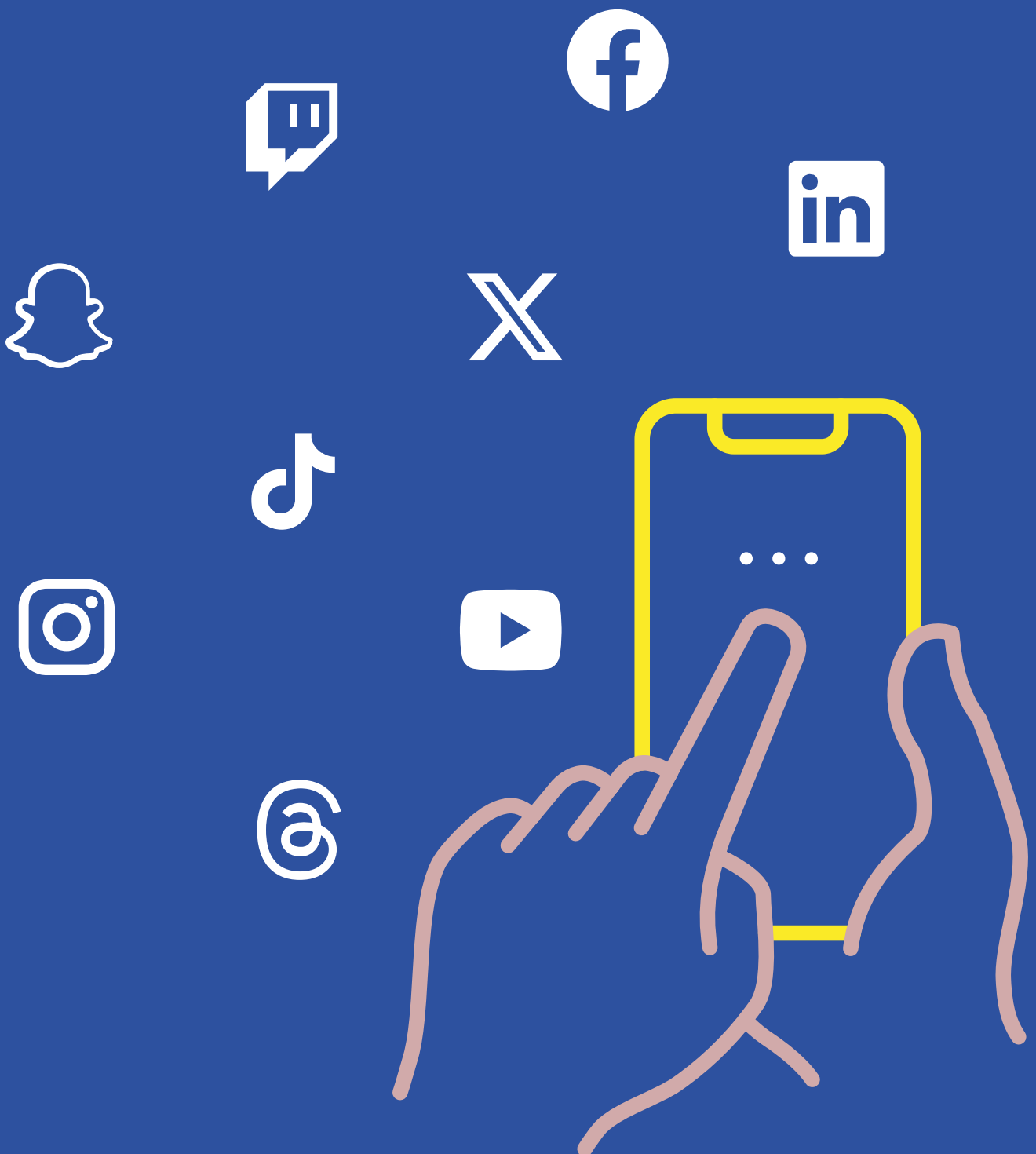


Zwischen Klick und Konsequenz: Eine Evaluation der Melde- verfahren von Plattformen nach dem Digital Services Act





INHALTSÜBERSICHT

5	1. Auf dem Prüfstand: Meldewege und Plattformverantwortung im digitalen Raum
9	2. Digitale Sorgfaltspflichten: Der DSA im Kontext politischer und gesellschaftlicher Erwartungen
13	3. Forschungsstand
19	4. Methodisches Vorgehen
23	5. Befunde der Untersuchung
24	5.1. Wahrnehmung, Nutzung, Wirkung: Was die Daten über Meldeprozesse verraten
24	5.2. Gründe für das Melden problematischer Inhalte
27	5.3. Erfahrung mit der Meldung von Inhalten
28	5.4. Simulation der Meldewege
39	6. Limitationen
41	7. Handlungsoptionen für die Politik: DSA-Wirkung stärken, Meldewege klären
45	8. Fazit – Zwischen Regulierung und Realität
47	9. Literaturverzeichnis



1. Auf dem Prüfstand: Meldewege und Plattform- verantwortung im digitalen Raum



Wer auf sozialen Plattformen unerlaubte Inhalte zu sehen bekommt, sollte diese eigentlich unkompliziert an die Betreiber*innen rückmelden können. So verlangt es zumindest das EU-Gesetz über digitale Dienste (engl. Digital Services Act, DSA). Mit dem DSA gibt es klare Regelungen für den Umgang mit rechtswidrigen Inhalten. Inhalte, die nach nationalem Recht illegal sind, müssen bei Meldung von Plattformen unverzüglich gelöscht werden. Artikel 16 Absatz 1 ist explizit darauf ausgerichtet, dass Plattformen „leicht zugängliche und benutzerfreundliche“ Meldemechanismen einrichten und es für Nutzer*innen einfach und nachvollziehbar sein soll, wie sie Beiträge melden können. Das soll helfen, strafbare Inhalte von sehr großen Online-Plattformen (VLOPs) und Suchmaschinen (VLOSEs) prüfen und gegebenenfalls löschen zu lassen.

Das Gesetz über digitale Dienste sieht dementsprechend vor, dass die Wege, über die Nutzer*innen potenziell rechtswidrige oder gegen die Community-Richtlinien verstoßende Plattformbeiträge melden können, nutzerfreundlich und leicht zugänglich sind. Die Meldewege sollten dabei leicht auffindbar und gut verständlich gestaltet sein, damit sie für Nutzer*innen keine Hürde darstellen, Inhalte zu melden. Bis dato gibt es auf keiner Plattform einheitliche Melde- und Abhilfeverfahren für Beiträge.

Die vorliegende Studie von Das NETTZ, in Kooperation mit Dr. Pablo Jost (Johannes-Gutenberg-Universität Mainz) und Ursula Kristin Schmid (Ludwig-Maximilians-Universität München), widmet sich der umfassenden Bewertung der Nutzerfreundlichkeit und Wirksamkeit der Meldewege von Plattformen qua DSA in Deutschland und der Beweggründe von Plattformnutzer*innen, Inhalte zu melden. Dabei steht für uns als Forscher*innen die Perspektive der Plattformnutzer*innen bewusst im Fokus, um einen Beitrag zu einer konsequenten Aufsicht im Sinne der Nutzer*innen zu leisten.

Die Studie konzentriert sich auf vier Forschungsfragen:

1. Was motiviert und was hemmt Nutzer*innen im Zuge einer Meldung?
2. Wie verständlich sind die Meldewege gestaltet und welche Schwierigkeiten haben Nutzer*innen?
3. Wie werden die Meldewege verschiedener Plattformen wahrgenommen und von Nutzer*innen bewältigt?
4. Ergeben sich Unterschiede zwischen Meldewegen nach DSA und nach Community-Richtlinien?

Abschließend richten wir unsere Empfehlungen an die Europäische Kommission, die nationalen Koordinierungsbehörden sowie direkt an die genannten Plattformen im Zuge der Umsetzung von Artikel 16 des

Digital Services Act. Die Erkenntnisse der vorliegenden Studie sind entscheidend, um Plattformen stärker in die Pflicht zu nehmen. Sie sollen sicherstellen, dass die Sicherheit und die Rechte der Nutzer*innen im Mittelpunkt der Gestaltung und Weiterentwicklung von Meldewegen stehen.

-
- 1 Europäische Kommission (2025). DSA: Very large online platforms and search engines. Europäische Kommission. <https://digital-strategy.ec.europa.eu/en/policies/dsa-vlops>.
 - 2 Während die EU-Kommission die ausschließliche Zuständigkeit für die Durchsetzung und Überwachung der Einhaltung der erweiterten Sorgfaltspflichten durch VLOPs und VLOSEs hat, die ihnen das Gesetz über digitale Dienste zur Bewältigung systemischer Risiken auferlegt, teilen sich die Kommission und die nationalen Behörden die Zuständigkeit für alle anderen Verpflichtungen, die VLOPs und VLOSEs im Rahmen des Gesetzes über digitale Dienste auferlegt werden.

2. Digitale Sorgfaltspflichten: Der DSA im Kontext politischer und gesellschaft- licher Erwartungen



Lange wurde der 17. Februar 2024 erwartet – an diesem Tag trat der DSA, die EU-weite Verordnung für digitale Dienste, vollständig in Kraft. Am 21. März 2024, knapp einen Monat später, stimmte der Bundestag für die nationale Umsetzung des DSA, das Digitale-Dienste-Gesetz (DDG).³ Konkret betrifft das Gesetz Internetprovider, Hostinganbieter, Cloud-Dienste, soziale Netzwerke, Messengerdienste und Online-marktplätze. Bereits seit dem 25. August 2023 gilt der europäische Rechtsakt für VLOPs und VLOSEs, die jeweils mehr als 45 Millionen monatliche Nutzer*innen in der EU haben. Darunter fallen zum Beispiel alle bekannten sozialen Netzwerke und Video-Hostingplattformen, Google Search, Microsoft Bing, aber auch Online-Versandhändler (Zalando, Amazon), Reiseportale (Booking.com) und auch Pornoplattformen (Pornhub).

Der DSA soll europaweit einheitliche Regeln für das Entfernen und Einschränken von Online-Inhalten einführen, die gegen rechtswidrige und plattforminterne Richtlinien verstoßen. Zudem müssen die Betreiber*innen proaktiv deren Entstehungs- und Verbreitungsrisiko abmildern und transparent begründen, wie ihre Inhalte moderiert werden. Bereits im April 2024 hat die Kommission ein förmliches Verfahren gegen die Facebook- und Instagram-Mutter Meta bezüglich ihrer Melde-mechanismen eingeleitet.⁴ Die Kommission vermutet, dass der Melde- und Abhilfemechanismus von Meta gegen die im DSA festgelegten Verpflichtungen verstößt. Sie nimmt hier Bezug auf die Anforderungen an das Unternehmen, laut derer dieser Mechanismus leicht zugänglich und nutzerfreundlich sein muss. Das Verfahren läuft derzeit noch. Auch Coimisiún na Meán, die irische Aufsichts- und Koordinierungs-behörde, hat inzwischen förmliche Auskunftersuchen an verschiedene Plattformen gerichtet, um detaillierte Informationen über deren Meldeoptionen zu erhalten.⁵

Art. 16(1) Die Hostingdiensteanbieter richten Verfahren ein, nach denen Personen oder Einrichtungen ihnen das Vorhandensein von Einzelinformationen in ihren Diensten melden können, die die betreffende Person oder Einrichtung als rechtswidrige Inhalte ansieht. Diese Verfahren müssen leicht zugänglich und benutzerfreundlich sein und eine Übermittlung von Meldungen ausschließlich auf elektronischem Weg ermöglichen.

Meldungen von potenziell rechtswidrigen Inhalten und von Verstößen gegen Community-Richtlinien sind auf allen VLOPs und VLOSEs anders gestaltet: Sind Meldungen nach DSA oftmals mit zusätzlichen Bestätigungen, der Angabe von persönlichen Daten und juristischen Terminen verbunden, sind Meldungen wegen Verstößen gegen Community-Richtlinien hingegen mit ein paar Klicks einzureichen. Meldewege für

illegale Inhalte erfordern beim Ausfüllen ein gewisses Maß an Rechtsverständnis. So müssen Nutzer*innen eine Erklärung abgeben, „warum die betreffende Person oder Einrichtung die fraglichen Informationen als rechtswidrige Inhalte ansieht“ (Art. 16 (2,a) und „dass die meldende Person oder Einrichtung in gutem Glauben davon überzeugt ist, dass die in der Meldung enthaltenen Angaben und Anführungen richtig und vollständig sind“ (Art. 16 (2,d).

