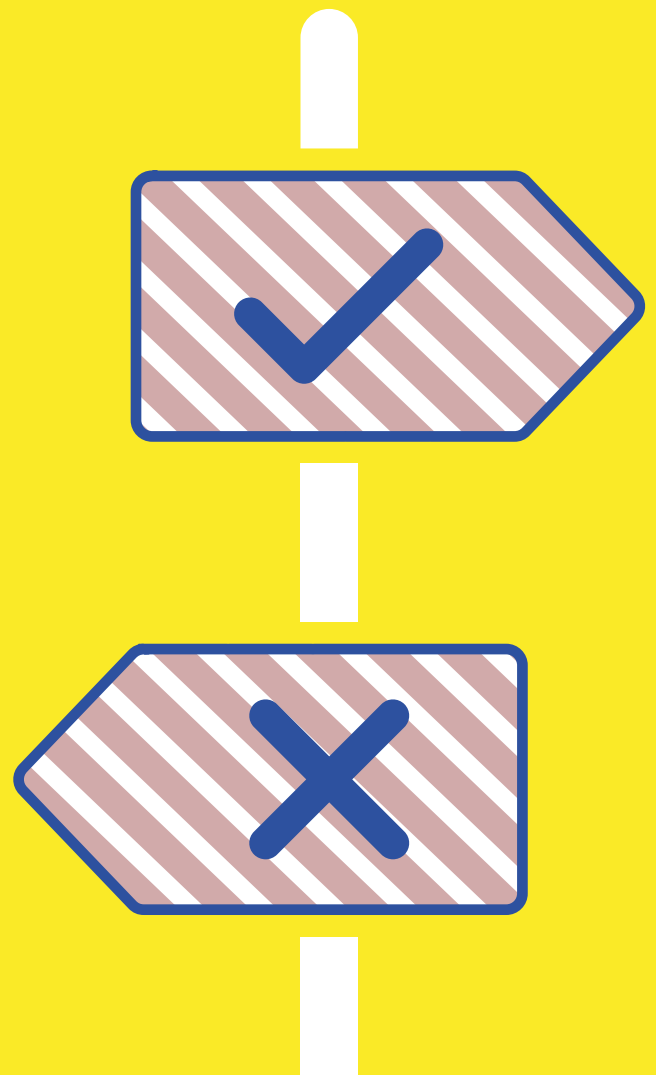
Standardization of reporting channels

The DSA requires that the channels through which users can report potentially unlawful content or violations of community guidelines on platforms must be user-friendly and easily accessible. So far, however, platforms have no standardized reporting procedures – and depending on the platform, there are also differences in when, how and where a post or comment can be reported. This not only makes things confusing for users, it can also mean that users end up reporting more content on platforms where the reporting channels are easier to find and understand than on others. We therefore recommend standardizing the reporting channels for DSA-relevant content on VLOPSEs. In addition, it must be clarified what exactly is meant by reporting mechanisms that are "easy to access and user-friendly" under Article 16(1) – especially in light of our study's finding that education level has a positive influence on past reporting behavior. To this end, the European Commission could promote the development of industry standards for VLOPSEs to standardize reporting channels.

Enforcement of community guidelines

Our findings show that reporting violations of community guidelines is seen as simpler and less complicated than reporting illegal content. Respondents in the qualitative study said that, given a choice, they would be more likely to report under community guidelines. We therefore call on VLOPSEs to enforce their community guidelines clearly.

Public awareness

Our data shows that content often goes unreported because users expect no platform response or consider reporting pointless. Users are also cautious about classifying content as potentially illegal. The reporting procedure under Article 16 DSA in its current form relies heavily on user motivation to flag harmful content and on their legal knowledge to use the mechanisms effectively. The European Commission and the national DSCs should therefore invest in large-scale awareness campaigns to explain that reporting illegal content does not entail legal consequences for the reporting individual. Such campaigns should also make clear how citizens can use their reporting rights to hold platforms accountable.

