



# NetzDG Transparenzbericht

Januar 2023

## 1. ALLGEMEINE AUSFÜHRUNGEN

Allgemeine Ausführungen dazu, welche Anstrengungen der Anbieter des sozialen Netzwerks unternimmt, um strafbare Handlungen auf den Plattformen zu unterbinden.

Meta Platforms Ireland Limited, welche das soziale Netzwerk Instagram für Nutzer\*innen in Deutschland bereitstellt, möchte eine sichere Plattform bieten, der die Menschen vertrauen und auf der sie sich frei äußern können. Dabei erlauben wir Nutzer\*innen jedoch nicht, Inhalte zu posten, die rechtswidrig sind oder zu kriminellen Verhalten anregen. Auch jegliche Art von Bullying oder Belästigung ist nicht gestattet. Um diese Balance zu halten, nutzen wir einen vielschichtigen Ansatz zur Bekämpfung potenziell strafbarer Handlungen bzw. schädlicher Aktivitäten auf der Instagram-Plattform.

Zum einen haben wir weltweit gültige Gemeinschaftsrichtlinien definiert, die festlegen, was auf Instagram zulässig ist und was nicht, und die in vielen Fällen abdecken, was gemäß deutschem Recht rechtswidrig ist. Diese Richtlinien gelten für Inhalte auf der ganzen Welt und sind von entscheidender Bedeutung für den Schutz der freien Meinungsäußerung sowie der persönlichen Sicherheit auf Instagram. Unsere Gemeinschaftsrichtlinien verbieten eine Vielzahl von anstößigen bzw. schädlichen Inhalten, zum Beispiel Inhalte, die:

- gewalttätiges und kriminelles Verhalten fördern,
- die Sicherheit Anderer bedrohen,
- als Hassrede gelten,
- als drastische Gewaltdarstellung gelten,
- als Spam eingestuft werden,
- als Mobbing oder Belästigung eingestuft werden.

Die [Gemeinschaftsrichtlinien](#) werden durch ein globales Team erstellt, dessen Mitglieder eine Vielzahl unterschiedlicher Fachgebiete abdecken. Dazu gehören Expert\*innen in den Bereichen Kindersicherheit, Hassrede und Terrorismus. Dieses Team wird regelmäßig von externen Expert\*innen und Organisationen beraten, um eine Balance aus unterschiedlichen Sichtweisen zu den Themen Meinungsfreiheit und Sicherheit zu wahren und die Auswirkungen unserer Richtlinien auf verschiedene Personengruppen weltweit besser zu verstehen. Unsere Prüfer\*innen setzen diese Standards mithilfe umfassender Richtlinien durch, die sicherstellen, dass die Entscheidungen so konsistent wie möglich getroffen werden. Dabei ist zu beachten, dass eine Gemeinschaftsrichtlinien-Meldung nicht dazu führt, dass wir den Inhalt einer rechtlichen Prüfung



unterziehen. Wir prüfen entsprechend der Gemeinschaftsrichtlinien-Meldung eine Verletzung unserer Gemeinschaftsrichtlinien. Wir stellen separate Meldemechanismen bereit, mit denen Nutzer\*innen Inhalte melden können, die ihrer Ansicht nach gegen nationales Recht verstoßen (siehe dazu nachfolgend im Detail).

Angemeldete Instagram-Nutzer\*innen können Verstöße gegen die Gemeinschaftsrichtlinien auf verschiedene Arten melden, u. a. über ein Auswahlmenü, das zu jedem Inhalt angezeigt wird. Wir prüfen Inhalte (automatisiert oder manuell), die über diese Tools gemeldet werden, um festzustellen, ob die gemeldeten Inhalte gegen unsere Gemeinschaftsrichtlinien verstoßen. Wenn dies der Fall ist, entfernen wir den gemeldeten Inhalt. Darüber hinaus setzen wir Technologie ein, um dabei zu helfen, potenziell gegen die Gemeinschaftsrichtlinien verstoßende Inhalte zu erkennen, bevor Menschen diese sehen und uns melden (siehe Abschnitt 2 für Details hierzu).

Zum anderen stellen wir verschiedene Möglichkeiten bereit, mit denen Personen Inhalte melden können, die nach ihrer Ansicht gegen nationales Recht verstoßen. Wenn uns ein Inhalt auf Instagram wegen Verstoßes gegen deutsches Recht gemeldet wird, er aber nicht gegen unsere Gemeinschaftsrichtlinien verstößt, können wir den Zugang zu dem Inhalt in Deutschland sperren. Personen in Deutschland können verschiedene Kanäle zur Meldung vermeintlicher Verstöße gegen deutsches Recht nutzen, so zum Beispiel:

- Das [NetzDG-Beschwerdeformular](#) – Mithilfe dieses Formulars können Personen in Deutschland Inhalte melden, die nach ihrer Ansicht gegen eine oder mehrere der im NetzDG aufgeführten Bestimmungen des deutschen Strafgesetzbuchs verstoßen (für weitere Informationen zu diesem Formular siehe Abschnitt 3 dieses Transparenzberichts).
- Die [Meldeformulare für Verletzungen geistiger Eigentumsrechte](#) – Diese Formulare erlauben es Rechteinhabern und deren autorisierten Vertretern, Inhalte zu melden, die nach ihrer Ansicht Urheberrechte oder Markenrechte verletzen.
- Das [Formular zum Melden von Verleumdung](#) – Mit diesem Formular können Betroffene und bestimmte autorisierte Vertreter Inhalte melden, die ihrer Ansicht nach diffamierenden Charakter haben. Dies umfasst auch Inhalte, die unwahre Tatsachenbehauptungen darstellen, welche zu einer Rufschädigung führen. Wenn Personen Deutschland als das Land auswählen, für das sie Rechte geltend machen wollen, werden sie auf die Option hingewiesen, ihre Beschwerde stattdessen über unser NetzDG-Beschwerdeformular zu übermitteln, wenn es sich ihrer Ansicht nach bei dem Inhalt um Beleidigung, üble Nachrede oder Verleumdung im strafrechtlichen Sinne handelt.
- Das [Antragsformular zur Entfernung von Inhalten aufgrund von Rechtsverstößen](#) – Mit diesem Formular können Personen in den Mitgliedstaaten der Europäischen Union Inhalte melden, die ihrer Ansicht nach ihre Persönlichkeitsrechte verletzen oder gegen national geltendes Recht verstoßen. Wenn Personen Deutschland als das Land auswählen, für das sie Rechte geltend machen wollen, erhalten sie zunächst die Möglichkeit, eine Beschwerde über unser NetzDG-Beschwerdeformular einzureichen. Entscheiden sie sich dagegen, gelangen sie zu den übrigen oben genannten Formularen sowie zu einem separaten Formular für das Recht auf Datenschutz/Löschung. Zudem bietet das Formular Personen die Möglichkeit, Inhalte zu melden, die ihrer Ansicht nach gegen andere Gesetze verstoßen und nicht von den oben genannten Kategorien abgedeckt sind.



## 2. VERFAHREN ZUR AUTOMATISIERTEN ERKENNUNG VON INHALTEN, DIE ENTFERNT ODER GESPERRT WERDEN SOLLEN

Art, Grundzüge der Funktionsweise und Reichweite von gegebenenfalls eingesetzten Verfahren zur automatisierten Erkennung von Inhalten, die entfernt oder gesperrt werden sollen, einschließlich allgemeiner Angaben zu verwendeten Trainingsdaten und zu der Überprüfung der Ergebnisse dieser Verfahren durch den Anbieter, sowie Angaben darüber, inwieweit Kreise der Wissenschaft und Forschung bei der Auswertung dieser Verfahren unterstützt werden und diesen zu diesem Zweck Zugang zu Informationen des Anbieters gewährt wurde.

### 2A. ART, FUNKTIONSWEISE UND REICHWEITE DER AUTOMATISIERTEN ERKENNUNG VON INHALTEN, DIE ENTFERNT ODER GESPERRT WERDEN SOLLEN

Zur Durchsetzung unserer Gemeinschaftsrichtlinien setzen wir eine Kombination aus menschlicher Prüfung und Technologie ein (siehe schon Abschnitt 1). Jeden Tag entfernen wir Millionen von unzulässigen Inhalten und Konten von Facebook und Instagram. In den meisten Fällen geschieht dies automatisch mithilfe von Technologie wie künstlicher Intelligenz, die im Hintergrund arbeitet, um Inhalte zu erkennen und zu entfernen, die gegen die Gemeinschaftsrichtlinien verstoßen. In anderen Fällen wählt unsere Technologie Inhalte für eine menschliche Überprüfung aus. Unsere Review-Teams prüfen eine Mischung aus den in Abschnitt 1 beschriebenen Nutzer\*innen-Meldungen und Beiträgen, die von unseren Systemen künstlicher Intelligenz identifiziert wurden. Ferner unterstützt unsere Technologie die Review-Teams auch, die kritischsten zu prüfenden Inhalte nach Schweregrad, Viralität und Wahrscheinlichkeit eines Verstoßes zu priorisieren. Unsere Prüfsysteme verwenden Technologien, um Inhalte mit hohem Schweregrad, die Schaden abseits des Internets anrichten können (etwa Beiträge im Zusammenhang mit Terrorismus und Selbstmord), sowie virale Inhalte, die sich schnell verbreiten und eine große Reichweite haben könnten, zu priorisieren, um so viel Schaden wie möglich zu verhindern.

Unsere Technologie ist darauf ausgerichtet, Verstöße gegen unsere Gemeinschaftsrichtlinien zu erkennen. Sie ist nicht darauf ausgerichtet, rechtswidrige Inhalte im Sinne von § 1 Abs. 3 NetzDG unmittelbar zu erkennen, auch wenn es selbstverständlich Überschneidungen zwischen unzulässigen Inhalten im Sinne unserer Gemeinschaftsrichtlinien und rechtswidrigen Inhalten im Sinne von § 1 Abs. 3 NetzDG gibt (z. B. im Bereich der Hassrede). Zur Erkennung von Verstößen gegen die Gemeinschaftsrichtlinien werden in erster Linie drei Techniken eingesetzt.

- Erstens begrenzen wir die Geschwindigkeit, mit der Konten Aktionen, inklusive des Postens von Beiträgen, auf unserer Plattform durchführen können (Rate Limits), um den Einsatz von Bots zu verhindern.
- Zweitens verwenden wir eine Abgleichtechnologie, die identische oder nahezu identische Kopien von URLs, Text, Bildern, Audio- und Videobeiträgen identifiziert, die wir zuvor als Verstoß gegen unsere Gemeinschaftsrichtlinien identifiziert haben. Diese Abgleichtechnologie kann auch dann funktionieren, wenn der zu untersuchende Inhalt geringfügige Unterschiede zum ursprünglichen Inhalt aufweist. Wenn sich aus dem Abgleich ergibt, dass ein Inhalt identisch oder nahezu identisch mit einem vorigen Verstoß ist, entfernen wir den Inhalt typischerweise. Diese Technologie wird manchmal als Content Hashing oder Content Digital



Fingerprinting bezeichnet.

- Schließlich verwenden wir auch künstliche Intelligenz im engeren Sinne (d. h. maschinelles Lernen (Machine Learning) und regelbasierte Systeme, im Folgenden: künstliche Intelligenz, KI), und zwar auf zweierlei Weise: Wie bei der Abgleichtechnologie entfernt die künstliche Intelligenz einen Inhalt typischerweise, wenn sie sich sicher genug ist, dass er gegen unsere Gemeinschaftsrichtlinien verstößt. Wir verwenden künstliche Intelligenz außerdem, um auf der Grundlage von Schweregrad, Viralität und Wahrscheinlichkeit eines Verstoßes die Inhalte für die menschliche Überprüfung auszuwählen. Wie die Abgleichtechnologie wird auch die künstliche Intelligenz für URLs, Text, Bilder, Audio und Videos eingesetzt. Anders als Technologien, die nur Verstöße abgleichen können, die sie bereits gesehen haben, hat künstliche Intelligenz das Potenzial, auch bestimmte Verstöße zu erkennen, die sie noch nie gesehen hat.

Wir setzen unsere Technologie bei Inhalten, die für alle Nutzer\*innen zugänglich sind, in allen Bereichen der Instagram-Plattform und für praktisch sämtliche Verstöße gegen unsere Gemeinschaftsrichtlinien ein. Da sich die Erwartungen der Nutzer\*innen je nach Produktbereich und Publikumsgröße (z.B. Gruppen, Freunde) verändern, begrenzen wir den Kreis der Richtlinien, die mit automatisierten Mitteln durchgesetzt werden, nach ihrem Schweregrad.

## 2B. VERWENDETE TRAININGSDATEN

Rate Limits: Wir legen unsere Begrenzungen der Geschwindigkeit (Rate Limits) fest, indem wir beobachten, wie Menschen die Instagram-Plattform nutzen, und dann konservative Grenzwerte festlegen, die es uns ermöglichen, gegen die schwerwiegendsten Botverhalten vorzugehen. Diese Grenzwerte sind gleichzeitig so gestaltet, dass eine zulässige Nutzung nur sehr selten beeinträchtigt wird.

Abgleichtechnologie: Unsere Listen bekannter verletzender Inhalte, die der Abgleichtechnologie zugrunde liegen, werden typischerweise dann erstellt, wenn der gleiche Inhalt von unseren menschlichen Prüfer\*innen mehrfach als unzulässig eingestuft wurde.

Maschinelles Lernen (Machine Learning): Die Entwicklung von Techniken zum Trainieren von Machine-Learning-Modellen ist ein sich schnell entwickelnder Forschungsbereich in Industrie und Wissenschaft. Meta verwendet hauptsächlich zwei Techniken, um seine Machine-Learning-Modelle zu trainieren.

Die erste Technik ist gemeinhin unter der Bezeichnung „überwachtes Lernen“ bekannt. Meta verwendet für das Training seiner überwachten Modelle zur Inhaltsmoderation Varianten der gleichen allgemeinen Technik. Wir wählen eine statistische Zufallsstichprobe aller Inhalte aus, die von den Nutzer\*innen angesehen wurden (dies ist die gleiche Methode, die wir zur Berechnung der von uns veröffentlichten Messungen der Verbreitung von Verstößen verwenden) oder eine statistische Zufallsstichprobe aus allen Meldungen, die wir von Nutzer\*innen erhalten. Die ausgewählten Inhalte werden von menschlichen Prüfer\*innen entweder als unbedenklich oder als Verstoß gegen einen oder mehrere unserer Gemeinschaftsrichtlinien eingestuft. Im Rahmen dieses Vorgangs kann derselbe Inhalt zur Qualitätskontrolle mehrfach überprüft werden. Anschließend kombinieren wir diese Beispiele für unbedenkliche und verletzende Inhalte als Input für unsere Trainingsalgorithmen für maschinelles Lernen. Das Ergebnis dieser Trainingsalgorithmen ist ein Modell, das auch als „Classifier“ bezeichnet wird. Wir können diesen



Classifier dann einsetzen, um festzustellen, ob ein Beitrag wahrscheinlich gegen unsere Gemeinschaftsrichtlinien verstößt.

