Projektpartner:

Co-funded by the European Union

Handout für Jugendorganisationen

Wie man den Status eines “Trusted Flagger” für Hassrede erhält.

Einleitung: Was ist der Status „Trusted Flagger“ und warum ist er wichtig?

Gemäß dem Digital Services Act (DSA) können sich zivilgesellschaftliche Organisationen als **„Trusted Flaggers“ (vertrauenswürdige Melder) bewerben** – offiziell anerkannte Einrichtungen, deren Meldungen über illegale Inhalte (einschließlich Hassrede) von Online-Plattformen **vorrangig behandelt werden müssen** . „Trusted Flaggers“ profitieren von verbesserten Meldekanälen, schnelleren Bearbeitungszeiten und einer strukturierten Zusammenarbeit mit den Plattformen und dem nationalen **Koordinator für digitale Dienste (DSC)** .

Für Jugendorganisationen, insbesondere solche, die sich mit Hassrede, Cybermobbing, Diskriminierung, Online-Belästigung oder dem Schutz gefährdeter Gruppen befassen, kann der Status als „Trusted Flagger“ ihre Wirkung deutlich steigern. Viele NGOs, die sich für Jugendliche engagieren, melden bereits Vorfälle in sozialen Netzwerken, doch ohne besondere Anerkennung werden diese Meldungen nur langsam oder uneinheitlich geprüft. Als „Trusted Flagger“ genießen die Meldungen Ihrer Organisation höheres Ansehen und werden mit besonderer Dringlichkeit behandelt. Dadurch können schädliche Inhalte, die junge Menschen betreffen, schneller und zuverlässiger entfernt werden.

Jugendorganisationen spielen dabei eine besonders wichtige Rolle: Sie haben Einblick in die realen Erfahrungen junger Nutzer, verstehen die sich wandelnde Online-Sprache (Slang, Geheimcodes, Memes) und gehören oft zu den Ersten, die neue Formen von Hassrede oder Belästigung erkennen. Der Status als vertrauenswürdiger Melder trägt dazu bei, diese Expertise zu formalisieren und in das umfassendere Ökosystem der Online-Sicherheit zu integrieren.

Dieses Merkblatt erläutert die Kriterien, wie Sie Ihre Organisation vorbereiten können, wie Sie den Antrag stellen und auf welche Verantwortlichkeiten Sie sich nicht nach Erfolg einstellen sollten.

Die Kriterien für vertrauenswürdige Melder verstehen

Das Gesetz über digitale Dienste legt klare Kriterien für die Anerkennung als vertrauenswürdiger Melder fest. Jugendorganisationen sollten vor der Antragstellung prüfen, ob sie die erforderlichen Kapazitäten besitzen oder aufbauen können.

Co-funded by the European Union

1. Nachgewiesene Fachkompetenz in einem relevanten Bereich

Ihre Organisation muss nachweislich über langjährige Erfahrung in Bereichen wie illegale Hassrede, schädliche Inhalte, Jugendschutz, Online-Sicherheit und digitale Kompetenz verfügen. Dazu gehören beispielsweise der Betrieb eines Meldeportals, die Durchführung von Programmen gegen Hassrede, Monitoring oder Forschung, die Bearbeitung von Fällen von Online-Belästigung oder die Unterstützung marginalisierter Gruppen, die von Hassrede betroffen sind.

2. Unabhängigkeit und politische Neutralität

Vertrauenswürdige Flagger müssen unabhängig von politischem Einfluss sein und nachweisen, dass ihre Entscheidungen auf Fachwissen und nicht auf politischen Agenden beruhen. Jugendorganisationen sollten transparente Eigentumsverhältnisse, klare Führungsstrukturen und keine parteipolitische Ausrichtung aufweisen.

3. Strenge Berichts- und Verifizierungsverfahren

Ihre Organisation muss über interne Systeme verfügen, um zu überprüfen, ob die gemeldeten Inhalte nach EU- oder nationalen Gesetzen gegen Hassrede illegal oder schädlich sind. Dies erfordert geschultes Personal, einen klaren Eskalationsmechanismus und die Möglichkeit, Beweise zu dokumentieren (URLs, Screenshots, Kontextinformationen, Zeitstempel).