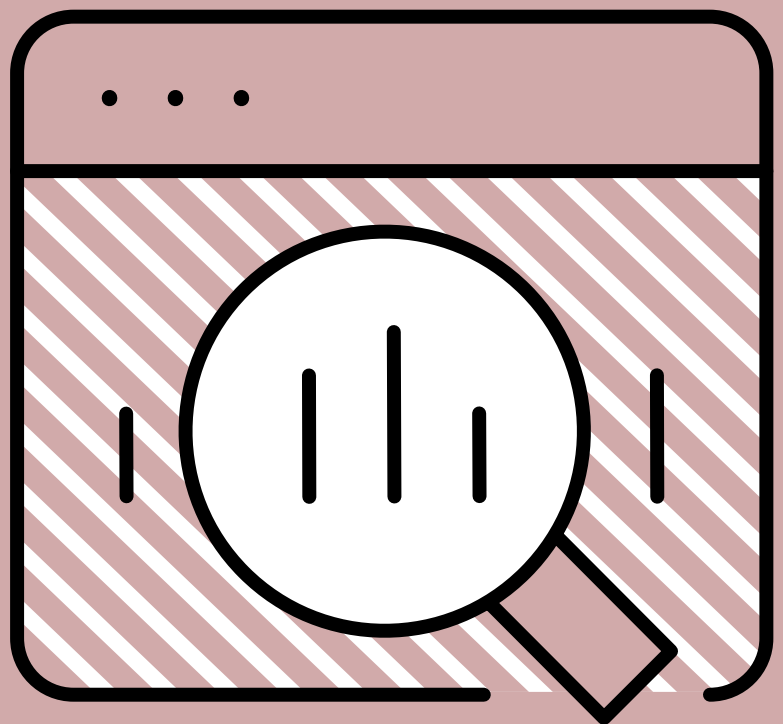
Die gemeinnützige Organisation HateAid reichte im Mai 2024 eine formelle Beschwerde gegen TikToks unzureichende Meldewege nach DSA Artikel 16 bei der Bundesnetzagentur (BNetzA), die als nationale Koordinierungsbehörde in Deutschland für die Umsetzung des Gesetzes über digitale Dienste zuständig ist, ein.⁶ Dass die BNetzA Plattformmeldewege als relevantes Instrument betrachtet, zeigte sich unter anderem im Februar 2025, als die Bundesbehörde in ihrem Plattformstresstest vor der Bundestagswahl 2025 die „Meldewege und Mechanismen der Plattformen“ testete.⁷

Laut HateAid seien auch TikToks Beschwerdemechanismen nicht ausreichend nutzerfreundlich gestaltet. In Zusammenarbeit mit Jurist Daniel Holznagel, Richter am Kammergericht Berlin, wurden das Design und die Funktionslogik der Meldewege auf allen großen Online-Plattformen als „Verstoß nach DSA Art. 16(1)“ bezeichnet.⁸ Die Annahme: Viele sehr große Online-Plattformen und Suchmaschinen (VLOPSEs) zwingen potenziell meldende Nutzer*innen übermäßig dazu, schwache, weitgehend unregulierte Meldungen zu Community-Richtlinien einzureichen. So denken Nutzer*innen möglicherweise nicht über die Unterschiede zwischen Meldungen nach Community-Richtlinien der Plattformen oder DSA nach, geht es für sie doch primär um die aktive Entscheidung der Meldung statt um die rechtlichen Konsequenzen. Dennoch – so die Einschätzung von HateAid – hielten die Plattformen Nutzer*innen davon ab, (starke) Hinweise einzureichen, die unter Meldungen nach dem DSA fallen würden. Die Folge wäre, dass diese Meldungen höchstwahrscheinlich weder in die für die Evaluation des DSA wichtige Transparenzberichterstattung (Art. 15, 24 DSA) aufgenommen würden, noch dass ein Anfechten von Moderationsentscheidungen über den plattforminternen Beschwerdemechanismus ermöglicht würde (Art. 20, 21 DSA). Darüber hinaus ist anzumerken, dass die im DSA (Art. 17) vorgesehenen „Statements of Reasons“ (die Pflicht der Plattformen, Nutzer*innen über die Gründe einer Moderationsentscheidung zu informieren), in der praktischen Inhaltsmoderation bislang nur eine untergeordnete Rolle spielen.

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 Europäische Kommission (2024). Gesetz über digitale Dienste: Kommission leitet förmliches Verfahren gegen Facebook und Instagram ein. EU-Repräsentanz Deutschland. 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. Forschungsstand



Die Frage, wie Nutzer*innen mit den vorhandenen Meldewegen umgehen, ist bereits seit einiger Zeit Gegenstand wissenschaftlicher und zivilgesellschaftlicher Debatten. Während der DSA neue rechtliche Rahmenbedingungen für Melderichtlinien geschaffen hat, bleibt unklar, ob diese tatsächlich die angestrebte Wirkung entfalten. Frühere Regelwerke wie das Netzwerkdurchsetzungsgesetz (NetzDG), das durch den Digital Services Act abgelöst wurde, zeigten bereits, dass die Umsetzung erhebliche Lücken aufweisen kann. Dennoch liegen bisher nur wenige empirische, systematische Daten zu den konkreten Erfahrungen und Wahrnehmungen der Nutzer*innen vor.

Im Kontext dessen, was Nutzer*innen im Zuge einer Meldung motiviert bzw. hemmt, bietet eine Studie des bayerischen Forschungsinstituts für digitale Transformation⁹ zu Beleidigungen in den sozialen Medien von 2024 Zahlen: Demnach gaben 54 % der direkt von Online-Beleidigungen betroffenen Befragten (N = 890) an, dass sie beleidigende Beiträge direkt auf Plattformen oder bei einer staatlichen Stelle gemeldet haben. Als Gründe für die Nichtmeldung von Beleidigungen nannten die Befragten mangelndes Meldeinteresse ihrerseits (66 %), die empfundene Zwecklosigkeit der Meldung auf Plattformen (57 %), fehlende Kenntnis hinsichtlich vorhandener Meldewege (knapp 44 %) und die Unsicherheit über die Rechtswidrigkeit der Inhalte (knapp 35 %). Wu (2024) konnte die folgenden sechs Motivationsfaktoren für oder gegen eine Meldung von Inhalten kategorisieren: die Art des Beitrags, die Art der Quelle, Reaktionen von anderen, das Interesse und Wissen der*des Betroffenen, kulturelle Normen und mögliche Konsequenzen des Meldevorgangs.¹⁰ Letzteres deckt sich mit den Ergebnissen unserer qualitativen Umfrage, bei der Befragte angaben, den Meldevorgang abubrechen, wenn juristische Konsequenzen bei Falschmeldung für sie folgen könnten. Eine Studie der TU München, gemeinsam mit der Meldestelle REspect!, konnte wiederum feststellen, dass Menschen, die Inhalte auf Plattformen meldeten, wenig Vertrauen in Plattformen haben, online bereits negative Erfahrungen gemacht haben und primär aus einer zivilen Pflicht heraus melden.¹¹ Zhang, Montague und Jhaver (2018) stellten in ihren Interviews zur Meldemotivation fest, dass Inhalte gemeldet wurden, wenn sie den eigenen Wertvorstellungen widersprechen, unabhängig von den von der Plattform vorgegebenen Community-Richtlinien.¹² Mehr als die Hälfte der Teilnehmer*innen war der Meinung, dass jede*r Nutzer*in für die Meldung von Normverstößen verantwortlich ist. Diese Überzeugung stand in engem Zusammenhang mit Überlegungen, Inhalte zu kennzeichnen, um andere zu schützen. Vom weiteren Flagging, d.h. der Kennzeichnung/Meldung von Inhalten, hielten die Teilnehmer*innen vor allem ausbleibende oder unzureichende Reaktionen der Plattformen auf ihre Meldungen ab. Inhalte blieben selbst Wochen nach der Meldung online. Die ausbleibende Reaktion stärkte die Frustration und das Gefühl der Melder*innen, im Unrecht zu sein.

Die Frage, wie verständlich Meldewege gestaltet sind, behandelt ein Blick in die Praxis: Bereits 2019 hatte Deutschland von Facebook ein Bußgeld gefordert, da die Plattform Beschwerden zu illegalen Inhalten über den plattforminternen Meldeweg nicht ausreichend an die entsprechenden Behörden meldete und damit gegen die Transparenzverpflichtungen des NetzDG verstieß.¹³ Trotz der vielfach vermuteten Mängel der Meldewege, sind empirische Untersuchungen und Daten rar. Frühere Studien zur Gestaltung und Durchsetzung der Melde-mechanismen nach dem NetzDG konzentrieren sich darauf, wie Facebook und X ihre Nutzer*innen dazu bringen, eher den Meldeweg via Community-Richtlinien zu nutzen als die gesetzliche Meldepflicht, indem die Plattformen die notwendigen Klicks zur Meldung nach dem Gesetz erschweren.¹⁴ Shim und Jhaver (2024) resümierten in ihrer Studie zum fairem Flagging, dass die rechtliche Komplexität von Flagging-Systemen auf sozialen Plattformen häufig die Nutzerfreundlichkeit überlagert. Dies führt zu Schnittstellen, die für viele Nutzer*innen nur schwer handhabbar sind.¹⁵

Die gemeinnützige Organisation HateAid analysiert seit Anfang 2024 die Umsetzung der Nutzer*innenrechte gemäß Artikel 16, 20 und 21 DSA auf Facebook, Instagram, TikTok, YouTube und X.¹⁶ Ein besonderes Augenmerk liegt dabei – wie bei der vorliegenden Studie – auf der Umsetzung von Nutzer*innenrechten aus der Perspektive von Nutzer*innen. So konnte mittels einer quantitativen Auswertung der Moderationsreaktionen festgestellt werden, dass nur 44 % der 151 als rechtswidrig gemeldeten Inhalte entfernt wurden; eine interne Beschwerde bei der Plattform konnte diese Quote lediglich auf 54 % steigern. Außerdem kann – ähnlich wie bei einer Studie von Alliance4Europe auf X¹⁷ – angenommen werden, dass die Inhaltsmoderation häufig automatisch erfolgt, da auf Meldungen in weniger als zwei Minuten reagiert wurde. Zudem resümiert HateAid in ihrem Abschlussbericht, dass Plattformen in der Kennzeichnung der Meldewege nicht eindeutig zwischen Verstößen nach Community-Richtlinien und rechtswidrigen Inhalten unterscheiden. Mehr direkte Einblicke in die Meldewege nach Artikel 16 DSA bieten die Ergebnisse semi-strukturierter Interviews von Sekwenz et al. (2025).¹⁸ So stellten die Befragten fest, dass es zahlreiche Überschneidungen zwischen den Kategorien der Meldewege nach DSA und den Community-Richtlinien gibt, was die Meldung illegaler Inhalte selbst bei jenen Inhalten, deren Legalität unklar ist, erschwert. Eine primär auf die Einhaltung gesetzlicher Vorgaben ausgerichtete Meldestruktur kann zudem die Motivation von Nutzer*innen verringern, Inhalte zu melden, obwohl Meldewege gerade dem Schutz von Nutzer*innenrechten dienen sollen.

Vor dem Hintergrund, welche Schwierigkeiten Nutzer*innen bei der Meldung nach DSA und nach Community-Richtlinien haben, lassen

sich Ergebnisse einer Investigativrecherche der spanischen Nicht-regierungsorganisation Maldita deuten. Laut dieser ist die Einschätzung berechtigt, dass unzureichende Reaktionen der Plattformen auf Flagging zu Frustration unter den Nutzer*innen führt. So wurden 93 % der illegalen Inhalte (N = 58), die von Maldita über den offiziellen DSA-Meldeweg auf Facebook gemeldet wurden, nach einer Woche nicht von der Plattform genommen. Anders sah es bei Beiträgen aus, die über die Community-Richtlinien gemeldet wurden: So löschte Meta nur 6,8 % der als illegal über den DSA-Meldeweg gemeldeten Beiträge, entfernte aber 17 % der als Verstoß gegen die internen Regeln gemeldeten Beiträge.¹⁹ Meta war also bei der Anwendung seiner plattform-internen Regeln aufmerksamer als bei der Einhaltung seiner rechtlichen Verpflichtungen in der EU. Auch laut einer Analyse von Suzor et al. (2019) herrscht bei Nutzer*innen, die vor einer Meldeentscheidung stehen, große Frustration über die Durchsetzung von Meldemechanismen durch Plattformen.²⁰

Es zeigt sich zudem, dass die Wahrnehmung und Bewältigung der Meldewege auf verschiedenen Plattformen entscheidend davon abhängen, wie verständlich und wirksam diese gestaltet sind. Neben diesen Aspekten spielt auch die Motivation der Nutzer*innen, Inhalte tatsächlich zu melden, eine zentrale Rolle in unserer Studie. Aktuelle Zahlen dazu liefert eine Studie der TU München und HateAid, deren Ergebnisse zeigen, dass mehr als die Hälfte der befragten politisch aktiven Personen bereits Inhalte auf Plattformen gemeldet haben.²¹ Laut der repräsentativen Umfrage „Lauter Hass, leiser Rückzug“ des Kompetenznetzwerks gegen Hass im Netz aus dem Jahr 2024 haben nur 34 % aller Befragten Hass im Netz direkt bei einer Plattform gemeldet, jedoch bereits 74 % der von Hass im Netz Betroffenen.²² Betroffene wurden für die jeweils genutzten Plattformen gefragt, wie sie die Meldewege dieser bewerten: Nutzer*innen von Facebook, Instagram und TikTok gaben mehrheitlich (51–56 %) an, dass die Meldewege auf der jeweiligen Plattform eher bzw. sehr einfach sind. Die Meldewege auf YouTube, X und Dating-Plattformen bewerteten die Nutzer*innen ähnlich oft als einfach (44–46 %).