Die zweite Technik wird als „selbstüberwachtes Lernen“ bezeichnet. Bei dieser Trainingstechnik entfernt das Machine-Learning-Modell ein Wort aus einem Satz und prüft dann, ob es das fehlende Wort vorhersagen kann. Es handelt sich hierbei um eine erst vor Kurzem entwickelte Technik, die Meta in begrenzteren Rahmen verwendet, vor allem zum Trainieren von Machine-Learning-Modellen für die Auswertung von Sprache.

## 2C. ÜBERPRÜFUNG DER ERGEBNISSE

Zur Sicherstellung und Verbesserung der Qualität der vorstehend beschriebenen Technologien (d. h. wie präzise sie Verstöße gegen unsere Gemeinschaftsrichtlinien erkennen) gibt es laufende Qualitätsprüfverfahren. Meta verwendet einander überschneidende Techniken und Systeme, um insgesamt eine hohe Präzision unserer automatisierten Inhaltsmoderation zu gewährleisten.

Bevor wir neue Rate Limits, Abgleichtechnologien oder künstliche Intelligenz endgültig einführen, verwenden wir die Technologie nur, um das zu protokollieren, was sie entfernt *hätte*, nicht für die tatsächliche Entfernung des Inhalts. Im Anschluss setzen wir menschliche Prüfer\*innen ein, um den Grad der Präzision anhand von Echtzeit-Inhalten zu bestimmen (und nicht nur anhand von Inhalten aus der Vergangenheit, wie wir es beim Training der Technologie getan haben). Die Technologie ist oft in der Lage, einen höheren Präzisionsgrad zu erreichen als menschliche Prüfer\*innen.

Nach der Einführung von Rate Limits, Abgleichtechnologien oder künstlicher Intelligenz überwachen wir die Anzahl der Entfernungen und der Widersprüche durch den/die Nutzer\*in, der/die den Inhalt gepostet hat, sowie den Anteil erfolgreicher Widersprüche. Falls irgendeine der von uns überwachten Kennzahlen Auffälligkeiten zeigt, führen unsere Entwicklungsteams eine Untersuchung durch. Bei Rate Limits prüfen die Einwickler\*innen, ob die Beschränkung dafür sorgt, dass Botverhalten unterbunden wird. Wenn ein Eintrag in unserer Liste der Inhalte, die bekanntermaßen gegen die Gemeinschaftsrichtlinien verstoßen, auffällige Signale zeigt, überprüfen wir den Eintrag erneut, um zu verifizieren, dass er gegen unsere Gemeinschaftsrichtlinien verstößt. Wenn eines unserer KI-Tools auffällige Signale zeigt, senden wir entweder eine Stichprobe der aktuellsten Ergebnisse dieses Tools an menschliche Prüfer\*innen zu einer erneuten Prüfung des Präzisionsgrads oder wir stellen das Tool ab.

Darüber hinaus werden viele unserer Machine-Learning-Classifer nach jeder menschlichen Überprüfung im Rahmen unseres Systems zur Reduzierung der Verbreitung unzulässiger Inhalte (severity weighted viewership prevalence reduction system) automatisch auf ihre Präzision geprüft, unter Verwendung eines sogenannten "[mehrarmiger Bandit-Algorithmus](#)" (multi armed bandit algorithm). Diese Überprüfung der Classifier ist ein Beispiel für die generelle Feedback-Schleife zwischen menschlicher Prüfung und Technologie. Die Entscheidungen, die menschliche Prüfer\*innen bei der Kennzeichnung von Inhalten treffen, werden benutzt, um unsere Technologie zu trainieren und zu verbessern. Im Rahmen dieses Vorgangs kennzeichnen die Review-Teams manuell die jeweilige Richtlinie, die ihrer Entscheidung zugrunde liegt, also die Richtlinie, gegen die der Inhalt, das Konto oder das Verhalten verstößt. Dies trägt bei zur Qualitätsverbesserung unserer KI-Algorithmen und der für unsere Abgleichtechnologie verwendeten Listen bekannter gegen die Gemeinschaftsrichtlinien verstößender Inhalte.



Bei unseren automatischen Entfernungen vermeiden wir Diskriminierung vor allem durch Qualitätskontrollen, um die Einhaltung unserer Gemeinschaftsrichtlinien zu gewährleisten. Unsere Gemeinschaftsrichtlinien vermeiden Diskriminierung durch Prozesse wie die bereits in Abschnitt 1 beschriebene Konsultation von Stakeholdern und eine umfassende Analyse unserer internen Signale. Zu unseren internen Signalen gehören Nutzer\*innenanalysen, groß angelegte Umfragen unter unseren Nutzer\*innen und detaillierte Analysen dessen, was unsere Nutzer\*innen über die in Abschnitt 1 beschriebenen Mechanismen melden. Unsere Konsultation von Stakeholdern umfasst die aktive Einbindung von Nichtregierungsorganisationen, Regierungen, einzelnen Aktivist\*innen und der Wissenschaft.

Neben der Einhaltung unserer Gemeinschaftsrichtlinien verfügen wir über ein Responsible AI-Team (Team für verantwortungsbewusste künstliche Intelligenz) – ein multidisziplinäres Team aus Ethiker\*innen, Sozial- und Politikwissenschaftler\*innen, Politikexpert\*innen und KI-Forscher\*innen sowie -Ingenieur\*innen. Das übergeordnete Ziel dieses Teams besteht darin, Richtlinien, Tools und Prozesse zu entwickeln, um Fragen der KI-Verantwortlichkeit anzugehen und dazu beizutragen, dass diese systemischen Ressourcen für Meta umfassend verfügbar sind. Auf diese Weise wollen wir den mit dem Einsatz von künstlicher Intelligenz in Meta-Technologien verbundenen Fragen hinsichtlich Fairness und Inklusion begegnen. Schließlich hat das unabhängige [Oversight Board](#) die Möglichkeit, Empfehlungen zur Durchsetzung unserer Richtlinien abzugeben, die auch die automatisierte Durchsetzung betreffen können.

## 2D. UNTERSTÜTZUNG UND ZUGANG FÜR KREISE DER WISSENSCHAFT UND FORSCHUNG

Meta ist davon überzeugt, dass Offenheit und Kooperation mit Kreisen der Wissenschaft die Forschung und Entwicklung anregen, neue Wege zur Erkennung und Vermeidung gefährlicher Inhalte eröffnen und zum Schutz der Nutzer\*innen beitragen. Wir arbeiten mit diesen Kreisen auf verschiedene Weise zusammen – im Rahmen der Beschränkungen zum Schutz der Daten unserer Nutzer\*innen. Wir zeigen bei unserer Moderation von Inhalten in drei wichtigen Punkten Offenheit und Transparenz: (1) den Algorithmen, (2) den individuellen Ergebnissen dieser Algorithmen und (3) den Gesamtergebnissen unserer Bemühungen zur Inhaltsmoderation.

Seit der Gründung unseres Facebook KI-Forschungslabors (Facebook AI research (FAIR) lab) (heute: Meta AI) im Jahr 2013 haben wir uns zu einem Open-Science-Ansatz verpflichtet. Unser Forschungsmodell dreht sich um die Veröffentlichung von Quellcode und Methoden, die Zusammenarbeit mit anderen Forscher\*innen aus Industrie und Wissenschaft und um die Schaffung offener Benchmarks und Challenges. Darüber hinaus veröffentlichen uns angegliederte Forscher\*innen häufig die Ergebnisse unserer Quellcodes und Methoden angewendet auf unseren Plattformen. Hier einige Beispiele unserer Methoden und Technologien, die in den letzten Jahren veröffentlicht wurden:

- [XLM-R](#), [LInformer](#) und [RoBERTa](#): Wir haben unsere Modelle und unseren Code als Open Source zur Verfügung gestellt, sodass die Forschungsgemeinde unsere Machine-Learning-Modelle zum Verständnis natürlicher Sprache, insbesondere solche mit mehrsprachigem Sprachverständnis, bewerten kann.
- [Faiss](#), [PDQ](#) und [TMK+PDQF](#): Wir haben unsere Forschung veröffentlicht und den Code für drei unserer Algorithmen freigegeben, die für den Abgleich identischer und nahezu identischer Inhalte mit bereits bekannten Foto- und Videoinhalten eingesetzt werden.



- [CLARA](#) (Confidence of Labels and Raters): Wir haben die Methodik und Evaluierung veröffentlicht, die wir zur Verbesserung der Präzision der Einstufungen menschlicher Prüfer\*innen einsetzen.

Informationszugang zum Zwecke der Auswertung der Verfahren zur automatisierten Erkennung von Inhalten speziell auf Grundlage des neuen § 5a Abs. 2 NetzDG wurde im Berichtszeitraum nicht gewährt.

Wir veröffentlichen unsere Gesamtzahlen zur Durchsetzung von diversen Bereichen der Facebook-Gemeinschaftsstandards und der Instagram-Gemeinschaftsrichtlinien in unserem vierteljährlichen [Bericht zur Durchsetzung der Gemeinschaftsstandards](#) (Community Standards Enforcement Report (CSER)). Darüber hinaus veröffentlichen wir für viele Verstöße Ergebnisse von Messungen der Verbreitung. Diese Verbreitungsmessung zeigt an, wie viele Aufrufe unzulässiger Inhalte wir nicht verhindert haben – entweder weil wir nicht früh genug gegen die Verstöße vorgegangen sind oder weil sie uns auf Facebook oder Instagram gänzlich entgangen sind. Nähere Angaben zur Messung der Verbreitung sind [hier](#) verfügbar. Allerdings ist Transparenz nur dann hilfreich, wenn die von uns zur Verfügung gestellten Informationen relevant und präzise sind. Im Zusammenhang mit dem Bericht zur Durchsetzung der Gemeinschaftsstandards bedeutet dies, dass die von uns berichteten Kennzahlen auf einer soliden Methodik beruhen und präzise widerspiegeln, was auf unserer Plattform passiert. Zu diesem Zweck haben wir mit internationalen Expert\*innen für Messung, Statistik, Recht, Wirtschaft und Governance zusammengearbeitet, damit diese eine unabhängige, öffentliche Bewertung darüber abgeben, ob die von uns in dem Bericht zur Durchsetzung der Gemeinschaftsstandards angegebenen Kennzahlen präzise und nützliche Messgrößen darstellen für die Herausforderungen im Zusammenhang mit unserer Inhaltsmoderation und den von uns ergriffenen Maßnahmen zu deren Bewältigung. Die Expert\*innen waren sich weitgehend einig, dass wir die richtigen Kennzahlen analysieren, und sie haben einige Empfehlungen zu Verbesserungen abgegeben. Der vollständige Bericht ist [hier](#) verfügbar.

Im August 2020 haben wir uns außerdem verpflichtet, eine unabhängige Bewertung unseres Berichts zur Durchsetzung der Gemeinschaftsstandards durch Dritte vornehmen zu lassen und zu veröffentlichen. In diesem Jahr sind wir dieser Verpflichtung durch [Veröffentlichung der unabhängigen Ergebnisse von EY](#) nachgekommen.

### 3. BESCHWERDEMECHANISMEN/ENTSCHEIDUNGSKRITERIEN

[Darstellung der Mechanismen zur Übermittlung von Beschwerden über rechtswidrige Inhalte, Darstellung der Entscheidungskriterien für die Entfernung und Sperrung von rechtswidrigen Inhalten und Darstellung des Prüfungsverfahrens einschließlich der Reihenfolge der Prüfung, ob ein rechtswidriger Inhalt vorliegt oder ob gegen vertragliche Regelungen zwischen Anbieter und Nutzer\\*in verstoßen wird.](#)

Wenn jemand der Ansicht ist, dass Inhalte auf Instagram gemäß einer oder mehrerer der im NetzDG aufgeführten Bestimmungen des deutschen Strafgesetzbuchs rechtswidrig sind, kann die Person diese Inhalte über das [NetzDG-Beschwerdeformular](#) von Instagram melden. Dieses Formular ist ausschließlich für Beschwerden wegen Verstößen gegen die im NetzDG aufgeführten Bestimmungen des deutschen Strafgesetzbuchs vorgesehen und steht seit dem 1. Januar 2018 für Nutzer\*innen in Deutschland zur Verfügung.



Das NetzDG-Beschwerdeformular ist über einen Link, der neben dem Inhalt auf der Instagram-Plattform verfügbar ist, erreichbar.

Damit wir NetzDG-Beschwerden korrekt bewerten und § 2 des NetzDG (die „Berichtspflicht“) erfüllen können, erscheint die Aufforderung, folgende Informationen in unserem NetzDG-Beschwerdeformular anzugeben:

- vollständige Kontaktinformationen, einschließlich der Angabe, ob der/die Beschwerdeführer\*in eine Beschwerdestelle im Sinne des NetzDG ist und/oder den Inhalt im Auftrag eines/einer Mandant\*in meldet,
- den/die Paragraphen des deutschen Strafgesetzbuchs, der/die angeblich durch den gemeldeten Inhalt verletzt wird/werden,
- die Gründe, warum diese Aussagen oder Bilder angeblich „rechtswidrig“ im Sinne des NetzDG sind,
- ggf. eine gerichtliche Entscheidung.

Da Nutzer\*innen das Beschwerdeformular direkt von einem Inhalt aus erreichen können, ist es nicht notwendig, Links zum gemeldeten Inhalt (sofern der/die meldende Nutzer\*in bei seinem/ihrem Instagram-Konto angemeldet ist) anzugeben. Das entsprechende Feld wird automatisch ausgefüllt. In dem Beschwerdeformular wird außerdem deutlich hervorgehoben, welche Informationen optional abgefragt werden.

Alle Nutzer\*innen in Deutschland (also sowohl registrierte als auch nicht registrierte Nutzer\*innen) können auch auf anderem Wege zum NetzDG Beschwerdeformular gelangen: Das Beschwerdeformular ist nämlich zusätzlich im Instagram-Hilfereich über eine gesonderte Seite, die [NetzDG-Hilfereich-Seite](#), erreichbar, die darüber informiert, wie man NetzDG-Beschwerden einreichen kann. Auf dieser Seite erreicht man unser Beschwerdeformular, indem man auf den Button „Meldung einreichen“ klickt. Zur NetzDG-Hilfereich-Seite selbst können die Nutzer\*innen auf verschiedenen Wegen gelangen. Die NetzDG-Hilfereich-Seite steht im [Instagram-Hilfereich](#) unter dem Tab „Nutzungsbedingungen und Richtlinien“, Unterpunkt „Netzwerkdurchsetzungsgesetz („NetzDG“)" zur Verfügung. Darüber hinaus ist sie über die Hilfereich-Suchleiste zu finden. Eine weitere Möglichkeit, die NetzDG-Hilfereich-Seite zu erreichen, ist, auf der Instagram Startseite auf den Link „NetzDG/UrhDaG/Ranking von Inhalten“ zu klicken und anschließend im Abschnitt „Netzwerkdurchsetzungsgesetz („NetzDG“)" den Link „NetzDG-Hilfereich“ auszuwählen.

Das über die NetzDG-Hilfereich-Seite erreichbare Beschwerdeformular ist an das direkt am Inhalt zugängliche Formular angeglichen. Der einzige Unterschied besteht darin, dass Nutzer\*innen, die eine Beschwerde nicht über das direkt am Inhalt verfügbare NetzDG-Beschwerdeformular einreichen, (zusätzlich zu den vorgenannten Punkten) gebeten werden, den Link zu dem konkreten Inhalt auf Instagram, der Gegenstand der Beschwerde ist, anzugeben.