4. Fähigkeit zur systematischen Bearbeitung von Berichten

Vertrauenswürdige Melder müssen eine große Anzahl von Meldungen professionell und zuverlässig bearbeiten. Sie benötigen einen nachhaltigen Arbeitsablauf für die Annahme, Bewertung und Weiterleitung von Fällen. Dies kann Fallmanagement-Tools, interne Richtlinien, Qualitätssicherungsprozesse und in der Klassifizierung von Hassrede geschultes Personal umfassen.

5. Einhaltung der Datenschutzbestimmungen

Da Ihre Organisation sensible Daten von jungen Menschen verarbeitet, müssen Sie die Einhaltung der DSGVO nachweisen. Dies umfasst Datenminimierung, Anonymisierung, sichere Speicherung und klare Aufbewahrungsrichtlinien.

6. Integrität und Genauigkeit der Berichterstattung

Plattformen und digitale Servicezentren erwarten hohe Genauigkeit. Sie müssen in der Lage sein, illegale Hassrede (gemäß gesetzlicher Definition) von schädlichen, aber legalen Äußerungen zu unterscheiden. Klare interne Kriterien, Rechtsberatung und Mitarbeiterschulungen sind unerlässlich.

Wenn Ihre Organisation diese Kriterien erfüllen kann – oder bereit ist, sie zu entwickeln – können Sie mit der Antragstellung fortfahren.

Vorbereitung Ihrer Organisation auf den Antrag

Der Bewerbungsprozess erfordert den Nachweis von Leistungsfähigkeit, Fachkompetenz und internen Strukturen. Vor der Bewerbung sollten Jugendorganisationen folgende Unterlagen vorbereiten:

1. Dokumentieren Sie Ihre Erfahrung und Ihr Fachwissen.

Erstellen Sie ein Dossier, das Ihre Arbeit im Bereich Hassrede und Jugendschutz zusammenfasst. Dies kann Folgendes beinhalten:

- Beschreibungen relevanter Programme (Anti-Mobbing, Prävention von Hassrede, digitale Bürgerbildung)
- Jahresberichte
- Statistiken zu bearbeiteten Fällen
- vorherige Zusammenarbeit mit Plattformen oder Behörden
- Forschung oder Veröffentlichungen über Online-Gefahren

Eine klare Dokumentation stärkt Ihre Glaubwürdigkeit und zeigt dem DSC, dass Sie die digitale Risikolandschaft verstehen.

2. Berichtsprotokolle erstellen oder aktualisieren

Ein vertrauenswürdiger Melder muss strukturierte Meldemechanismen verwenden. Erstellen oder optimieren Sie interne Protokolle, die Folgendes beinhalten:

- ein Verifizierungsprozess für jeden Bericht
- Kategorien illegaler Inhalte
- Dokumentationsstandards (Screenshots, Metadaten, Kontext)
- Qualitätskontrollverfahren
- ein Arbeitsablauf für die Bearbeitung dringender Fälle mit Beteiligung von Jugendlichen

Dies gewährleistet Konsistenz und zeigt, dass Ihre Organisation bereit ist, priorisierte Meldekanäle verantwortungsvoll zu handhaben.

3. Schulung der Mitarbeiter zu rechtlichen Rahmenbedingungen

Mitarbeiter, die Berichte einreichen, müssen Folgendes verstehen:

- nationale Gesetze gegen Hassrede
- einschlägige EU-Rechtsvorschriften
- Unterscheidungen zwischen illegaler Hassrede und schädlichen, aber legalen Inhalten
- Kontextfaktoren (geschützte Merkmale, Anstiftung, Bedrohungen)

Schulungen können intern, mit externen Experten oder in Partnerschaft mit Organisationen für digitale Rechte durchgeführt werden.

