



Unsere gemeinsame
Agenda

Kurzdossier 8

**Informations-
integrität
auf digitalen
Plattformen**

JUNI 2023



**Vereinte
Nationen**



Einleitung

VORWORT

Die Herausforderungen, vor denen wir stehen, lassen sich nur durch stärkere internationale Zusammenarbeit bewältigen. Der Zukunftsgipfel 2024 bietet Gelegenheit, multilaterale Lösungen für eine bessere Zukunft zu vereinbaren und die globale Ordnungspolitik für die heutigen und die kommenden Generationen zu stärken (Resolution [76/307](#) der Generalversammlung). In meiner Eigenschaft als Generalsekretär bin ich gebeten worden, zu den Gipfelvorbereitungen Beiträge in Form von handlungsorientierten Empfehlungen zu leisten, die auf den Vorschlägen in meinem Bericht Unsere gemeinsame Agenda ([A/75/982](#)), aufbauen, der seinerseits eine Antwort auf die Erklärung zum fünfundsiebzigsten Jahrestag des Bestehens der Vereinten Nationen (Resolution [75/1](#) der Generalversammlung) war. Dieses Kurzdossier ist ein solcher Beitrag. Es entwickelt die in Unserer gemeinsamen Agenda erstmals vorgestellten Ideen weiter, berücksichtigt spätere Handlungsanleitungen der Mitgliedstaaten der Vereinten Nationen und mehr als ein Jahr zwischenstaatlicher Konsultationen sowie Konsultationen mit einer Vielzahl von Interessenträgern und beruht auf den Zielen und Grundsätzen der Charta der Vereinten Nationen, der Allgemeinen Erklärung der Menschenrechte und anderen internationalen Rechtsinstrumenten.

ZWECK DIESES KURZDOSSIER

In diesem Kurzdossier geht es vorrangig darum, wie sich die Bedrohungen der Informationsintegrität auf die Fortschritte bei der Bewältigung globaler, nationaler und lokaler Probleme auswirken. In Unserer gemeinsamen Agenda forderte ich einen empirisch untermauerten Konsens auf der Grundlage von Fakten, Wissenschaft und Wissen. Zu diesem Zweck werden im vorliegenden Kurzdossier mögliche Prinzipien für einen Verhaltenskodex skizziert, der den Mitgliedstaaten, den digitalen Plattformen und anderen Interessenträgern bei ihren Bemühungen um einen inklusiveren und sichereren digitalen Raum für alle als Richtschnur dienen soll, wobei auch das Recht auf Meinungsfreiheit und freie Meinungsäußerung sowie das Recht auf Zugang zu Information vehement verteidigt wird. Der Verhaltenskodex für die Informationsintegrität auf digitalen Plattformen wird derzeit im Rahmen der Vorbereitungen für den Zukunftsgipfel erarbeitet. Ich hoffe, dass er als höchster Maßstab für Maßnahmen zur Stärkung der Informationsintegrität dienen wird.

Digitale Plattformen sind wichtige Instrumente, die die sozialen, kulturellen und politischen Interaktionen weltweit verändert haben. Überall auf der Welt verbinden sie besorgte Weltbürgerinnen und -bürger, die sich mit wichtigen Anliegen befassen. Die Plattformen unterstützen die Vereinten Nationen dabei, die Menschen in ihrem Streben nach Frieden, Würde und Gleichheit auf einem gesunden Planeten unmittelbar zu informieren und einzubeziehen. Sie haben Menschen in Krisen- und Konfliktzeiten Hoffnung gegeben,

Stimmen, die bislang kein Gewicht hatten, Gehör verschafft und globale Bewegungen in Gang gesetzt.

Doch dieselben Plattformen haben auch die Schattenseite des digitalen Ökosystems zutage gefördert. Sie haben die rasante Verbreitung von Lügen und Hass ermöglicht und damit weltweit erheblichen Schaden verursacht. Der Optimismus, den das Potenzial der sozialen Medien für die Vernetzung und Einbeziehung der Menschen auslöste, wurde gedämpft, als Fehl- und Desinformation sowie Hetze sich von Randerscheinungen zum Mainstream des digitalen Raums entwickelten. Die Gefahr kann nicht hoch genug eingeschätzt werden. Über soziale Medien verbreitete

Hetze und Desinformation kann zu Gewalt und Tod führen.¹ Die Möglichkeit, in großem Stil Desinformation zu betreiben und wissenschaftliche Erkenntnisse zu untergraben, stellt eine existenzielle Bedrohung für die Menschheit dar ([A/75/982](#), Ziff. 26) und gefährdet demokratische Institutionen und grundlegende Menschenrechte. Durch den raschen technologischen Fortschritt, etwa im Bereich der generativen künstlichen Intelligenz, haben sich diese Risiken weiter erhöht. Überall auf der Welt beobachten die Vereinten Nationen, wie Fehl- und Desinformation sowie Hetze die Fortschritte bei den Nachhaltigkeitszielen bedrohen können. Eines steht fest: Wir können nicht einfach weitermachen wie bisher.

IN DER PRAXIS SIND DIE UNTERSCHIEDE
ZWISCHEN FEHL- UND DESINFORMATION

OFT SUBTIL UND

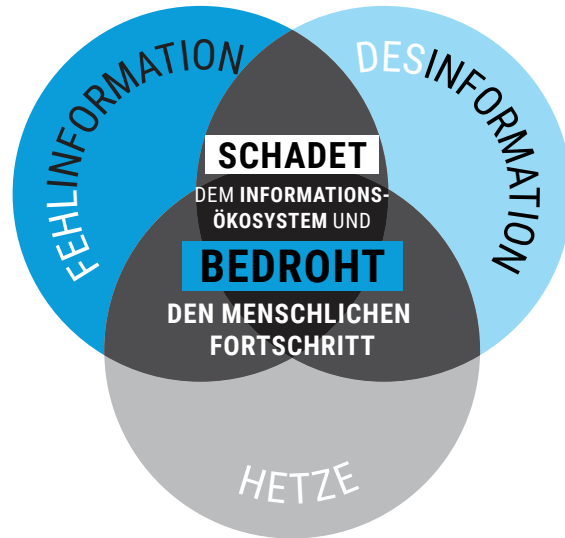
SCHWER

AUSZUMACHEN

Was bedeutet Informationsintegrität?

Informationsintegrität bezieht sich auf die Richtigkeit, Beständigkeit und Verlässlichkeit von Informationen. Desinformation, Fehlinformationen und Hetze stellen eine Bedrohung für sie dar. Zwar gibt es keine universell anerkannten Definitionen dieser Begriffe, die Institutionen der Vereinten Nationen haben jedoch Arbeitsdefinitionen erarbeitet.

Die Sonderberichterstatterin über die Förderung und den Schutz des Rechts auf Meinungsfreiheit und freie Meinungsäußerung definiert Desinformation als eine „Falschinformation, die mit der Absicht verbreitet wird, schweren sozialen Schaden anzurichten“.² Die Organisation der Vereinten Nationen für Bildung, Wissenschaft und Kultur (UNESCO) beschreibt Desinformation als falsche oder irreführende Inhalte, die ungeachtet der Motivation, des Bewusstseins oder des Verhaltens spezifischen Schaden verursachen können.³



Für die Zwecke dieses Kurzdossiers liegt der Unterschied zwischen Fehlinformation und Desinformation in der Absicht. Eine Desinformation ist eine Information, die nicht nur falsch ist, sondern bewusst verbreitet wird, um zu täuschen und Schaden anzurichten. Desinformation kann von staatlichen oder nichtstaatlichen Akteuren in unterschiedlichen Kontexten verbreitet werden, auch in bewaffneten Konflikten, und kann alle Entwicklungsbereiche betreffen, von Frieden und Sicherheit bis hin zu den Menschenrechten, dem öffentlichen Gesundheitswesen, zu humanitärer Hilfe und Klimamaßnahmen.

Fehlinformation bezieht sich auf die unbeabsichtigte Verbreitung falscher Informationen, die in gutem Glauben von Personen geteilt wurden, denen nicht bewusst war, dass sie Unwahrheiten verbreiten. Fehlinformationen können in Desinformation

wurzeln, da bewusste Lügen und irreführende Narrative mit der Zeit instrumentalisiert werden, in den öffentlichen Diskurs einfließen und unwissentlich weitergegeben werden.⁴ Eine Unterscheidung zwischen Fehl- und Desinformation kann sich in der Praxis schwierig gestalten.

Gemäß der Arbeitsdefinition der Strategie und des Aktionsplans der Vereinten Nationen gegen Hetze umfasst sie „jede Art von gesprochener oder geschriebener Kommunikation oder Verhaltensweise, durch die eine Person oder Gruppe aufgrund ihrer Identität – mit anderen Worten aufgrund ihrer Religion, Ethnizität, Staatsangehörigkeit, „Rasse“, Hautfarbe, Herkunft, Geschlecht oder eines anderen identitätsstiftenden Faktors – angegriffen oder aus diesen Gründen sprachlich abgewertet und diskriminiert wird.“⁵

Fehl- und Desinformation sowie Hetze sind verwandte, aber unterschiedliche Phänomene. Bei ihrer Ermittlung, Abschwächung und Bekämpfung gibt es gewisse Überlappungen, jedoch auch Unterschiede. Alle drei schaden dem Informationsökosystem und stellen eine Bedrohung für den Fortschritt der Menschheit dar.⁶

Bedrohungen der Informationsintegrität sind nichts Neues. Aus der Verbreitung von Lügen und Hass wird seit jeher politisches oder finanzielles Kapital geschlagen. Im digitalen Zeitalter kann

dieser Vorgang jedoch bisher unvorstellbare Ausmaße annehmen. Durch Kommunikationsmittel mit enormer Reichweite lassen sich Inhalte heutzutage in Sekundenschnelle über den gesamten Planeten verbreiten. Das dadurch entstehende Problem ist so weitreichend, dass selbst Online-Plattformen zeitweise nicht in der Lage sind, sein volles Ausmaß zu erkennen. Das Fehlen staatlich vereinbarter Definitionen für diese Phänomene sollte uns nicht zur Untätigkeit verleiten. Wir müssen alles daransetzen, den von ihnen verursachten Schaden einzudämmen.

Informationsintegrität und digitale Plattformen

Digitale Plattformen sollten ein integraler Bestandteil der Anstrengungen zur Aufrechterhaltung der Informationsintegrität sein. Zwar können auch bestimmte traditionelle Medien Fehl- und Desinformation verbreiten, doch die Geschwindigkeit, das Volumen und die Viralität ihrer Verbreitung über digitale Kanäle erfordern eine dringende und gezielte Reaktion. Für die Zwecke dieses Kurzdossiers bezeichnet der Begriff „digitale Plattform“ eine digitaler Dienstleister, der eine Interaktion zwischen zwei oder mehreren Nutzerinnen und Nutzern ermöglicht und unterschiedlichste Funktionen anbietet – von sozialen Medien über Suchmaschinen bis hin zu Kommunikationsanwendungen. In der Regel sammeln sie Daten über ihre Nutzerinnen und Nutzer und deren Interaktionen.⁷

Fehl- und Desinformation werden von einem breiten Spektrum an Akteurinnen und Akteuren mit unterschiedlichen Absichten erstellt, denen es weitgehend gelingt, ihre Anonymität zu wahren. Staatliche und nichtstaatliche Akteurinnen und Akteure haben sich die Schwachstellen in digitalen Systemen zunutze gemacht, um durch koordinierte Desinformationskampagnen schädliche Narrative zu verbreiten – mit schwerwiegenden Folgen.