Sobald jemand das NetzDG-Beschwerdeformular (über einen der zuvor beschriebenen Wege) ausgefüllt hat und auf „Senden“ klickt, erhalten wir die Beschwerde automatisch über ein internes Prüf-Tool, ohne dass eine weitere Handlung des/der Beschwerdeführer\*in erforderlich ist. Wir verfolgen dann einen zweistufigen Ansatz zur Überprüfung von Inhalten, die uns über das NetzDG-Beschwerdeformular gemeldet werden. Zunächst prüfen wir den gemeldeten Inhalt auf eine Verletzung unserer Gemeinschaftsrichtlinien hin. Wenn er gegen unsere Gemeinschaftsrichtlinien verstößt, stellen wir sicher, dass er weltweit von der Instagram-Plattform entfernt wird. Wenn



der gemeldete Inhalt nicht gegen unsere Gemeinschaftsrichtlinien verstößt, prüfen wir anschließend anhand der in der Beschwerde bereitgestellten Informationen seine Rechtmäßigkeit. Konkret prüfen wir, ob der gemeldete Inhalt gegen die entsprechenden im NetzDG aufgeführten Bestimmungen des deutschen Strafgesetzbuchs verstößt. Wenn wir den gemeldeten Inhalt als rechtswidrig im Sinne des NetzDG erachten, sperren wir den Zugang zu dem Inhalt in Deutschland (siehe Abschnitt 5 für weitere Informationen zu unserem Umgang mit NetzDG-Beschwerden). Außerdem kommunizieren wir mit dem/der Beschwerdeführer\*in, um ihm/ihr Updates zu seiner/ihrer Beschwerde zu geben, ggf. zusätzliche Informationen anzufordern und ihn/sie über unsere Entscheidung zu informieren, sobald wir die Überprüfung der Beschwerde abgeschlossen haben (siehe Abschnitt 10 für weitere Informationen hierzu).

## 4. BESCHWERDEAUFKOMMEN

Anzahl der im Berichtszeitraum eingegangenen Beschwerden über rechtswidrige Inhalte, aufgeschlüsselt nach Beschwerden von Beschwerdestellen und Beschwerden von Nutzer\*innen und nach dem Beschwerdegrund.

Die nachfolgenden Tabellen zeigen die Anzahl der zwischen 1. Juli 2022 und 31. Dezember 2022 über Instagrams NetzDG-Beschwerdeformular übermittelten Beschwerden. Pro NetzDG-Beschwerde kann grundsätzlich nur ein Inhalt gemeldet werden. In einigen Fällen führen Nutzer\*innen jedoch mehrere Inhalte in einer einzigen NetzDG-Beschwerde an (z. B. durch Nennung mehrerer URLs in einem Anhang, der zur Beschwerde hochgeladen wurde). Die in den nachfolgenden Tabellen dargestellten Zahlen beziehen sich auf übermittelte Beschwerden und nicht auf einzelne in den Beschwerden beanstandete Inhalte. Erwähnenswert ist, dass im Zeitraum zwischen 1. Juli 2022 und 31. Dezember 2022 57.541 NetzDG-Beschwerden eingegangen sind, in denen insgesamt 36.990 Inhalte genannt wurden. Wenn uns ein und derselbe Inhalt mehrfach gemeldet wurde, haben wir ihn als einen Inhalt gezählt.

Die Tabellen enthalten zwei unterschiedliche Kategorien von Zahlen:

### 4A. NETZDG-BESCHWERDEN, AUFGESCHLÜSSELT NACH ART DES/DER BESCHWERDEFÜHRER\*IN

- In diesem Abschnitt ist die Anzahl der Beschwerden nach Beschwerden von Beschwerdestellen und von Nutzer\*innen aufgeschlüsselt.

Tabelle 1. NetzDG-Beschwerden, aufgeschlüsselt nach Art des/der Beschwerdeführer\*in

Beschwerden von Beschwerdestellen	14.864
Beschwerden von Nutzer*innen	42.677
Gesamt	57.541

### 4B. NETZDG-BESCHWERDEN, AUFGESCHLÜSSELT NACH ANGEFÜHRTEN PARAGRAFEN DES STRAFGESETZBUCHS

- In diesem Abschnitt ist die Anzahl der Beschwerden nach den Bestimmungen des



Strafgesetzbuchs aufgeschlüsselt, die durch den/die Beschwerdeführer\*in angeführt werden.

- Dabei ist zu beachten, dass in einer NetzDG-Beschwerde möglicherweise mehrere Gründe für die Rechtswidrigkeit angeführt werden. Daher übersteigt die Gesamtzahl der Beschwerden, die in der nachfolgenden Tabelle in den Zeilen pro Beschwerdeführertyp aufgeführt sind, die Gesamtzahl der in Abschnitt 4A pro Beschwerdeführertyp übermittelten Beschwerden.

Tabelle 2. NetzDG-Beschwerden, aufgeschlüsselt nach angeführten Paragraphen des Strafgesetzbuchs

Paragraf des Strafgesetzbuchs	Beschwerdeführer*in		
	Beschwerden von Beschwerdestellen	Beschwerden von Nutzer*innen	Gesamt
Verbreiten von Propagandamitteln verfassungswidriger und terroristischer Organisationen (§ 86)	4.958	8.066	13.024
Verwenden von Kennzeichen verfassungswidriger und terroristischer Organisationen (§ 86a)	3.848	5.424	9.272
Vorbereitung einer schweren staatsgefährdenden Gewalttat (§ 89a)	2.839	3.226	6.065
Anleitung zur Begehung einer schweren staatsgefährdenden Gewalttat (§ 91)	2.578	3.038	5.616
Landesverräterische Fälschung (§ 100a)	2.584	3.195	5.779
Öffentliche Aufforderung zu Straftaten (§ 111)	2.851	5.707	8.558
Störung des öffentlichen Friedens durch Androhung von Straftaten (§ 126)	2.809	5.017	7.826
Bildung krimineller Vereinigungen (§ 129)	2.842	4.718	7.560
Bildung terroristischer Vereinigungen (§ 129a)	2.401	2.703	5.104
Bildung krimineller und/oder terroristischer Vereinigungen im Ausland (§ 129b)	2.462	2.681	5.143
Volksverhetzung (§ 130)	4.754	12.120	16.874
Gewaltdarstellung (§ 131)	3.113	5.568	8.681
Belohnung und Billigung von Straftaten (§ 140)	2.623	6.140	8.763
Beschimpfung von Bekenntnissen, Religionsgesellschaften und Weltanschauungsvereinigungen (§ 166)	4.131	10.268	14.399
Verbreitung, Erwerb und Besitz kinderpornographischer Inhalte (§ 184b)	2.794	3.535	6.329



Paragraf des Strafgesetzbuchs	Beschwerdeführer*in		
	Beschwerden von Beschwerdestellen	Beschwerden von Nutzer*innen	Gesamt
Beleidigung (§ 185)	5.304	14.124	19.428
Üble Nachrede (§ 186)	4.659	11.100	15.759
Verleumdung (§ 187)	3.776	8.787	12.563
Verunglimpfung des Andenkens Verstorbener (§ 189)	2.713	4.203	6.916
Verletzung des höchstpersönlichen Lebensbereichs und von Persönlichkeitsrechten durch Bildaufnahmen (§ 201a)	3.440	5.236	8.676
Bedrohung (§ 241)	3.041	6.385	9.426
Fälschung beweisheblicher Daten (§ 269)	3.167	4.806	7.973

## 5. ORGANISATION, PERSONELLE AUSSTATTUNG, KOMPETENZEN, SCHULUNG UND BETREUUNG ZUSTÄNDIGER MITARBEITER\*INNEN

Organisation, personelle Ausstattung, fachliche und sprachliche Kompetenz der für die Bearbeitung von Beschwerden zuständigen Arbeitseinheiten und Schulung und Betreuung der für die Bearbeitung von Beschwerden zuständigen Personen.

### 5A. ORGANISATION

NetzDG-Beschwerden werden in zwei Schritten durch Teams bestehend aus geschulten Fachkräften und Jurist\*innen geprüft, die sowohl für die Facebook- als auch für die Instagram-Plattform zuständig sind.

Zunächst werden die über das NetzDG-Beschwerdeformular von Instagram gemeldeten Inhalte durch Mitglieder unseres Global Operations-Teams geprüft. Unser Global Operations-Team setzt sich aus Vollzeitmitarbeiter\*innen sowie Mitarbeiter\*innen von Unternehmen zusammen, mit denen wir zusammenarbeiten. Jede NetzDG-Beschwerde wird durch ein Mitglied des Global Operations-Teams geprüft. Diese Prüfung erfolgt bei unseren Vertragspartnern Majorel in Berlin und Telus/CCC in Essen. Dort wird ermittelt, ob die gemeldeten Inhalte gegen die Instagram-Gemeinschaftsrichtlinien verstoßen (im Gegensatz zu einer Überprüfung des Inhalts auf eine potenzielle Rechtswidrigkeit, die, wie weiter unten erläutert, durch andere Teams durchgeführt wird). Wenn die Überprüfung ergibt, dass ein Inhalt gegen die Gemeinschaftsrichtlinien verstößt, wird er weltweit entfernt.



Sämtliche NetzDG-Beschwerden zu Inhalten, die nicht wegen Verstößen gegen unsere Gemeinschaftsrichtlinien entfernt wurden, werden anschließend einem rechtlichen Prüfprozess unterzogen, der aus mehreren Stufen bestehen kann (siehe unten) und durch unser Scaled Regulatory Operations-Team durchgeführt wird.

Dieses Team umfasst zwei Gruppen – eine aus Mitarbeiter\*innen in Irland und Sunnyvale, Kalifornien bestehende Gruppe (inklusive einer Mitarbeiterin, welche remote aus Austin, Texas arbeitet) und eine Gruppe, die sich aus Mitarbeiter\*innen von Partnerunternehmen in Dublin und Austin, Texas zusammensetzt. NetzDG-Beschwerden zu Inhalten, die nicht bereits wegen Verstößen gegen unsere Gemeinschaftsrichtlinien entfernt wurden, werden zunächst von einem/einer Mitarbeiter\*in unserer Partnerunternehmen geprüft. Zweck dieser Stufe der Überprüfung ist sicherzustellen, dass offensichtlich rechtswidrige Inhalte innerhalb von 24 Stunden gesperrt werden. Jede Beschwerde wird durch ein Mitglied dieses Teams geprüft, dessen Aufgabe darin besteht, offensichtlich rechtswidrige Inhalte zu identifizieren und zu sperren sowie mit dem/der Beschwerdeführer\*in zu kommunizieren, unter anderem wenn weitere Informationen zur Bearbeitung der Beschwerde notwendig sind. All dies erfolgt nach Maßgabe von Orientierungshilfen, die durch unsere unternehmensinternen Jurist\*innen und externe Rechtsanwält\*innen entwickelt wurden. Sollte eine tiefergehende Untersuchung der Beschwerde erforderlich sein, so wird die Beschwerde zur Überprüfung an einen/eine unserer Mitarbeiter\*innen aus dem Scaled Regulatory Operations-Team weitergeleitet. Dieser/diese Mitarbeiter\*in prüft die Beschwerde sorgfältig und ergreift geeignete Maßnahmen, falls auf der Grundlage der oben genannten Orientierungshilfen die Rechtmäßigkeit oder Rechtswidrigkeit festgestellt werden kann.

Wenn die Rechtswidrigkeit eines gemeldeten Inhalts weiterhin unklar ist, wird die Beschwerde zur Überprüfung an unsere unternehmensinternen Jurist\*innen weitergeleitet. Bei besonders komplexen Sachverhalten holen unsere unternehmensinternen Jurist\*innen eine rechtliche Beurteilung von externen deutschen Rechtsanwält\*innen ein.

Sobald eine Entscheidung über den Umgang mit dem Inhalt getroffen worden ist, ergreift das Scaled Regulatory Operations-Team die empfohlenen Maßnahmen bezüglich des Inhalts und kommuniziert direkt mit dem/der Beschwerdeführer\*in und (im Falle einer Sperrung) mit dem/der Nutzer\*in, der/die den gemeldeten Inhalt gepostet hat.

In jeder Stufe des Prüfprozesses findet eine offene Kommunikation zwischen den Personen statt, die die Inhalte prüfen.

Im Global Operations-Team arbeiten die Teams von Majorel und Telus/CCC eng mit den Kolleg\*innen des Global Operations-Teams in Dublin zusammen. Die Teams von Majorel und Telus/CCC erhalten die erforderlichen Schulungen, zusätzliche Hilfestellungen und Fachwissen zu Gemeinschaftsrichtlinien in Verbindung mit NetzDG-Fällen von speziell darauf geschulten Partnerunternehmensmitarbeiter\*innen, die ihrerseits durch verschiedene Meta-Teams geschult werden.

Die Partnerunternehmensmitarbeiter\*innen im Scaled Regulatory Operations-Team arbeiten eng mit unseren Mitarbeiter\*innen des Scaled Regulatory Operations-Teams zusammen, die Schulungen, Anleitungen und Hilfestellungen bei schwierigen oder außergewöhnlichen Beschwerden bereitstellen. Eine ähnliche Kommunikation findet auch zwischen den Mitarbeiter\*innen des Scaled Regulatory Operations-Teams und einem Team aus unseren



unternehmensinternen Jurist\*innen statt. Diese beiden Teams treffen sich mehrmals pro Woche. So stehen sie in ständigem Austausch miteinander, um rechtlich komplexe NetzDG-Beschwerden zu diskutieren. Jeden Monat senden Mitarbeiter\*innen des Scaled Regulatory Operations-Team interne Updates an ein funktionsübergreifendes Team. Diese Updates enthalten Informationen zu Beschwerdeaufkommen, Trends, eventuellen Fragen, die bei der Bearbeitung von Beschwerden aufgetreten sind, sowie zu Trainings unserer Teams.