4. Datenschutzvorkehrungen treffen

Erstellen oder aktualisieren:

- Datenschutzrichtlinien
- Datenaufbewahrungsverfahren
- sichere Speicherlösungen
- interne DSGVO-Konformitätsrichtlinien
- Vertraulichkeitsvereinbarungen für Mitarbeiter

Dies ist besonders wichtig für Organisationen, die mit Minderjährigen arbeiten.

5. Aufbau einer Fallmanagement-Infrastruktur

Für kleine Organisationen mag eine einfache Tabellenkalkulation ausreichen, aber diejenigen, die den Status eines trusted Flaggers anstreben, sollten Folgendes beachten:

- Fallmanagementsysteme
- Ticketing-Tools
- Freigegebene Ordner mit kontrolliertem Zugriff
- sichere Plattformen zur Vorfallverfolgung

Dies wird Ihnen helfen, das Berichtsvolumen zu bewältigen und Prüfprotokolle zu führen.

Der Bewerbungsprozess: Zusammenarbeit mit dem Koordinator für digitale Dienste

Sobald Ihre Organisation bereit ist, reichen Sie einen Antrag bei Ihrem nationalen **Koordinator für digitale Dienste ein** . Der Prozess kann sich von Land zu Land geringfügig unterscheiden, die Kernelemente sind jedoch in der gesamten EU gleich.

Schritt 1: Reichen Sie einen formellen Antrag ein.

Ihre Bewerbung umfasst in der Regel Folgendes:

- Organisationsinformationen
- Governance-Struktur
- dokumentierte Expertise in den Bereichen Hassrede und Jugendschutz
- Beschreibung der Berichts-Workflows
- Dokumentation zum Datenschutz
- Qualifikation und Schulung des Personals
- Beispiele für frühere Berichterstattung oder Überwachung

Einige DSCs verlangen möglicherweise auch Referenzen oder Nachweise über die Zusammenarbeit mit Plattformen oder anderen Regulierungsbehörden.

Schritt 2: An der Evaluierung und Beratung teilnehmen

Die DSC kann zusätzliche Informationen anfordern, um Klärung bitten oder ein Interview führen. Sie kann auch andere Behörden oder zivilgesellschaftliche Akteure konsultieren, um die bisherige Arbeit der Organisation zu überprüfen. Jugendorganisationen sollten bereit sein, Folgendes nachzuweisen:

- ihr rechtliches Verständnis von Definitionen von Hassrede
- ihre Fähigkeit, den Kontext zu analysieren
- ihre Fähigkeit, falsche Berichterstattung zu vermeiden
- ihre Unabhängigkeit und Neutralität

Schritt 3: Ernennung erhalten und Zusammenarbeit beginnen

Bei Genehmigung wird Ihre Organisation offiziell als **trusted Flagger gelistet** . Das DSC benachrichtigt die sozialen Netzwerke und stellt Ihnen Folgendes zur Verfügung:

- Prioritätsmeldekanäle
- plattformspezifische Anweisungen
- Kommunikationspunkte
- Erwartungen an die Genauigkeit der Berichterstattung und das Fallvolumen

Trusted-Flagger -Teilnehmer müssen dann gemäß den vereinbarten Regeln über die dafür vorgesehenen Kanäle Berichte einreichen.

Schritt 4: Laufende Überwachung und Rechenschaftspflicht

Der Trusted-Flagger-Status ist weder dauerhaft noch bedingungslos. DSCs können die Leistung, Genauigkeit und Unabhängigkeit Ihrer Organisation regelmäßig überprüfen. Dies gewährleistet, dass das Trusted-Flagger-Ökosystem glaubwürdig, effektiv und missbrauchsresistent bleibt.

Jugendorganisationen, die eine konsequente, genaue und ethische Berichterstattung gewährleisten, behalten ihren Status langfristig und erhalten möglicherweise zusätzliche Möglichkeiten zur Zusammenarbeit, wie z. B. die Teilnahme an systemischen Risikoanalysen oder politischen Konsultationen.