Viele Staaten haben bereits Initiativen zur Regulierung digitaler Plattformen gestartet. In den letzten vier Jahren wurden mindestens 70 entsprechende Gesetze verabschiedet oder geprüft.⁸ Den Kern dieser rechtlichen Regelungen bildet für gewöhnlich ein enger Rahmen von Abhilfemaßnahmen

zur Definition und Entfernung schädlicher Inhalte. Da staatliche Maßnahmen vorwiegend auf die Entfernung schädlicher Inhalte abzielen, haben einige Staaten fehlerhafte und zu weit gefasste Rechtsvorschriften eingeführt, die den Schutz der Meinungsfreiheit, die nach dem Völkerrecht erlaubt ist, de facto beschnitten haben. Andere Gegenmaßnahmen wie die Pauschalabschaltungen des Internets und Verbote von Plattformen haben möglicherweise keine rechtliche Grundlage und könnten Menschenrechtsverletzungen darstellen.

Viele Staaten und Politikverantwortliche haben vermeintliche Bedenken in Bezug auf Informationsintegrität als Vorwand genutzt, um den Zugang zu Informationen zu beschränken, mediale Berichterstattung zu diskreditieren und einzuschränken und journalistisch Tätige sowie Gegnerinnen und Gegner ins Visier zu nehmen.⁹ Auch staatliche Akteurinnen und Akteure haben sich Plattformen unter dem Deckmantel der Bekämpfung von Falsch- und Desinformation gefügig gemacht.¹⁰ Expertinnen und Experten für das Recht auf freie Meinungsäußerung haben betont, dass staatlichen Akteurinnen und Akteuren in diesem Zusammenhang eine besondere Pflicht zukommt und dass sie „keine Fehlinformationen in die Welt setzen, unterstützen, begünstigen oder fördern sollten“ (A/77/287, Ziff. 45).

Die Risiken, die mit der Regulierung der Meinungsäußerung einhergehen, erfordern einen sorgfältig ausgearbeiteten Ansatz, der den Anforderungen der Rechtmäßigkeit, Notwendigkeit und

Verhältnismäßigkeit nach den Menschenrechtsnormen gerecht wird, selbst wenn aus legitimen Gründen des öffentlichen Interesses gehandelt wird (ebd., Ziff. 42).

Desinformation ist auch ein florierendes Geschäft. Firmen, die von Staaten, Politikverantwortlichen und dem Privatsektor engagiert werden, um „Dark PR“ oder reguläre Öffentlichkeitsarbeit für sie zu leisten, sind wesentliche Quellen für falsche und irreführende Inhalte.¹¹ Eine Taktik besteht darin, Inhalte auf gefälschten Website Klonen von Nachrichtenseiten zu veröffentlichen, um ihnen den Anstrich einer legitimen Quelle zu geben.¹² Da sich dieses undurchsichtige Geschäft schwer verfolgen und erforschen lässt, ist das tatsächliche Ausmaß des Problems nicht bekannt. Auch Einzelpersonen verbreiten falsche Behauptungen, um Produkte oder Dienstleistungen gegen Gewinn feilzubieten, und wählen dafür als Zielgruppe oft geschwächte Gruppen in unsicheren oder Krisenzeiten.

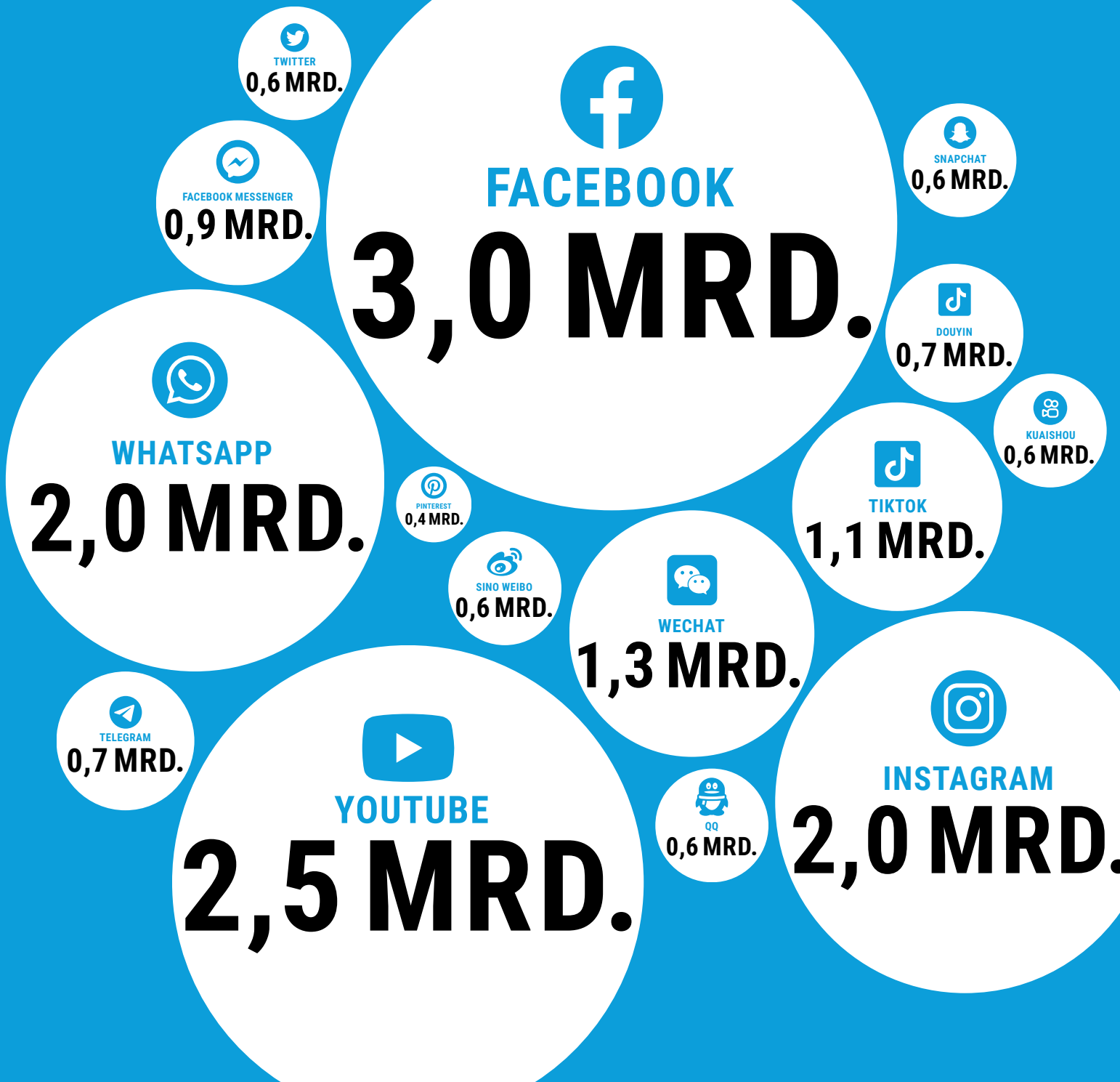
Neuerdings beruhen die Geschäftsmodelle der meisten digitalen Plattformen auf der „Aufmerksamkeitsökonomie“. Die Algorithmen sind so konzipiert, dass sie Inhalte priorisieren, die die Aufmerksamkeit der Nutzerinnen und Nutzer fesseln und so ihre Bindung und die Werbeeinnahmen maximieren können. Falsche und hasserfüllte Inhalte, die polarisieren und starke Emotionen hervorrufen sollen, werden oft am häufigsten konsumiert und geteilt. Dies hat zur Folge, dass Algorithmen nachweislich Falsch- und Desinformation sowie Hetze belohnen und verbreiten.¹³

Angesichts sinkender Werbeeinnahmen suchen digitale Plattformen nach alternativen Einnahmequellen, die über die Aufmerksamkeitsökonomie hinausgehen. So planen manche von ihnen eine kostenpflichtige Verifizierung, die Konten ein Gütesiegel zum Kauf anbietet, das vormals als Zeichen der Authentizität galt. Dieses Konzept hat ernsthafte Bedenken bezüglich der Informationsintegrität hervorgerufen, da es Missbrauch durch Desinformationsakteure zulässt.¹⁴

MILLIARDEN

NUTZEN SOZIALE MEDIEN

Quelle: Kepios, Digital 2023: Global Overview Report (2023)



Welcher völkerrechtliche Rahmen ist maßgeblich?

Die Förderung der Informationsintegrität muss vollständig in den einschlägigen völkerrechtlichen Normen und Standards verankert sein, einschließlich der Menschenrechtsnormen und der Grundsätze der Souveränität und Nichteinmischung in innere Angelegenheiten. Im August 2022 übermittelte ich der Generalversammlung einen Bericht mit dem Titel „Desinformationsbekämpfung zur Förderung und zum Schutz der Menschenrechte und Grundfreiheiten“.¹⁵ In dem Bericht habe ich die internationalen Menschenrechtsnormen dargelegt, die auf Desinformation zur Anwendung kommen, einschließlich der Allgemeinen Erklärung der Menschenrechte und des Internationalen Paktes über bürgerliche und politische Rechte. Nach diesen völkerrechtlichen Übereinkünften hat jede und jeder ein Recht auf freie Meinungsäußerung.¹⁶

Artikel 19 der Allgemeinen Erklärung der Menschenrechte und Artikel 19 Absatz 2 des Paktes verbiefen das Recht auf freie Meinungsäußerung, einschließlich der Freiheit, ohne Rücksicht auf Staatsgrenzen Informationen und Gedankengut jeder Art über Mittel jeglicher Art sich zu beschaffen, zu empfangen und zu verbreiten. Das Menschenrecht auf freie Meinungsäußerung beschränkt sich nicht auf Informationen mit positiver Resonanz ([A/77/287](#), Ziff. 13). Die Informationsfreiheit, die mit dem Recht auf freie Meinungsäußerung verknüpft ist, stellt ein eigenständiges Recht dar. Die Generalversammlung erklärte: „Die Informationsfreiheit ist ein

grundlegendes Menschenrecht und stellt den Maßstab für alle Freiheiten dar, denen sich die Vereinten Nationen verschrieben haben“ (ebd., Ziff. 14). Das Recht auf freie Meinungsäußerung und der Zugang zu Informationen können bestimmten Einschränkungen unterliegen, die den Kriterien in Artikel 19 Absatz 3 des Paktes entsprechen.¹⁷ Staaten können keine zusätzlichen Gründe anführen oder die Meinungsäußerung anhand von Kriterien, die über die nach dem Völkerrecht zulässigen hinausgehen, einschränken.