## 5B. PERSONELLE AUSSTATTUNG UND KOMPETENZEN

Stand 31. Dezember 2022 stehen in drei Teams 178 Personen zur Bearbeitung von NetzDG-Beschwerden zur Verfügung und wurden entsprechend geschult (wie in Abschnitt 5C näher erläutert). Sie sind auch in anderen Bereichen als der Bearbeitung von NetzDG-Beschwerden tätig, was einen flexiblen Personaleinsatz ermöglicht. Bei einem geringen Beschwerdeaufkommen bearbeitet nur ein kleiner Teil dieser Mitarbeiter\*innen NetzDG-Beschwerden. Wenn dagegen die Anzahl steigt, können auch zusätzliche geschulte Teammitglieder für die Bearbeitung der Beschwerden eingesetzt werden. Wir geben nachfolgend einen Überblick über die personelle Ausstattung und die Kompetenzen der einzelnen Teams:

- **Global Operations-Team:** Am 31. Dezember 2022 standen in diesem Team 130 Personen zur Bearbeitung von NetzDG-Beschwerden zur Verfügung. Bevor ihnen diese Aufgabe anvertraut wurde, mussten alle diese Personen mindestens 3-6 Monate lang ihre Kompetenz bei der Bearbeitung anderer Arten von Anfragen zur Entfernung von Inhalten unter Beweis stellen. Außerdem sprechen alle fließend Deutsch und haben eine NetzDG-Schulung erhalten (weitere Erläuterungen hierzu in Abschnitt 5C unten). Wie bereits oben erwähnt, sind sie auch in anderen Bereichen der Moderation von Inhalten als der Bearbeitung von NetzDG-Beschwerden tätig, so dass die Arbeitslast je nach NetzDG-Beschwerdeaufkommen umverteilt werden kann.
- **Scaled Regulatory Operations-Team:** Am 31. Dezember 2022 waren insgesamt 46 zur Bearbeitung von NetzDG-Beschwerden berechnigte Personen im Scaled Regulatory Operations-Team tätig. Davon waren 31 Mitarbeiter\*innen von Partnerunternehmen und 15 unsere Mitarbeiter\*innen. Die 31 Partnerunternehmensmitarbeiter\*innen sprechen fließend Deutsch und haben eine NetzDG-Schulung erhalten (weitere Erläuterungen hierzu in Abschnitt 5C unten). Die zur Bearbeitung von NetzDG-Beschwerden berechtigten Mitarbeiter\*innen des Scaled Regulatory Operations-Teams beherrschen eine Vielzahl von Sprachen fließend, darunter Deutsch, Englisch, Französisch und Türkisch. Für die Bearbeitung von Beschwerden in anderen Sprachen, z. B. Bosnisch, Polnisch, Spanisch, Russisch und Niederländisch, erhalten sie Unterstützung durch andere Teammitglieder. Grundsätzlich werden Beschwerden durch Teammitglieder bearbeitet, die sowohl Deutsch als auch Englisch fließend beherrschen. Hierbei gibt es vereinzelt Ausnahmen, in denen Beschwerden in anderen Sprachen eingereicht werden oder in denen der beanstandete Inhalt keine Kenntnis der deutschen Sprache erfordert. Um sowohl den rechtlichen als auch den operativen Komplexitäten von NetzDG-Beschwerden Rechnung zu tragen, kommen unsere 15 Mitarbeiter\*innen aus unterschiedlichen Fachbereichen. 5 Teammitglieder haben einen juristischen Hochschulabschluss, und 14 Mitarbeiter\*innen brachten operative Erfahrung mit in ihre Rolle. Alle haben eine NetzDG-Schulung erhalten (weitere Erläuterungen hierzu in Abschnitt 5C unten).
- **Jurist\*innen:** Am 31. Dezember 2022 waren 2 unternehmensinterne Jurist\*innen in die Bearbeitung von NetzDG-Beschwerden eingebunden (neben anderen Tätigkeiten). Diese 2 Jurist\*innen sind Spezialist\*innen für die Beurteilung potenziell problematischer Inhalte und haben umfassende Erfahrung im Umgang mit rechtlichen Anfragen zur Entfernung von



Inhalten. Sie arbeiten eng mit Mitarbeiter\*innen des Scaled Regulatory Operations-Teams zusammen und stehen in regelmäßigem Kontakt mit externen deutschen Rechtsanwält\*innen, die Beratung zu konkreten NetzDG-Beschwerden bieten.

## 5C. SCHULUNG UND BETREUUNG

Die Teams, die NetzDG-Beschwerden bearbeiten, erhalten entsprechend ihrer jeweiligen Tätigkeiten verschiedene Arten von Schulungen .

Da die Mitglieder des Global Operations-Teams, die NetzDG-Beschwerden auf Instagram prüfen, diese Beschwerden nur im Hinblick auf Verstöße gegen die Gemeinschaftsrichtlinien überprüfen, steht im Mittelpunkt ihrer Schulungen die Entwicklung von operativen Fähigkeiten und Kompetenzen zur Umsetzung der Gemeinschaftsrichtlinien (und nicht die Schulung zur Beurteilung der Rechtmäßigkeit von Inhalten). Sie müssen eine mehrwöchige Schulung in der Überprüfung von Inhalten nach Maßgabe der Instagram-Gemeinschaftsrichtlinien absolvieren. Wie bereits oben in Abschnitt 5B erwähnt, werden NetzDG-Beschwerden nur von einer ausgewählten Gruppe von Mitgliedern des Global Operations-Teams bearbeitet. Diese müssen bereits eine bestimmte Zeit im Team sein und ein konstant hohes Maß an operativer Kompetenz zeigen, bevor sie für diese Arbeit infrage kommen. Bevor sie diese Rolle einnehmen, erhalten sie Schulungen mit Hintergrundinformationen zum NetzDG, die ihnen das erforderliche Verständnis für ihre Tätigkeit vermitteln.

Die Mitglieder des Scaled Regulatory Operations-Teams prüfen Inhalte auf ihre Rechtswidrigkeit und werden deshalb intensiver im Hinblick auf das NetzDG und die Bestimmungen des deutschen Strafgesetzbuchs geschult, die das Gesetz als „rechtswidrige Inhalte“ definiert. Jedes zur Bearbeitung von NetzDG-Beschwerden berechnigte Mitglied dieses Teams erhält mehrwöchige Schulungen mit Schwerpunkt auf operativer Kompetenz und NetzDG vor Aufnahme der Prüfungstätigkeit an NetzDG-Beschwerden. Die speziellen NetzDG-Schulungen umfassen Hintergrundmaterial zum Gesetz, detaillierte Aufschlüsselungen zu jeder im NetzDG angeführten Bestimmung des Strafgesetzbuchs sowie Anleitungen für die Kommunikation mit Beschwerdeführer\*innen und Nutzer\*innen, deren Inhalt gemeldet wurde. Außerdem erhalten die Teammitglieder mindestens einmal pro Halbjahr Auffrischungsschulungen. Diese beinhalten Wiederholungen und Aktualisierungen zu operativen Best Practices sowie Aufschlüsselungen gängiger Beschwerdearten und häufig angeführter Bestimmungen des Strafgesetzbuchs. Die Auffrischungsschulungen für Partnerunternehmensmitarbeiter\*innen werden durch Mitarbeiter\*innen aus dem Scaled Regulatory Operations-Team durchgeführt. Die Auffrischkurse für die Mitarbeiter\*innen des Scaled Regulatory Operations-Teams erfolgen durch unsere unternehmensinternen Jurist\*innen.

Es besteht außerdem ein umfassendes und vielfältiges Programm zur Unterstützung der Mitglieder unserer Global Operations- und Scaled Regulatory Operations-Teams, die NetzDG-Beschwerden prüfen. Dieses Programm bietet Teammitgliedern aktuell z.B. folgende Dienstleistungen an:

- Psychologische Betreuung
- Einzelsitzungen mit einem/einer unternehmensinternen Vollzeitpsycholog\*in
- Gruppentherapie
- Vor-Ort-Beratungsgespräche zur Förderung des emotionalen Wohlbefindens



Wir engagieren uns dafür, Menschen, die Inhalte überprüfen, ein qualitativ hochwertiges und vielfältiges Betreuungsprogramm zur Verfügung zu stellen, und werden auch in Zukunft neue Angebote hinzufügen und bestehende Angebote verbessern.

## 6. BRANCHENVERBÄNDE

Mitgliedschaft in Branchenverbänden mit Hinweis darauf, ob in diesen Branchenverbänden eine Beschwerdestelle existiert.

Unternehmen des Meta Konzerns sind Mitglied in folgenden Branchenverbänden in Deutschland:

- Eco – Verband der Internetwirtschaft e.V. (hier existiert eine interne Beschwerdestelle)
- Bitkom – Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V. (hier existiert keine interne Beschwerdestelle)
- BVDW – Bundesverband Digitale Wirtschaft e.V. (hier existiert keine interne Beschwerdestelle)

## 7. EXTERNE KONSULTATION

Anzahl der Beschwerden, bei denen eine externe Stelle konsultiert wurde, um die Entscheidung vorzubereiten.

Im Zeitraum zwischen 1. Juli 2022 und 31. Dezember 2022 haben wir 8 NetzDG-Beschwerden an die Freiwillige Selbstkontrolle Multimedia-Diensteanbieter e.V. (FSM) zur Entscheidung über die Rechtmäßigkeit der gemeldeten Inhalte weitergeleitet.

In 2 der 57.541 Beschwerden, die zwischen 1. Juli 2022 und 31. Dezember 2022 eingegangen sind, haben wir externe Rechtsanwält\*innen konsultiert (die wir nicht als „externe Stelle“ sondern als Erweiterung unserer Rechtsabteilung betrachten), um die Entscheidung hinsichtlich der Beschwerden vorzubereiten. Alle diese Konsultationen erfolgten mit externen deutschen Rechtsanwält\*innen, mit denen wir bei Bedarf zusammenarbeiten, um die Rechtslage im Hinblick auf konkrete Beschwerden zu klären.

## 8. ANZAHL VON ENTFERNUNGEN/SPERRUNGEN

Anzahl der Beschwerden, die im Berichtszeitraum zur Löschung oder Sperrung des beanstandeten Inhalts führten, nach der Gesamtzahl sowie aufgeschlüsselt nach Beschwerden von Beschwerdestellen und von Nutzer\*innen, nach dem Beschwerdegrund, ob ein Fall des § 3 Abs. 2 Nr. 3 Buchstabe a vorlag, ob in diesem Fall eine Weiterleitung an den/die Nutzer\*in erfolgte, welcher Schritt der Prüfungsreihenfolge nach Nummer 3 zur Entfernung oder Sperrung geführt hat sowie ob eine Übertragung an eine anerkannte Einrichtung der Regulierten Selbstregulierung nach § 3 Abs. 2 Nr. 3 Buchstabe b erfolgte.

Die nachfolgende Tabelle stellt dar, wie oft Inhalte entfernt oder gesperrt wurden, nachdem zwischen 1. Juli 2022 und 31. Dezember 2022 Beschwerden über Instagrams NetzDG-



Beschwerdeformular übermittelt wurden. Dabei ist Folgendes hinsichtlich dieser Tabelle zu beachten:

- In dieser Tabelle ist aufgeschlüsselt, wie häufig eine Beschwerde gemäß den durch den/die Beschwerdeführer\*in angeführten Bestimmungen des deutschen Strafgesetzbuchs zur Entfernung oder Sperrung von Inhalten geführt hat.
- Pro NetzDG-Beschwerde kann grundsätzlich nur ein Inhalt gemeldet werden. In einigen Fällen führen Nutzer\*innen jedoch mehrere Inhalte in einer einzigen NetzDG-Beschwerde an (z. B. durch Nennung mehrerer URLs in einem Anhang, der zur Beschwerde hochgeladen wurde). Die in den nachfolgenden Tabellen dargestellten Zahlen beziehen sich auf übermittelte Beschwerden und nicht auf einzelne in den Beschwerden beanstandete Inhalte. Erwähnenswert ist, dass im Zeitraum zwischen 1. Juli 2022 und 31. Dezember 2022 5.141 NetzDG-Beschwerden zu einer Entfernung oder Sperrung von Inhalten führten. Dies machte insgesamt 4.273 entfernte oder gesperrte Inhalte aus. Die Gesamtzahl der entfernten oder gesperrten Inhalte ist niedriger als die Gesamtzahl der NetzDG-Beschwerden, da derselbe Inhalt teilweise in mehreren NetzDG-Beschwerden gemeldet wurde und wir in diesen Fällen bei der Statistik der Entfernung/Sperrung von Inhalten den mehrfach gemeldeten Inhalt nur einmal gezählt haben.
- Nutzer\*innen können mehrere Gründe für Rechtswidrigkeit in einer einzigen NetzDG-Beschwerde anführen. Wenn wir infolge einer Beschwerde Maßnahmen in Bezug auf einen Inhalt ergriffen haben, wird der entsprechende Inhalt in der Tabelle unter jeder in der Beschwerde genannten Bestimmung aufgeführt. Daher übersteigt die Gesamtzahl der in der nachfolgenden Tabelle aufgeführten Entfernungen/Sperrungen die Gesamtzahl der Beschwerden, die zur Entfernung oder Sperrung von Inhalten führten.

Tabelle 3. Anzahl von Beschwerden, die zu Entfernungen/Sperrungen führten

Paragraf des Strafgesetzbuchs	Beschwerdeführer*in		
	Beschwerden von Beschwerdestellen	Beschwerden von Nutzer*innen	Gesamt
Verbreiten von Propagandamitteln verfassungswidriger und terroristischer Organisationen (§ 86)	209	574	783
Verwenden von Kennzeichen verfassungswidriger und terroristischer Organisationen (§ 86a)	174	586	760
Vorbereitung einer schweren staatsgefährdenden Gewalttat (§ 89a)	151	230	381
Anleitung zur Begehung einer schweren staatsgefährdenden Gewalttat (§ 91)	134	241	375
Landesverräterische Fälschung (§ 100a)	143	393	536
Öffentliche Aufforderung zu Straftaten (§ 111)	165	798	963
Störung des öffentlichen Friedens durch Androhung von Straftaten (§ 126)	157	475	632
Bildung krimineller Vereinigungen (§ 129)	165	501	666
Bildung terroristischer Vereinigungen (§ 129a)	118	158	276



Paragraf des Strafgesetzbuchs	Beschwerdeführer*in		
	Beschwerden von Beschwerdestellen	Beschwerden von Nutzer*innen	Gesamt
Bildung krimineller und/oder terroristischer Vereinigungen im Ausland (§ 129b)	143	185	328
Volksverhetzung (§ 130)	237	1.155	1.392
Gewaltdarstellung (§ 131)	165	563	728
Belohnung und Billigung von Straftaten (§ 140)	162	744	906
Beschimpfung von Bekenntnissen, Religionsgesellschaften und Weltanschauungsvereinigungen (§ 166)	194	851	1.045
Verbreitung, Erwerb und Besitz kinderpornographischer Inhalte (§ 184b)	167	492	659
Beleidigung (§ 185)	275	1.227	1.502
Üble Nachrede (§ 186)	231	850	1.081
Verleumdung (§ 187)	206	679	885
Verunglimpfung des Andenkens Verstorbener (§ 189)	140	388	528
Verletzung des höchstpersönlichen Lebensbereichs und von Persönlichkeitsrechten durch Bildaufnahmen (§ 201a)	173	466	639
Bedrohung (§ 241)	155	518	673
Fälschung beweiserheblicher Daten (§ 269)	172	444	616

Von den 4.273 entfernten oder gesperrten Inhalten wurden 4.155 Inhalte wegen eines Verstoßes gegen unsere Gemeinschaftsrichtlinien weltweit entfernt. 118 Inhalte haben nicht gegen unsere Gemeinschaftsrichtlinien verstoßen, wurden jedoch wegen eines Verstoßes gegen eine im NetzDG aufgeführte Bestimmung des deutschen Strafgesetzbuchs in Deutschland gesperrt.

Anzahl der Beschwerden, in denen ein Fall gemäß § 3 Abs. 2 Nr. 3 Buchstabe a NetzDG vorlag und wir Kontakt mit dem/der Nutzer\*in, der/die den gemeldeten Inhalt gepostet hat, aufgenommen haben, um weitere Fakten einzuholen: 1.

Wir haben 8 Beschwerden an eine anerkannte Einrichtung der Regulierten Selbstregulierung übertragen.