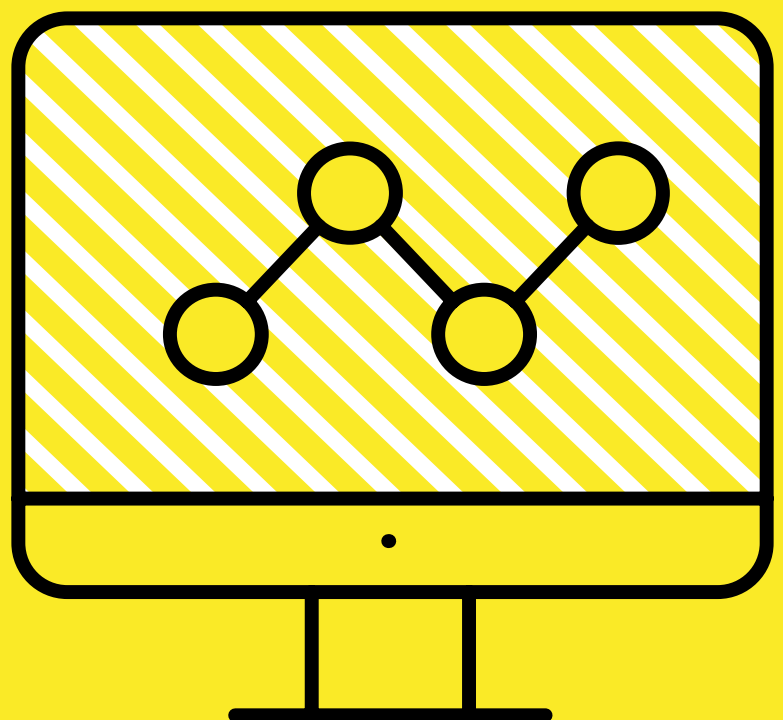
Im Zentrum des Digital Services Act steht die Verpflichtung von Plattformen, klare Strukturen für den Umgang mit problematischen Inhalten bereitzustellen und ihre Verantwortung bei Moderation und Transparenz wahrzunehmen. Wie wichtig solche Zuschreibungen von Verantwortung durch die Nutzer*innen selbst sind, zeigt eine vergleichende Studie zum Umgang mit problematischen Nutzerkommentaren in Deutschland und den USA von Riedl et al. (2021). Sie macht deutlich, dass Nutzer*innen die Verantwortung bei den Plattformen, aber auch beim Staat und bei sich selbst sehen. Während in Deutschland allen Akteur*innen – einschließlich der Nutzer*innen selbst – in

höherem Maße Verantwortung zugesprochen wird, fällt die Zuschreibung in den USA insgesamt zurückhaltender aus. Zugleich zeigt sich, dass ein starkes Bekenntnis zur Meinungsfreiheit die Erwartung an staatliches Eingreifen reduziert, ohne jedoch die Verantwortung von Plattformen, Medien oder Nutzer*innen zu schmälern. Diese Ergebnisse unterstreichen, dass Nutzer*innen ein integriertes Verständnis von Regulierung entwickeln, in dem verschiedene Akteur*innen ergänzend wirken – ein Befund, der die Mehr-Ebenen-Logik des DSA bestätigt.²³

-
- 9 Steliyana Doseva, Fay Carathanassis, Hannah Schmid-Petri und Dirk Heckmann (2024). Beleidigungen auf Social Media: Wie betroffen sind Nutzerinnen und Nutzer auf Social Media und wie wehren sie sich? Bayerisches Forschungsinstitut für Digitale 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 Siehe Friederike Quint, Yannis Theocharis, Spyros Kosmidis und 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 und 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 und Jatinder Singh (2020). Regulating transparency? Facebook, Twitter and the German Network Enforcement Act. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20)*. 261–271. <https://doi.org/10.1145/3351095.3372856>.
 - 15 Yunhee Shim und 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 und 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 und 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 und 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. *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency (FAccT '25)*, 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 und 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 und Janina Steinert (2025). Angegriffen und alleingelassen – Wie sich digitale Gewalt auf politisches Engagement auswirkt. Ein Lagebild. Technische Universität München in Kooperation mit HateAid. <https://hateaid.org/wp-content/uploads/2025/01/hateaid-tum-studie-angegriffen-und-alleingelassen-2025.pdf>.
- 22 Das NETTZ, Gesellschaft für Medienpädagogik und Kommunikationskultur, HateAid, und Neue deutsche Medienmacher*innen (Hrsg.) (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, und Neue deutsche Medienmacher*innen. https://tonesshift.org/wp-content/uploads/2024/02/Studie_Lauter-Hass-leiser-Rueckzug.pdf.
- 23 Martin J. Riedl, Teresa K. Naab, Gina M. Masullo, Pablo Jost und Marc Ziegele (2021). Who is responsible for interventions against problematic comments? Comparing user attitudes in Germany and the United States. *Policy und Internet* 13, 3 (2021), 433–451. <https://doi.org/10.1002/poi3.257>.

4. Methodisches Vorgehen



Um ein umfassenderes Verständnis für die Nutzerfreundlichkeit und Effektivität von plattforminternen Meldewegen zu erlangen, wählten wir einen Mixed-Methods-Ansatz für die vorliegende Studie. Zur Erfassung der Meldemotivation und -bereitschaft sowie der Verständlichkeit und Wirksamkeit der Meldewege wurden folgende Befragungen durchgeführt: vom 13. bis 21. Mai 2025 eine qualitative explorative Befragung, rekrutiert durch S&O Market Research Germany, und vom 10. bis 24. Juli 2025 eine repräsentative Bevölkerungsbefragung für alle deutschsprachigen Nutzer*innen von sozialen Medien ab 18 Jahren von Bilendi GmbH.

Methodisch erfolgte die qualitative Befragung in Anlehnung an remote self-confrontation interviews.²⁴ 20 Nutzer*innen wurden zunächst nach ihrem Nutzungsverhalten auf Plattformen und ihrem bisherigen Meldeverhalten befragt. Danach erhielten sie die Aufgabe, einen vorgegebenen Kommentar auf Facebook, Instagram oder TikTok exemplarisch zu melden. Während des Vorgangs äußerten die Teilnehmer*innen ihre Gedanken in Form des lauten Denkens, bei dem Überlegungen und Eindrücke spontan und ungefiltert mitgeteilt werden. Abschließend wurden die Meldeerfahrungen gemeinsam reflektiert und das zukünftige Meldeverhalten diskutiert.

Die quantitative Erhebung erfolgte mittels einer Online-Befragung. Im Rahmen der Rekrutierung wurden insgesamt 3.683 Personen rekrutiert. Befragt wurden ausschließlich Menschen, die angaben, „mindestens mehrmals im Monat“ eine digitale Plattform zu nutzen. Nach einer zusätzlichen Qualitätskontrolle (Bearbeitungszeit, Antwortmuster) blieben für die Datenauswertung insgesamt 2.040 Fälle. Die Befragten waren zwischen 18 und 69 Jahre alt. Das Durchschnittsalter der Stichprobe lag bei 44,5 Jahren. 53 % der Befragten waren Frauen und 47 % Männer, nur ein sehr kleiner Anteil (0,1 %) gab eine andere Geschlechtskategorie an. Knapp 53 % der Befragten verfügten über einen höheren Bildungsabschluss, das heißt mindestens Fachabitur. Besonders aussagekräftig wäre eine Studie mit einer Befragung, in dem Moment, in dem die Nutzer*innen einen schädlichen Inhalt auffinden und motiviert sind, ihn zu melden. So könnte der gesamte Prozess bis zur gegebenenfalls erfolgenden Moderation durch die entsprechende Plattform verfolgt werden. Da dies aufgrund forschungspraktischer und -ethischer Restriktionen kaum möglich ist, simulierten wir die Struktur und den Ablauf der Meldevorgänge der Meta-Plattformen sowie von TikTok und X. Entsprechend wurden die funktionalen Kernaspekte der verschiedenen Meldevorgänge erhalten, wobei sich die Benutzeroberfläche von den originalen Meldewegen der Plattformen unterschied. Darüber hinaus verzichteten wir aus datenschutzrechtlichen Gründen auf die Angabe der für die DSA-Meldung erforderlichen persönlichen Daten.

Der Fragebogen wurde von Das NETTZ, in Zusammenarbeit mit Dr. Pablo Jost und Ursula Kristin Schmid, erstellt. Dabei wurde Feedback von der Bundesnetzagentur (BNetzA), der Bundeszentrale für Kinder- und Jugendmedienschutz (BzKJ) sowie zivilgesellschaftlichen und wissenschaftlichen Akteur*innen (Institute for Strategic Dialogue, HateAid, Agora Digitale Transformation, Technische Universität München, Technische Universität Delft) eingeholt.

-
- 24 Siehe Ursula K. Schmid, Anna S. Kümpel und 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. Befunde der Untersuchung



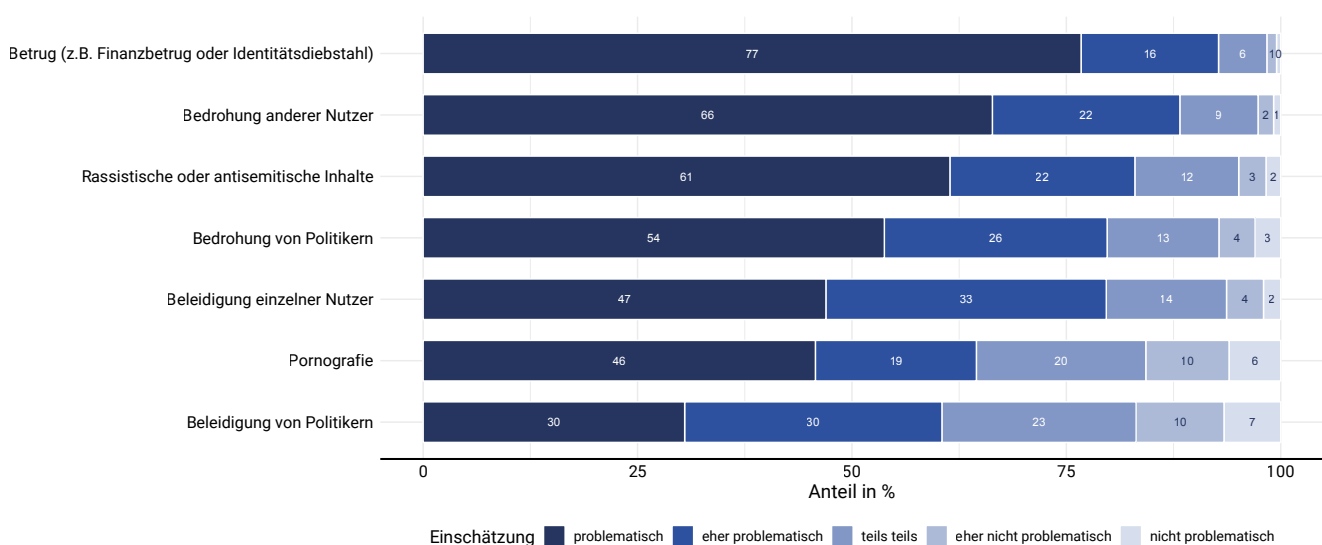
5.1 Wahrnehmung, Nutzung, Wirkung: Was die Daten über den Meldeprozess verraten

Das Rückgrat des folgenden Ergebnisteils bildet die quantitative Befragung. Diese gibt Aufschluss darüber, wie häufig bestimmte Phänomene in der Bevölkerung auftreten und lässt Aussagen über statistische Zusammenhänge zwischen diesen zu. Diese Befunde reichern wir mit den Ergebnissen der qualitativen Untersuchung an, um in der quantitativen Befragung gewonnene Erkenntnisse mit konkreten Aussagen veranschaulichen zu können. Zudem lassen sich so Vermutungen über die zugrundeliegenden Prozesse gewinnen und plausibilisieren.

5.2 Gründe für das Melden problematischer Inhalte

In der quantitativen Befragung wurden die Teilnehmer*innen gebeten, anzugeben, welche Inhalte sie auf Plattformen als problematisch wahrnehmen. Besonders deutlich zeigte sich dabei, dass Betrug – etwa in Form von Finanzbetrug oder Identitätsdiebstahl – fast einhellig als kritisch eingeschätzt wurde: 93 % bewerteten solche Inhalte als (eher) problematisch. Sehr hohe Zustimmung fanden auch Bedrohungen anderer Nutzer*innen (88 %) sowie rassistische oder antisemitische Inhalte (83 %). Ebenfalls eine große Mehrheit sah Bedrohungen von Politiker*innen (80 %) und Beleidigungen einzelner Nutzer*innen (80 %) als problematisch an. Etwas weniger stark, aber noch immer in der Mehrheit wurden Pornografie (65 %) und die Beleidigung von Politiker*innen (60 %) kritisch bewertet.

① Typen problematischer Inhalte



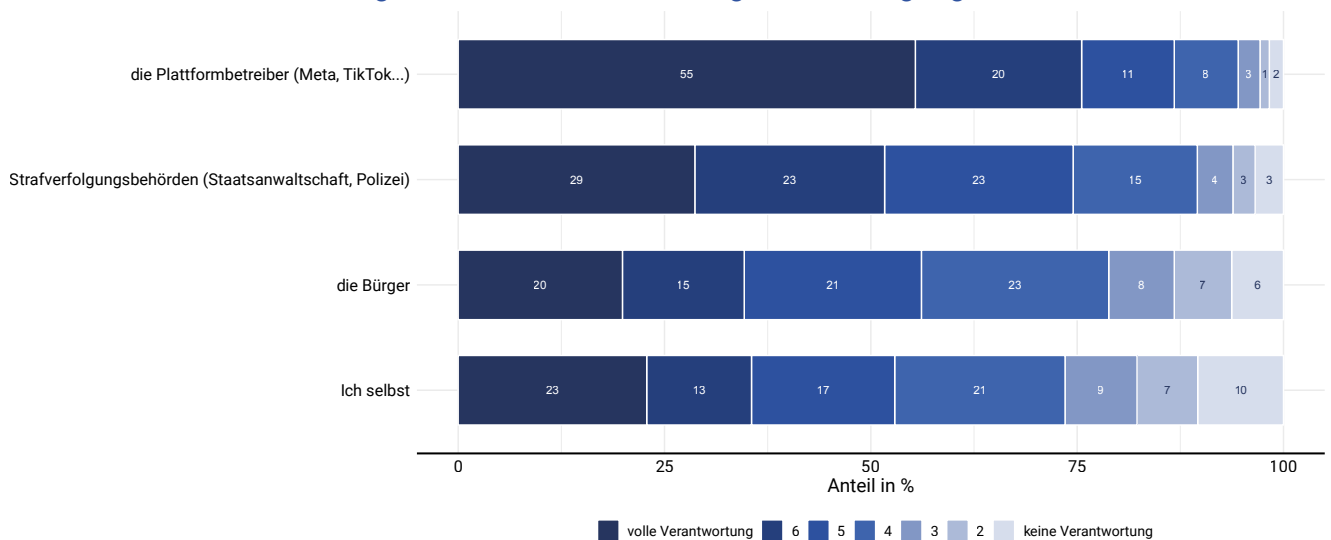
In der qualitativen Befragung wurden vor allem Inhalte als besonders problematisch bezeichnet, die Gewalt zeigen und verherrlichen oder

zu Gewalt aufrufen: „Also alles, was mit Gewalt halt zu tun hat, finde ich persönlich, sollte definitiv gemeldet werden“ (Marek, 28).