Design of reporting channels

When designing and updating reporting channels, platforms should focus on user-friendliness as defined by Article 16 and on child-friendliness as laid out in Article 28 at every step of the way. The qualitative study shows that the choice of the different categories – many of which only appear later in the reporting process – was seen as overwhelming and a key reason for aborting the process. Users tended to reject the requirement to make legal assessments. The high drop-off rate for DSA reporting compared with community guideline reporting also reflects a certain frustration with the design. Current reporting interfaces on VLOPSEs often require users to have prior legal knowledge to make competent reporting decisions. This makes it seem as though compliance with the law is prioritized over user-friendliness, with the legalistic language of the reporting process itself discouraging users from filing reports. Yet it should not be up to users to act as preliminary legal filters for content moderation.

We therefore urge the European Commission to simplify the process and remove the expectation of legal expertise. Based on our findings, we recommend that VLOPSEs clearly distinguish between reports under community guidelines and those under the DSA – for example by asking users directly about their reporting intention without intimidating them with legal language (e.g. by differentiating between "I want to report content because it violates the platform's community guidelines" and "I want to report content because I believe it is illegal").

In addition, intimidating notices suggesting that reporting illegal content could have legal consequences should be softened. Article 16 DSA does require users to provide "a sufficiently substantiated explanation of the reasons why the individual or entity alleges the information in question to be illegal content" (Article 16(2)(a)) and to make this statement in "the bona fide belief [...] that the information and allegations contained therein are accurate and complete" (Article 16(2)(d)); still, a detailed categorization or legal assessment by the users, as is currently being demanded by platforms, is not warranted. We recommend allowing users to select legal reasons for a report from a drop-down menu, supplemented with an optional free-text field for additional explanations.

Support by Design

How can legally recognized user rights be embedded into the design of interfaces without sacrificing usability?

Our qualitative survey allowed us to identify "pain points" for respondents and to determine to what extent problems of understanding or the presentation of unclear choices could be addressed with targeted support. We recommend that reporting channels incorporate built-in support from experts who can best address user difficulties. This could include references to legal aid organizations, reporting offices or out-of-court dispute settlement bodies.

8. Conclusion – Between Regulation and Reality

The findings of our study show that, on the platforms analyzed, the reporting process under the DSA is perceived as noticeably more difficult, confusing and complicated, as well as somewhat less clear and less transparent than reporting under community guidelines. Notices submitted via the DSA process were rated far more often as burdensome and discouraging compared to reports made under community guidelines. As a result, respondents – when faced with the choice – were more inclined to report content under community guidelines. Moreover, the reporting mechanisms of the VLOPSEs examined generally do not make it clear when user reports are being handled under the DSA and when they are being handled under community guidelines. Fear of personal legal consequences could also discourage reporting and result in too little illegal content being flagged. It could also pressure users willing to report content to submit their notices via the largely unregulated community guidelines instead of choosing the more complex, regulated DSA channel. The design of the interfaces can also influence the reporting process, the user experience, and the expectations placed on users. The DSA promises to create safer online spaces by providing user-friendly reporting mechanisms. Yet the complexity of the legal requirements and the untransparent practices of the platforms risk undermining efforts to encourage users to report problematic content. It remains questionable whether reporting systems can genuinely enable non-experts to flag content effectively. Reporting mechanisms were created as a tool for users – but they are not always designed in a way that meets users' needs.

9. References
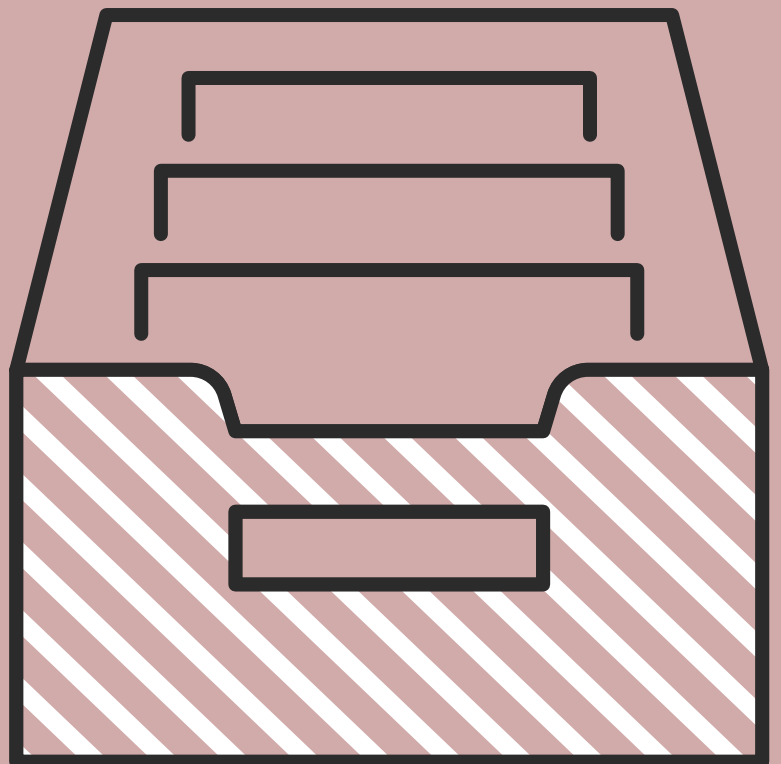
Bundesnetzagentur (2025). "Bundesnetzagentur testet Abläufe und Maßnahmen bei Verstößen gegen den Digital Services Act." Bundesnetzagentur (Digital Services Coordinator Germany). Retrieved 12 August 2025 at https://www.bundesnetzagentur.de/SharedDocs/Pressemitteilungen/DE/2025/20250131_DSC.html.