Der 2012 verabschiedete Aktionsplan von Rabat über das Verbot des Eintretens für nationalen, rassistischen oder religiösen Hass, durch das zu Diskriminierung, Feindseligkeit oder Gewalt aufgestachelt wird, enthält praktische rechtliche und politische Leitlinien dafür, wie die Staaten Artikel 20 Absatz 2 des Paktes und Artikel 4 des Internationalen Übereinkommens zur Beseitigung jeder Form von Rassendiskriminierung, die bestimmte Formen der Hetze verbieten, am besten anwenden können. Die Mitgliedstaaten haben den Aktionsplan von Rabat bereits in unterschiedlichen Kontexten angewendet.¹⁸

HETZE
IST EIN WEGBEREITER FÜR
GRÄUELSTATEN
DARUNTER
VÖLKERMORD

Hetze ist ein Wegbereiter für Gräueltverbrechen, darunter Völkermord. Die Konvention über die Verhütung und Bestrafung des Völkermordes von 1948 verbietet „die unmittelbare und öffentliche Aufstachelung zur Begehung von Völkermord“.

In ihrer 2021 verabschiedeten Resolution [76/227](#) unterstrich die Generalversammlung, dass sich alle Formen von Desinformation negativ auf die

Ausübung der Menschenrechte und Grundfreiheiten sowie die Erreichung der Ziele für nachhaltige Entwicklung auswirken können. Gleichmaßen bekräftigte der Menschenrechtsrat in seiner Resolution [49/21](#) von 2022, dass Desinformation den Genuss und die Verwirklichung aller Menschenrechte beeinträchtigen kann.

75 %

**DER VN-FRIEDENSSICHERUNGSKRÄFTE
GABEN AN, DASS**

**FEHL- UND
DESINFORMATION**

IHRE

SICHERHEIT

BEEINTRÄCHTIGTE

INTERNE UMFRAGE BEI DEN VN (2022)

Welchen Schaden verursachen Fehl- und Desinformation sowie Hetze im Internet?

Online verbreitete Fehl- und Desinformation sowie Hetze stellen eine ernsthafte Besorgnis für die Weltöffentlichkeit dar. Die Analyse einer in 142 Ländern durchgeführten Umfrage ergab, dass 58,5 Prozent der regulären Nutzerinnen und Nutzer des Internets und sozialer Medien weltweit besorgt darüber sind, im Internet Fehlinformationen ausgesetzt zu sein, wobei sich junge Menschen und Bevölkerungsschichten mit niedrigem Einkommen besonders verwundbar fühlen.¹⁹ Die Jugend von heute ist mit der digitalen Welt aufgewachsen und eher über das Internet verbunden als die restliche Bevölkerung. Damit gehören sie der am stärksten digital vernetzten Generation in der Geschichte an.²⁰ Weltweit benutzt jede halbe Sekunde ein Kind zum ersten Mal das Internet und läuft damit Gefahr, Hetze zu begegnen und Schaden zu erleiden. In manchen Fällen führt dies zur Beeinträchtigung ihrer psychischen Gesundheit.²¹

Die Auswirkungen von Fehl- und Desinformation und Hetze im Internet lassen sich überall auf der Welt beobachten, auch in den Bereichen

EINE UMFRAGE IN

142 LÄNDERN ERGAB, DASS

58,5 %

DER REGULÄREN NUTZERINNEN UND NUTZER
DES INTERNETS UND SOZIALER MEDIEN
WELTWEIT BESORGT DARÜBER SIND, IM

INTERNET FEHLINFORMATIONEN
AUSGESETZT ZU SEIN.

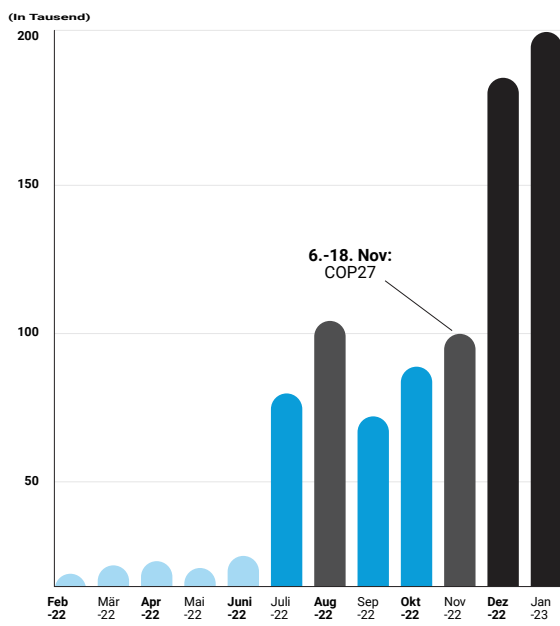
Gesundheit, Klimamaßnahmen, Demokratie und Wahlen, Gleichstellung der Geschlechter, Sicherheit und humanitäre Maßnahmen. Laut einer Umfrage aus dem Jahr 2021 sehen 75 Prozent der Landesbüros des Entwicklungsprogramms der Vereinten Nationen Informationsverschmutzung als erhebliches Problem an. Sie hat schwerwiegende Folgen für Vertrauen, Sicherheit, Demokratie und nachhaltige Entwicklung, wie eine kürzlich von der UNESCO in Auftrag gegebene Überprüfung von über 800 wissenschaftlichen, zivilgesellschaftlichen, journalistischen und Unternehmensdokumenten ergab.²²

Fehl- und Desinformation können gefährlich und unter Umständen tödlich sein, insbesondere in Zeiten von Krisen, Notfällen oder Konflikten. Während der COVID-19-Pandemie begann eine Flut an Fehl- und Desinformation über den Virus, Maßnahmen der öffentlichen Gesundheit und Impfungen im Internet zu zirkulieren.²³ Bestimmte Akteurinnen und Akteure machten sich die Verwirrung für ihre eigenen Zwecke zunutze. So lenkten Impfgegnerinnen und -gegner die Nutzerinnen und Nutzer auf Websites, auf denen gefälschte Heilmittel oder Präventivmaßnahmen verkauft wurden.²⁴ Viele Opfer von COVID-19 verweigerten die Impfung oder grundlegende gesundheitliche Vorsorgemaßnahmen, nachdem sie Fehl- und Desinformation im Internet ausgesetzt gewesen waren.²⁵

Auch in bereits unbeständigen gesellschaftlichen und politischen Kontexten kann Desinformation tödliche Folgen haben. In einem 2022 veröffentlichten Bericht untersuchte die Sonderberichterstatterin über die Förderung und den Schutz des Rechts auf Meinungsfreiheit und freie Meinungsäußerung, welche Rolle die Instrumentalisierung von Information dabei spielt, Verwirrung zu stiften, Hass zu schüren, zu Gewalt aufzustacheln und Konflikte zu verlängern.²⁶ In einem weiteren Bericht von 2022 wurde festgestellt, dass Desinformation auch „Intoleranz und Hetze gegen Minderheiten, Frauen und alle sogenannten „anderen“ beinhalten kann und somit nicht nur diejenigen gefährdet, gegen die sie sich unmittelbar richtet, sondern auch die Inklusion und den sozialen Zusammenhalt. Sie kann Spannungen und Spaltungen in Zeiten von Notfällen, Krisen, bewaffneten Konflikten oder zu politischen Wendepunkten verstärken“.²⁷

ABBILDUNG I

MONATLICHE VERWENDUNG DES HASHTAGS #CLIMATESCAM AUF TWITTER



Quelle: Hauptabteilung Globale Kommunikation, Daten von Talkwalker

Den größten Schaden verursacht die Nutzung von Online-Plattformen mitunter in Gebieten, die von den Plattformen trotz hoher Nutzerzahlen vernachlässigt werden. Länder, in denen aktive Konflikte wüten oder die aus anderen Gründen ein unbeständiges Umfeld darstellen, sind oft weniger lukrative Märkte. Daher werden sie nicht mit ausreichenden Ressourcen für die Inhaltsmoderation und Benutzerhilfe versorgt. Obwohl traditionelle Medien für die meisten Menschen in Konfliktregionen nach wie vor eine wichtige Nachrichtenquelle sind, hat über digitale Plattformen verbreiteter Hass bereits Gewalt entfacht und geschürt.²⁸ Manche digitale Plattformen wurden für ihre Rolle in Konflikten kritisiert, auch während des andauernden Krieges in der Ukraine.²⁹

Gleichermaßen führen Fehl- und Desinformation über den Klimanotstand zu Verzögerungen bei Maßnahmen, die dringend erforderlich sind, um eine lebenswerte Zukunft auf unserem Planeten zu sichern. Zu Fehl- und Desinformation über das Klima können falsche oder irreführende Inhalte gezählt werden, die die wissenschaftlich anerkannte Grundlage für die Existenz des vom Menschen verursachten Klimawandels, seiner Ursachen und Auswirkungen untergraben. Koordinierte Kampagnen zielen darauf ab, den wissenschaftlichen Konsens des Zwischenstaatlichen Ausschusses für Klimaänderungen zu leugnen, zu schmälern oder davon abzulenken und Maßnahmen, die zur Erfüllung der Ziele des Übereinkommens von Paris (2015) dringend erforderlich sind, zu Fall zu bringen. Eine kleine, aber lautstarke Minderheit von Leugnerinnen und Leugnern der Klimawissenschaften³⁰ lehnen den Konsens weiterhin ab und haben auf einigen digitalen Plattformen eine überragende Präsenz. So ergaben 2022 durchgeführte zufällige Simulationen zivilgesellschaftlicher Organisationen, dass der Algorithmus von Facebook Inhalte, die den Klimawandel leugnen, zulasten klimawissenschaftlicher Beiträge empfiehlt.³¹ Auf Twitter stieg die Verwendung des Hashtags #climatescam von weniger als 2.700 pro Monat in der

70%

DER VN-FRIEDENSSICHERUNGSKRÄFTE GABEN AN, DASS

FEHL- UND DESINFORMATIONEN

SICH **GRAVIEREND,
ERHEBLICH**

ODER MÄßIG AUF IHRE ARBEIT AUSWIRKEN.

ersten Jahreshälfte 2022 auf 80.000 im Juli und auf 199.000 im Januar 2023. Bei der Suche nach dem Stichwort „Klima“ tauchte der Begriff auf Twitter als eines der wichtigsten Suchergebnisse auf.³² Im Februar 2022 hat der Zwischenstaatliche Ausschuss für Klimaänderungen die Desinformation über den Klimawandel zum ersten Mal angeprangert und erklärt, dass ein „absichtliches Untergraben der Wissenschaft“ zu einer „falschen Wahrnehmung des wissenschaftlichen Konsenses, zu Unsicherheit und zu Dissens beiträgt und zu einer Vernachlässigung der Risiken und der Dringlichkeit führt“.³³

Manche Unternehmen für fossile Brennstoffe betreiben häufig „Grünfärberei“. Dabei wird der Öffentlichkeit vorgegaukelt, dass ein Unternehmen oder eine Einrichtung sich stärker für den Umweltschutz engagiert und die Umwelt weniger schädigt, als in Wirklichkeit der Fall ist. Dabei handeln die Unternehmen nicht allein. Werbe- und PR-Firmen, werbetreibende Technologieunternehmen, Nachrichtenagenturen und digitale Plattformen ermöglichen und unterstützen die Bemühungen, die Öffentlichkeit in die Irre zu führen und die Aufmerksamkeit von der Verantwortung der Fossilbrennstoffindustrie abzulenken.³⁴ Werbeunternehmen und Unternehmen für Öffentlichkeitsarbeit, die „grünfärbende“ Inhalte verfassen, und Dritte, die diese Inhalte verbreiten, verdienen gemeinsam Milliarden daran, die Fossilbrennstoffindustrie der Prüfung und Rechenschaftspflicht zu entziehen. PR-Firmen haben Hunderte von Kampagnen für Kohle-, Öl- und Gasunternehmen durchgeführt.³⁵