Daneben wurden Suizid-Ankündigungen und Kindesmissbrauch oder Kinderpornographie als vorrangig meldewürdig eingestuft. Ob zum Beispiel Beleidigungen aber tatsächlich gemeldet werden, hing laut Aussagen der Befragten häufig von der persönlichen Betroffenheit ab – insbesondere, wenn sie selbst oder nahestehende Personen direkt angegriffen werden, waren Nutzer*innen bereit, den Meldeprozess zu durchlaufen.

Zudem wurden die Teilnehmer*innen gefragt, wem sie die Verantwortung zuschreiben, problematische Inhalte auf Plattformen zu entfernen. Am deutlichsten zeigte sich dabei, dass die große Mehrheit der Befragten die Plattformbetreiber*innen selbst in der Pflicht sah: 86 % gaben an, dass Meta, TikTok und andere Anbieter*innen eine höhere oder gar die volle Verantwortung tragen. Ebenfalls als stark verantwortlich betrachtet wurden Strafverfolgungsbehörden, denen 75 % der Befragten eine zentrale Rolle zugeschrieben. Etwas seltener, aber dennoch von einer Mehrheit wurde die Verantwortung auch bei den Bürger*innen allgemein (56 %) sowie bei den Nutzer*innen selbst (53 %) gesehen. Damit wird deutlich, dass aus Sicht der Befragten vor allem die Plattformen und – in zweiter Linie – staatliche Stellen für die Entfernung problematischer Inhalte verantwortlich sind, während individuelle Verantwortung eher als ergänzend verstanden wird.

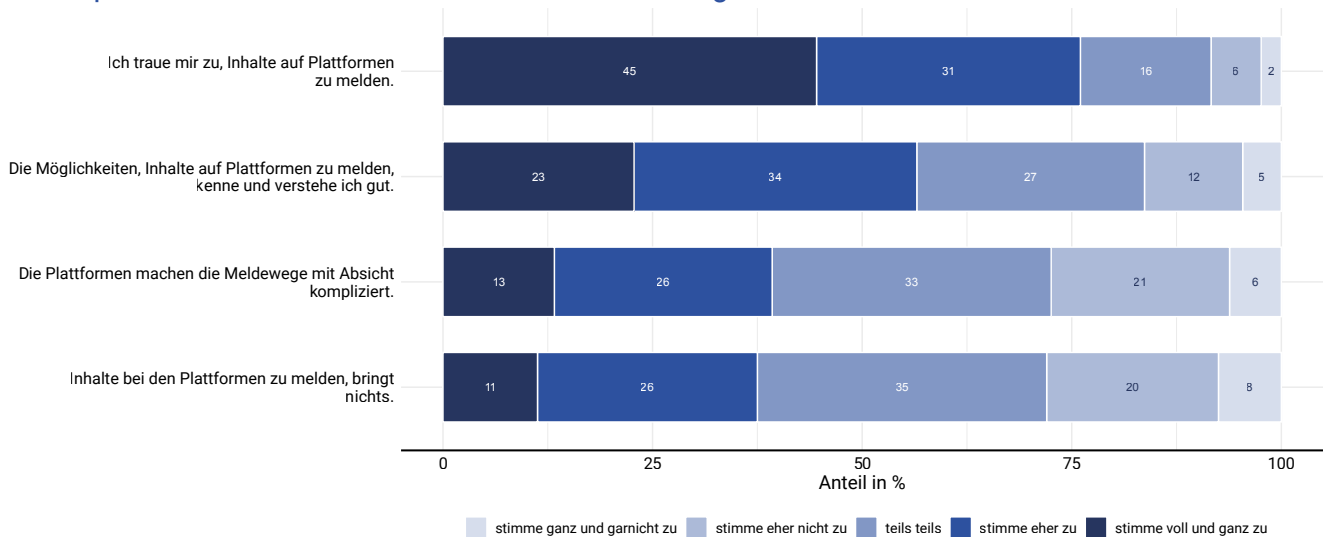
② Wahrgenommene Verantwortung für den Umgang mit Inhalten



Außerdem wurde erfasst, wie die Teilnehmer*innen die Möglichkeiten einschätzen, problematische Inhalte auf Plattformen zu melden. Dabei zeigte sich ein überwiegend positives Bild der eigenen Handlungskompetenz: Rund drei Viertel (76 %) der Befragten gaben an, sich (eher) zuzutrauen, Inhalte auf Plattformen zu melden. Mehr als die Hälfte

(57 %) erklärte zudem, die Meldewege gut zu kennen und zu verstehen. Kritischer fiel dagegen die Einschätzung über Plattformen selbst aus: 39 % stimmten der Aussage zu, dass Plattformen die Meldewege absichtlich kompliziert gestalten, und 37 % meinten, dass Meldungen ohnehin nichts brächten. Damit wird deutlich, dass die Nutzer*innen ihre eigene Fähigkeit zum Melden überwiegend positiv bewerten, zugleich aber signifikante Zweifel an der Wirksamkeit und Transparenz der Plattformmechanismen bestehen.

3 Empfundene (Selbst-)Wirksamkeit bei der Meldung

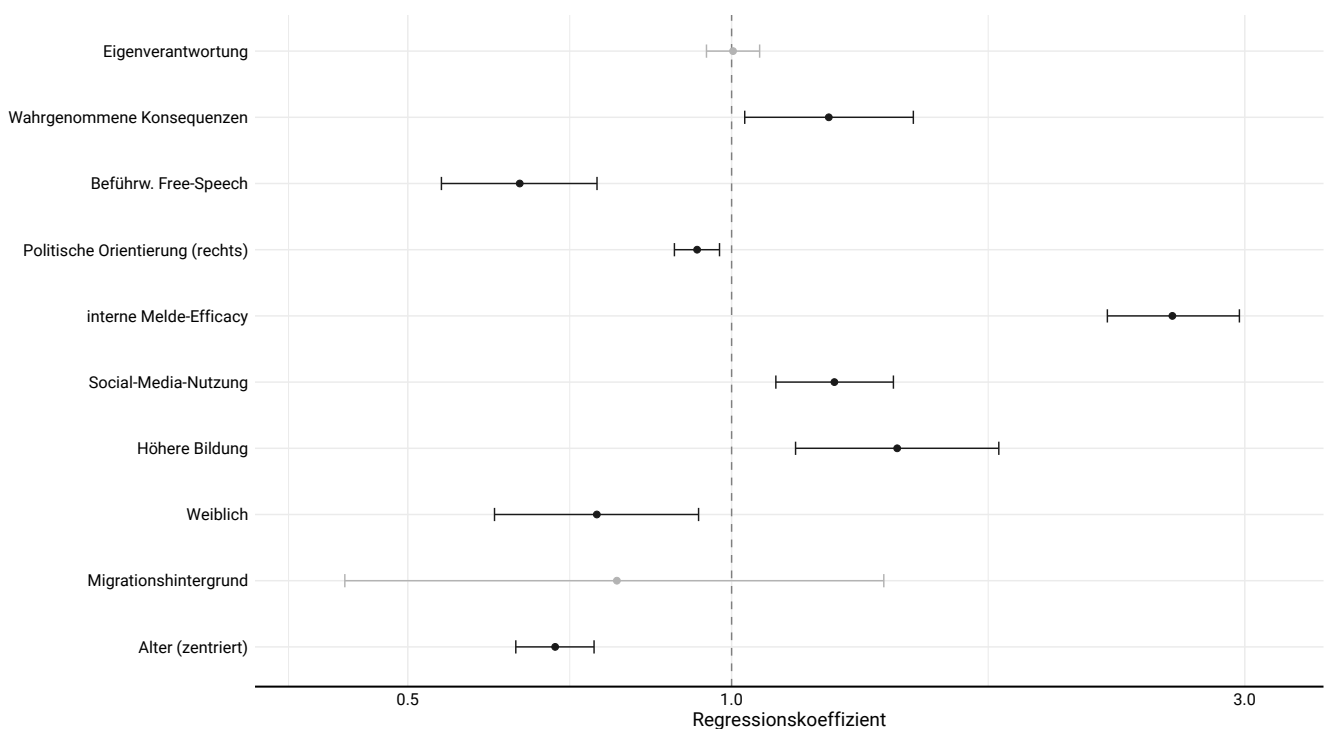


Auch in der qualitativen Studie hielten die Teilnehmer*innen Moderation und Regulierung grundsätzlich für wichtig – besonders die Möglichkeit, Inhalte zu melden. Danach sahen die meisten Nutzer*innen die Verantwortung bei den Plattformen. Häufig zeigte sich aber, dass sie „dem ganzen Prozess nicht wirklich hundertprozentig vertrauen“ (Marek, 28). Gleichzeitig wurde betont, dass der Einsatz solcher Funktionen stark vom Kontext und der Schwere des Vorfalls abhängen sollte. Einige Nutzer*innen sahen Gegenrede als wirksameres Mittel als sofortiges Löschen oder Anzeigen – insbesondere bei weniger gravierenden Inhalten. Ein Teilnehmer fasste zusammen, dass es „gewisse Kommentare [gibt], da tut man sich einfach eher einen Gefallen, wenn man die Leute stellt, anspricht, und versucht, in eine inhaltliche Diskussion zu kommen. Alles, was das [Gegenrede] abdecken kann, muss erstmal ausgehalten werden. Auch das ist Kennzeichen einer Demokratie“ (Lucas, 45). Kritisch betrachtet wurde zudem, dass Meldesysteme teils zu leicht zu bedienen sind und dadurch missbraucht oder zu streng ausgelegt werden könnten.

5.3 Erfahrung mit der Meldung von Inhalten

Sowohl in der quantitativen Befragung (45 %) als auch unter den Teilnehmer*innen der qualitativen Studie gaben rund die Hälfte der Personen an, in der Vergangenheit problematische Inhalte gemeldet zu haben. Welche Faktoren vergangenes Meldeverhalten begünstigten bzw. hemmten, haben wir mittels einer logistischen Regression geprüft. Demnach haben jüngere Personen und häufig hochgebildete Personen in der Vergangenheit eher Inhalte bei den Plattformen gemeldet. Personen, die sich politisch eher links verorten und die negative Konsequenzen problematischer Inhalte für die Gesellschaft befürchten, gaben ebenfalls häufiger an, Inhalte gemeldet zu haben. Migrationshintergrund und das Geschlecht der Befragten hingen hingegen nicht mit der Meldung von Inhalten zusammen.

4 Hemmende und begünstigende Faktoren vergangenen Meldeverhaltens



Ein Regressionskoeffizient > 1 zeigt einen positiven Zusammenhang mit vergangenem Meldeverhalten, ein Wert < 1 einen negativen.

Auch unter den Teilnehmer*innen der qualitativen Untersuchung berichteten besonders jüngere, politisch interessierte Nutzer*innen mit hoher Online-Aktivität davon, häufiger auf problematische Inhalte zu stoßen und diese zu melden. Dabei wurde in der Vergangenheit meist nach Community-Richtlinien gemeldet, vor allem, weil die Meldung nach DSA nicht bekannt war. Die Teilnehmer*innen der qualitativen Untersuchung berichteten insbesondere von Erfahrungen mit gängige Beleidigungen sowie von Hass gegenüber Minderheiten. Besonders oft stießen sie auf Fake-Profilen und Betrugsversuche, die häufig mit nega-

tiven persönlichen Erfahrungen verbunden waren. Häufigste Gründe, solche Inhalte zu melden, waren die wahrgenommenen Schweregrade der Inhalte sowie der generelle Wunsch, das (eigene) Online-Umfeld zu verbessern und zu erreichen, „dass dadurch weniger Hass gepostet wird“ (Lina, 26).

Bei den Gründen dafür, Inhalte in der Vergangenheit nicht gemeldet zu haben, treten verschiedene Muster hervor. Mehr als ein Drittel der Befragten gab an, dass die Meldung von Inhalten bei der Plattform als zwecklos angesehen wurde (36 %). Ebenfalls oft genannt wurde, dass die problematischen Inhalte die Befragten schlicht nicht interessiert hätten (32 %). Knapp 28 % erklärten, dass sie noch nie auf problematische Inhalte gestoßen seien, und 26 % fanden keinen passenden Meldeweg. Weitere genannte Gründe waren Zeitmangel (19 %), schlechte Erfahrungen mit dem Meldesystem (19 %) sowie Unklarheit über die Konsequenzen einer Meldung (18 %). Deutlich seltener spielte die Befürchtung einer emotionalen Belastung durch das Melden (13 %) eine Rolle.