Coimisiún na Meán (2024). "Coimisiún na Meán opens review of online platforms' compliance with EU Digital Services Act." Coimisiún na Meán. Retrieved 18 July 2025 at https://www.cnam.ie/coimisiun-na-mean-opens-review-of-online-platforms-compliance-with-eu-digital-services-act/.

Deutscher Bundestag (2024). "Bundestag stimmt für die Umsetzung des Digital Services Act (DSA)." Deutscher Bundestag. Retrieved 2 July 2025 at https://www.bundestag.de/dokumente/textarchiv/2024/kw12-de-digitale-dienste-994508.

Steliyana Doseva, Fay Carathanassis, Hannah Schmid-Petri and Dirk Heckmann (2024). "Beleidigungen auf Social Media: Wie betroffen sind Nutzerinnen und Nutzer auf Social Media und wie wehren sie sich?" Bavarian Research Institute for Digital Transformation. Retrieved 20 August 2025 at https://www.bidt.digital/publikation/beleidigungen-auf-social-media-wie-betroffen-sind-nutzerinnen-und-nutzer-auf-social-media-und-wie-wehren-sie-sich/.

Thomas Escritt (2019). "Germany fines Facebook for under-reporting complaints." Reuters. Retrieved 15 July 2025 at https://www.reuters.com/article/us-facebook-germany-fine-idUSKCN1TX1IC/.

European Commission (2024). "Gesetz über digitale Dienste: Kommission leitet förmliches Verfahren gegen Facebook und Instagram ein." Representation of the European Commission in Germany. Retrieved 5 August 2025 at https://germany.representation.ec.europa.eu/news/gesetz-uber-digitale-dienste-kommission-leitet-formliches-verfahren-gegen-facebook-und-instagram-ein-2024-04-30_de.

European Commission (2025). "DSA: Very large online platforms and search engines." European Commission. Retrieved 9 July 2025 at https://digital-strategy.ec.europa.eu/en/policies/dsa-vlops.

Daniel Holznagel (2024). "Follow Me to Unregulated Waters! Are Major Online Platforms Violating the DSA's Rules on Notice and Action?" Verfassungsblog. Retrieved 19 August 2025 at https://verfassungsblog.de/follow-me-to-unregulated-waters/.

Luise Koch, Angelina Voggenreiter and Janina Steinert (2025). "Angegriffen & alleingelassen: Wie sich digitale Gewalt auf politisches Engagement auswirkt. Ein Lagebild." Technical University of Munich in cooperation with HateAid. Retrieved 6 July 2025 at https://hateaid.org/wp-content/uploads/2025/01/hateaid-tum-studie-angegriffen-und-alleingelassen-2025.pdf.

Maldita (2025). "Denunciamos a Meta 58 publicaciones fraudulentas en Facebook utilizando los mecanismos legales de la DSA y, una semana después, el 93 % siguen activas." Maldita.es. Retrieved 1 July 2025 at https://maldita.es/investigaciones/20250619/denunciamos-meta-publicaciones-fraudulentas-facebook-dsa/.