Fehl- und Desinformation haben schwerwiegende Auswirkungen auf die Demokratie, denn sie schwächen das Vertrauen in demokratische Institutionen und die unabhängigen Medien und schwächen die Beteiligung am politischen und öffentlichen Leben. Wird die Wählerschaft während eines Wahlzyklus falschen und irreführenden Informationen ausgesetzt, kann sie der Möglichkeit beraubt werden, fundierte Entscheidungen zu treffen. Die Verbreitung von Fehl- und Desinformation kann das Vertrauen der Öffentlichkeit in die Wahlinstitutionen und -vorgänge selbst untergraben, etwa die Wählerregistrierung, Stimmabgabe und Wahlergebnisse, und möglicherweise zu Wahlverdrossenheit oder zur Ablehnung glaubwürdiger Wahlergebnisse führen. Auch Staatsoberhäupter und politische Führungsverantwortliche haben sich als ergiebige Quelle für Desinformation erwiesen, indem sie bewusst und strategisch Unwahrheiten verbreiten, um an der Macht zu bleiben oder sie zu sichern oder um demokratische Prozesse in anderen Ländern zu untergraben.³⁶

Marginalisierte und schutzbedürftige Gruppen sind auch häufig Ziel von Fehl- und Desinformation und Hetze, durch die sie noch weiter in die soziale, wirtschaftliche und politische Ausgrenzung getrieben werden. Kandidatinnen, Wählerinnen, Mitarbeiterinnen der Wahlaufsicht, Journalistinnen und Vertreterinnen der Zivilgesellschaft werden online gezielt mit geschlechtsspezifischer Desinformation angegriffen.³⁷ Solche Angriffe untergraben die politische Teilhabe und schwächen die demokratischen Institutionen und die Menschenrechte, unter anderem das Recht auf freie Meinungsäußerung und den Zugang zu Informationen dieser Gruppen.³⁸ Dies muss verstärkt zu einer vordringlichen Priorität der internationalen Gemeinschaft werden, nicht zuletzt da 2024 mehr als 2 Milliarden Wählerinnen und Wähler weltweit an die Urnen gerufen werden.

Fehl- und Desinformation treiben auch zwischen und in den Online-Plattformen und traditionellen Medien weitere Blüten und sind daher immer schwerer nachzuverfolgen und zu bekämpfen,

sofern sie nicht an der Quelle aufgedeckt werden. Desinformation kann eine bewusste Taktik von Medien mit ideologischer Prägung sein, die von politischen und wirtschaftlichen Interessen vereinnahmt wurden.³⁹ Zugleich hat das Aufkommen digitaler Plattformen einen dramatischen Rückgang vertrauenswürdiger, unabhängiger Medien ausgelöst. Die informierte Öffentlichkeit und die Werbetreibenden sind scharenweise zu den Internet-Plattformen übergelaufen – ein Trend, der durch die COVID-19-Pandemie noch verschärft wurde. In manchen Regionen oder Ländern spricht man vom „Aussterben der Medien“ oder so genannten „Nachrichtenwüsten“⁴⁰, durch die Gemeinwesen ihre verlässlichen lokalen Nachrichtenquellen verlieren. Diese Phänomene tragen zur Belastung des Informationsökosystems bei. „News-washing“ – von Sponsoren platzierte Inhalte, die im Gewand seriös recherchierter Nachrichtenmeldungen daherkommen – wird auf digitalen Plattformen häufig nicht hinreichend kenntlich gemacht und verleiht solchen Inhalten eine falsche Legitimität. Werden diese Informationen von anderen Medien aufgegriffen, von Politikverantwortlichen zitiert oder auf einer Vielzahl von Plattformen verbreitet, wird ihre ursprüngliche Quelle immer

ungewisser, und die Nachrichtenkonsumentinnen und Konsumenten sind nicht mehr in der Lage, sie von objektiven Tatsachen zu unterscheiden.

Desinformation wirkt sich auch unmittelbar auf die Arbeit der Vereinten Nationen aus. Residierende Koordinatorinnen und Koordinatoren, Mittelspersonen und Friedenssicherungskräfte haben ihre Besorgnis über die Auswirkungen von Desinformation auf die operative Sicherheit, Wirksamkeit und Leistungsfähigkeit der Organisation geäußert. In einer Umfrage aus dem Jahr 2022 erklärten 70 Prozent der Friedenssicherungskräfte der Vereinten Nationen, dass Fehlinformationen und Desinformation sich gravierend, erheblich oder mäßig auf ihre Arbeit auswirken, während 75 Prozent angaben, ihre Sicherheit werde dadurch beeinträchtigt. Fehlinformationen und Desinformation können auch dazu verwendet werden, humanitäre Akteure anzugreifen und lebensrettende Einsätze in Konfliktgebieten zu behindern.

ABBILDUNG II INFORMATIONSIINTEGRITÄT UND DIE ZIELE FÜR NACHHALTIGE ENTWICKLUNG

Wie in Abbildung III zu erkennen, kann sich eine Bedrohung der Informationsintegrität unter Umständen auch negativ auf die Verwirklichung der Ziele für nachhaltige Entwicklung auswirken.

	<p>Fehl- und Desinformation wirken sich nach wie vor auf die Anstrengungen zur Beseitigung der Armut und auf die Weltwirtschaft aus. Wirtschaftliche Härten können zudem die Verbreitung polarisierender und hasserfüllter Lügen anfachen, unter anderem über marginalisierte Gruppen. Die Krise der Lebenshaltungskosten war ein besonders fruchtbarer Nährboden für Desinformation, darunter die Behauptung, steigende Energiekosten oder Arbeitsplatzverluste seien auf den Umstieg auf erneuerbare Energien zurückzuführen.</p>		<p>Im Internet verbreitete Hetze und Fehl- und Desinformation polarisieren Gesellschaften und richten sich gegen ohnehin bereits marginalisierte und gefährdete Gemeinschaften, wodurch sich ihre soziale, wirtschaftliche und politische Ausgrenzung weiter verschärfen kann.</p>
	<p>Die Bedrohung der Informationsintegrität kann den Hunger weltweit verstärken, unter anderem durch die Verschärfung der Konflikte, des Klimawandels, der Naturkatastrophen, Armut und Ungleichheit. Desinformation kann von konfliktbedingten Herausforderungen für die globale Ernährungssicherheit ablenken.</p>		<p>Nachhaltigkeitsbemühungen von Städten und Gemeinden können durch Desinformation untergraben werden, die von den Folgen menschlichen Handelns für die Umwelt ablenkt oder diese leugnet.</p>
	<p>Während der COVID-19-Pandemie untergrub eine „Infodemie“ miteinander zusammenhängender Fehl- und Desinformation Gesundheitsmaßnahmen und Impfkampagnen. Das Risiko für die Gesundheit und das Wohlergehen von Kindern infolge impfskeptischer Fehl- und Desinformation ist noch nicht gebannt.</p>		<p>Aktivistinnen und Aktivisten von Initiativen für Kreislaufwirtschaft und Nullverschwendung wurden zur Zielscheibe von Hass und Desinformation im Internet.</p>
	<p>Fehl- und Desinformation sowie Hetze können den Zugang zu hochwertiger Bildung beeinträchtigen, insbesondere für marginalisierte Gruppen, darunter junge Frauen und Mädchen. Der Zugang zu Informationen und Initiativen zur Förderung der digitalen Kompetenz sowie der Resilienz werden bei der Eindämmung der gesellschaftlichen Folgen der durch Internetnutzung verursachten Schäden eine Schlüsselrolle spielen.</p>		<p>Desinformation über das Klima und die daraus folgende Untätigkeit torpedieren dringend notwendige Maßnahmen zur Bewältigung der Klimakrise, unter anderem den entscheidenden Umstieg von umweltschädlichen fossilen Brennstoffen auf erneuerbare Energie sowie dringende Investitionen in die Klimaresilienz.</p>
	<p>Mit geschlechtsspezifischer Hetze und Desinformation sollen Frauen systematisch unterdrückt werden, indem sie zum Verstummen gebracht und aus dem öffentlichen Raum verdrängt werden. Dies kann verheerende Folgen haben, von der Unterdrückung weiblicher Stimmen über verstärkte Selbstzensur bis hin zu beruflichen Nachteilen, Rufschädigung und körperlicher Gewalt.</p>		<p>Fehl- und Desinformation können nachteilige Auswirkungen auf die Initiativen zur Erhaltung und nachhaltigen Nutzung der Ozeane, Meere und Meeresressourcen haben.</p>
	<p>Zwei Milliarden Menschen haben keine sicher verwaltete Trinkwasserversorgung. Fehl- und Desinformation über die Sicherheit von Trinkwasser und Sanitäreinrichtungen können schwere Folgen für die Gesundheit haben.^a</p>		<p>Umweltaktivistinnen und -aktivisten, die sich für den Schutz des Lebens an Land einsetzen, wurden Gegenstand von Hass- und Desinformationskampagnen im Internet, was Folgen im wirklichen Leben nach sich zog. Fehl- und Desinformation über das Klima behindern Klimamaßnahmen.^c</p>
	<p>Fehl- und Desinformation über das Klima, die zum Großteil von der Fossilenergiewirtschaft gestreut werden, sabotieren den dringend notwendigen Übergang zu saubereren Formen der Energieerzeugung und verengen damit das sich schließende Zeitfenster für die Herbeiführung einer nachhaltigen Zukunft für alle.</p>		<p>Desinformation und Hetze werden genutzt, um Wahlen und öffentliche Narrative zu beeinflussen und Verwirrung zu stiften. Sie dienen dazu, Gegner auszuschalten, friedensschaffende Maßnahmen zu durchkreuzen, zu Gewalt anzustacheln, Konflikte zu verlängern und das Vertrauen in die Rechtsstaatlichkeit zu zerrütten. Bemühungen um die Förderung friedlicher und inklusiver Gesellschaften sowie die Rolle der Vereinten Nationen bei der Förderung von Frieden und Sicherheit wurden dadurch erheblich beeinträchtigt.^d</p>
	<p>Wissenschaftliche Studien belegen die nachteiligen Auswirkungen von Hetze und Fehl- und Desinformation auf die Wirtschaft.^b</p>		<p>Hetze, Fehl- und Desinformation können konstruktive Partnerschaften daran hindern, die Nachhaltigkeitsziele zu erreichen, derweil Ressourcen, die zur Problembehandlung umgewidmet wurden, unter Umständen die Bestrebungen schwächen, niemanden zurückzulassen.</p>
	<p>Fehl- und Desinformation sowie Hetze und Pauschalantworten können sich schädlich auf die Innovation auswirken, etwa indem sie das Potenzial marginalisierter Gruppen einschränken und die Gleichberechtigung und Inklusion in digitalen Räumen verringern.</p>		

a World Health Organization und United Nations Children's Fund (UNICEF), *Progress on Household Drinking Water, Sanitation and Hygiene 2000–2020: Five Years into the SDGs* (Genf 2021).

b Siehe Roberto Cavazos und CHEQ, "The economic cost of bad actors on the Internet: fake news, 2019"; und London Economics, "The cost of lies: assessing the human and financial impact of COVID-19 related online misinformation on the UK", Dezember 2020.

c Global Witness, *Last Line of Defence: The Industries Causing the Climate Crisis and Attacks against Land and Environmental Defenders* (2021).

d A/77/288.