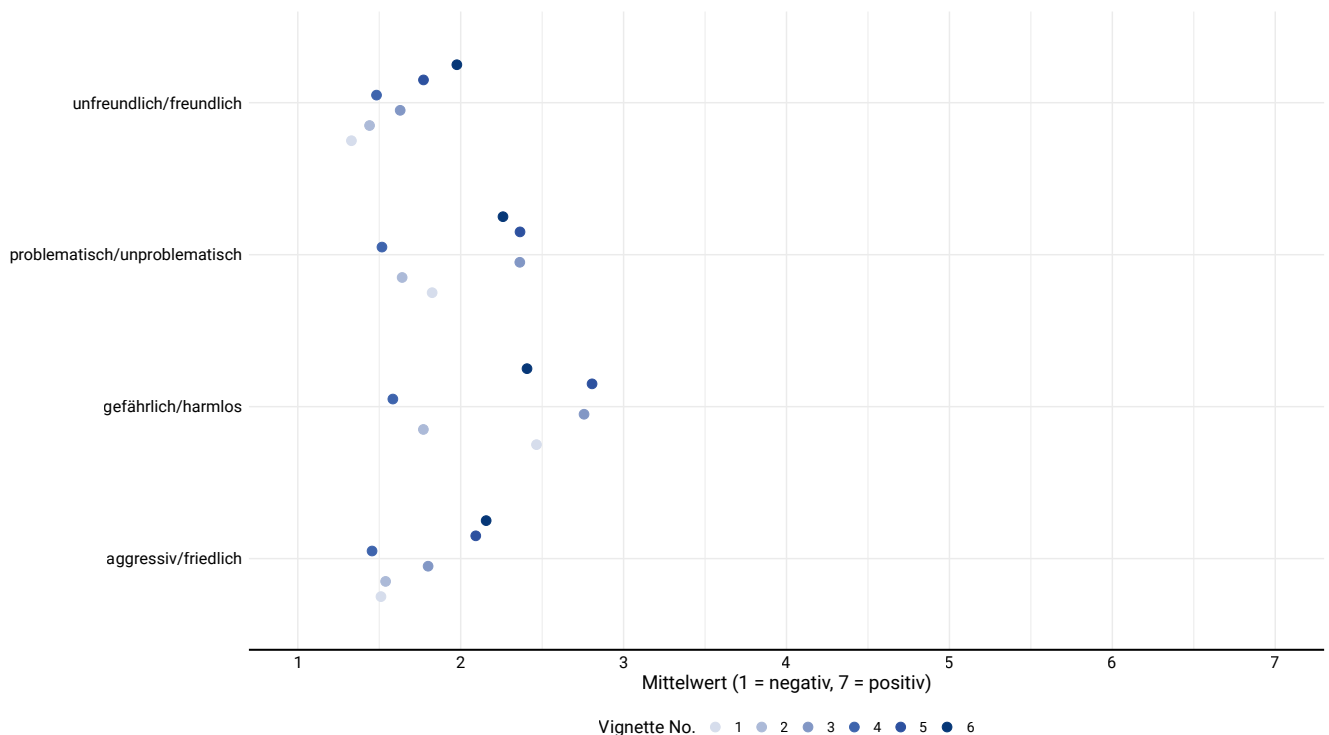
Ähnliches wurde auch von den Teilnehmer*innen der qualitativen Befragung ohne Vorgabe von Gründen genannt. Besonders negativ wirkte sich auch bei ihnen das fehlende Vertrauen in die Plattformen auf ihr Meldeverhalten aus. Insbesondere, wenn sie bereits die Erfahrung gemacht hatten, dass ihre Meldung nicht zur Löschung der Inhalte führte, wie eine Teilnehmerin ausführte: „Wenn ich dann sehe, hey, es bringt gar nichts, weil TikTok zurückmeldet, dass sie es prüfen und nichts daraus machen ... Dann denkst du dir halt irgendwann auch, hey, warum soll ich das machen?“ (Lina, 26). Neben der fehlenden Notwendigkeit, zu hoch eingeschätztem Aufwand und dem Vermeiden von eigener Verantwortungsübernahme, fanden es einige Nutzer*innen schlichtweg angenehmer, problematische Inhalte zu ignorieren, ihnen keine weitere Aufmerksamkeit zu schenken oder sich zu sagen: „Sollen die halt machen! Das ist nicht mein Thema“ (Marek, 28).

5.4 Simulation der Meldewege

Vor dem simulierten Meldevorgang wurden die Befragten darauf hingewiesen, dass sie im Folgenden einen Inhalt sehen, auf deren Basis sie anschließend einen Meldevorgang durchspielen sollten. Dabei handelte es sich um strafbewehrte Inhalte (Beleidigungen, Androhung von Gewalt, rassistische bzw. volksverhetzende Kommentare), die bereits von Gerichten als strafrechtlich relevant eingestuft wurden und allesamt im Kontext des DSA hätten gemeldet werden können. Die Befragten hatten zu jeder Zeit die Möglichkeit, den Meldevorgang abzubrechen.

Zunächst wurden die Befragten gebeten, den Inhalt entlang verschiedener Dimensionen zu bewerten. Dabei zeigte sich, dass die Stimuli überwiegend negativ wahrgenommen wurden: So wurden sie im Mittel eher als aggressiv und unfreundlich eingeschätzt, zugleich als problematisch und in Teilen auch als gefährlich bewertet. Die Werte auf den jeweiligen Skalen verdeutlichen damit, dass die präsentierten Inhalte nicht als neutrale Beiträge aufgefasst wurden, sondern von den Teilnehmer*innen klar in den Bereich negativer und problematischer Kommunikation eingeordnet wurden.

5 Bewertung der gezeigten Inhalte



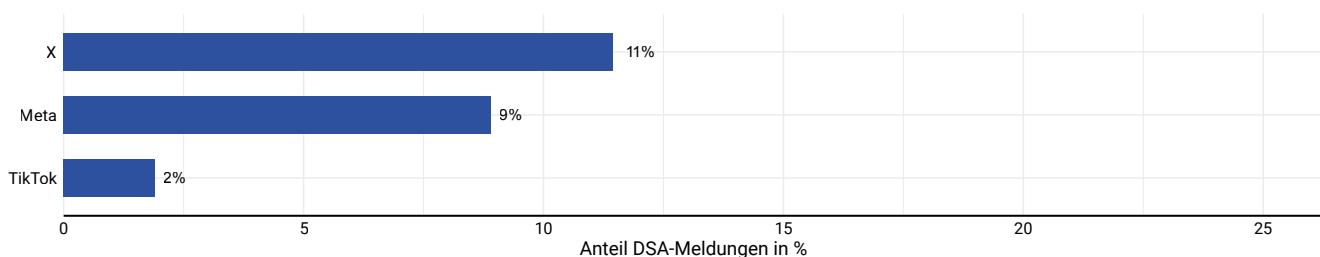
Anschließend folgte eine zufällige Zuteilung zum Meldeweg einer Plattform und die Befragten durchliefen entweder den Meldeweg von Meta (Facebook oder Instagram), TikTok oder X. Wir wählten diese vier Plattformen, um mehrere Altersschichten abzudecken und da sich das Interface zur DSA-Meldung bei allen vier Plattformen gravierend unterscheidet. Während bei Facebook bzw. Instagram Paragraphen nach Strafgesetzbuch aufgelistet werden, fordert TikTok eine Nennung des Gesetzes und eine eigenständige Meldungserklärung. X bietet im Vergleich zu den Meta-Plattformen und TikTok zwei im ersten Meldeschritt voneinander getrennte Meldewege für illegale Inhalte und solche, die gegen die Community-Richtlinien verstoßen. Die folgenden Aussagen über die Funktionsweise der Meldewege können deshalb nur für die behandelten Plattformen getätigt werden.

Im ersten Schritt interessierte uns, ob die Teilnehmer*innen die Inhalte korrekt zuordneten. Die Zuordnung dieser Inhalte bewerteten wir als

inhaltlich korrekt, wenn sie in den ersten beiden Meldeschritten in die entsprechenden Kategorien nach den Community-Richtlinien der Plattformen (bspw. Gewaltverherrlichung, Hassrede, Beleidigungen) oder nach DSA eingeteilt wurden.²⁵ Wählten die Teilnehmer*innen andere Kategorien, die von der Kernbotschaft des Inhalts stark abwichen (bspw. Betrug), so wurde der Meldevorgang abgebrochen, um ein besseres Verständnis dafür zu erlangen, ob die Kategorisierung selbst eine Hürde darstellt.

Insgesamt knapp drei Viertel der Befragten (73 %) ordneten die Inhalte der inhaltlich korrekten Kategorie zu. Zwischen den Plattformen zeigten sich dabei Unterschiede: Bei Meta lag die korrekte Zuordnung bei 61%, bei TikTok bei 75 % und bei X sogar bei 82 %. Dies könnte damit zusammenhängen, dass die Befragten im weiteren Verlauf auch unterschiedlich häufig den DSA-Meldeweg wählten. So nutzten bei Meta im Schnitt rund 9 % der Teilnehmer*innen den DSA-Meldeweg, bei TikTok waren es lediglich 2 %, während bei X mit 11 % ein vergleichsweise hoher Anteil der Befragten auf diesen Weg zurückgriff.

6 Anteil der Meldungen nach DSA



Eine mögliche Erklärung dafür könnte sein, dass bei X gleich zu Beginn des Meldeprozesses eine klare Entscheidung zwischen einer Meldung nach Community-Richtlinien und einer Meldung als rechtswidriger Inhalt (DSA) getroffen werden muss. Auf anderen Plattformen hingegen ist der DSA-Meldeweg lediglich eine von vielen Unterkategorien, was seine Sichtbarkeit und Nutzung erschwert.

Darauf deuten auch die Ergebnisse der qualitativen Studie hin. Die meisten Befragten entschieden sich intuitiv für eine der ersten Kategorien aus den Community-Richtlinien – ohne überhaupt bis zur Kategorie „illegale Inhalte“ zu kommen. Das lag häufig daran, dass die Kategorien schon den konkreten Fall abdeckte. Ein Teilnehmer beschrieb seine Gründe, warum er einen Hasskommentar Community-Richtlinien melden würde, folgendermaßen:

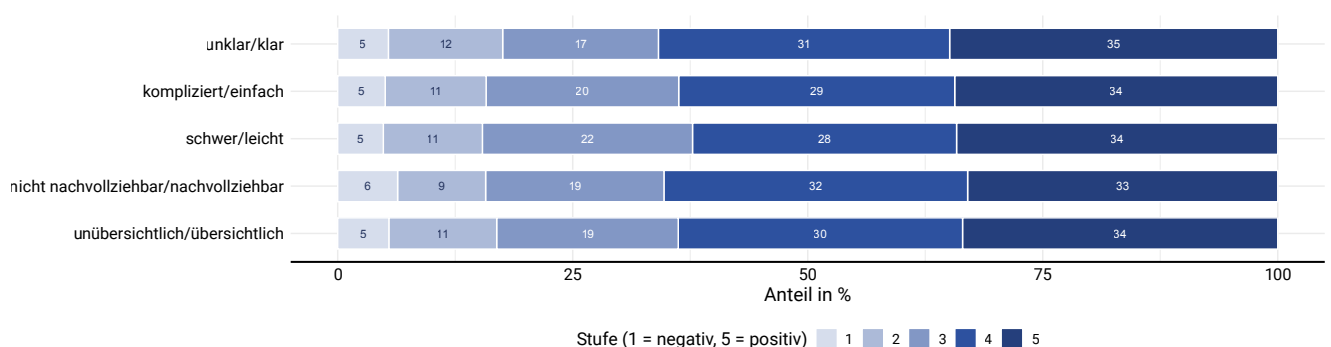
Weil wenn da irgendwas mit Hassrede in dem Kommentar wäre, dann hätte ich das schon davor gemeldet und wäre nicht hier

drauf [DSA-Meldung] gegangen. Also den Punkt hätte ich nur dann ausgewählt, wenn ich mich gar nicht mehr in diesen Themen [Community-Richtlinien] wiedergefunden hätte.
(Salim, 35)

Ein weiterer Grund, der verhindert, dass Nutzer*innen nach DSA melden, ist ihre Unsicherheit bezüglich der Einstufung der Inhalte als illegal. Häufig entschieden sie sich dann dagegen, weil damit weniger Unsicherheit einherging oder eine schnelle Entscheidung getroffen werden musste: „Ich wüsste jetzt aber nicht explizit, worin der Unterschied besteht, weil für mich der Punkt ist, ab wann ist etwas tatsächlich illegal? Das könnte ich jetzt nicht ad hoc einschätzen“ (Laura, 43).