## 9. BEARBEITUNGSZEIT BEI ENTFERNUNGEN/SPERRUNGEN

Jeweils die Anzahl der Beschwerden über rechtswidrige Inhalte, die nach ihrem Eingang innerhalb von 24 Stunden, innerhalb von 48 Stunden, innerhalb einer Woche oder zu einem späteren Zeitpunkt zur Entfernung oder Sperrung des rechtswidrigen Inhalts geführt haben, zusätzlich aufgeschlüsselt nach Beschwerden von Beschwerdestellen und von Nutzer\*innen sowie jeweils aufgeschlüsselt nach dem Beschwerdegrund.



Der folgende Abschnitt enthält einen Überblick der Zeit, die nach Eingang einer NetzDG-Beschwerde benötigt wurde, um einen Inhalt zu entfernen oder zu sperren, der als rechtswidrig im Sinne des NetzDG oder als Verstoß gegen unsere Gemeinschaftsrichtlinien eingestuft wurde, nach Gesamtzahl und aufgeschlüsselt nach der Art des/der Beschwerdeführer\*in (Beschwerdestelle oder Nutzer\*in) und dem Grund der Beschwerde. Dieser Abschnitt betrifft die Anzahl der zwischen 1. Juli 2022 und 31. Dezember 2022 über das Instagram NetzDG-Beschwerdeformular übermittelten Beschwerden.

Dabei ist Folgendes hinsichtlich dieses Abschnitts zu beachten:

- Wie in § 2 Abs. 2 Nr. 9 NetzDG festgelegt, sind unsere bis zur Entfernung benötigten Zeiträume folgendermaßen aufgeteilt: (a) innerhalb von 24 Stunden, (b) innerhalb von 48 Stunden, (c) innerhalb einer Woche und (d) zu einem späteren Zeitpunkt.
- Die in diesem Abschnitt dargestellten Zahlen beziehen sich auf übermittelte Beschwerden und nicht auf einzelne in den Beschwerden beanstandete Inhalte.
- Die Zeiträume beziehen sich auf die Zeit zwischen der Übermittlung der Beschwerde und der letzten von uns als Reaktion auf die Beschwerde ergriffenen Maßnahme. Wenn in einer Beschwerde beispielsweise zwei Inhalte angeführt werden und wir den einen Inhalt innerhalb von 24 Stunden und den anderen innerhalb von 7 Tagen bearbeitet haben, dann würde diese Beschwerde in der Tabelle als innerhalb von 7 Tagen bearbeitet aufgeführt.
- Nutzer\*innen können mehrere Gründe für Rechtswidrigkeit in einer einzigen NetzDG-Beschwerde anführen. Wenn wir infolge einer Beschwerde Maßnahmen in Bezug auf einen Inhalt ergriffen haben, wird der entsprechende Inhalt in Tabelle 5 unter jeder in der Beschwerde genannten Bestimmung aufgeführt. Daher übersteigt die Gesamtzahl der in Tabelle 5 aufgeführten Entfernungen/Sperrungen die Gesamtzahl der Beschwerden, die zur Entfernung oder Sperrung von Inhalten führten.

Bearbeitungszeit für die Gesamtzahl der NetzDG-Beschwerden, die zu einer Sperrung oder Entfernung geführt haben: Bei den 5.141 Beschwerden, die zu einer Sperrung oder Entfernung geführt haben, erfolgte unsere letzte Maßnahme zur Sperrung/Entfernung 4.793 Mal innerhalb von 24 Stunden, 105 Mal innerhalb von 48 Stunden, 214 Mal innerhalb von 7 Tagen und 29 Mal später als nach 7 Tagen.

Des Weiteren zeigt Tabelle 4 die folgenden Details für jede Beschwerde, bei der ein Inhalt entfernt oder gesperrt wurde: (1) die Art des/der Beschwerdeführer\*in, der die Beschwerde eingereicht hat, und (2) die Zeit, die für die Entfernung oder Sperrung des gemeldeten Inhalts benötigt wurde.

Tabelle 4. Bearbeitungszeit für Beschwerden bei Entfernungen/Sperrungen aufgeschlüsselt nach Art des/der Beschwerdeführer\*in

	<b>24 Stunden</b>	<b>48 Stunden</b>	<b>7 Tage</b>	<b>&gt; 7 Tage</b>
Beschwerden von Beschwerdestellen	692	4	14	1
Beschwerden von Nutzer*innen	4.101	101	200	28
<b>Gesamt</b>	<b>4.793</b>	<b>105</b>	<b>214</b>	<b>29</b>

Darüber hinaus gibt die nachstehende Tabelle 5 folgende Details für jede Beschwerde an, infolgeder ein Inhalt entfernt oder gesperrt wurde: (1) Die Art des/der Beschwerdeführer\*in, der/die



die Beschwerde übermittelt hat, (2) die bis zur Entfernung oder Sperrung des gemeldeten Inhalts benötigte Zeit und (3) die Bestimmung(en) des deutschen Strafgesetzbuchs, die durch den/die Beschwerdeführer\*in in der Beschwerde angeführt wurde(n).

Tabelle 5. Bearbeitungszeit für Beschwerden bei Entfernungen/Sperrungen aufgeschlüsselt nach Art des/der Beschwerdeführer\*in und dem Grund der Beschwerde

Paragraf des Strafgesetzbuchs	Beschwerdeführer*in							
	Beschwerden von Beschwerdestellen				Beschwerden von Nutzer*innen			
	24 Std.	48 Std.	7 Tage	>7 Tage	24 Std.	48 Std.	7 Tage	>7 Tage
Verbreiten von Propagandamitteln verfassungswidriger und terroristischer Organisationen (§ 86)	206	1	2	0	544	13	16	1
Verwenden von Kennzeichen verfassungswidriger und terroristischer Organisationen (§ 86a)	171	1	2	0	548	14	22	2
Vorbereitung einer schweren staatsgefährdenden Gewalttat (§ 89a)	148	1	2	0	222	3	5	0
Anleitung zur Begehung einer schweren staatsgefährdenden Gewalttat (§ 91)	132	0	2	0	236	2	3	0
Landesverräterische Fälschung (§ 100a)	140	1	2	0	375	8	6	4
Öffentliche Aufforderung zu Straftaten (§ 111)	162	1	2	0	776	10	7	5
Störung des öffentlichen Friedens durch Androhung von Straftaten (§ 126)	154	1	2	0	463	5	3	4
Bildung krimineller Vereinigungen (§ 129)	163	1	1	0	492	1	7	1
Bildung terroristischer Vereinigungen (§ 129a)	116	0	2	0	155	0	3	0
Bildung krimineller und/oder terroristischer Vereinigungen im Ausland (§ 129b)	141	1	1	0	179	3	2	1
Volksverhetzung (§ 130)	233	1	3	0	1.093	22	35	5
Gewaltdarstellung (§ 131)	162	1	2	0	526	11	19	7
Belohnung und Billigung von Straftaten (§ 140)	160	1	1	0	722	10	8	4
Beschimpfung von Bekenntnissen, Religionsgesellschaften und Weltanschauungsvereinigungen (§ 166)	191	0	3	0	809	14	21	7



Paragraf des Strafgesetzbuchs	Beschwerdeführer*in							
	Beschwerden von Beschwerdestellen				Beschwerden von Nutzer*innen			
	24 Std.	48 Std.	7 Tage	>7 Tage	24 Std.	48 Std.	7 Tage	>7 Tage
Verbreitung, Erwerb und Besitz kinderpornographischer Inhalte (§ 184b)	164	1	2	0	383	27	74	8
Beleidigung (§ 185)	267	4	3	1	1.159	30	31	7
Üble Nachrede (§ 186)	222	4	5	0	795	18	28	9
Verleumdung (§ 187)	196	4	6	0	632	15	23	9
Verunglimpfung des Andenkens Verstorbener (§ 189)	139	0	1	0	368	16	4	0
Verletzung des höchstpersönlichen Lebensbereichs und von Persönlichkeitsrechten durch Bildaufnahmen (§ 201a)	169	1	3	0	429	10	22	5
Bedrohung (§ 241)	152	1	2	0	501	6	7	4
Fälschung beweisheblicher Daten (§ 269)	171	0	1	0	422	7	10	5

## 10. KORRESPONDENZ

Maßnahmen zur Unterrichtung des/der Beschwerdeführer\*in sowie des/der Nutzer\*in, für den/die der beanstandete Inhalt gespeichert wurde, über die Entscheidung über die Beschwerde.

Wir ergreifen eine Reihe an Maßnahmen, um mit Beschwerdeführer\*innen, die Inhalte über das NetzDG-Beschwerdeformular melden, und mit registrierten Nutzer\*innen, deren Inhalte gemäß NetzDG entfernt oder gesperrt werden, zu kommunizieren.

### 10A. KOMMUNIKATION MIT DEM/DER BESCHWERDEFÜHRER\*IN

Wenn eine Person eine Beschwerde übermittelt, kommunizieren wir per E-Mail mit ihr. Unmittelbar nach dem Erhalt einer Beschwerde senden wir dem/der Beschwerdeführer\*in eine automatische Antwort, in der wir ihn/sie darüber informieren, dass seine/ihre Beschwerde geprüft wird. Zugleich erhält er/sie eine Referenznummer sowie einen Link zu unserem NetzDG-Hilfereich. Wenn wir 24 Stunden nach der Übermittlung noch mit einer Beschwerde befasst sind, unterrichten wir den/die Beschwerdeführer\*in darüber, dass wir die Beschwerde noch weiter prüfen. Sollten wir zu irgendeinem Zeitpunkt für die Prüfung der Beschwerde zusätzliche Informationen von dem/der Beschwerdeführer\*in benötigen, fordern wir diese an.

Sobald wir entscheiden, ob eine bzw. welche Maßnahme bezüglich des gemeldeten Inhalts ergriffen wird, senden wir dem/der Beschwerdeführer\*in eine individuell zugeschnittene Benachrichtigung, in der wir die Person darüber informieren, warum wir eine bzw. keine Maßnahme bezüglich des gemeldeten Inhalts ergriffen haben. Diese Benachrichtigungen werden



je nach Ergebnis unserer Prüfung unterschiedlich gestaltet und lassen sich in die folgenden Kategorien unterteilen:

- **Der beanstandete Inhalt wird wegen eines Verstoßes gegen unsere Gemeinschaftsrichtlinien entfernt.** In diesem Fall informieren wir den/die Beschwerdeführer\*in, dass der Inhalt entfernt wurde, da er gegen unsere Gemeinschaftsrichtlinien verstoßen hat. Außerdem stellen wir Links zu unserer NetzDG-Hilfereich-Seite und unseren Gemeinschaftsrichtlinien für den Fall bereit, dass der/die Beschwerdeführer\*in zusätzliche Informationen wünscht.
- **Der beanstandete Inhalt wird wegen eines Verstoßes gegen eine im NetzDG aufgeführte Bestimmung des deutschen Strafgesetzbuchs gesperrt.** In diesem Fall informieren wir den/die Beschwerdeführer\*in, dass der Inhalt in Deutschland nicht mehr zugänglich ist. Außerdem führen wir die konkrete Bestimmung des Strafgesetzbuchs an, gemäß der wir die Rechtswidrigkeit des Inhalts festgestellt haben. Wir stellen auch einen Link zu unserer NetzDG-Hilfereich-Seite für den Fall bereit, dass der/die Beschwerdeführer\*in zusätzliche Informationen wünscht.
- **Der beanstandete Inhalt verstößt weder gegen unsere Gemeinschaftsrichtlinien noch gegen eine im NetzDG aufgeführte Bestimmung des deutschen Strafgesetzbuchs.** Hier informieren wir den/die Beschwerdeführer\*in, dass wir die Beschwerde zwar geprüft haben, allerdings den Inhalt nicht entfernen, da wir nicht festgestellt haben, dass er rechtswidrig im Sinne des NetzDG ist.
- **Mehrere Maßnahmen wurden bezüglich der beanstandeten Inhalte ergriffen.** Nutzer\*innen können in einigen Fällen mehrere Inhalte in einer einzigen NetzDG-Beschwerde anführen. Wir prüfen jeden Inhalt individuell und handeln entsprechend dem Ergebnis der Prüfung. Um die Menge der von uns an den/die Beschwerdeführer\*in gesendeten Benachrichtigungen gering zu halten, prüfen wir in der Regel erst jeden gemeldeten Inhalt und handeln entsprechend, bevor wir den/die Beschwerdeführer\*in von unseren Entscheidungen in Kenntnis setzen. Wenn wir unterschiedliche Maßnahmen für in einer Beschwerde angeführte Inhalte ergreifen (Beispiel: Wir sperren einen Inhalt und ergreifen keine Maßnahme bezüglich eines anderen Inhalts), lassen wir dem/der Beschwerdeführer\*in mithilfe einer Kombination aus den oben erläuterten Antworten konkrete Informationen zu unseren Handlungen bzw. Maßnahmen zukommen.

Wir haben einige wenige Beschwerden erhalten, in denen E-Mail-Adressen als Kontaktinformationen angegeben wurden, die mit an Sicherheit grenzender Wahrscheinlichkeit nicht dem/der Beschwerdeführer\*in gehören. Um den/die wahren Inhaber\*innen dieser E-Mail-Konten vor Spam zu schützen, haben wir die Beschwerden geprüft und den gemeldeten Inhalt entsprechend des Ergebnisses unserer Prüfung behandelt, aber keine Benachrichtigungen über die Ergebnisse der Prüfung an diese E-Mail-Adressen versendet.

Zusätzlich zu den Informationen, die wir als Reaktion auf bestimmte Beschwerden bereitstellen, finden Nutzer\*innen umfassende weiterführende Informationen zum NetzDG auch in unserem [NetzDG-Hilfereich](#). Der NetzDG-Hilfereich enthält Informationen zu Folgendem: (1) Zur Art der Inhalte, die Nutzer\*innen gemäß NetzDG melden können, (2) dazu, was geschieht, nachdem Nutzer\*innen eine NetzDG-Beschwerde übermitteln, (3) wie Nutzer\*innen Inhalte melden können, die nach ihrer Ansicht unter das NetzDG fallen, (4) welche Informationen in einer NetzDG-Beschwerde enthalten sein müssen, (5) in welchen Fällen man sich an den Zustellungsbevollmächtigten gemäß § 5 Abs. 1 NetzDG wenden kann, (6) was die Unterschiede



zwischen dem NetzDG und den Instagram-Gemeinschaftsrichtlinien sind und (7) wo die NetzDG-Transparenzberichte von Instagram zu finden sind.

## **10B. KOMMUNIKATION MIT DEM/DER NUTZER\*IN, DER/DIE DEN INHALT GEPOSTET HAT**

Wir kommunizieren auch mit den Nutzer\*innen, deren Inhalte aufgrund einer NetzDG-Beschwerde entfernt oder gesperrt werden. Wenn wir einen Inhalt wegen eines Verstoßes gegen unsere Gemeinschaftsrichtlinien entfernen, setzen wir den/die Nutzer\*in, der/die den Inhalt gepostet hat, über eine persönliche Benachrichtigung auf Instagram von unserer Maßnahme in Kenntnis und informieren die Person darüber, dass der Inhalt gegen unsere Gemeinschaftsrichtlinien verstoßen hat. Wenn wir einen Inhalt wegen eines Verstoßes gegen eine im NetzDG aufgeführte Bestimmung des deutschen Strafgesetzbuchs sperren, informieren wir den/die Nutzer\*in, der/die den Inhalt gepostet hat, per E-Mail über unsere Maßnahme sowie darüber, gegen welche konkrete Bestimmung des deutschen Strafgesetzbuchs der Inhalt verstoßen hat. Wenn eine Beschwerde nicht dazu führt, dass eine Maßnahme bezüglich des Inhalts einer Person ergriffen wird, benachrichtigen wir den/die Nutzer\*in nicht.