ZEITSPANNE, IN DER WELTWEIT

100 MILLIONEN

**AKTIVE NUTZERINNEN
UND NUTZER IM MONAT
VERZEICHNET WURDEN**

Quelle: Similarweb unter Verwendung der Daten von Sensor Tower



INSTAGRAM

30 MONATE



TIKTOK

9 MONATE



CHATGPT

2 MONATE

Wie können wir die Informationsintegrität stärken?

Fehl- und Desinformation und Hetze entstehen nicht im luftleeren Raum. Sie breiten sich aus, wenn Menschen sich ausgegrenzt fühlen, kein Gehör finden, den Folgen der wirtschaftlichen Ungleichheit ausgesetzt sind und Politikverdrossenheit empfinden. Diesen realen Herausforderungen sollten wir mit geeigneten Maßnahmen begegnen. Die Bemühungen um die Ziele für nachhaltige Entwicklung sind entscheidend, wenn wir künftig in einer Welt leben wollen, in der Vertrauen wiederhergestellt werden kann.

Beim Entwurf einer Gegenstrategie ist es wichtig, den enormen Wert digitaler Plattformen für unsere Welt nicht außer Acht zu lassen. Sie haben die in Echtzeit erfolgende Massenkommunikation revolutioniert und ermöglichen die Verbreitung lebensrettender Informationen während Naturkatastrophen und Pandemien. Sie konnten auch Unterstützung für die von den Vereinten Nationen angestrebten Ziele mobilisieren und haben sich häufig als positive Kräfte für Inklusion und Teilhabe am öffentlichen Leben erwiesen. Sie haben Gruppen von Menschen miteinander verbunden, die geografisch weit entfernt leben und von ihren Gesellschaften ausgegrenzt werden, unter anderem Menschen mit seltenen Erkrankungen, und ein breites Spektrum von Aktivistinnen und Aktivisten vernetzt, die sich für eine bessere Welt einsetzen.

REGULIERUNGSMASSNAHMEN

Die Frage, ob digitale Plattformen für die von ihnen bereitgestellten Inhalte gesetzlich haftbar gemacht werden können und sollten, war bereits Gegenstand einer ausführlichen Debatte. In manchen Fällen wurden bestehende Gesetze

zu Tatbeständen wie Verleumdung, Cybermobbing und Belästigung dazu herangezogen, Bedrohungen der Informationsintegrität wirksam abzuwehren, ohne dabei dem Recht auf freie Meinungsäußerung weitere Beschränkungen aufzuerlegen ([A/77/287](#), Ziff. 44).

Zudem sind in jüngerer Zeit auf regionaler und nationaler Ebene einige Gesetzesinitiativen zur Behandlung dieser Problematik angestoßen worden. Hierzu zählt der von der Europäischen Union im Jahr 2022 verabschiedete Rahmen, der das Gesetz über digitale Dienste, den Entwurf für eine Verordnung über die Transparenz und das Targeting politischer Werbung und den Verhaltenskodex für den Bereich der Desinformation umfasst. Das Gesetz über digitale Dienste legt neue Regeln für Nutzerinnen und Nutzer, digitale Plattformen und in der Europäischen Union tätige Online-Unternehmen fest. Die Maßnahmen zielen auf illegale Online-Inhalte, -Waren und -Dienstleistungen ab und sehen einen Mechanismus für Nutzerinnen und Nutzer vor, mit dem diese zum einen auf illegale Inhalte aufmerksam machen, zum anderen zu ihren Ungunsten gefallene Moderationsentscheidungen anfechten können. Mit den Maßnahmen werden die digitalen Plattformen zu mehr Transparenz verpflichtet, insbesondere im Hinblick auf den Einsatz und die Beschaffenheit von Empfehlungsalgorithmen, während größere Plattformen dazu verpflichtet werden, Wissenschaftlerinnen und Wissenschaftlern Datenzugang zu gewähren.

Der Verhaltenskodex für den Bereich der Desinformation legt Grundsätze und Verpflichtungen für Online-Plattformen und Werbeunternehmen fest, zu deren Umsetzung sich die Unterzeichner verpflichtet haben, um der Online-Verbreitung

von Desinformation in der Europäischen Union entgegenzuwirken.⁴¹ Dazu zählen freiwillige Verpflichtungen, um der Desinformation die Finanzierung zu entziehen, indem einerseits die Verbreitung von Werbeanzeigen mit Desinformation verhindert wird, und andererseits vermieden werden soll, dass Anzeigen neben Inhalten platziert werden, die ihrerseits Desinformation enthalten. Die Unterzeichner haben sich auch dazu bereit erklärt, politische Werbung deutlicher zu kennzeichnen und Angaben zu Sponsoren, Werbungskosten und Anzeigezeitraum zu machen sowie Datenbanken mit einer Suchfunktion für politische Werbung einzurichten. Zudem verpflichteten sie sich, Informationen über auf ihren Plattformen festgestelltes böswilliges und manipulatives Verhalten zwecks Verbreitung von Desinformation (wie beispielsweise Fake-Accounts, Bot-gesteuerte Vervielfältigungen, Annahme falscher Identitäten und böswillige Deepfakes) auszutauschen und regelmäßig Maßnahmen zu dessen Abwehr zu ergreifen und sie zu aktualisieren. Weitere Verpflichtungen dienen schwerpunktmäßig dazu, Nutzerinnen und Nutzer in die Lage zu versetzen, Desinformation zu erkennen, zu durchschauen und zu melden, die Zusammenarbeit mit Faktenprüfern zu stärken und Wissenschaftlerinnen und Wissenschaftlern besseren Zugang zu Daten zu verschaffen. Die eigentliche Bewährungsprobe für diese neuen Mechanismen liegt in ihrer Umsetzung.

Eines der Hauptziele des Verhaltenskodex für den Bereich der Desinformation besteht darin, die Transparenz der Plattformen zu verbessern. Im Februar 2023 veröffentlichten die Unterzeichner des Verhaltenskodex ihre ersten Grundlagenberichte zur schrittweisen Erfüllung ihrer Verpflichtungen. Die Berichte gewährten Aufschluss darüber, inwieweit der Rückfluss von Werbeeinnahmen an Desinformationsakteure sowie zugunsten anderer festgestellter manipulativer Handlungen verhindert wurde, unter anderem bei einer breit angelegten und konzertierten Aktion zur

Manipulation der öffentlichen Meinung in mehreren europäischen Ländern bezüglich des Krieges in der Ukraine.⁴²

REAKTIONEN DER DIGITALEN PLATTFORMEN

Digitale Plattformen unterscheiden sich sehr stark in ihrer Größe, Funktion und Struktur und haben eine Vielzahl von Maßnahmen ergriffen, um schädliches Verhalten zu unterbinden. Einige der größeren Plattformen haben sich öffentlich verpflichtet, die Leitprinzipien für Wirtschaft und Menschenrechte⁴³ zu wahren, bei den Maßnahmen, der Transparenz und der Umsetzung bestehen jedoch nach wie vor Lücken. Manche Plattformen versäumen es, die eigenen Standards durchzusetzen, und tolerieren oder verstärken in unterschiedlichem Maße Lügen und Hetze.⁴⁴ Algorithmen, die geschaffen wurden, um das Gewinnmodell der Plattformen zum Erfolg zu führen, sind darauf ausgelegt, eine größtmögliche Bindung der Nutzerinnen und Nutzer zu erzielen und ihre Aufmerksamkeit zu monopolisieren; tendenziell lenken sie sie in Richtung polarisierender oder provozierender Inhalte.

Die Mehrzahl der digitalen Plattformen verfügt über eine Art Selbstregulierungssystem, Moderation oder Aufsichtsinstanz, jedoch erweist es sich weiterhin als schwierig, im Hinblick auf ihre Richtlinien zur Löschung von Inhalten und deren praktische Umsetzung Transparenz zu wahren.⁴⁵ Die Investitionen in solche Mechanismen sind je nach Region und Sprache extrem lückenhaft und weitgehend auf den globalen Norden konzentriert; gleiches gilt für die Durchsetzung der Regeln, die sich die Plattformen selbst auferlegt haben. Eine jüngst durchgeführte Erhebung ergab, dass die Übersetzung von Moderationsinstrumenten und Aufsichtsinstanzen in lokale Sprachen über alle Plattformen hinweg unvollständig ist.⁴⁶ Zugleich wird die Moderation häufig an Fremdfirmen

ausgelagert und ist in allen Sprachen mit Ausnahme des Englischen beklagenswert schlecht ausgestattet.⁴⁷ Die Aussagen von Moderatorinnen und Moderatoren werfen beunruhigende Fragen im Hinblick auf Misshandlung, Arbeitsnormen und Sekundärtraumatisierung auf.⁴⁸ Sie berichten, dass sie ständig gewaltsamen und verstörenden Inhalten ausgesetzt sind und binnen weniger Sekunden entscheiden müssen, ob Beiträge gegen Unternehmensrichtlinien verstoßen oder nicht. Automatisierte Systeme für die Inhaltsmoderation können eine wesentliche Rolle spielen, bergen jedoch das Risiko der Voreingenommenheit, je nachdem, mit welchen Daten oder Strukturen sie angelern wurden. Sie weisen zudem hohe Fehlerquoten in englischer Sprache und noch weitaus geringere Erfolgsraten in anderen Sprachen auf. Einige digitale Plattformen beschäftigen Teams für Vertrauen und Sicherheit, Menschenrechte und Informationsintegrität, doch diese Fachleute werden zumeist nicht in die Frühphase der Produktentwicklung eingebunden. Zudem fallen ihre Stellen häufig als erste eventuellen Kosteneinsparungen zum Opfer.

DATENZUGANG

Der Datenzugang für Wissenschaftlerinnen und Wissenschaftler ist weltweit ebenfalls eine vordringliche Priorität. Bisherige Forschungsarbeiten und Ressourcen konzentrieren sich ganz überwiegend auf die Vereinigten Staaten von Amerika und Europa. Von einigen besonderen Ausnahmen abgesehen, darunter die Berichte über die Friedenssicherung durch die Vereinten Nationen in Afrika und der unabhängigen internationalen Ermittlungsmission für Myanmar⁴⁹ sowie einige Fälle investigativer Recherche und Berichterstattung durch Journalistinnen und Journalisten⁵⁰, wurden bislang nur wenige Forschungsergebnisse zu den Auswirkungen auf andere Weltregionen publiziert. Dies liegt teilweise daran, dass die Wissenschaftlerinnen und Wissenschaftler

keinen Zugang zu Plattformen und deren Daten haben. Die Instrumente, die für eine aussagekräftige wissenschaftliche Untersuchung der in begrenztem Umfang von den Plattformen freigegebenen Daten erforderlich sind, wurden in der Regel für Marketingzwecke erstellt und sind zum Großteil unerschwinglich. Ein Sinneswandel auf Seiten der Plattformen – weg von der Devise „Zugang nur auf Anfrage“ hin zu einer Offenlegung im Regelfall unter Wahrung der erforderlichen Datenschutzbestimmungen – würde es der Wissenschaft ermöglichen, Schäden und Risiken angemessen zu bewerten.

