



Pressemitteilung HateAid gGmbH

HateAid wird Trusted Flagger

Gute Nachrichten für Betroffene von digitaler Gewalt: Am heutigen Tag wird die gemeinnützige Organisation HateAid, sowie zwei weitere Stellen, vom Digital Services Koordinator bei der Bundesnetzagentur als Trusted Flagger gemäß dem Digital Services Act (DSA) zugelassen. Diese Zulassung als vertrauenswürdiger Hinweisgeber stärkt die Arbeit von HateAid gegen digitale Gewalt im Netz und für die Rechte von betroffenen Personen durch effizientere Meldewege.

Berlin, 02. Juni 2025

Holocaustleugnungen, Mord- oder Vergewaltigungsandrohungen stehen oft lange im Netz bis sie entfernt werden. Denn Plattformen müssen Inhalte erst überprüfen, wenn diese ihnen gemeldet werden. Hier sollen Trusted Flagger jetzt unterstützen: HateAid und andere qualifizierte Organisationen, Verbände und Stellen können den Plattformen jetzt strafbare Inhalte melden. Wichtig ist: Die Entscheidung, ob die Inhalte als rechtswidrig einzuschätzen sind und entfernt werden sollen, trifft dann allein die jeweilige Social Media Plattform nach eigener interner Prüfung. Trusted Flagger sind lediglich qualifizierte und vertrauenswürdige Hinweisgeber nach dem DSA. Plattformen sind jetzt aber gesetzlich verpflichtet, Meldungen solcher durch die Aufsichtsbehörde zertifizierten Stellen vorrangig und unverzüglich zu bearbeiten.

Dazu **Josephine Ballon**, Geschäftsführerin von HateAid

„Ob Meta, X oder Tiktok: Social Media Plattformen lassen zu, dass auf ihnen strafbare Inhalte wie Morddrohungen oder antisemitische Volksverhetzung von radikalen Kräften verbreitet werden. Mit dem DSA werden den Plattformen nun erstmals demokratisch legitimierte Nutzendenrechte vorgeschrieben, die sie einhalten müssen. Es ist der Versuch, unsere Demokratie, die EU-Mitgliedstaaten und jeden einzelnen Nutzenden gegen die Willkür der Tech-Plattformen zu verteidigen. Dazu wollen wir als Trusted Flagger einen Beitrag leisten.“

Die Dringlichkeit zeigen aktuelle Zahlen des Bundeskriminalamtes: In der aktuellen Kriminalstatistik erhebt das BKA die Zahl der im Internet – vor allem in sozialen Medien – verbreiteten strafbaren Inhalte. So hat es für das Jahr 2024 einen Anstieg von 34 Prozent von Straftaten im Zusammenhang mit sogenannten „Hasspostings“ erfasst. Bei der erfassten Hasskriminalität steigen antisemitische Straftaten auf einen neuen Höchststand. Die Zahlen dokumentieren, laut BKA, eine wachsende Polarisierung und Radikalisierung innerhalb der Gesellschaft. Soziale Medien spielten dabei aufgrund der rasant und enorm angestiegenen Verbreitung von Hass, Hetze, Desinformation und Propaganda eine entscheidende Rolle.

Diese Entwicklungen verdeutlichen die Notwendigkeit, Unterstützung für betroffene Personen beim Umgang mit digitaler Gewalt zu schaffen. Für Betroffene von digitaler Gewalt ist es wichtig, dass strafbare Inhalte zügig entfernt werden. Aber auch für uns als Gesellschaft ist das wichtig – gerade, wenn es sich um antisemitische oder rassistische illegale Inhalte handelt. Als Trusted Flagger handelt HateAid mit behördlicher Zertifizierung für mehr Sicherheit und respektvollen Umgang im Netz.

Über HateAid gGmbH



Die gemeinnützige Organisation HateAid wurde 2018 gegründet und hat ihren Hauptsitz in Berlin. Sie setzt sich für Menschenrechte im digitalen Raum ein und engagiert sich auf gesellschaftlicher wie politischer Ebene gegen digitale Gewalt und ihre Folgen. HateAid unterstützt Betroffene von digitaler Gewalt konkret durch Beratung und Prozesskostenfinanzierung. Geschäftsführerinnen sind Anna-Lena von Hodenberg und Josephine Ballon.

HateAid ist Trägerin der Theodor-Heuss-Medaille 2023, des Rothenburger Preises für Erinnerung und Zukunft, des Wertepreises für Demokratie der Werte-Stiftung und des For..Net Awards der Technischen Universität München.

Für weitere Informationen über HateAid gGmbH besuchen Sie unsere Website: <https://hateaid.org/>

Pressekontakt: presse@hateaid.org, Tel. 030 25208837

Background-Sheet für redaktionellen Hintergrund:

1. Was ist der Digital Services Act (DSA)?

Der Digital Services Act (DSA) ist eine EU-Verordnung, die seit Februar 2024 auch in Deutschland vollständig zur Anwendung kommt und durch das nationale Digitale-Dienste-Gesetz seit Mai 2024 verbindliche Regeln für Online-Dienste festlegt (Quelle: Bundesregierung). Ziel des DSA ist es, Online-Plattformen stärker in die Verantwortung zu nehmen, um strafbare Inhalte, Desinformation und Risiken für Nutzer*innen zu minimieren. Plattformen müssen unter anderem transparente Moderationsprozesse gewährleisten, Nutzerrechte wahren und illegale Inhalte effektiv entfernen.