## **11. GEGENVORSTELLUNGEN NACH § 3B ABSATZ 1 SATZ 2**

Anzahl der im Berichtszeitraum eingegangenen Gegenvorstellungen nach § 3b Abs. 1 Satz 2, nach der Gesamtzahl sowie aufgeschlüsselt nach Gegenvorstellungen von Beschwerdeführern und von Nutzern, für die der beanstandete Inhalt gespeichert wurde, jeweils mit Angaben, in wie vielen Fällen der Gegenvorstellung abgeholfen wurde.

Wir bieten die Möglichkeit, gegen die von uns getroffenen Entscheidungen über die Entfernung oder Sperrung des Zugangs zu Inhalten nach einer Beschwerde über rechtswidrige Inhalte eine Gegenvorstellung einzulegen. Ein/e Nutzer\*in, der/die Inhalte über das NetzDG-Beschwerdeformular gemeldet hat (meldende/r Nutzer\*in) und die daraufhin nicht entfernt oder gesperrt wurden, kann gegen die Entscheidung über die ursprüngliche Beschwerde eine Gegenvorstellung einlegen. Wenn Inhalte aufgrund einer NetzDG-Beschwerde wegen eines Verstoßes gegen unsere Gemeinschaftsrichtlinien entfernt werden, kann der/die Nutzer\*in, der/die diese gepostet hat (postende/r Nutzer\*in), in der Regel eine Überprüfung dieser Entscheidung beantragen. In bestimmten Fällen wird diese Möglichkeit nicht zur Verfügung gestellt, etwa bei einem Terrorismuszusammenhang oder bei sexuellem Kindesmissbrauch und Kindesausbeutung. Wird ein Inhalt aufgrund einer Entscheidung gesperrt, die besagt, dass der Inhalt gegen die im NetzDG aufgeführten einschlägigen Bestimmungen des deutschen Strafgesetzbuchs verstößt, kann der/die postende Nutzer\*in gegen diese Entscheidung eine Gegenvorstellung einlegen.

Die folgende Tabelle zeigt die Anzahl der Gegenvorstellungen, die nach einer über das NetzDG-Beschwerdeformular von Instagram eingereichten Beschwerde zwischen dem 1. Juli 2022 und dem 31. Dezember 2022 eingegangen sind, sowie die Anzahl der Fälle, in denen der Gegenvorstellung abgeholfen wurde. Grundsätzlich kann pro NetzDG-Beschwerde nur ein Inhalt gemeldet werden, so dass auch jede Gegenvorstellung nur einen Inhalt betrifft. In einigen Fällen nennen meldende Nutzer\*innen jedoch mehrere Inhalte in einer einzigen NetzDG-Beschwerde (z. B. durch Nennung mehrerer URLs in einem zur Beschwerde hochgeladenen Anhang). In diesen



Fällen kann der/die meldende Nutzer\*in nur eine Gegenvorstellung gegen die Entscheidung einlegen, die wir in Bezug auf alle genannten Inhalte getroffen haben, und wir zählen diese Gegenvorstellung als eine eingegangene Gegenvorstellung. Wenn wir unsere ursprüngliche Entscheidung in Bezug auf alle oder einige der genannten URLs abändern, wird dies als eine Gegenvorstellung gezählt, der abgeholfen wurde.

Tabelle 6 . NetzDG-Gegenvorstellungen aufgeschlüsselt nach Antragsteller\*in

	<b>Anzahl der eingegangenen Gegenvorstellungen</b>	<b>Anzahl der Gegenvorstellungen denen abgeholfen wurde</b>
Gegenvorstellungen des/der meldenden Nutzer*in	6.576	833
Gegenvorstellungen des/der postenden Nutzer*in	132	50
Gesamt	6.708	883

## 12. GEGENVORSTELLUNGEN NACH § 3B ABSATZ 3 SATZ 1

Anzahl der im Berichtszeitraum eingegangenen Gegenvorstellungen nach § 3b Abs. 3 Satz 1, jeweils mit Angaben, in wie vielen Fällen von einer Überprüfung gemäß § 3b Abs. 3 Satz 3 abgesehen wurde und in wie vielen Fällen der Gegenvorstellung abgeholfen wurde.

In der Regel können uns auch Nutzer\*innen, die Inhalte auf andere Weise als durch eine NetzDG-Beschwerde melden (meldende Nutzer\*innen) und die daraufhin nicht entfernt oder gesperrt wurden, sowie Nutzer\*innen, dessen/deren Inhalte als Folge einer solchen Meldung oder ohne Meldung entfernt oder gesperrt wurden (postende Nutzer\*innen), mitteilen, dass sie mit der von uns getroffenen Entscheidung zur Entfernung oder Sperrung des Inhalts nicht einverstanden sind.

Zwischen dem 1. Juli 2022 und dem 31. Dezember 2022 gingen bei uns für 211.887 einzelne Inhalte Widersprüche zu Entscheidungen über die Entfernung oder Sperrung von Inhalten mit Bezug zu Deutschland ein, die nicht auf einer NetzDG-Beschwerde beruhten. Bezüglich 37.623 einzelnen Inhalten haben wir unsere Entscheidung infolge des Widerspruchs abgeändert.

## 13. INFORMATIONSZUGANG FÜR KREISE DER WISSENSCHAFT UND FORSCHUNG

Angaben darüber, ob und inwieweit Kreisen der Wissenschaft und Forschung im Berichtszeitraum Zugang zu Informationen des Anbieters gewährt wurde, um ihnen eine anonymisierte Auswertung zu ermöglichen, inwieweit a) entfernte oder gesperrte rechtswidrige Inhalte an Eigenschaften im Sinne des § 1 des Allgemeinen Gleichbehandlungsgesetzes vom 14. August 2006 (BGBl. I S. 1897), das zuletzt durch Artikel 8 des Gesetzes vom 3. April 2013 (BGBl. I S. 610) geändert worden ist, in der jeweils geltenden Fassung anknüpfen, b) die Verbreitung von rechtswidrigen Inhalten zu spezifischer



Betroffenheit bestimmter Nutzerkreise führt und c) organisierte Strukturen oder abgestimmte Verhaltensweisen der Verbreitung zugrunde liegen.

Mit unseren Bemühungen um Datentransparenz unterstützt Meta qualifizierte Forscher\*innen, indem wir mit ihnen Facebook- und Instagram-Daten (in geschützter Form, d. h. unter Sichtung und Risikominimierung für die Privatsphäre durch eine Kombination aus programmatischen (Zugangskriterien, Vertragsbedingungen) und technischen Kontrollen) teilen, um die Untersuchung wichtiger sozialer Fragen zu unterstützen. Unsere Initiativen zur Bereitstellung von Daten haben die Forschung z.B. zu Inhalten erleichtert, die Falschmeldungen/Desinformation enthalten, die gegen unsere Gemeinschaftsrichtlinien verstoßen, sowie zu anderen wichtigen gesellschaftlichen Themen. Im Folgenden nennen wir einige Beispiele für unsere Bemühungen um Datentransparenz. Weitere Informationen sind im [Meta Transparency Center](#) verfügbar.

- **IO Research Archive:** Wir haben das IO (Influence Operations) Research Archive (Forschungsarchiv zu Einflussnahmehandlungen) gegründet, um einen Informationsaustausch mit unabhängigen Forscher\*innen und Wissenschaftler\*innen, die sich mit Einflussnahmehandlungen von Netzwerken befassen, die Meta wegen Coordinated Inauthentic Behavior (CIB) entfernt hat, zu ermöglichen. „CIB“ definieren wir als „koordinierte Bemühungen, die öffentliche Debatte für ein strategisches Ziel zu manipulieren, bei denen gefälschte Konten im Mittelpunkt der Operation stehen“. Über das IO Research Archive können Forscher\*innen öffentliche Daten über CIB-Netzwerke auf Facebook und Instagram untersuchen, die von Meta-Teams identifiziert und unterbunden wurden. Ende 2020 startete Meta ein Beta-Archiv mit einer kleinen Gruppe von Forscher\*innen, die CIB-Aktivitäten untersuchen und Gegenmaßnahmen entwickeln. Zu dieser Beta-Gruppe gehörten das Stanford Internet Observatory, das Digital Forensic Lab des Atlantic Council, Graphika, die Cardiff University und das Australian Strategic Policy Institute. Zum 31. Dezember 2022 hatte Meta Daten von mehr als 100 CIB-Netzwerken, die seit Juni 2020 gelöscht wurden, zur Verfügung gestellt, um damit unabhängige Analysen und Studien von deren Aktivitäten, Taktiken, Techniken und Verfahren zu ermöglichen.
- **CrowdTangle:** CrowdTangle ist eine Plattform zum Auffinden von Inhalten und zur Beobachtung von Social-Media-Aktivitäten, die einen Zugang zu einem kleinen Teil der öffentlichen Daten auf Facebook und Instagram bereitstellt. Am 31. Dezember 2022 bestanden rund 144 registrierte aktive deutsche Konten, inklusive Konten von Universitäten, mit Zugang zu CrowdTangle. Forscher\*innen nutzen CrowdTangle, um eine Vielzahl von wichtigen Themen von gesellschaftlichem Interesse zu untersuchen, unter anderem Falschmeldungen, Wahlen, Covid-19 und Rassismus.
- **Berichte:** Ferner veröffentlichen wir vierteljährlich Berichte zur Durchsetzung der Gemeinschaftsstandards ([Community Standards Enforcement Report](#)), die auch Informationen zu Instagram enthalten, um unsere Fortschritte effektiver zu verfolgen und unser kontinuierliches Engagement für Transparenz und Unterstützung der Forschung durch Facebook- und Instagram-Daten zu zeigen.

Zusätzlich zu den Daten, die wir zur Verfügung stellen, waren wir, wie in dem Transparenzbericht zu dem vorangegangenen Berichtszeitraum 1. Januar bis 30. Juni 2022 (nachfolgend “H1/2022 Transparenzbericht”) ausgeführt, aktive Teilnehmer der Arbeitsgruppe der Europäischen Beobachtungsstelle für digitale Medien (EDMO) zum Zugang von Forscher\*innen zu Plattformdaten. Wir haben erhebliche Ressourcen für diese Arbeit bereitgestellt, weil wir es als wichtig erachten, die richtige Balance zwischen der Privatsphäre des Einzelnen und dem gesellschaftlichen Nutzen von Forschung zu finden. Die Arbeitsgruppe hat im Jahr 2022 einen



[Bericht](#) darüber veröffentlicht, wie dieses Gleichgewicht im Rahmen der rechtlichen Verpflichtungen der Datenschutz-Grundverordnung (DSGVO) erreicht werden kann.

## 14. SCHUTZMAßNAHMEN

[Sonstige Maßnahmen des Anbieters zum Schutz und zur Unterstützung der von rechtswidrigen Inhalten Betroffenen.](#)

Zusätzlich zu den in diesem Bericht bereits beschriebenen Maßnahmen (z.B. den Maßnahmen zur Erkennung von Inhalten, die gegen unsere Gemeinschaftsrichtlinien verstoßen und leicht zugänglichen Meldekanälen) ergreift Meta eine Reihe weiterer Maßnahmen zum Schutz und zur Unterstützung von Personen, die von im Internet verbreiteten schädlichen Inhalten - einschließlich rechtswidrigen Inhalten - betroffen sind.

Unternehmen des Meta-Konzerns arbeiten mit einer Reihe von Organisationen und Institutionen zusammen, die sich für die Bekämpfung rechtswidriger Inhalte und der Förderung der Sicherheit im Internet einsetzen. Dazu gehören z. B. der FSM (Freiwillige Selbstkontrolle Multimedia-Diensteanbieter e. V.) sowie der DsiN (Deutschland sicher im Netz e. V.). Unternehmen des Meta-Konzerns unterhalten zudem eine enge Arbeitsbeziehung mit [Jugendschutz.net](#). Wir arbeiten ferner mit der Jugendschutzorganisation [Troubled Desire](#) zusammen. Ein Hauptziel des Programms ist die Verhinderung von sexuellem Kindesmissbrauch und der Verwendung von Bildern von Kindesmissbrauch.

Wir melden dem National Center for Missing and Exploited Children (NCMEC) alle offensichtlichen Fälle von sexuellem Kindesmissbrauch, die wir auf unseren Seiten in irgendeinem Teil der Welt identifizieren, einschließlich Inhalten, von denen wir durch Regierungsanfragen erfahren. Das NCMEC arbeitet mit Strafverfolgungsbehörden auf der ganzen Welt zusammen.

Darüber hinaus arbeiten Unternehmen des Meta-Konzerns eng mit Strafverfolgungsbehörden zusammen. Diese können unser speziell zu diesem Zweck eingerichtetes Law Enforcement Online Request System (LEORS) für die Einreichung, Verfolgung und Bearbeitung von Anfragen, einschließlich der Beantwortung dringender Anfragen, nutzen.

Meta bietet außerdem umfassende Informationen und Tools zur praktischen Unterstützung für Betroffene an, unter anderem:

- [Anti-Mobbing-Hub](#), mit Informationen rund um Funktionen und Tools, die wir zum Schutz vor Mobbing entwickelt haben, z.B. wie man unerwünschte Interaktionen limitiert, Personen blockiert, Inhalte meldet oder die Sichtbarkeit von Kommentaren unter eigenen Beiträgen einschränkt.
- [Familienbereich](#), der Eltern Tools und Ressourcen bietet, die ihnen dabei helfen, die Online-Erfahrung ihrer Kinder zu begleiten. Dazu gehören ein Medienkompetenz-Hub und eine Reihe von Elternaufsichts-Tools. Teenager\*innen zwischen 13 und 18 Jahren können Eltern darum bitten, ihr Konto zu beaufsichtigen, sodass Eltern sich keine Sorgen machen müssen. Gleichzeitig wird die Unabhängigkeit ihrer Kinder respektiert. Eltern können sich hier beispielsweise darüber informieren, wenn ihr Kind einen Account meldet, etwa wegen Mobbing oder Hassrede.
- [Instagram Leitfaden für Eltern und Betreuer\\*innen von Jugendlichen und Kindern](#), der in



Zusammenarbeit mit den EU-Initiativen [Klicksafe.de](https://www.klicksafe.de) und [SaferInternet.at](https://www.saferinternet.at) sowie dem Projekt [Stop Hate Speech](#) entwickelt wurde, und z. B. Erklärungen zu Sicherheitsfunktionen auf Instagram enthält, etwa der Möglichkeit die Kommentarfunktion zu deaktivieren, unangenehme Kommentare und Nachrichten zu filtern und einzustellen, ob man in einem Kommentar, einer Bildunterschrift oder einer Story markiert oder erwähnt werden kann.

Wir führen auch Veranstaltungen in Form von Workshops oder Diskussionen durch, sowohl persönlich als auch virtuell, wie beispielsweise unser SafeTeens Event in Berlin im November 2022, zu dem wir politische Entscheidungsträger, Medien und Eltern-Creator\*innen eingeladen haben, um über Medienkompetenz und Jugendmedienschutz zu diskutieren und über unsere Sicherheits-Tools zu informieren.