STÄRKUNG DER SELBSTBESTIMMUNG VON NUTZERINNEN UND NUTZERN

Zivilgesellschaftliche Gruppen und Hochschulen haben umfangreiche Forschungsarbeit zu der Frage geleistet, wie Fehlinformationen, Desinformation und Hetze am besten bekämpft werden können, ohne den Schutz des Rechts auf freie Meinungsäußerung zu beeinträchtigen. Einige von ihnen haben betont, dass es Lösungen bedarf, die an der Basis ansetzen und Internetnutzerinnen und Nutzer dazu ermächtigen, die Schädigung ihrer eigenen Gemeinschaften im Internet zu begrenzen und die Konzentration der Macht in den Händen einiger weniger Plattformen aufzulösen.

Nutzerinnen und Nutzer von Plattformen, darunter marginalisierte Gruppen, sollten zur Teilnahme an der Politikgestaltung ermutigt, einbezogen und beteiligt werden. Insbesondere Jugendliche verfügen über einen breiten Wissens- und Erfahrungsschatz. Da sie mit digitalen Technologien aufwachsen, geraten junge Menschen, insbesondere junge Frauen und Kinder, bereits häufig ins Visier von Hetze und Fehl- und Desinformation und sind damit von neu aufkommenden und aktuellen Plattformen unmittelbar betroffen. Jüngere

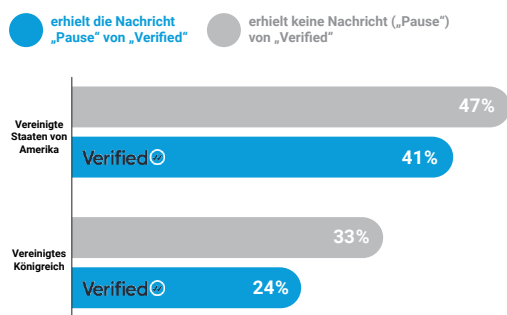
Nutzerinnen und Nutzer können aus eigener Erfahrung von den unterschiedlichen Auswirkungen der verschiedenen Angebote und ihren potenziellen Mängeln berichten. Sie haben sich zudem aktiv an der Interessenvertretung im Internet sowie an Initiativen zur Faktenüberprüfung beteiligt.⁵¹

Durch ein geschultes kritisches Denkvermögen können Nutzerinnen und Nutzer digitale Manipulationsversuchen besser abwehren. Konkret können sie durch digitale Kompetenz lernen, Informationen, denen sie im Internet begegnen, besser einzuordnen und sie verantwortungsbewusst weiterzugeben. Eine ganze Reihe von Institutionen der Vereinten Nationen haben in diesem Bereich wertvolle Erfahrungen gesammelt. Die „Verified“-Initiative der Vereinten Nationen⁵² hat erfolgreich eine große Bandbreite von Taktiken eingesetzt, unter anderem gezielte Nachrichtendienste für Nutzerinnen und Nutzer, das so genannte Prebunking – an Nutzerinnen und Nutzer gerichtete Warnungen vor Falschmeldungen, noch bevor diese ihre Adressatinnen und Adressaten erreichen – und Aktionen zur Förderung der digitalen Kompetenz.

ABBILDUNG III

WIRKSAME KAMPAGNEN DER VEREINTEN NATIONEN ZUR BEKÄMPFUNG VON FEHL- UND DESINFORMATION

Wahrscheinlichkeit der Weitergabe von Falschmeldungen (2021)^a



^a Basierend auf Forschungsarbeiten des Massachusetts Institute of Technology (März 2021).

NEGATIVE ANREIZE SETZEN

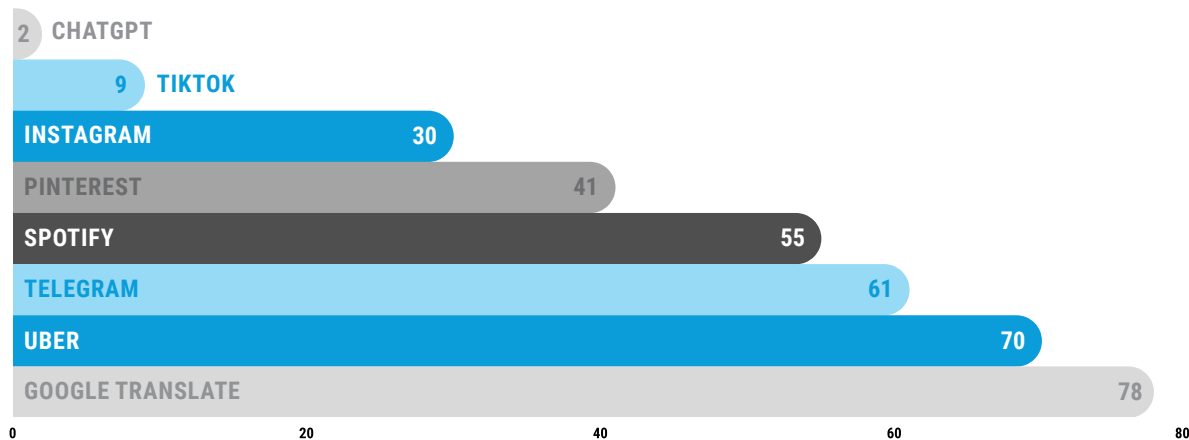
Die aktuellen Geschäftsmodelle der meisten digitalen Plattformen geben der Steigerung der Nutzerzahlen den Vorrang vor Menschenrechten, Datenschutz und Sicherheit. Dies umfasst auch die Verwertung personenbezogener Daten zur Gewinnerzielung trotz zunehmender Belege für den gesellschaftlichen Schaden, der durch dieses Geschäftsmodell entsteht.

Einige zivilgesellschaftliche Gruppen sowie Wissenschaftlerinnen und Wissenschaftler haben verschiedene Optionen untersucht, wie der Erstellung und Verbreitung von Fehlinformationen, Desinformation und Hetze im Internet die Finanzierung und damit Anreize entzogen werden können, und in diesem Zusammenhang darauf verwiesen, dass die freie Meinungsäußerung ein grundlegendes Menschenrecht ist, nicht jedoch die mit ihr verbundene Gewinnerzielung.⁵³ Die Vorschläge zielen darauf ab, die Rentabilität der Desinformation zu zerschlagen, vollständige Transparenz bei der Monetarisierung von Inhalten sowie eine unabhängige Risikobewertung sicherzustellen und Online-Werbetreibende davon abzuhalten, Desinformation zu begünstigen.

Marken, die im Umfeld von Fehl- und Desinformation und Hetze Werbeanzeigen schalten, riskieren damit, die Wirkung ihrer Anzeigenkampagnen zu untergraben und damit letztlich ihr Ansehen zu beschädigen. Werbetreibende können klare Richtlinien aufstellen, um die unbeabsichtigte Finanzierung und Legitimierung von Fehlinformationen, Desinformation und Hetze zu vermeiden und daran mitzuwirken, diesen die Finanzierungsgrundlage zu entziehen. Zu den praktischen Maßnahmen zählen die Nutzung aktualisierter Einschluss- und Ausschlusslisten und der Einsatz von Instrumenten zur Überprüfung von Anzeigen. Werbetreibende können zudem Druck auf digitale Plattformen ausüben, damit diese verstärkt Maßnahmen zum Schutz der Informationsintegrität ergreifen.

ABBILDUNG IV

**ZEITSPANNE (IN MONATEN), IN DER CHATGPT UND ANDERE BELIEBTE APPS
JE 100 MILLIONEN AKTIVE NUTZERINNEN UND NUTZER DAZUGEWANNEN**



Quelle: Similarweb unter Verwendung der Daten von Sensor Tower.

Sie können auch davon Abstand nehmen, Anzeigen bei Medienunternehmen zu schalten, die zu Hass anstacheln und Desinformation verbreiten.⁵⁴

UNABHÄNGIGE MEDIEN

Nach wie vor untergraben neu eingeführte Maßnahmen in Dutzenden von Ländern die Pressefreiheit. Laut dem Global Report 2022 der UNESCO aus der richtungsweisenden Reihe World Trends in Freedom of Expression and Media Development (Weltweite Trends zum Recht der freien Meinungsäußerung und zur Medienentwicklung) erlebten 85 Prozent der Weltbevölkerung in dem fünfjährigen Berichtszeitraum einen Rückgang der Pressefreiheit in ihrem jeweiligen Land.⁵⁵ Da 2,9 Milliarden Menschen immer noch keinen Internetanschluss haben⁵⁶, besteht eine weitere Priorität darin, unabhängige Medien zu stärken, die Verbreitung von Initiativen zur Faktenüberprüfung zu fördern und verlässliche und genaue Berichterstattung im öffentlichen Interesse zu unterstützen. Eine echte öffentliche Debatte beruht auf Fakten, die klar vermittelt werden und aus einer unabhängigen

Berichterstattung auf der Grundlage ethischer Grundsätze stammen. Ethisch handelnde Reporterinnen und Reporter mit hochwertiger Ausbildung und guten Arbeitsbedingungen verfügen über die beruflichen Kompetenzen, um ein Gegengewicht zu Fehl- und Desinformation zu schaffen. Sie können eine wesentliche Leistung erbringen: genaue, objektive und verlässliche Informationen über wichtige Themen.

FÜR DIE ZUKUNFT GERÜSTET

Wenn wir nach Lösungen suchen, um unsere Informationsintegrität im aktuellen Umfeld zu schützen, müssen wir zugleich sicherstellen, dass die entsprechenden Empfehlungen zukunftstauglich sind und neu aufkommende genauso wie zukünftige Technologien berücksichtigen. Die im November 2022 in Betrieb genommene Plattform ChatGPT-3 von Open AI verzeichnete im Januar 2023 bereits 100 Millionen Nutzerinnen und Nutzer und wurde damit zur am schnellsten wachsenden Verbraucher-App aller Zeiten⁵⁷. Etliche weitere Unternehmen arbeiten mit Hochdruck an der

Entwicklung von Konkurrenzprodukten. Trotz des schier unvorstellbaren Potenzials künstlicher Intelligenz für die Bewältigung globaler Herausforderungen bestehen ernsthafte und dringende Bedenken angesichts des ebenso großen Potenzials, dass KI die Informationsintegrität durch ihre jüngsten Entwicklungen, beispielsweise Bildgenerator und Video-Deepfakes, bedroht. Jüngste Berichte und Forschungsergebnisse haben gezeigt, dass generative künstliche Intelligenz auch Fehlinformationen, Desinformation und Hetze erzeugt, die den Nutzerinnen und Nutzern gegenüber in überzeugender Weise als Tatsachen ausgegeben werden.⁵⁸

Unter der Federführung des Gesandten des Generalsekretärs für Technologie wird derzeit der Versuch unternommen, die Auswirkungen der generativen künstlichen Intelligenz und anderer neu aufkommender Plattformen zu bewerten. Dabei müssen wir aus den Fehlern der Vergangenheit lernen. Digitale Plattformen wurden in die Welt gesetzt, ohne dass ein hinreichendes Bewusstsein dafür vorhanden war oder ausreichend überprüft wurde, welche potenziellen Schäden der Gesellschaft und jedem einzelnen Menschen drohen. Wir haben jetzt die Chance, dafür zu sorgen, dass sich die Geschichte bei den neuesten technologischen Entwicklungen nicht wiederholt. Die Zeit des im Silicon Valley gängigen Wahlspruchs „Move fast and break things“ (Schnell sein, ohne Rücksicht auf Verluste) muss jetzt ein Ende finden. Es kommt

ganz entscheidend darauf an, den Schutz der Nutzerdaten, ihre Sicherheit und Transparenz sowie das Konzept „Sicherheit durch Design“ von Beginn an in alle neuen Technologien und Produkte zu integrieren.