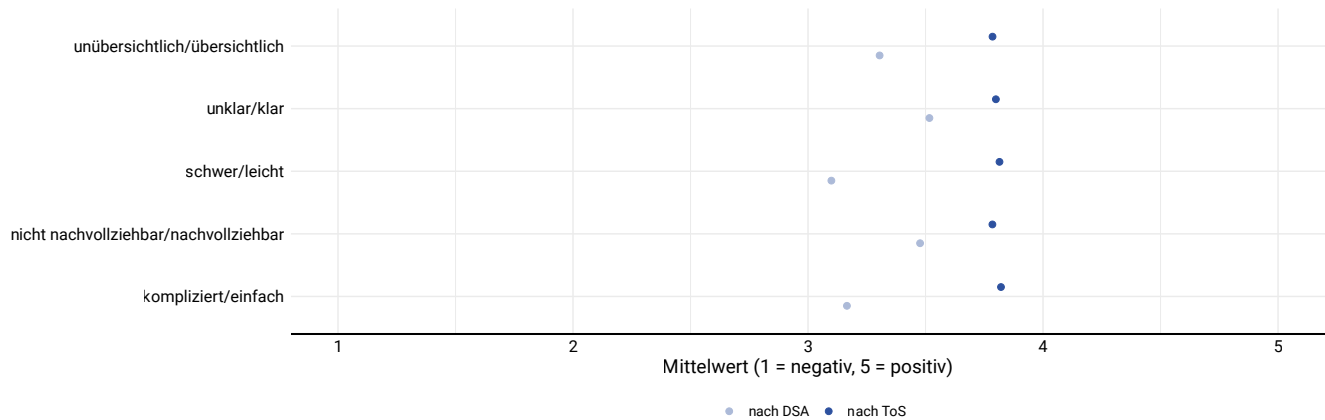
Nach Beendigung bewerteten die Befragten den Meldevorgang anhand verschiedener Gegensatzpaare. Insgesamt überwog eine (eher) positive Einschätzung: 66 % empfanden die Meldewege als (eher) klar, 65 % als nachvollziehbar und 64 % als (eher) übersichtlich. 63 % beschrieben den Vorgang als (eher) einfach und 62 % fanden ihn (eher) leicht. Damit zeigt sich, dass eine Mehrheit der Nutzer*innen die Meldeprozesse grundsätzlich als verständlich und handhabbar wahrnahm, auch wenn zugleich ein erheblicher Anteil neutral oder kritisch urteilte. Unterschiede zwischen den Plattformen waren marginal, wobei der Meldeprozess von X über alle Kategorien hinweg etwas kritischer wahrgenommen wurde. Das ist aufgrund der höheren Bereitschaft, auf X nach DSA zu melden, überraschend, könnte aber darauf hindeuten, dass Nutzer*innen X als weniger vertrauenswürdige Plattform einstufen. Motivationsgründe für die Bewertung des Meldeprozesses wurden nicht gemessen.

7 Bewertung des Meldevorgangs



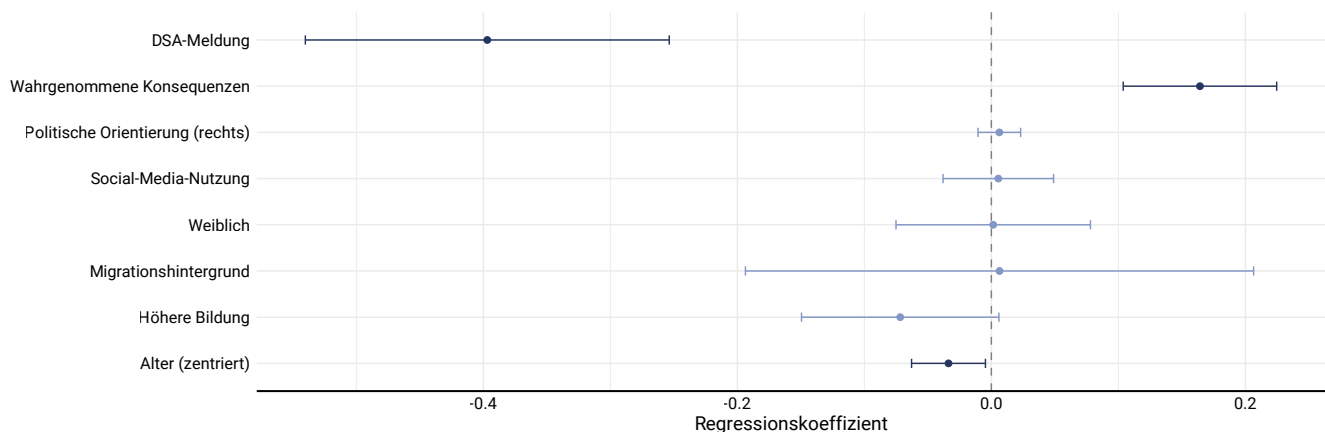
Deutliche Unterschiede in der Bewertung des Meldeprozesses zeigen sich hinsichtlich der Meldewege. Der Meldeprozess via DSA wurde als deutlich schwerer, unübersichtlicher und komplizierter sowie etwas weniger klar und weniger nachvollziehbar wahrgenommen als die Meldung via Community-Richtlinien.

8 Bewertung des Meldevorgangs nach Meldeweg



Die Regressionsanalyse zur Bewertung des Meldevorgangs bestätigt diesen Befund statistisch: Eine Meldung nach DSA ist mit einer deutlich kritischeren Bewertung der Meldewege verbunden. Überdies bewerteten jüngere Personen die Meldewege etwas positiver ebenso diejenigen, die stärkere Konsequenzen problematischer Inhalte auf die Gesellschaft wahrnehmen. Bildungsniveau, Migrationshintergrund, Geschlecht, Social-Media-Nutzung und politische Orientierung zeigten dagegen keine signifikanten Effekte auf die Bewertung des Meldevorgangs.

9 Prädiktoren für die Bewertung der Meldung



Ein Regressionskoeffizient > 1 zeigt einen positiven Zusammenhang mit der Bewertung einer Meldung, ein Wert < 1 einen negativen.

Dies deckt sich auch mit den qualitativen Befunden. Rückblickend beschrieben viele Teilnehmer*innen den DSA-Meldeprozess als insgesamt machbar und teilweise sogar unkomplizierter als angenommen. Vor allem Personen, die ohnehin häufiger Meldungen erstatteten, empfanden die Schritte als nachvollziehbar und gut zu bewältigen. Sie beschrieben, dass es „im Endeffekt auch nicht so viel“ ist (Elif, 29). Die genauen Angaben innerhalb der Meldung hinterließen bei einigen Teilnehmer*innen den Eindruck, eine differenzierte Meldung abzugeben,

die dann auch ernst genommen wird und die wahrscheinlicher zum Erfolg führt. Ein junger Teilnehmer fasste zusammen, dass er den DSA-Meldeweg „dafür, wofür er ist, auch einfach [findet]. Weil es ist ja wirklich gegen illegale Sachen.“ Dennoch ergänzte er: „Aber trotzdem ... für unsere Generation, da würden sich viele, glaube ich, nicht die Mühe machen, jetzt dazu noch was zu schreiben. [...] Also es ist so ein Zwischending“ (Jonas, 18).

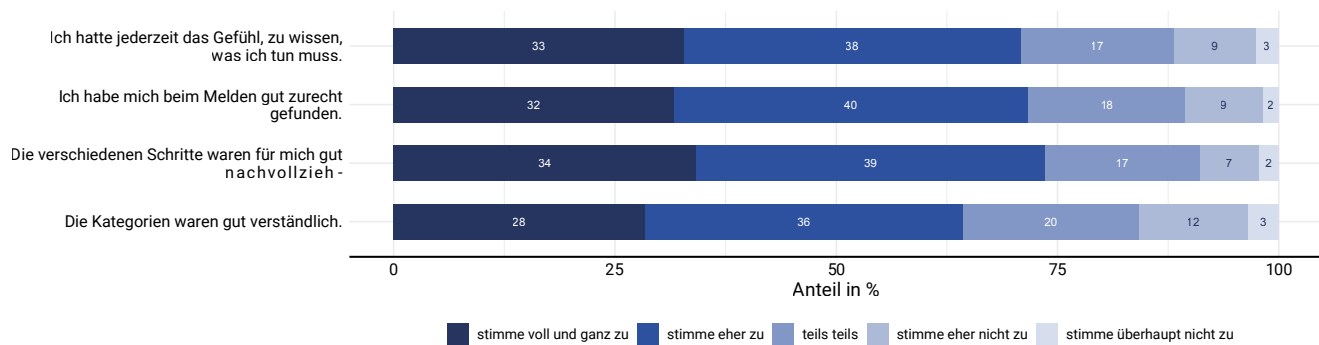
Tatsächlich wurde die Meldung von Verstößen gegen die Community-Richtlinien als deutlich unkomplizierter und einfacher angesehen als die Meldung illegaler Inhalte. Viel häufiger wurde der Meldeweg nach DSA im Vergleich zu den Community-Richtlinien als aufwändig und abschreckend wahrgenommen, weshalb Befragte – wenn vor der Wahl stehend – eher nach Community-Richtlinien melden würden: „Wenn ich mir nicht ganz sicher bin, würde ich die eher erst über diesen normalen Weg melden“ (Jonas, 18). Außerdem könnte die Angst vor individuellen rechtlichen Konsequenzen von der Meldung abhalten und dazu führen, dass zu wenig illegale Inhalte gemeldet werden. Dabei spielte es auch eine Rolle, dass sich viele Nutzer*innen nicht wohl damit fühlten, persönliche Daten anzugeben oder zu hinterlegen. Eine Teilnehmerin beschrieb beispielsweise, wie die Abgabe ihrer Unterschrift abschreckend auf sie wirkte: „Da hätte ich jetzt gleich schon Angst, irgendwas falsch zu machen. Nicht, dass ich am Ende was irgendwie Illegales tue“ (Elif, 29).

Es ergibt sich das Paradoxon, dass durch die „Verrechtlichung“ des Melderechts der Nutzer*innen, die juristische Expertise voraussetzt, das Verfahren möglicherweise zu komplex und abschreckend für Nutzer*innen geworden ist. Dies führte einerseits zu Abbrüchen, andererseits zu weniger Bereitschaft, in Zukunft zu melden:

Ich glaube, das wird tendenziell Leute – oder mich persönlich – dann auch eher dazu führen, das Melden hier nicht vorzunehmen oder gegebenenfalls nicht so häufig, weil das ja mit einem sehr, sehr hohen zeitlichen Aufwand verbunden ist, wenn man das gewissenhaft machen möchte. (Dennis, 31)

An die Bewertung des Meldevorgangs schlossen sich konkrete Aussagen zu den Erfahrungen im Prozess an. Auch hier zeigt sich insgesamt eine überwiegend positive Einschätzung: 74 % der Befragten fanden die einzelnen Schritte gut nachvollziehbar, 72 % gaben an, sich beim Melden gut zurechtgefunden zu haben, und 71 % hatten jederzeit das Gefühl, zu wissen, was zu tun ist. Etwas zurückhaltender fiel die Zustimmung zur Verständlichkeit der Kategorien aus, die von 64 % positiv bewertet wurde. Insgesamt bestätigen diese Ergebnisse, dass die Mehrheit der Nutzer*innen den Meldevorgang als handhabbar und verständlich wahrnahm.

10 Aussagen zum Meldevorgang



In unserer Studie haben 6 % der Befragten den Meldevorgang abgebrochen. Dabei zeigten sich deutliche Unterschiede zwischen den Plattformen: Bei Meta kam es nur selten zu Abbrüchen (3 %), bei TikTok lag der Anteil bereits bei 6 %, während er bei X mit 9 % am höchsten ausfiel. Dies deutet darauf hin, dass die Ausgestaltung und Handhabung der Meldewege je nach Plattform unterschiedlich stark zu Abbrüchen führen. Noch gewichtiger als die Ausgestaltung der Plattformen scheint die Art des Meldeweges zu sein: Über ein Viertel der Meldungen nach DSA (27 %) wurde abgebrochen, während die Meldungen nach den Community-Richtlinien in lediglich 4 % der Fälle abgebrochen wurden.

Als Gründe für den Abbruch gab rund ein Viertel der 117 Abbrecher*innen an, „mit der Auswahl überfordert“ gewesen zu sein. Ein weiteres Drittel gibt an, nicht mehr gewusst zu haben, „wie ich weitermachen kann“ (32 %).