Wir setzen uns nachdrücklich für die Sicherheit von Frauen im Internet und für den Schutz von Frauen ein, die auf unseren Plattformen Angriffen ausgesetzt sind. In diesem Zusammenhang haben wir bereits das Programm „Starke Frauen, starke Politik“ ins Leben gerufen, das wir zusammen mit der Gegenrede-Initiative #ichbinhier sowie einer spezialisierten Psychologin konzipiert haben, um Frauen in der Politik und deren Teams, Aktivistinnen und andere politisch aktive Frauen zu unterstützen. Gemeinsam mit unseren Partnern bietet es z.B. Mitgliedern des Programms die Möglichkeit, persönliche Coaching-Sitzungen zu buchen oder Informationsmaterial, z.B. zur Moderation von Inhalten, anzufordern.

Meta hat außerdem in Partnerschaft mit der Generaldirektion Justiz und Verbraucher der Europäischen Kommission den [„Courage Against Hate“ \(Zivilcourage gegen Hass\)-Bericht](#) initiiert. Courage Against Hate (CAH) ist eine von Meta (seinerzeit noch “Facebook”) ins Leben gerufene Initiative mit dem Ziel, bereichsübergreifend und europaweit zu Gesprächen und Aktionen gegen Hassrede und Extremismus anzuregen. In dem am 13. Juli 2021 veröffentlichten [CAH-BERICHT](#) haben sich vier Forschungseinrichtungen (das [Centre for the Analysis of the Radical Right, CARR](#); [HOPE Not Hate](#); das [Institut für Demokratie und Zivilgesellschaft, IDZ-Jena](#); und die [Swedish Defence Research Agency/Universität Uppsala](#)) sowie acht auf dem Gebiet tätige Nichtregierungsorganisationen/Unternehmen ([Iamhere International](#); [Galop UK](#); [Moonshot](#); das [Media Diversity Institute/Textgain](#); das [Institut für Strategischen Dialog, ISD](#); und [Zivilcourage & Anti-Rassismus-Arbeit, ZARA](#)) zusammengetan, um dazu beizutragen, eine Bestandsaufnahme sowohl der Trends bei den Themen Hassrede und Extremismus als auch der wirksamen Programme und Initiativen gegen Hass in Europa zu erstellen. Der Bericht regt einen fachbereichsübergreifenden Dialog darüber an, welche politischen Maßnahmen, weiteren Analysen und Programme für eine effiziente Bekämpfung von Hass, Extremismus und Terrorismus erforderlich sind. Die Artikelsammlung bringt europäische akademische Analysen mit Praktikern zusammen, die aktiv gegen Extremismus in der Gesellschaft vorgehen, und zeigt unser kontinuierliches Engagement bei der Bekämpfung dieser Probleme.

Darüber hinaus haben wir die Einrichtung einer Website der britischen Nichtregierungsorganisation „The Revenge Porn Helpline (RPH)“ ermöglicht, die in einer Reihe von Ländern, unter anderem auch in Deutschland, zugänglich ist (s. bereits den H1/2022 Transparenzbericht). RPH ist eine Organisation, die erwachsenen Opfern von Missbrauch intimer Aufnahmen (d. h. solchen über 18 Jahren) hilft und eine führende NGO in diesem Bereich ist. Die Plattform ([StopNCII.org](https://www.stopncii.org)) ist die erste ihrer Art, und RPH hat sie mit dem Ziel entwickelt, den Opfern ein privates und sicheres Instrument an die Hand zu geben, mit dem sie die Verbreitung ihrer nicht einvernehmlich geteilten intimen Aufnahmen (non-consensual intimate image - NCII) im Internet proaktiv stoppen können. Es verwendet eine Technologie, die Bilder und Videos direkt



auf dem Gerät des Plattformnutzers mit einem Hashcode versieht, so dass die Opfer die Originalinhalte weder an die NGO noch an die StopNCII-Plattform weitergeben müssen. Sobald die Hashes mit Facebook und Instagram geteilt werden, verwenden wir eine Technologie zur Erkennung von identischen oder ähnlichen Inhalten während sie auf die Plattform hochgeladen werden, um entsprechend gegen diese Inhalte vorzugehen.

## 15. ZUSAMMENFASSUNG: BESCHWERDEN & GEGENVORSTELLUNGEN

Eine Zusammenfassung mit einer tabellarischen Übersicht, die die Gesamtzahl der eingegangenen Beschwerden über rechtswidrige Inhalte, den prozentualen Anteil der auf diese Beschwerden hin entfernten oder gesperrten Inhalte, die Anzahl der Gegenvorstellungen jeweils nach § 3b Abs. 1 Satz 2 und nach § 3b Abs. 3 Satz 1 sowie jeweils den prozentualen Anteil der auf diese Gegenvorstellungen hin abgeänderten Entscheidungen den entsprechenden Zahlen für die beiden vorangegangenen Berichtszeiträume gegenüberstellt, verbunden mit einer Erläuterung erheblicher Unterschiede und ihrer möglichen Gründe.

Tabelle 7. Anzahl der Beschwerden, Gegenvorstellungen/Widersprüche und deren Ergebnisse

		Juli bis Dezember 2022 ("H2/2022")	Januar bis Juni 2022 ("H1/2022")	Juli bis Dezember 2021 ("H2/2021")
Beschwerden	Anzahl	57.541	63.696	56.208
	Beschwerden mit Entfernungen/Sperrungen	8,9%	9,5%	14,6%
Gegenvorstellungen gemäß § 3b Abs. 1 Satz 2 NetzDG (Abschnitt 11)	Anzahl	6.708	2.109	N/A
	Gegenvorstellungen denen abgeholfen wurde	13,2%	10,3%	N/A
Inhalte, für die ein Widerspruch infolge einer Erstentscheidung eingelegt wurde, der keine NetzDG-Beschwerde vorausging (Abschnitt 12)	Anzahl	211.887	237.680	N/A
	Widersprüchen denen abgeholfen wurde	17,8%	22,4%	N/A

Wie aus der Tabelle ersichtlich ist, ist der bedeutendste Unterschied, dass die Zahl der Beschwerden von H2/2021 auf H2/2022 gestiegen ist, während in H1/2022 noch höhere Zahlen zu verzeichnen waren. Der Anstieg von H2/2021 auf H2/2022 steht im Einklang mit dem stetigen Anstieg der Beschwerden, der auch aus den Transparenzberichten vor den in dieser Tabelle dargestellten Berichtszeiträumen ersichtlich ist. Dieser stetige Anstieg lässt sich durch verschiedene Faktoren erklären, wie z.B. die zunehmende Gewöhnung der Nutzer\*innen an die NetzDG-Meldewege und die Aufnahme von § 189 StGB als Vorschrift, für die das NetzDG ab dem 1. Februar 2022 eine Meldemöglichkeit vorsieht. Die Differenz zu dem Volumen in H1/2022 lässt sich ebenfalls durch verschiedene Faktoren erklären, wie z.B. aktuelle Ereignisse, die ein höheres Beschwerdevolumen auslösen. Der Anstieg der Anzahl der Gegenvorstellungen nach § 3b Abs. 1 Satz 2 NetzDG (Ziffer 11) von H1/2022 zu H2/2022 lässt sich durch verschiedene Faktoren erklären, wie z.B. die zunehmende Gewöhnung der Nutzer\*innen an die Option der



Gegenvorstellung. Wir weisen darauf hin, dass im Hinblick auf die Anzahl der Gegenvorstellungen/Widersprüche kein Vergleich zu dem Berichtszeitraum Juli bis Dezember 2021 gezogen werden kann, da diese gemäß § 6 Abs. 3 NetzDG seit dem Berichtszeitraum Januar bis Juni 2022 erstmalig berichtet werden. Wir weisen ferner darauf hin, dass wir (ursprüngliche) Beschwerden, bei denen die ursprüngliche Entscheidung darin bestand, den Inhalt zu entfernen oder zu sperren (siehe Abschnitt 8), getrennt von einer möglichen Gegenvorstellung zählen, die zu einer Änderung der ursprünglichen Entscheidung führen kann. Das bedeutet, dass sich die Anzahl der Beschwerden, die wir hier und in Abschnitt 8 angeben, nicht ändert, wenn die ursprüngliche Entscheidung nach einer Gegenvorstellung geändert wurde.

## 16. ALLGEMEINE GESCHÄFTSBEDINGUNGEN

Erläuterung der Bestimmungen in den Allgemeinen Geschäftsbedingungen des Anbieters über die Zulässigkeit der Verbreitung von Inhalten auf dem sozialen Netzwerk, die der Anbieter für Verträge mit Verbraucher\*innen verwendet.

Wir verwenden eine Strategie mit dem Namen „Entfernen, Reduzieren, Informieren“. Das beinhaltet, dass wir Inhalte, die gegen unsere Richtlinien verstoßen, entfernen, wenn sie uns gemeldet oder von unseren Systemen als problematisch erkannt werden, sowie, dass wir die Verbreitung bestimmter Inhalte reduzieren oder Menschen durch die Bereitstellung von zusätzlichem Kontext informieren, damit sie entscheiden können, was sie anklicken, lesen oder teilen möchten. Zur Unterstützung dieser Strategie haben wir Richtlinien eingeführt, aus denen sich je nach Nutzergruppe und der Art der Nutzung Vorgaben für die Verbreitung von Inhalten auf Instagram ergeben können. Für Verbraucher\*innen sind die Folgenden von zentraler Bedeutung:

- [Nutzungsbedingungen](#)
- [Gemeinschaftsrichtlinien](#)

Je nachdem, wie Instagram genutzt wird, sind darüber hinaus auch die [Musik-Richtlinien](#) einschlägig, welche die Verantwortlichkeit der Nutzer\*innen unterstreichen, die Inhalte mit Musik posten, und beispielsweise festhalten, dass Inhalte mit Musik entfernt werden können, wenn die Nutzung der Musik nicht ordnungsgemäß autorisiert ist.

Daneben gibt es eine Reihe weiterer Vorgaben, aus denen sich ebenfalls Beschränkungen hinsichtlich der zulässigen Inhalte ergeben können, die sich allerdings primär an gewerbliche Nutzer und Unternehmen richten oder den Verkauf von Waren betreffen (z.B. [Branded Content-Richtlinien](#), [Handelsrichtlinien](#), [Werberichtlinien](#)).

Im Folgenden liefern wir einen Überblick über die Nutzungsbedingungen und die Gemeinschaftsrichtlinien.

Unsere Nutzungsbedingungen, Gemeinschaftsrichtlinien und sonstigen Richtlinien werden von Zeit zu Zeit aktualisiert. So verwenden wir seit dem 26. Juli 2022 beispielsweise eine überarbeitete Fassung der Nutzungsbedingungen. Diese Fassung liegt dem vorliegenden Bericht zugrunde, da sie im relevanten Berichtszeitraum (Juli bis Dezember 2022) zuletzt gültig war (s. auch der H1/2022 Transparenzbericht zur vorherigen Fassung).



## 16A. NUTZUNGSBEDINGUNGEN

Die [Nutzungsbedingungen](#) von Instagram verbieten unter der Überschrift „Deine Verpflichtungen“ eine Reihe von Verhaltensweisen, die auch das Teilen von Inhalten betreffen. Dazu gehört unter anderem:

- Die Nutzer\*innen dürfen sich nicht als jemand oder etwas anderes ausgeben oder ungenaue bzw. falsche Informationen bereitstellen.
- Es darf auf Instagram nichts Rechtswidriges, Irreführendes oder Betrügerisches oder etwas für einen illegalen oder unberechtigten Zweck getan werden.
- Es darf nicht gegen die Nutzungsbedingungen oder Richtlinien von Instagram verstoßen werden, insbesondere die Gemeinschaftsrichtlinien, die Meta-Plattform-Nutzungsbedingungen oder die Musik-Richtlinien. Andere dürfen weder bei Verstößen unterstützt noch dazu ermutigt werden. Wer sogenannten Branded Content postet, muss unsere Branded Content-Richtlinien einhalten.
- Es dürfen nicht ohne Erlaubnis private oder vertrauliche Informationen einer anderen Person gepostet werden oder etwas getan werden, das die Rechte einer anderen Person verletzt, einschließlich des geistigen Eigentums (z.B. Urheberrechtsverletzung, Markenverletzung).

Diese Grundregeln werden insbesondere in den Gemeinschaftsrichtlinien, aber je nach Nutzung auch in weiteren Richtlinien konkretisiert.

Wir sind nach den Nutzungsbedingungen dazu berechtigt, Inhalte, die auf dem Dienst geteilt werden, im Falle von Verstößen gegen die Nutzungsbedingungen oder Richtlinien oder im Falle einer entsprechenden gesetzlichen Verpflichtung zu entfernen. Wir können die Bereitstellung des Dienstes ganz oder teilweise verweigern oder einstellen (einschließlich der Beendigung oder Deaktivierung des Zugriffs), wenn Nutzer\*innen gegen unsere Nutzungsbedingungen oder Richtlinien verstoßen, wiederholt die Rechte am geistigen Eigentum anderer verletzen oder wenn wir gesetzlich dazu verpflichtet sind.

## 16B. GEMEINSCHAFTSRICHTLINIEN

Die [Gemeinschaftsrichtlinien](#) sind für die Nutzer\*innen das zentrale Regelwerk über die Zulässigkeit der Verbreitung von Inhalten auf Instagram. Ziel der Gemeinschaftsrichtlinien ist die Schaffung und Erhaltung einer sicheren und offenen Umgebung für alle Nutzer\*innen. Wir möchten, dass Instagram ein authentischer und sicherer Ort bleibt, der inspiriert und es ermöglicht, sich selbst auszudrücken. In einigen Fällen lassen wir Inhalte zu, die andernfalls gegen unsere Gemeinschaftsrichtlinien verstoßen würden. Das betrifft berichtenswerte Inhalte, deren Veröffentlichung im öffentlichen Interesse liegt. Dies geschieht allerdings erst, nachdem wir das öffentliche Interesse gegen das Schadensrisiko abgewogen haben. Außerdem berücksichtigen wir bei diesen Beurteilungen auch internationale Menschenrechtsstandards.

Die Gemeinschaftsrichtlinien enthalten im Einzelnen folgende Vorgaben:



### **1) Teile nur Fotos und Videos, die du selbst aufgenommen hast oder zu deren Weitergabe du berechtigt bist**

Die Gemeinschaftsrichtlinien untersagen das Posten von Inhalten, zu deren Veröffentlichung die Nutzer\*innen nicht berechtigt sind.

### **2) Poste Fotos und Videos, die für eine vielfältige Zielgruppe geeignet sind**

Die Darstellung von Nacktheit ist auf Instagram nicht zulässig. Das gilt beispielsweise auch für Fotos, Videos und einige digital erstellte Inhalte, auf denen Geschlechtsverkehr, Genitalien und Nahaufnahmen nackter Gesäße zu sehen sind. Aus Sicherheitsgründen können wir auch Bilder entfernen, auf denen nackte oder halbnackte Kinder zu sehen sind.

Bestimmte Spielräume für die Darstellung von Nacktheit bestehen etwa in künstlerischen Zusammenhängen oder – speziell bei der Darstellung unbedeckter weiblicher Brustwarzen – beim Stillen, im Kontext einer Entbindung, in gesundheitsbezogenen Kontexten oder bei Protestaktionen.