MASSNAHMEN DER VEREINTEN NATIONEN

Die Friedensmissionen der Vereinten Nationen und die Landesbüros haben bereits Schritte gesetzt, um die Bedrohungen, die von Fehl- und Desinformation für die Durchführung der Mandate der Vereinten Nationen ausgehen, zu überwachen und zu analysieren und ihnen entgegenzuwirken. Die Strategie und der Aktionsplan der Vereinten Nationen gegen Hetze enthalten strategische Richtlinien dafür, wie die Organisation Hetze auf nationaler und globaler Ebene bekämpfen kann. Im Februar 2023 veranstaltete die UNESCO die Konferenz über Internet und Vertrauen, auf der ein Entwurf einer Reihe von Richtlinien für die Regulierung digitaler Plattformen erörtert wurde, die gegen Ende des Jahres fertiggestellt werden sollen.⁵⁹

Zusammen tragen diese Initiativen und Ansätze dazu bei, den Weg für grundlegende Prinzipien eines Verhaltenskodex der Vereinten Nationen zu ebnen.

MONATLICHE NUTZUNG DES HASHTAGS

#CLIMATESCAM

AUF TWITTER

100K

NOVEMBER 2022

COP27 SPIKE

200

FEBRUAR 2022

HÖCHSTSTAND DER
GEPOSTETEN BEITRÄGE

199.300

(**JANUAR 2023**)

Erarbeitung eines Verhaltenskodex der Vereinten Nationen

Mein Vorschlag für einen Verhaltenskodex der Vereinten Nationen für die Informationsintegrität auf digitalen Plattformen würde auf den folgenden Grundsätzen beruhen:

- Verpflichtung zur Informationsintegrität
- Achtung der Menschenrechte
- Förderung unabhängiger Medien
- erhöhte Transparenz
- Stärkung der Selbstbestimmung der Nutzerinnen und Nutzer
- verstärkte Forschung und erweiterter Datenzugang
- verschärfte Gegenmaßnahmen
- stärkere negative Anreize
- gefestigtes Vertrauen und erhöhte Sicherheit

Diese Grundprinzipien wurden aus den Kerngedanken in diesem Kurzdossier herausgefiltert und korrespondieren inhaltlich mit meinem Kurzdossier zum Globalen Digitalpakt. Den Mitgliedstaaten wird nahegelegt, den Verhaltenskodex auf nationaler Ebene umzusetzen. Die Beratungen mit den Interessenträgern werden fortgesetzt, um den Inhalt des Verhaltenskodex weiter auszufeilen und konkrete Methoden zu bestimmen, um seine Prinzipien in die Tat umzusetzen.

Folgende Empfehlungen können als Anregung für den Verhaltenskodex dienen:

Verpflichtung zur Informationsintegrität

- a) Alle Interessenträger sollten es unterlassen, Desinformation und Hetze für wie auch immer geartete Zwecke, unter anderem zur Verfolgung politischer, militärischer und anderer strategischer Ziele, zur Aufstachelung zu Gewalt, zur Untergrabung demokratischer Prozesse oder für gezielte Angriffe auf die Zivilbevölkerung, schutzbedürftige Gruppen, Gemeinschaften oder Einzelpersonen, zu verwenden, zu unterstützen oder zu verstärken;

Achtung der Menschenrechte

- b) Die Mitgliedstaaten sollten:
 - i) sicherstellen, dass Maßnahmen gegen Fehl- und Desinformation und Hetze mit dem Völkerrecht, unter anderem mit den internationalen Menschenrechtsnormen, vereinbar sind und nicht dazu missbraucht werden, legitime Meinungsäußerungen zu unterdrücken, beispielsweise durch flächendeckende Abschaltungen des Internets oder das Verbot von Plattformen und Medienunternehmen;
 - ii) Regulierungsmaßnahmen ergreifen, darunter auch Durchsetzungsmechanismen, um

die grundlegenden Rechte der Nutzerinnen und Nutzer digitaler Plattformen zu schützen, zugleich aber volle Transparenz bezüglich der für Technologieunternehmen geltenden Anforderungen zu wahren;

- c) Alle Interessenträger sollten die Leitprinzipien für Wirtschaft und Menschenrechte befolgen;

Förderung unabhängiger Medien

- d) Die Mitgliedstaaten sollten eine freie, zukunftsfähige, unabhängige und pluralistische Medienlandschaft mit starkem Schutz für journalistisch tätige Personen und unabhängige Medien gewährleisten und die Einrichtung, Finanzierung und Ausbildung unabhängiger Organisationen zur Faktenüberprüfung in Lokalsprachen unterstützen;
- e) Nachrichtenmedien sollten eine korrekte, ethische und unabhängige Berichterstattung sicherstellen, die durch eine hochwertige Ausbildung und angemessene Arbeitsbedingungen im Einklang mit den internationalen Normen und Standards auf dem Gebiet der Arbeitnehmer- und der Menschenrechte unterstützt wird;

Erhöhte Transparenz

- f) Digitale Plattformen sollten
 - i) wirkliche Transparenz im Hinblick auf Algorithmen, Daten, Moderation von Inhalten und Werbung sicherstellen;
 - ii) zugängliche Richtlinien zu Fehlinformationen, Desinformation und Hetze veröffentlichen und bekannt machen und darüber Bericht erstatten, wie stark über ihre Dienstleistungen koordinierte Desinformation verbreitet wird und wie wirksam ihre Gegenmaßnahmen sind;

- g) Nachrichtenmedien sollten im Hinblick auf Finanzierungsquellen und Werberichtlinien wirkliche Transparenz sicherstellen und klar zwischen redaktionellen Inhalten und bezahlter Werbung unterscheiden, unter anderem bei Veröffentlichung der Inhalte auf digitalen Plattformen;

Stärkung der Selbstbestimmung der Nutzerinnen und Nutzer

- h) Die Mitgliedstaaten sollten den Zugang der Öffentlichkeit zu korrekten und transparenten staatlichen Informationen aus glaubwürdigen Quellen sicherstellen, insbesondere Informationen, die dem öffentlichen Interesse dienen, einschließlich sämtlicher Aspekte der Ziele für nachhaltige Entwicklung;
- i) Digitale Plattformen sollten auf transparente Weise die Selbstbestimmung und den Schutz der Nutzerinnen und Nutzer gewährleisten und den Menschen zugleich mehr Wahlmöglichkeiten bezüglich der angezeigten Inhalte und der Verwendung ihrer Daten einräumen. Die Nutzerinnen und Nutzer sollten ihre Identität oder Authentizität unter Beweis stellen können, ohne dafür bezahlen oder Einschränkungen beim Datenschutz in Kauf nehmen zu müssen. Die Plattformen sollten transparente Beschwerde- und Meldeverfahren einführen, die durch unabhängige, hinlänglich publik gemachte und zugängliche Beschwerdeüberprüfungsmechanismen unterstützt werden;
- j) Alle Interessenträger sollten in robuste Initiativen zur Förderung der digitalen Kompetenz investieren, um Nutzerinnen und Nutzern aller Altersgruppen ein besseres Verständnis davon zu vermitteln, wie digitale Plattformen funktionieren und wie ihre personenbezogenen Daten verwendet werden könnten, sowie sie in die Lage zu versetzen,

Fehlinformationen, Desinformation und Hetze zu erkennen und darauf zu reagieren. Besonderes Augenmerk sollte darauf gelegt werden, dass junge Menschen, Jugendliche und Kinder sich ihrer Rechte im Online-Umfeld vollkommen bewusst sind;

Verstärkte Forschung und erweiterter Datenzugang

- k)** Die Mitgliedstaaten sollten in unabhängige Forschung zur Verbreitung und Wirkung von Fehlinformationen, Desinformation und Hetze in unterschiedlichen Ländern und Sprachen investieren, insbesondere in unterversorgten Bereichen und in allen Sprachen mit Ausnahme des Englischen, und eine solche Forschung unterstützen, um der Zivilgesellschaft und der Wissenschaft eine freie und sichere Tätigkeit zu ermöglichen;
- l)** Digitale Plattformen sollten
 - i)** der Forschung und der Wissenschaft Zugang zu Daten gewähren, dabei jedoch den Schutz der Daten von Nutzerinnen und Nutzern achten. Die in Wissenschaft und Forschung Tätigen sollten die Möglichkeit haben, Stichproben zu nehmen und qualitative Daten zu Einzelpersonen und Gruppen zu erheben, die Zielscheibe von Fehlinformationen, Desinformation und Hetze waren, um das Ausmaß und die Art der Schäden besser zu verstehen, unter Achtung des Datenschutzes und der Menschenrechte;
 - ii)** die volle Mitwirkung der Zivilgesellschaft an den Bemühungen um die Bekämpfung von Fehlinformationen, Desinformation und Hetze sicherstellen;

Ausweitung der Gegenmaßnahmen

- m)** Alle Interessenträger sollten
 - i)** Ressourcen bereitstellen, um die Ursprünge, Verbreitung und Wirkung von Fehlinformationen, Desinformation und Hetze zu bekämpfen und darüber Bericht zu erstatten, sowie die Menschenrechtsnormen und -standards wahren und die Kapazitäten zur Überprüfung von Fakten in allen Ländern und Bereichen weiter ausbauen;
 - ii)** breite Bündnisse für Informationsintegrität eingehen, die unterschiedliche Fachkenntnisse und Ansätze zusammenführen, um die Kluft zwischen den lokalen Organisationen und auf globaler Ebene operierenden Technologieunternehmen zu überbrücken;
 - iii)** Ausbildung und Kapazitätsaufbau fördern, um ein Verständnis dafür zu entwickeln, wie sich Fehlinformationen, Desinformation und Hetze manifestieren, und die Strategien zur Prävention und Risikominderung zu stärken;

Stärkere negative Anreize

- n)** Digitale Plattformen sollten von Geschäftsmodellen Abstand nehmen, die der Steigerung der Nutzerzahlen den Vorrang vor Menschenrechten, Datenschutz und Sicherheit geben;
- o)** Werbetreibende und digitale Plattformen sollten sicherstellen, dass keine Anzeigen neben Fehlinformationen, Desinformation oder Hetze im Internet platziert werden und dass keine Werbung, die Desinformation enthält, gefördert wird;

- p) Nachrichtenmedien sollten sicherstellen, dass alle bezahlten Werbe- und Advertorial-Inhalte klar als solche gekennzeichnet sind und keine Fehlinformationen, Desinformation und Hetze enthalten;

Gestärktes Vertrauen und erhöhte Sicherheit

- q) Digitale Plattformen sollten

- i) das Konzept „Sicherheit und Datenschutz durch Design“ in allen Produkten umsetzen, unter anderem durch die Bereitstellung ausreichender Ressourcen für firmeninterne Sachkompetenz in Vertrauens- und Sicherheitsfragen sowie durch eine konsequente Anwendung der Richtlinien in allen Ländern und Sprachen;

- ii) in menschliche und KI-gestützte Systeme zur Inhaltsmoderation investieren, die in allen Sprachen verfügbar sind, die in ihren Tätigkeitsländern gesprochen werden, und sicherstellen, dass Mechanismen zur Meldung von Inhalten transparent sind und über eine beschleunigte Rücklaufquote verfügen, insbesondere in Konfliktsituationen;

- r) Alle Interessenträger sollten dringende und unverzügliche Maßnahmen ergreifen, um die sichere, verantwortliche, ethische und mit den Menschenrechten vereinbare Nutzung künstlicher Intelligenz zu gewährleisten und den Auswirkungen der jüngsten Entwicklungen in diesem Bereich auf die Verbreitung von Fehlinformationen, Desinformation und Hetze entgegenzuwirken.