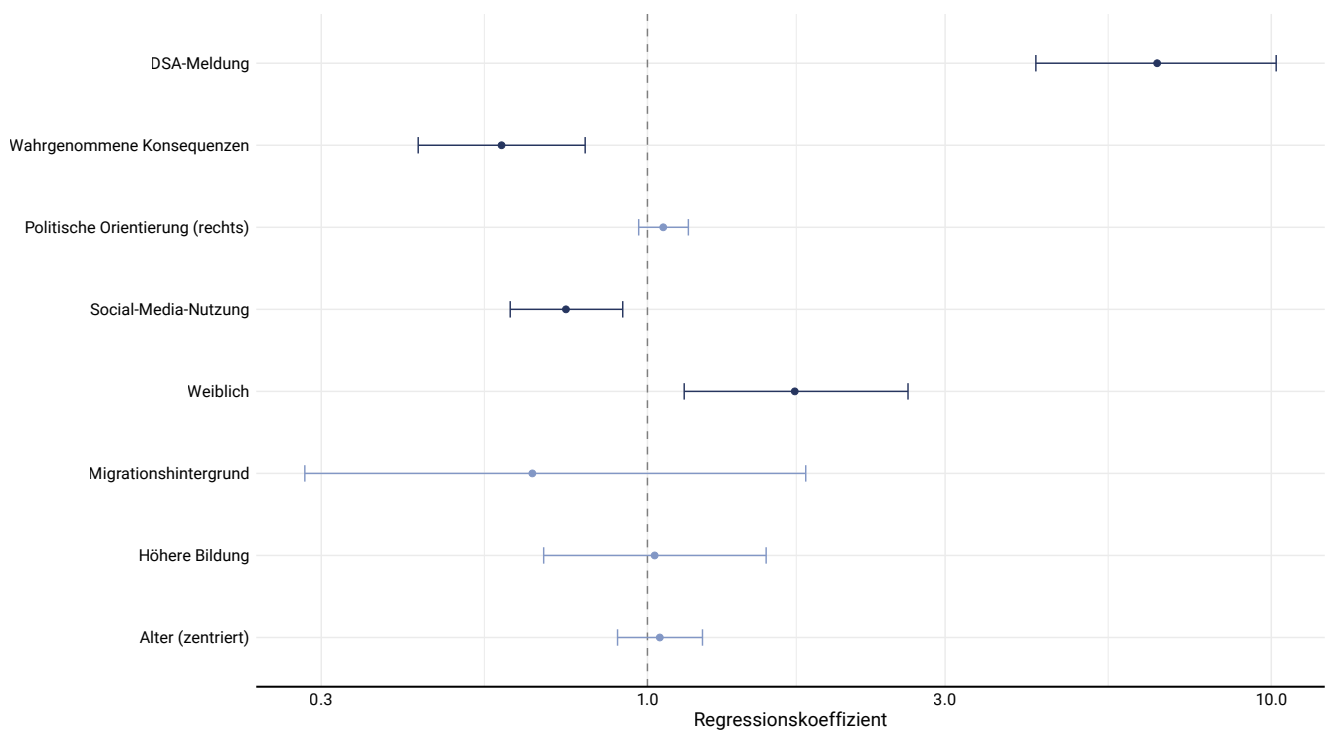
Dies zeigt sich auch in den qualitativen Befunden. Insbesondere die Auswahl der unterschiedlichen Kategorien, welche erst im späteren Verlauf des Meldeprozesses erscheint, wurde von vielen Befragten als zu komplex und als potenzieller Grund zum Abbrechen der Meldung wahrgenommen. Viele Teilnehmer*innen empfanden die Kategorisierung als „überfordernd, um ehrlich zu sein. [...] das ist halt irgendwie sehr überladen. [...] und dann würde man sich denken, ach komm. Oder im Zweifel jetzt hier ‚Sonstiges‘ anklicken“ (Leon, 26). Die Komplexität des Meldeprozesses fiel vielen Teilnehmer*innen vor allem an dieser Stelle negativ auf. Insbesondere die Notwendigkeit des Detailgrades wurde angezweifelt und kritisiert: „Kann ja nicht ernst gemeint sein. [...] was macht das für einen Unterschied jetzt? Also, warum muss man das so detailliert aufdröseln, um das Leuten so schwer zu machen, solche Sachen zu melden? Also, kann ich nicht nachvollziehen“ (Florian, 38).

Statt an dieser Stelle abzubrechen, gaben sich einige Befragte aber auch damit zufrieden, eine der Kategorien zu wählen, die am besten passt – Hauptsache die Inhalte werden gemeldet: „Und ich denke mir

immer so, es ist ja vielleicht auch egal, was es ist, was man anklickt. Hauptsache, man klickt es im Zweifel an" (Leon, 26).

Mittels eines Regressionsmodells haben wir die Faktoren untersucht, die den Abbruch des Vorgangs begünstigten. Demnach beendeten Frauen den Prozess häufiger vorzeitig, während eine intensive Nutzung sozialer Medien sowie die Wahrnehmung negativer Konsequenzen problematischer Inhalte eher dazu beitrugen, den Meldevorgang durchzuhalten. Besonders auffällig war, dass Abbrüche deutlich wahrscheinlicher wurden, wenn eine Meldung nach dem DSA erfolgte. Alter, Bildung, Migrationshintergrund und politische Orientierung spielten keine nennenswerte Rolle.

11 Faktoren für den Abbruch des Meldevorgangs



Ein Regressionskoeffizient > 1 zeigt einen positiven Zusammenhang mit dem Abbruch einer Meldung, ein Wert < 1 einen negativen.

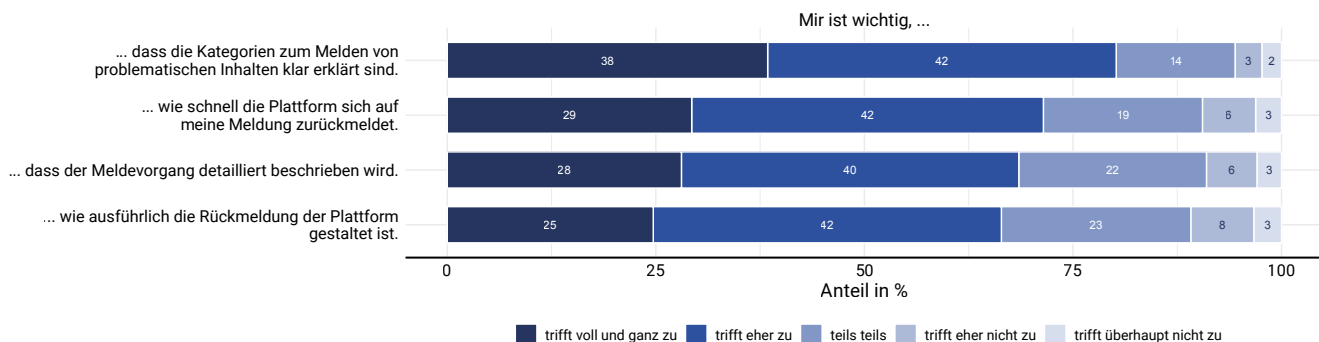
5.5 Zukünftiges Meldeverhalten

Abschließend interessierte uns, welche Faktoren für die Teilnehmer*innen ausschlaggebend wären, in Zukunft tatsächlich problematische Inhalte zu melden. Grundsätzlich sehen sich die Befragten auch zukünftig in der Verantwortung, problematische Inhalte zu melden – jedoch vor allem als Hinweisgeber*innen. Sie erwarten, dass die Plattformen die weitere Prüfung, Bewertung und Einordnung übernehmen und betonen, dass vor allem die Frage nach dem tatsächlichen Delikt, „gleich auf der nächsthöheren Ebene stattfinden [sollte]. Von Leuten

wie mir, also vom Ottonormalverbraucher, sollte nur der Impuls kommen, hier ist irgendetwas nicht in Ordnung" (Lucas, 43).

Die Befunde der quantitativen Befragung verweisen auf Faktoren, die künftiges Engagement fördern können. Vor allem Transparenz und Verständlichkeit des Prozesses spielen tragende Rollen. Am wichtigsten war, dass die Kategorien zum Melden klar erklärt sind – dies hielten 80 % für entscheidend. Ebenfalls betonten viele Befragte, dass es darauf ankomme, wie schnell die Plattform auf eine Meldung reagiert (71 %). Zwei weitere Faktoren wurden von einer Mehrheit ebenfalls als wichtig eingestuft: dass der Meldevorgang detailliert beschrieben wird (68 %) und wie ausführlich die Rückmeldung der Plattform ausfällt (67 %). Insgesamt wird deutlich, dass Nutzer*innen künftig vor allem dann bereit sind, Inhalte zu melden, wenn der Vorgang verständlich, transparent und mit einer klaren Rückmeldung verbunden ist.

12 Potenzial für Verbesserung der Meldewege



Verbesserungswürdig erscheint auch die Aufmachung des Meldeprozesses, die einigen Befragten nicht intuitiv, sondern überladen oder halbherzig beschrieben. Neben der Aufmachung erweckt auch die Struktur den Eindruck, dass die Plattformen sich wenig Mühe geben, denn:

Dieses ganze Ding sieht ein bisschen old school aus. Also, das ist vor zehn Jahren das letzte Mal aktualisiert wurde. Und dann ist das halt auch so ein bisschen halbherzig gemacht, weil hier oben ist dann Englisch. Also, da kommen irgendwelche Verlinkungen und unten soll ich dann entscheiden, welche Art von Rechtsverletzung das jetzt ist. Weiß ich, ob ich, also ob das jetzt sinnvoll ist. (Markus, 28)

Eine detaillierte Kategorisierung oder rechtliche Einschätzung durch die Nutzer*innen lehnten die meisten ab; sie fühlten sich dafür nicht zuständig oder ausreichend kompetent. Zu viel Verantwortung bei den Nutzer*innen wurde eher als abschreckend empfunden und konnte nicht nachvollzogen werden:

Ich bin der ganz normale User. [...] Mir ist durchaus auch gar nicht so ganz genau klar, warum das überhaupt so im Detail bewertet werden muss. [...] Ja, ich denke, die Plattform selber, die sollte ja eigentlich auch wollen, dass da keine rechtswidrigen Inhalte auf der Plattform sind. (Markus, 28)

-
- 25 Alle den Teilnehmer*innen angezeigten Inhalte wurden als strafrechtlich relevant eingestuft. Illegale Inhalte sollten laut DSA idealerweise über den DSA-Meldeweg statt über Community-Richtlinien gemeldet werden. Die meisten Teilnehmer*innen entschieden sich in diesem Fall – äquivalent zur qualitativen Befragung – eher für den Weg über die Community-Richtlinien. Motivationsgründe hierfür wurden nicht gemessen.

6. Limitationen



Um die Wahrnehmung und Hürden der Meldewege zu untersuchen, ist es unabdinglich, kontrollierte Bedingungen zu schaffen. Die Untersuchung des Meldeverhaltens unter kontrollierten Bedingungen birgt gleichsam einige Limitationen, die bei der Interpretation der Befunde berücksichtigt werden müssen. Unser Nachbau der Plattformmeldestrukturen von TikTok, Facebook bzw. Instagram und X war realitätsnah konzipiert. Dennoch mussten einige rechtliche Schritte – wie zum Beispiel das Nennen des Klarnamens und eine Einwilligung vor dem Absenden – in unserer Simulation ausgelassen werden, um eine Anonymisierung der Umfrage nach Datenschutzgrundverordnung (DSGVO) zu ermöglichen. Das reduziert die externe Validität der Studie, da es eine in der qualitativen Untersuchung destillierte Hürde der Meldung nach DSA – die Angabe persönlicher Daten – aus dem Meldeweg entfernt. Zudem wurden nur Meldewege von vier VLOPSEs rekonstruiert. So können zwar pro gewählte Plattform individuelle Aussagen über die Funktionsweise des Meldeweges getroffen werden, aber nicht über andere VLOPSEs. Hier können bei ähnlichem Aufbau nur Designempfehlungen abgeleitet werden. Darüber hinaus haben wir uns lediglich auf die Meldewege in Deutschland konzentriert, wobei die Meldewege für illegale Inhalte in unterschiedlichen EU-Staaten anders gestaltet sind; zum Beispiel gibt es in anderen EU-Ländern keine Kategorisierung von Inhalten, sondern von Anfang an Freitext.²⁶ Schließlich werden in dieser Studie nur Meldungen von Inhalten, nicht von auf Plattformen geschalteter Werbung, betrachtet. Die Beispiele, die wir Befragten zeigten, beruhen zudem mehr auf Nutzer*innenkommentaren als auf Inhalten, die von einzelnen Profilen veröffentlicht wurden – wobei bei fast allen Plattformen die Meldung von Inhalten leicht anders aufgebaut ist als die Meldung von Kommentaren.

26 Siehe Marie-Therese Sekwenz, Ben Wagner und 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. Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency (FAccT '25), 532–546. <https://doi.org/10.1145/3715275.3732036>.

7. Handlungsoptionen für die Politik: DSA-Wirkung stärken, Meldewege klären



Vereinheitlichung der Meldewege

Der DSA sieht vor, dass die Wege, über die Nutzer*innen potenziell rechtswidrige oder gegen die Community-Richtlinien verstoßende Beiträge auf Plattformen melden können, nutzerfreundlich und leicht zugänglich sind. Bis jetzt gibt es auf den Plattformen keine einheitlichen Meldeverfahren – und je nach Plattform auch Unterschiede, wann, wie und wo ein Inhalt oder ein Kommentar gemeldet werden soll. Das ist nicht nur unübersichtlich für Nutzer*innen, es kann auch dazu führen, dass Nutzer*innen bei Plattformen mit leichter verständlichen und auffindbaren Meldewegen im Vergleich zu anderen Plattformen mehr Inhalte melden. Wir empfehlen deshalb eine Vereinheitlichung der Meldewege für DSA-relevante Inhalte bei VLOPSEs. Zusätzlich muss geklärt werden, was unter einem „leicht zugängliche[n] und benutzerfreundliche[n]“ Meldemechanismus nach Artikel 16 Absatz 1 zu verstehen ist – gerade in Anbetracht der Tatsache, dass unsere Studie zeigt, dass das Bildungsniveau einen positiven Einfluss auf das bisherige Meldeverhalten hat. Dafür könnte die EU-Kommission die Entwicklung von Industriestandards für VLOPSEs fördern, die eine Vereinheitlichung der Meldewege anvisieren.

Durchsetzung der Community-Richtlinien

Unsere Studienergebnisse zeigen, dass die Meldung von Verstößen gegen die Community-Richtlinien als unkomplizierter und einfacher angesehen werden als die Meldung illegaler Inhalte. Befragte der qualitativen Studie würden – wenn vor der Wahl stehend – auch eher nach Community-Richtlinien melden. Wir fordern deshalb von VLOPSEs eine klare Durchsetzung der plattforminternen Community-Richtlinien und eine konsequentere Aufsicht der Umsetzung.