Als Trusted-Flagger agieren: Verantwortlichkeiten und bewährte Vorgehensweisen

Nach Erhalt des Status „Trusted Flagger“ gewinnen Jugendorganisationen erheblichen Einfluss – übernehmen aber auch neue Verantwortlichkeiten. Vertrauenswürdige Melder müssen hohe professionelle Standards einhalten und konstruktiv mit Plattformen und Aufsichtsbehörden zusammenarbeiten.

1. Hochwertige, evidenzbasierte Berichte erstellen und pflegen

Jeder Bericht muss kontextbezogen, präzise und durch eindeutige Beweise untermauert sein. Berichte von geringer Qualität untergraben das System. Trusted Flagger sollten stets Folgendes dokumentieren:

- der genaue illegale Ausdruck
- das geschützte Merkmal

- der Kontext, in dem der Inhalt erscheint
- irgendwelche Muster oder wiederkehrende Verhaltensweisen
- potenzielle Risiken für den jungen Menschen

Dadurch wird sichergestellt, dass die Plattform schnell und sicher handeln kann.

2. Sicherheit und Wohlbefinden der Jugendlichen priorisieren.

Bei der Meldung von Hassrede gegen Minderjährige müssen Jugendorganisationen besonders behutsam vorgehen. Sie sollten die emotionale Belastung für die betroffene Person minimieren, den Meldevorgang erläutern und gegebenenfalls die psychologische Betreuung oder die Schule hinzuziehen. Die Pflichten von vertrauenswürdigen Meldern haben niemals Vorrang vor dem Kinderschutz.

3. Nutzen Sie Plattformen konstruktiv.

Trusted Flagger kommunizieren regelmäßig mit Online-Plattformen. Dies kann die Erläuterung des lokalen Kontexts, das Aufzeigen neuer Trends oder die Klärung unklarer Fälle umfassen. Ein konstruktiver Dialog hilft Plattformen, ihre Systeme zur Inhaltsmoderation zu verbessern.

4. Teilen Sie Ihre Erkenntnisse mit dem DSC und tragen Sie zur breiteren politischen Arbeit bei.

Trusted Flagger spielen nicht nur bei der Meldung einzelner Inhalte eine Rolle, sondern auch bei der Aufdeckung systemischer Mängel. Jugendorganisationen sollten Folgendes in Betracht ziehen:

- Vierteljährliche Zusammenfassungen der Trends bei Hassrede
- Jugendzentrierte Einblicke in verschlüsselte Sprache oder neue Belästigungsmuster
- Beobachtungen zur Reaktionsfähigkeit der Plattform

Diese Erkenntnisse helfen DSCs, Risiken einzuschätzen und die Durchsetzung im gesamten digitalen Bereich zu verbessern.

5. Transparenz, Unabhängigkeit und ethische Standards wahren.

Der Status als “Trusted Flagger” setzt Integrität voraus. Jugendorganisationen müssen politische Voreingenommenheit, Interessenkonflikte und Praktiken, die das Vertrauen schädigen könnten, vermeiden. Sie sollten klare ethische Richtlinien einhalten und bei der Bewertung von Inhalten unparteiisch bleiben.

Co-funded by the European Union

Abschluss

Die Anerkennung als „Trusted Flagger“ ist sowohl eine Verantwortung als auch eine Chance. Für Jugendorganisationen bietet sie ein wirksames Instrument, um junge Menschen vor Hassrede, Gewalt und Belästigung im Internet zu schützen. Die Auszeichnung stärkt den Einfluss Ihrer Organisation, verbessert die Zusammenarbeit mit Plattformen und Behörden und trägt zu einem sichereren digitalen Umfeld für alle jungen Nutzer bei.

Mit guter Vorbereitung, transparenter Führung, verlässlicher Expertise und robusten Verfahren können Jugendorganisationen erfolgreich den Status erlangen und zu zentralen Akteuren im neuen Online-Sicherheitsrahmen werden, der durch den Digital Services Act geschaffen wurde.

Co-funded by the European Union