Der DSA ist Teil der EU-Digitalstrategie und soll ein sicheres, transparentes und faires Online-Umfeld schaffen. Andererseits bietet er mehr Transparenz und niedrigschwellige Möglichkeiten, gegen Löschungen vorzugehen (Quelle: EU-Kommission).

2. Was ist ein Trusted Flagger?

Trusted Flagger sind vertrauenswürdige Hinweisgeber*innen, die von den zuständigen nationalen Behörden ernannt werden. In Deutschland ist die Bundesnetzagentur (BNetzA) als eine der zentralen Aufsichtsbehörden für die Anwendung des Digital Services Acts (DSA) zuständig (Quelle: Bundesnetzagentur). Als Digital Services Coordinator (DSC) ist sie gegenüber der Bundesregierung und insbesondere dem Bundeswirtschaftsministerium nicht weisungsgebunden. Sie nimmt Beschwerden entgegen, überwacht die Einhaltung der Vorschriften und ernennt Trusted Flagger. Derzeit sind EU-weit 31 Organisationen als vertrauenswürdige Hinweisgeber*innen gelistet (Quelle: EU-Kommission). Plattformen wie Meta, TikTok und X sind jeden Tag mit einer Vielzahl von Meldungen konfrontiert. Sie müssen überprüfen, ob Inhalte illegal sind und entfernt werden sollen oder nicht. Dies kann mitunter aufgrund der Vielzahl der gemeldeten Inhalte sehr lange dauern. Hier kommen Trusted Flagger ins Spiel: Ihre Rolle ist es, Online-Plattformen dabei zu unterstützen, die Entfernung von strafbaren Inhalten zu beschleunigen. Damit helfen sie Bürger*innen, die von digitaler Gewalt betroffen sind. Meldungen von Trusted Flaggern müssen von Plattformen vorrangig behandelt werden, da sie aufgrund ihrer Expertise als besonders zuverlässig gelten, illegale Inhalte zu identifizieren. Wichtig: Es geht hier nur um eine vorrangige Meldung. Die Entscheidung, ob die Inhalte wirklich illegal sind und entfernt werden müssen, obliegt immer der Plattform. Trusted Flagger sind somit lediglich Hinweisgeber und haben keinen Einfluss auf die Löscheidungen der Plattformen.

3. Nach welchen Kriterien werden Trusted Flagger ausgesucht?



Organisationen oder Einzelpersonen, die als Trusted Flagger tätig werden möchten, müssen einen Antrag bei der Bundesnetzagentur stellen. Dabei gibt es eine Reihe von Zulassungskriterien zu erfüllen, z. B.: eine nachgewiesene Sachkenntnis im Bereich der Erkennung und Meldung illegaler Inhalte (1), Neutralität und Unabhängigkeit (2), transparente und verlässliche Arbeits- und Monitoringprozesse (3), sowie die Einhaltung der EU-Grundrechte, insbesondere der Meinungsfreiheit (Quelle: Digital Services Coordinator, Zertifizierungsleitfaden).

4. Wie sieht die Arbeit von Trusted Flaggern aus?

Im Gegensatz zu allgemeinen Meldesystemen bieten Trusted Flagger einen spezialisierten, kontrollierten und priorisierten Meldemechanismus. Sie prüfen die Inhalte fachkundig, was die Trefferquote im Vergleich zu normalen Nutzer*innen in Bezug auf die Rechtswidrigkeit von Inhalten erhöht. Zudem erstellen sie professionelle Meldungen, da sie mit dem Prozess vertraut sind und wissen, welche Informationen Plattformen benötigen.

Wichtig ist: Die Meldung eines Trusted Flagger zwingt die Plattformen nicht zur Löschung der gemeldeten Inhalte. Die Bewertung und anschließende Entscheidung liegen weiterhin allein bei den Plattformen selbst. Sie können die gemeldeten Inhalte auf den Plattformen bestehen lassen, wenn sie nach einer Prüfung zu dem Schluss kommen, dass die gemeldeten Inhalte nicht rechtswidrig sind. Die gesetzlichen Grundlagen gewähren den Plattformen einen gewissen Freiraum.

Damit unterstützt HateAid das Ziel des Digital Services Acts, Plattformen sicherer zu machen und die Rechte der Betroffenen zu wahren. Dies hat hohe gesellschaftliche Relevanz über die Betroffenen hinaus: Zum Beispiel im Bereich Antisemitismus, der im digitalen Raum massiv angestiegen ist (Quelle: Bundeskriminalamt). Wird ein Post gelöscht, muss das nicht endgültig sein. Die Verfasser*innen haben das Recht, die Beanstandung überprüfen zu lassen. Sie werden informiert und können bei Schlichtungsstellen gegen die Löschung vorgehen (Quelle: Bundeszentrale für politische Bildung).

Auch Trusted Flagger informieren die Öffentlichkeit über ihre Arbeit: Sie sind verpflichtet, in einem leicht verständlichen und ausführlichen Jahresbericht die Anzahl und Art der gemeldeten Inhalte, die betroffenen Plattformen und die ergriffenen Maßnahmen darzulegen (Quelle: EU-Kommission).