Öffentliche Sensibilisierung

Unsere Daten zeigen, dass häufig nicht gemeldet wird, da keine Plattformreaktion erwartet oder eine Meldung als nutzlos erachtet wird. Zudem sind Nutzer*innen vorsichtig, Inhalte als potenziell illegal zu deklarieren. Das Meldeverfahren nach Artikel 16 DSA in seiner jetzigen Form hängt aber stark von der Motivation der Nutzer*innen ab, schädliche Inhalte zu melden, sowie von der Notwendigkeit rechtlicher Kenntnisse, um die Meldemechanismen adäquat nutzen zu können. Die EU-Kommission und die nationalen DSCs sollten in groß angelegte öffentliche Sensibilisierungskampagnen investieren, um Nutzer*innen darüber aufzuklären, dass eine Meldung von illegalen Inhalten keine rechtlichen Konsequenzen für die meldende Person hat. Diese Kampagnen sollten die Öffentlichkeit auch darüber informieren, wie Bür-

ger*innen durch ihre Meldung von ihren Rechten Gebrauch machen und Plattformen zur Verantwortung ziehen können.

Design der Meldewege

Bei der Entwicklung und Aktualisierung von Meldewegen sollten in allen Phasen eine nach Artikel 16 nutzerfreundliche und nach Artikel 28 kinderfreundliche Linse verwendet werden. Die Ergebnisse der qualitativen Studie zeigen auf, dass insbesondere die Auswahl der unterschiedlichen Kategorien, welche erst im späteren Verlauf des Meldeprozesses erscheint, als überfordernd und als potenzieller Grund für einen Abbruch des Meldevorgangs wahrgenommen werden kann. Eine rechtliche Einschätzung lehnen Nutzer*innen eher ab. Auch die quantitativ hohe Abbruchrate bei Meldungen über DSA statt über Community-Richtlinien zeigt eine Frustration mit dem Design. Die Navigation durch Schnittstellen zu rechtlichen Prozessen, so, wie sie bis dato auf VLOPSEs konzipiert sind, setzt ein vorheriges Verständnis des Gesetzes voraus, um die eigene Meldeentscheidungen kompetent treffen zu können. So wirkt es fast, als hätte die Einhaltung des Gesetzes Vorrang vor einer nutzerfreundlichen Gestaltung der Meldewege, wobei die rechtliche Sprache des Meldeprozesses einen starken Einfluss auf die Meldemotivation haben kann. Es sollte aber nicht Aufgabe der Nutzer*innen sein, rechtliche Entscheidungen als Vorfilter für die Inhaltsmoderation zu treffen.

Wir plädieren deshalb dafür, dass die EU-Kommission eine Vereinfachung des Prozesses anstrebt, die die Erwartung von Rechtskenntnissen aufhebt. Basierend auf unseren Ergebnissen würden wir VLOPSEs eine klare Abgrenzung von Meldungen nach Community-Richtlinien und nach DSA empfehlen, zum Beispiel durch das direkte Erfragen der Meldeabsicht, ohne Befragte mit juristischer Sprache einzuschüchtern (z.B. durch eine Differenzierung zwischen „Ich möchte einen Inhalt melden, weil er gegen die Community-Richtlinien der Plattform verstößt“ und „Ich möchte einen Inhalt melden, weil ich ihn als illegal einschätze“).

Darüber hinaus sollten abschreckende Hinweise, wie dass eine Meldung von illegalen Inhalten rechtliche Konsequenzen haben könne, abgemildert werden. Laut Artikel 16 DSA müssen Nutzer*innen eine „Erklärung abgeben, warum sie den Inhalt für illegal halten“ (Art. 16(2,c)), zudem muss diese Erklärung „nach bestem Wissen und Gewissen,“ abgegeben werden (Art. 16(2,d)). Eine detaillierte Kategorisierung oder rechtliche Einschätzung durch die Nutzer*innen, wie von den Plattformen gefordert, ist dadurch jedoch nicht legitimiert. Hier empfehlen wir die Möglichkeit, per Drop-Down-Menü rechtliche Gründe für die Meldung zur Auswahl zu stellen, inklusive eines zusätzlichen optionalen Freifelds für weitere Erläuterungen.

Support by Design

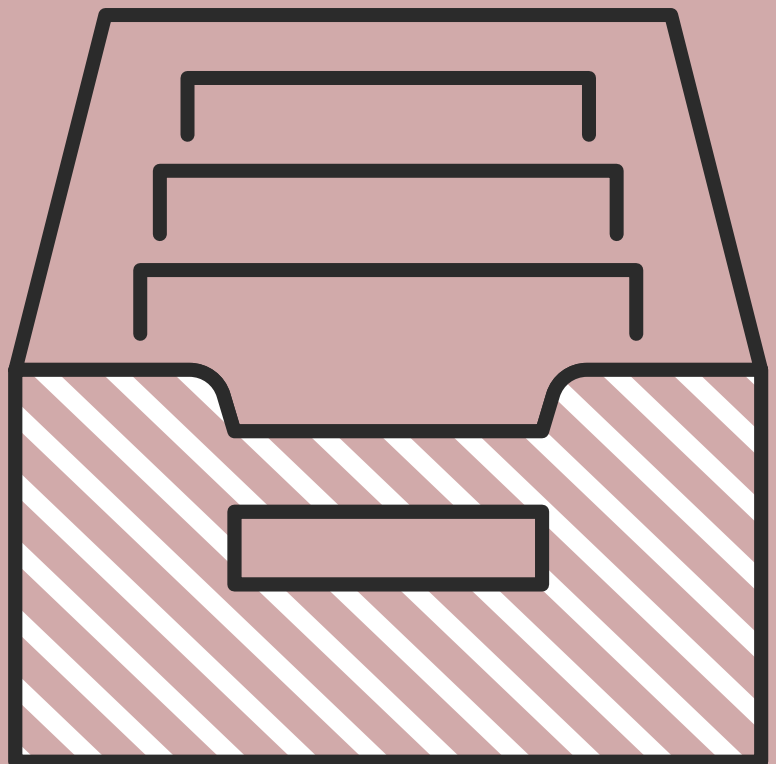
Durch unsere qualitative Umfrage konnten wir „Schmerzpunkte“ bei den Befragten identifizieren und herausfinden, inwieweit Verständnisprobleme oder unklare Auswahlmöglichkeiten durch gezielte Unterstützung geklärt werden können. Wir empfehlen eine gezielte, ins Design der Meldewege integrierte Unterstützung durch Expert*innen, die die Schwierigkeiten der Nutzer*innen am besten kompensieren kann. Das kann ein Verweis auf Rechtsschutzorganisationen, Meldestellen, aber auch außergerichtliche Streitbeilegungsstellen sein.

8. Fazit – Zwischen Regulierung und Realität



Die Ergebnisse unserer Studie zeigen, dass der Meldeprozess nach DSA bei den ausgewählten Plattformen als deutlich schwieriger, unübersichtlicher und komplizierter sowie etwas weniger klar und nachvollziehbar wahrgenommen wird als die Meldung nach Community-Richtlinien. Erheblich häufiger wurde der Meldeweg nach DSA im Vergleich zu den Community-Richtlinien als aufwändig und abschreckend bewertet, weshalb die Befragten — wenn vor der Wahl stehend — eher nach Community-Richtlinien melden würden. Zudem machen die Meldemechanismen der ausgewählten VLOPSEs im Durchschnitt nicht klar ersichtlich, wann die Meldungen von Nutzer*innen nach DSA und wann nach Community-Richtlinien behandelt werden. Außerdem könnte die Angst vor individuellen rechtlichen Konsequenzen von der Meldung abhalten und zusätzlich dazu führen, dass zu wenig illegale Inhalte gemeldet werden. Es könnte ebenso dazu führen, dass meldebereite Nutzer*innen dazu gedrängt werden, Inhalte über weitgehend unregulierte Community-Richtlinien zu melden, statt den aufwändigeren, regulierten Weg nach DSA zu wählen. Außerdem kann sich die Gestaltung der Schnittstellen auf den Meldeprozess, die Nutzererfahrung und die Erwartung an die Nutzer*innen auswirken. Der DSA verspricht, durch nutzerfreundliche Meldemechanismen sichere Online-Räume zu schaffen. Doch die Komplexität der rechtlichen Anforderungen und die undurchsichtigen Praktiken der Plattformen könnten die Bemühungen, Nutzer*innen zur Meldung zu ermutigen, zunichtemachen. Es bleibt fraglich, ob Meldesysteme auch Nicht-Expert*innen wirksam zur Meldung befähigen können. Meldemechanismen wurden als Instrument für Nutzer*innen geschaffen — sie sind aber nicht immer so gestaltet, dass sie den Bedürfnissen der Nutzer*innen gerecht werden.

9. Literaturverzeichnis



Bundesnetzagentur (2025). Bundesnetzagentur testet Abläufe und Maßnahmen bei Verstößen gegen den Digital Services Act. Bundesnetzagentur (Digital Services Coordinator Germany). Letzter Zugriff 12. August 2025 unter 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. Letzter Zugriff 18. Juli 2025 unter <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. Letzter Zugriff 02. Juli 2025 unter <https://www.bundestag.de/dokumente/textarchiv/2024/kw12-de-digitale-dienste-994508>.

Steliyana Doseva, Fay Carathanassis, Hannah Schmid-Petri und Dirk Heckmann (2024). Beleidigungen auf Social Media: Wie betroffen sind Nutzerinnen und Nutzer auf Social Media und wie wehren sie sich? Bayerisches Forschungsinstitut für Digitale Transformation. Letzter Zugriff 20. August 2025 unter <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. Letzter Zugriff 15. Juli 2025 unter <https://www.reuters.com/article/us-facebook-germany-fine-idUSKCN1TX1IC/>.

Europäische Kommission (2024). Gesetz über digitale Dienste: Kommission leitet förmliches Verfahren gegen Facebook und Instagram ein. EU-Repräsentanz Deutschland. Letzter Zugriff 05. August 2025 unter https://germany.representation.ec.europa.eu/news/gesetz-uber-digitale-dienste-kommission-leitet-formliches-verfahren-gegen-facebook-und-instagram-ein-2024-04-30_de.

Europäische Kommission (2025). DSA: Very large online platforms and search engines. Europäische Kommission. Letzter Zugriff 09. Juli 2025 unter <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. Letzter Zugriff 19. August 2025 unter <https://verfassungsblog.de/follow-me-to-unregulated-waters/>.

Luise Koch, Angelina Voggenreiter und Janina Steinert (2025). Angegriffen & alleingelassen – Wie sich digitale Gewalt auf politisches Engagement auswirkt. Ein Lagebild. Technische Universität München in Kooperation mit HateAid. Letzter Zugriff 06. Juli 2025 unter <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. Letzter Zugriff 01. Juli 2025 unter <https://maldita.es/investigaciones/20250619/denunciamos-meta-publicaciones-fraudulentas-facebook-dsa/>.

Das NETTZ, Gesellschaft für Medienpädagogik und Kommunikationskultur, HateAid und Neue deutsche Medienmacher*innen (Hrsg.) (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 und Neue deutsche Medienmacher*innen. Letzter Zugriff 21. August 2025 unter https://tonesift.org/wp-content/uploads/2024/02/Studie_Lauter-Hass-leiser-Rueckzug.pdf.

Friederike Quint, Yannis Theocharis, Spyros Kosmidis und 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 und 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, und 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 und 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. *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency (FAccT '25)*, 532–546. <https://doi.org/10.1145/3715275.3732036>.
- Yunhee Shim und 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. Letzter Zugriff 07. August 2025 unter <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, und 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 und Ewan Casandjian (2025). Flagged and Ignored: Testing X's Response to EU Sanction Violations. *Alliance4Europe*. Letzter Zugriff 14. August 2025 unter <https://alliance4europe.eu/flagged-and-ignored>.
- Anna-Lena von Hodenberg und Josephine Ballon (2025). 1 Jahr DSA: Sicherheit im digitalen Raum: eine Bilanz — und was jetzt nötig ist. *HateAid*. Letzter Zugriff 04. Juli 2025 unter 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 und Jatinder Singh (2020). Regulating transparency? Facebook, Twitter and the German Network Enforcement Act. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20)*. 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 und Shagun Jhaver (2024). Cleaning Up the Streets: Understanding Motivations, Mental Models, and Concerns of Users Flagging Social Media Content. *arXiv*. Letzter Zugriff 11. Juli 2025 unter <https://arxiv.org/pdf/2309.06688>.

Impressum

**Zwischen Klick und Konsequenz:
Eine Evaluation der Melde-
verfahren von Plattformen nach
dem Digital Services Act**

Herausgeber

Das NETTZ gGmbH
c/o bUm Berlin
Paul-Linke-Ufer 21
10999 Berlin
info@das-nettz.de
www.das-nettz.de

Autor*innen

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

Übersetzung

Christoph von Pohl (Deutsch-Englisch)

Lektorat

Franzi Prost

Titelbild | Design

Delia Keller | Gestaltung Berlin

Zitationshinweis

Das NETTZ (2025). Zwischen Klick und Konsequenz:
Eine Evaluation der Meldeverfahren von Plattformen
nach dem Digital Services Act. Berlin.

Die vorliegende Publikation wurde finanziell unter-
stützt durch Reset Tech. Die Veröffentlichung stellt
keine Meinungsäußerung von Reset Tech dar. Für
inhaltliche Aussagen tragen die Autor*innen die
Verantwortung.

Mit Unterstützung von

Reset • Tech

Geschäftsführung: Nadine Brömme, Hanna Gleiß
Registergericht: Amtsgericht Berlin Charlottenburg,
HRB 242638 B, Geschäftssitz: Berlin