Das NETTZ, Gesellschaft für Medienpädagogik und Kommunikationskultur, HateAid and Neue deutsche Medienmacher*innen (eds.) (2024). "Lauter Hass – leiser Rückzug. Wie Hass im Netz den demokratischen Diskurs bedroht. Ergebnisse einer repräsentativen Befragung." Das NETTZ, Gesellschaft für Medienpädagogik und Kommunikationskultur, HateAid and Neue deutsche Medienmacher*innen. Retrieved 21 August 2025 at https://toneshift.org/wp-content/uploads/2024/02/Studie_Lauter-Hass-leiser-Rueckzug.pdf.

Friederike Quint, Yannis Theocharis, Spyros Kosmidis and Margaret E. Roberts (2025). "From Bystanders to Reporters: Who Acts Against Illegal Online Content?" European Political Science Association (EPSA) Conference.

Martin J. Riedl, Teresa K. Naab, Gina M. Masullo, Pablo Jost and Marc Ziegele (2021). "Who is responsible for interventions against problematic comments? Comparing user attitudes in Germany and the United States." Policy & Internet 13, 3 (2021), 433–451. https://doi.org/10.1002/poi3.257.

Ursula K. Schmid, Anna S. Kümpel and Diana Rieger (2023). "The Remote Self-Confrontation Interview Method: Gaining Insights into Users' Perceptions of Digital Media Content from a Distance." American Behavioral Scientist 69, 10 (2023), 1221–1237. https://doi.org/10.1177/00027642231205768.

Marie-Therese Sekwenz, Ben Wagner and Simon Parkin (2025). "'It is unfair, and it would be unwise to expect the user to know the law!' Evaluating reporting mechanisms under the Digital Services Act." FAccT '25: Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency, 532–546. https://doi.org/10.1145/3715275.3732036.

Yunhee Shim and Shagun Jhaver (2024). "Incorporating Procedural Fairness in Flag Submissions on Social Media Platforms." arXiv. https://doi.org/10.48550/arXiv.2409.08498.

Der Spiegel (2024). "HateAid reicht Beschwerde gegen TikTok ein." Retrieved 7 August 2025 at https://www.spiegel.de/netzwelt/hateaid-reicht-beschwerde-gegen-tiktok-ein-a-2db19652-4f70-46dd-ac02-4c999661a53c.

Nicolas P. Suzor, Sarah Myers West, Andrew Quodling and Jillian York (2019). "What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation." International Journal of Communication 13 (2019), 1526–1543. https://ijoc.org/index.php/ijoc/article/view/9736.

Charles Terroille, Saman Nazari and Ewan Casandjian (2025). "Flagged and Ignored: Testing X's Response to EU Sanction Violations." Alliance4Europe. Retrieved 14 August 2025 at https://alliance4europe.eu/flagged-and-ignored.

Anna-Lena von Hodenberg and Josephine Ballon (2025). "1 Jahr DSA: Sicherheit im digitalen Raum: eine Bilanz – und was jetzt nötig ist." HateAid. Retrieved 4 July 2025 at https://hateaid.org/wp-content/uploads/2025/07/HateAid_Policy-Brief_1-Jahr-DSA_barrierefrei-2.pdf.

Ben Wagner, Krisztina Rozgonyi, Marie-Therese Sekwenz, Jennifer Cobbe and Jatinder Singh (2020). "Regulating transparency? Facebook, Twitter and the German Network Enforcement Act." FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 261–271. https://doi.org/10.1145/3351095.3372856.

Shangyuan Wu (2024). "What Motivates Audiences to Report Fake News? Uncovering a Framework of Factors That Drive the Community Reporting of Fake News on Social Media." Digital Journalism 12, 6 (2023), 790–807. https://doi.org/10.1080/21670811.2023.2243489.

Alice Q. Zhang, Kaitlin Montague and Shagun Jhaver (2024). "Cleaning Up the Streets: Understanding Motivations, Mental Models, and Concerns of Users Flagging Social Media Content." arXiv. Retrieved 11 July 2025 at https://arxiv.org/pdf/2309.06688.

Imprint

Between Click and Consequence:
An Evaluation of Platform
Reporting Procedures under the
Digital Services Act

Authors
Lena-Maria Böswald
Corinna Dolezalek
Dr. Pablo Jost
Ursula Kristin Schmid

Translation
Christoph von Pohl (German-English)

Editorial
Franzi Prost

Cover image | Design
Delia Keller | Gestaltung Berlin

DAS
NETTZ