### **3) Fördere bedeutungsvolle und echte Interaktionen**

Um Spam zu unterbinden, untersagen die Gemeinschaftsrichtlinien beispielsweise Versuche, künstlich „Gefällt mir“-Angaben, mehr Follower\*innen oder geteilte Inhalte zu erzielen, oder auch das wiederholte Posten derselben Kommentare oder Inhalte und das wiederholte Kontaktieren anderer Menschen für kommerzielle Zwecke ohne ihre Einwilligung. Die Gemeinschaftsrichtlinien untersagen auch Inhalte, die sich an gefälschten und irreführenden Nutzer\*innenbewertungen oder Ratings beteiligen, sie unterstützen, fördern oder ermöglichen oder die das Anbieten, Nachfragen oder einen Handel mit derartigen Bewertungen oder Ratings eingestehen. Es ist auch untersagt, andere Personen nachzuahmen oder Konten zu erstellen, um mit ihnen gegen unsere Richtlinien zu verstoßen oder andere in die Irre zu führen.

### **4) Halte dich an Recht und Gesetz**

Die Gemeinschaftsrichtlinien verlangen von den Nutzer\*innen ausdrücklich, sich an Recht und Gesetz zu halten. Dazu gehört beispielsweise, dass Instagram nicht benutzt werden darf, um Terrorismus, organisierte Kriminalität oder radikale Gruppen, die Hassrede verbreiten, zu unterstützen oder zu verherrlichen. Darüber hinaus untersagen wir bestimmte Inhalte im Zusammenhang mit bestimmten Geschäften oder Waren (z.B. sexuelle Dienstleistungen, Drogen, Medikamente, Schusswaffen, gefährdete Arten). Außerdem benötigen Konten, die Online-Glücksspiele, Online-Geschicklichkeitsspiele um echtes Geld oder Online-Lotterien bewerben, unsere schriftliche Genehmigung, bevor eines unserer Produkte verwendet werden darf. Ebenfalls untersagt sind sexuelle Inhalte, an denen Minderjährige beteiligt sind, oder Drohungen, intime Bilder anderer Personen zu posten.

### **5) Respektiere die anderen Mitglieder der Instagram-Community**

Inhalte, die glaubhafte An- oder Bedrohungen (z.B. körperliche Gewalt oder Vandalismus) oder Hassrede enthalten, werden von uns entfernt. Das Gleiche gilt für Inhalte, die darauf abzielen, Privatpersonen zu erniedrigen oder zu beschämen, Inhalte mit personenbezogenen Informationen zur Erpressung oder Belästigung von Personen sowie wiederholte unerwünschte Nachrichten. Offene und kontroverse Diskussionen über in der Öffentlichkeit stehende Personen



bleiben selbstverständlich erlaubt. Aufrufe zu Gewalt oder zu Angriffen gegen bestimmte Personen aufgrund bestimmter geschützter Eigenschaften (wie z.B. ethnische Zugehörigkeit, Nationalität, Geschlecht, sexuelle Orientierung, religiöse Zugehörigkeit oder aufgrund von Behinderungen oder Erkrankungen) sind nicht zulässig.

#### **6) Sorge dafür, dass unsere Gemeinschaft sicher bleibt. Verherrliche deshalb niemals Selbstverletzung**

Wir entfernen Inhalte, mit denen andere zu Selbstverletzungen ermutigt oder gedrängt werden, oder deaktivieren zugehörige Konten, wenn wir entsprechende Meldungen erhalten. Wir entfernen auch Inhalte, auf denen Opfer oder Überlebende von Selbstverletzungen dargestellt werden, wenn sie darauf abzielen, die betreffenden Personen anzugreifen oder zu verspotten.

#### **7) Denke nach, bevor du berichtenswerte Ereignisse postest**

Videos mit drastischen Gewaltdarstellungen werden möglicherweise von uns entfernt. So ist etwa das Teilen von Bildern mit deutlich sichtbarer Gewalt aus sadistischem Vergnügen oder zur Gewaltverherrlichung untersagt.

## **17. RECHTLICHE ZULÄSSIGKEIT DER ALLGEMEINEN GESCHÄFTSBEDINGUNGEN**

[Darstellung, inwiefern die Vereinbarung der Bestimmungen nach Nummer 16 mit den Vorgaben der §§ 307 bis 309 des Bürgerlichen Gesetzbuchs und dem sonstigen Recht in Einklang steht.](#)

Soweit wir mit den oben beschriebenen Nutzungsbedingungen und Richtlinien die Verbreitung bestimmter Inhalte auf Instagram untersagen oder beschränken, steht dies mit den §§ 307 bis 309 BGB und sonstigem anwendbarem Recht in Einklang.

Die Nutzungsbedingungen und die darin in Bezug genommenen Richtlinien wie die Gemeinschaftsrichtlinien sind wirksam in die vertraglichen Beziehungen zwischen uns und unseren Nutzer\*innen einbezogen. Bei der Registrierung auf der Plattform stimmen die Nutzer\*innen den Nutzungsbedingungen ausdrücklich zu. Dadurch werden die Nutzungsbedingungen und die weiteren Vorgaben wirksamer Vertragsbestandteil und regeln das zulässige Nutzerverhalten auf Instagram.

Die Bestimmungen hinsichtlich der zulässigen Inhalte berühren die Tatbestände der §§ 308 und § 309 BGB nicht. Soweit die Vorschriften überhaupt einen Anknüpfungspunkt hinsichtlich der zulässigen Inhalte haben, ist zunächst § 308 Nr. 3 BGB nicht einschlägig: Soweit uns nach den vertraglichen Bestimmungen das Recht zusteht, bei Verstößen Nutzerkonten zu deaktivieren und uns damit letztlich von der Leistungspflicht zu lösen, besteht dieses Recht von vornherein nicht ohne sachlich gerechtfertigten und in den Nutzungsbedingungen und den übrigen Richtlinien angegebenen Grund. Darüber hinaus begründet der Nutzungsvertrag ein Dauerschuldverhältnis, auf das § 308 Nr. 3 BGB von vornherein keine Anwendung findet. Auch § 308 Nr. 4 BGB ist nicht einschlägig: Die Sperrung oder Entfernung von Inhalten stellt von vornherein keine Änderung oder Abweichung von der versprochenen Leistung im Sinne der Vorschrift dar, sondern dient gerade der Durchsetzung der zuvor vereinbarten Kommunikationsstandards. Nichts anderes gilt für die mögliche Beschränkung oder Deaktivierung von Nutzerkonten im Falle von Verstößen.



Zudem wird unsere Leistungspflicht dabei auch nicht geändert, sondern im äußersten Fall ganz oder teilweise aufgehoben. Schließlich sind diese Maßnahmen unter Berücksichtigung unserer Interessen den betroffenen Nutzer\*innen jedenfalls zumutbar und wären damit selbst dann zulässig, wenn man den Anwendungsbereich der Vorschrift für eröffnet hielte.

Auch sonstige Gesetzesverstöße sind nicht ersichtlich. Im Gegenteil: Die Gemeinschaftsrichtlinien und andere Vorgaben zu den zulässigen Inhalten dienen insbesondere dazu, eine authentische und sichere Umgebung für die Nutzer\*innen zu schaffen. Dabei beugen sie zugleich der Verletzung von Gesetzen und von Rechten Dritter vor.

Dass Inhalte, die rechtswidrig oder sogar strafbar sind, auch auf Instagram unzulässig sind, versteht sich von selbst. Doch auch, soweit unsere Nutzungsbedingungen und Richtlinien Inhalte für unzulässig erklären, die die Schwelle zur Strafbarkeit oder sonstigen Rechtswidrigkeit nicht oder noch nicht überschreiten, ist dies zulässig. § 307 Abs. 1 BGB steht dem nicht entgegen. Nach § 307 Abs. 1 BGB sind Bestimmungen in Allgemeinen Geschäftsbedingungen unwirksam, wenn sie den Vertragspartner (hier: die Nutzer\*innen) des Verwenders (hier: wir, also die Meta Platforms Ireland Ltd.) entgegen den Geboten von Treu und Glauben unangemessen benachteiligen. Dabei kann sich eine unangemessene Benachteiligung auch daraus ergeben, dass die jeweilige Regelung nicht klar und verständlich ist.

Der Bundesgerichtshof hat mittlerweile geklärt, dass die Anbieter sozialer Netzwerke das Recht haben, die für ihre Plattform gültigen Kommunikationsstandards selbst festzulegen, und dass sie dabei auch Inhalte untersagen dürfen, die nicht strafbar oder sonst gesetzwidrig sind (BGH, Urteil vom 29. Juli 2021, Az. III ZR 192/20, Rn. 71 – abrufbar [hier](#); BGH, Urteil vom 29. Juli 2021, Az. III ZR 179/20, Rn. 59 – abrufbar [hier](#)). Richtigerweise hielt der Bundesgerichtshof die gegenteilige Ansicht für nicht überzeugend und hat einer staatsgleichen Bindung der Anbieter an die Grundrechte ausdrücklich eine Absage erteilt. Die Anbieter dürfen sich für den Fall der Verletzung ihrer Kommunikationsstandards unter anderem das Recht vorbehalten, einzelne Inhalte zu entfernen oder auch den Zugriff auf das Nutzerkonto zu sperren. Dies ist Ausdruck der Berufs- und Meinungsfreiheit der Anbieter und entspricht auch dem Interesse der übrigen Nutzer\*innen, denen in aller Regel ebenfalls an einer respektvollen Diskussionskultur und einem sicheren Kommunikationsraum gelegen ist. Dies gilt gerade auch mit Blick auf den Schutz minderjähriger Nutzer\*innen. Selbstverständlich müssen dabei die Meinungsfreiheit der Inhalteverfasser\*innen und die Bedeutung unserer Dienste für die Ausübung der Meinungsfreiheit gebührend berücksichtigt werden. Die Entfernung von Inhalten darf daher nicht willkürlich sein. Sie bedarf nach der Rechtsprechung des Bundesgerichtshofs eines sachlichen Grundes und muss zudem an objektive, überprüfbare Tatbestände anknüpfen.

Diese Voraussetzungen sind erfüllt. Die Nutzungsbedingungen sehen die Entfernung von Inhalten bei Verstößen gegen die Nutzungsbedingungen oder die Richtlinien, die für die Nutzung unserer Produkte gelten (einschließlich der Gemeinschaftsrichtlinien), vor, sowie im Falle einer rechtlichen Verpflichtung.

Dies sind objektive, überprüfbare Tatbestände, die ihrerseits jeweils auf sachliche Gründe abstellen.

Die Entfernung von Inhalten, die gegen die Nutzungsbedingungen verstoßen, ist von sachlichen Gründen getragen. So besteht beispielsweise an der Beibehaltung rechtswidriger, irreführender oder betrügerischer Inhalte schon kein schutzwürdiges Interesse. Ihre Entfernung oder



Sperrung ist vielmehr im Interesse aller Nutzer\*innen und in unserem eigenen Interesse geboten. Ebenso wenig müssen Anbieter es hinnehmen, wenn die Rechte einer anderen Person verletzt werden; insbesondere müssen sie rechtlichen Pflichten zur Entfernung nachkommen können. Auch hier bestehen ein sachlicher Grund und ein überwiegendes Interesse an der Entfernung oder Sperrung.

Auch die Entfernung von Inhalten, die gegen die Vorgaben in den Gemeinschaftsrichtlinien oder anderen Richtlinien verstoßen, stellt ein sachliches Anliegen dar. Die Richtlinien dienen legitimen Zwecken. Sie schützen die Interessen der Nutzer\*innen von Instagram, von Dritten und von uns selbst und können das Interesse an der Verbreitung bestimmter Inhalte überwiegen. Bei den Gemeinschaftsrichtlinien geht es insbesondere auch darum, Schäden in der Offline-Welt, Gefährdungen der Sicherheit und der Würde der Nutzer\*innen und Gesetzesverstöße zu verhindern. Zudem soll ein kommunikatives Umfeld geschaffen und erhalten werden, in dem sich alle Nutzer\*innen – auch Minderjährige – sicher fühlen können.

Dabei ist zu berücksichtigen, dass die Gemeinschaftsrichtlinien keineswegs starr und unnachgiebig formuliert sind, sondern vielfach Raum für die Berücksichtigung der Vielgestaltigkeit der Sachverhalte und von Nuancen in den Inhalten lassen und auch berechtigten Anliegen wie beispielsweise Information, Diskussion und Kritik ausdrücklich Rechnung tragen. Die Notwendigkeit einer differenzierten Betrachtung kommt in den einschlägigen Vorgaben vielfach zum Ausdruck. So ist es beispielsweise ein wesentliches Anliegen der Gemeinschaftsrichtlinien, die Verherrlichung oder Förderung von Selbstverletzungen, wie Essstörungen, zu verhindern. Gleichwohl möchten wir, dass Instagram ein Ort ist, an dem Menschen ihre Erfahrungen in diesen Bereichen teilen und Aufmerksamkeit auf diese Themen lenken können. Aus diesem Grund sind Diskussionen über diese wichtigen Themen selbstverständlich zulässig. Weiter entfernen wir bestimmte Gewaltdarstellungen, da wir die Verherrlichung von Gewalt auf Instagram nicht dulden. Mit gewissen Einschränkungen sind Gewaltdarstellungen demgegenüber etwa dann zulässig, wenn sie für entsprechende Themen sensibilisieren. Ähnlich gestatten wir bestimmte Darstellungen von Nacktheit, beispielsweise im Zusammenhang mit Kunst und Protestaktionen oder in gesundheitsbezogenen Zusammenhängen, obwohl die Gemeinschaftsrichtlinien hier grundsätzlich Beschränkungen vorsehen.

Soweit der Bundesgerichtshof im Zusammenhang mit der Plattform Facebook zusätzlich gefordert hat, das Erfordernis eines sachlichen Grundes für die Entfernung von Inhalten und für die Sperrung von Nutzerkonten durch verfahrensrechtliche Vorkehrungen direkt in den Nutzungsbedingungen abzusichern, namentlich durch bestimmte Informations-, Anhörungs- und Einspruchsrechte, haben wir die notwendigen Klarstellungen und Anpassungen im Zuge der jüngsten Überarbeitung unserer Nutzungsbedingungen vorgenommen (neue Fassung gültig seit dem 26. Juli 2022).

Die Bestimmungen sind auch jeweils klar und verständlich. Insbesondere sind unsere Nutzungsbedingungen und die Gemeinschaftsrichtlinien in einfacher und allgemeinverständlicher Sprache gehalten.

Die Nutzungsbedingungen zählen die wesentlichen Vorgaben auf und enthalten insbesondere Links zu den Gemeinschaftsrichtlinien. Die Übersichtlichkeit der zentralen Gemeinschaftsrichtlinien wird durch eine optisch übersichtliche Struktur und eingängige Formulierungen gewährleistet. Dabei ist den einzelnen Vorgaben jeweils ein einfach und treffend formuliertes „Gebot“ vorangestellt, das dann im Folgenden allgemeinverständlich näher erläutert



wird. Auf diese Weise kann sich jede/r Nutzer\*in ein belastbares Bild davon machen, welche Inhalte auf Instagram zulässig sind und welche nicht.