Die nächsten Schritte

- Das Sekretariat der Vereinten Nationen wird umfassende Konsultationen mit einer Reihe von Interessenträgern abhalten, um den Verhaltenskodex der Vereinten Nationen, einschließlich seiner Weiterverfolgungs- und Umsetzungsmechanismen, auszuarbeiten. Im Zuge dessen könnte eine unabhängige, mit anerkannten Fachleuten besetzte Beobachtungsstelle eingerichtet werden, deren Aufgabe es wäre, die Maßnahmen der an den Verhaltenskodex gebundenen Akteure zu bewerten. Denkbar sind auch weitere Berichtsmechanismen.
- Zur Unterstützung und inhaltlichen Anreicherung des Kodex könnte das Sekretariat der Vereinten Nationen eingehende Studien durchführen, um das Verständnis für Informationsintegrität weltweit zu erhöhen, insbesondere in den bislang noch unzureichend von Forschungsarbeiten erfassten Teilen der Welt.
- Der Generalsekretär wird im Sekretariat der Vereinten Nationen eine spezielle Kapazität einrichten, um die Reaktion auf solche Fehlinformationen, Desinformation und Hetze im Internet zu verstärken, die die Mandatserfüllung der Vereinten Nationen und ihre inhaltlichen Prioritäten beeinträchtigen. Auf der Grundlage sachverständiger Überwachung und Analyse würde diese Kapazität gezielte Kommunikationsstrategien entwickeln, um Bedrohungen vorauszusehen und/oder rasch zu bekämpfen, bevor sie so sehr eskalieren, dass sie online und offline Schäden verursachen. Zudem würde sie den Kapazitätsaufbau der Bediensteten und der Mitgliedstaaten der Vereinten Nationen fördern und die Mitgliedstaaten, digitalen Plattformen und anderen Interessenträger in ihren Bemühungen unterstützen, den Kodex nach dessen Fertigstellung einzuhalten und anzuwenden.

Schlussfolgerung

Die Stärkung der Informationsintegrität auf digitalen Plattformen ist von höchster Priorität für die internationale Gemeinschaft. Durch Maßnahmen zur Eindämmung der Auswirkungen von Fehlinformationen, Desinformation und Hetze werden wir den Bemühungen, eine nachhaltige Zukunft zu schaffen und niemanden zurückzulassen, Auftrieb verleihen. Dies gilt gleichermaßen für die Bereiche Gesundheit, Geschlechtergleichstellung, Frieden, Gerechtigkeit, Bildung und Klimaschutz. Nationale Maßnahmen sind hier nicht ausreichend. Letztlich lassen sich diese Probleme nur durch eine stärkere globale Zusammenarbeit zur Gänze lösen. Die in diesem Kurzdossier ausgeführten Kerngedanken zeigen, dass der Weg zu einer verstärkten Informationsintegrität an den Menschenrechten orientiert,

mehrdimensional gestaltet und unter Einbindung einer Vielzahl von Interessenträgern besprochen werden muss. Ausgehend von diesen zentralen Ideen wurde eine Reihe grundlegender Prinzipien formuliert, die bei der Erarbeitung eines Verhaltenskodex der Vereinten Nationen für die Informationsintegrität auf digitalen Plattformen in Betracht gezogen werden sollten. Mit diesem Kodex würde ein Modell zur Stärkung der Informationsintegrität unter strikter Achtung der Menschenrechte geschaffen. Ich freue mich darauf, mit den Mitgliedstaaten und anderen Akteuren darauf hinzuwirken, diese Grundsätze in greifbare Verpflichtungen umzuwandeln.

Anhang

KONSULTATIONEN MIT DEN MITGLIEDSTAATEN UND ANDEREN MASSGEBLICHEN INTERESSENTRÄGERN

Die Anregungen in diesem Kurzdossier beruhen auf den Vorschlägen, wie sie in Unserer gemeinsamen Agenda (A/75/982) skizziert wurden, die durch umfangreiche Konsultationen mit den Mitgliedstaaten, dem System der Vereinten Nationen, einflussreichen Denkerinnen und Denkern, jungen Menschen und Akteuren der Zivilgesellschaft aus aller Welt inhaltlich angereichert wurde. Das Kurzdossier geht insbesondere auf die reichhaltigen und ausführlichen Überlegungen der Mitgliedstaaten und anderer Interessenträger zu Unserer

gemeinsamen Agenda ein, die im Laufe von 25 Beratungen der Generalversammlung zusammengetragen wurden.

Im Vorfeld der Veröffentlichung dieses Kurzdossiers fanden Beratungen mit den Mitgliedstaaten statt, unter anderem in Form einer informellen Unterrichtung des Informationsausschusses, zu dem auch alle nicht im Ausschuss vertretenen Mitgliedstaaten eingeladen waren. Gespräche wurden zudem mit zivilgesellschaftlichen Partnern, der Wissenschaft, Sachverständigen und dem Privatsektor geführt, darunter auch mit Technologieunternehmen.

Vor dem Zukunftsgipfel werden umfassende Konsultationen zur Ausarbeitung des Verhaltenskodex stattfinden.

Endnoten

- 1 [A/HRC/42/50; A/77/287; A/HRC/51/53](#); United Nations, „[Statement by Alice Wairimu Nderitu, Special Adviser on the Prevention of Genocide, condemning the recent escalation of fighting in Ethiopia](#)“, press release, 19. Oktober 2022; Office of the United Nations High Commissioner for Human Rights (OHCHR), „[Myanmar: Social media companies must stand up to junta’s online terror campaign say UN experts](#)“, Pressemitteilung, 13. März 2023; OHCHR, „[Freedom of speech is not freedom to spread racial hatred on social media: UN experts](#)“, Erklärung, 6. Januar 2023; Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, „[#JournalistsToo: women journalists speak out](#)“, 24. November 2021 und OHCHR, „[Sri Lanka: Experts dismayed by regressive steps, call for renewed UN scrutiny and efforts to ensure accountability](#)“, Pressemitteilung, 5. Februar 2021.
- 2 [A/HRC/47/25](#), Ziff. 15.
- 3 Kalina Bontcheva und Julie Posetti, Hrsg., *Balancing Act: Countering Digital Disinformation While Respecting Freedom of Expression – Broadband Commission Research Report on „Freedom of Expression and Addressing Disinformation on the Internet“* (Genf, International Telecommunication Union (ITU); Paris, UNESCO, 2020).
- 4 Siehe United Nations, „[Countering disinformation](#)“, auf Englisch abrufbar unter www.un.org/en/countering-disinformation und [A/77/287](#).
- 5 Auf Englisch abrufbar unter www.un.org/en/genocideprevention/documents/advising-and-mobilizing/Action_plan_on_hate_speech_EN.pdf.
- 6 Die Vereinten Nationen prüfen derzeit in einer Untersuchung, welche Verknüpfungen und Beziehung zwischen Fehl- und Desinformation und Hetze bestehen und wo diese verwandten und doch unterschiedlichen Phänomene auf konzeptioneller und operativer Ebene aufeinanderzutreffen beziehungsweise auseinanderklaffen.
- 7 Die Europäische Kommission definiert Online-Plattformen in „[Shaping Europe’s digital future: online platforms](#)“, 7. Juni 2022. Auf Englisch abrufbar unter <https://digital-strategy.ec.europa.eu/en/policies/online-platforms>.
- 8 Siehe OHCHR, „[Moderating online content: fighting harm or silencing dissent?](#)“, 23. Juli 2021.
- 9 Siehe United Nations, „[Countering disinformation](#)“ und [A/77/287](#).
- 10 [A/HRC/47/25](#).
- 11 Stephanie Kirchgaessner et al., „[Revealed: the hacking and disinformation team meddling in elections](#)“, *The Guardian*, 14. Februar 2023.
- 12 Alexandre Alaphilippe et al., „[Doppelgänger – media clones serving Russian propaganda](#)“, EU DisinfoLab, 27. September 2022.
- 13 United Nations Economist Network, „[New economics for sustainable development: attention economy](#)“.
- 14 Twitter, „[About Twitter Blue](#)“ und Meta, „[Testing Meta Verified to help creators establish their presence](#)“, 17. März 2023.
- 15 [A/77/287](#).
- 16 Mit Stand vom Februar 2023 waren 173 Mitgliedstaaten Vertragsstaaten des Internationalen Paktes über bürgerliche und politische Rechte.
- 17 Einschränkungen des Rechts auf freie Meinungsäußerung müssen die folgenden bewährten Bedingungen erfüllen: Legalität: Einschränkungen müssen gesetzlich festgelegt werden und so formuliert sein, dass ausreichend klar zwischen einer rechtmäßigen und einer unrechtmäßigen Äußerung unterschieden wird; Notwendigkeit und Verhältnismäßigkeit: Die Einschränkung stellt nachweislich die geringstmögliche Beeinträchtigung der Ausübung des Rechtes dar und schützt die maßgeblichen rechtmäßigen Interessen des Staates tatsächlich oder wahrscheinlich; Rechtmäßigkeit: Um als rechtmäßig anerkannt zu werden, dürfen Einschränkungen nur die in Artikel 19 Absatz 3 des Internationalen Paktes über bürgerliche und politische Rechte genannten Interessen schützen.
- 18 Dazu zählen unter anderem die audiovisuelle Kommunikation in Côte d’Ivoire, Marokko und Tunesien und die Überwachung der Aufstachelung zu Gewalt durch die Mehrdimensionale integrierte Stabilisierungsmission der Vereinten Nationen in der Zentralafrikanischen Republik.
- 19 Aleksis Knuutila, Lisa-Maria Neudert und Philip N. Howard, „[Who is afraid of fake news? Modeling risk perceptions of misinformation in 142 countries](#)“, *Harvard Kennedy School (HKS) Misinformation Review*, Vol. 3, Nr. 3 (April 2022).
- 20 ITU, *Measuring the Information Society* (Genf, 2013).
- 21 United Nations Children’s Fund (UNICEF), „[Protecting children online](#)“, 23. Juni 2022. Auf Englisch abrufbar unter www.unicef.org/protection/violence-against-children-online.
- 22 UNESCO, Arbeitspapiere zu „[Digital governance and the challenges for trust and safety](#)“. Auf Englisch abrufbar unter www.unesco.org/en/internet-conference/working-papers.
- 23 Siehe Julie Posetti und Kalina Bontcheva, „[Disinfodemic: deciphering COVID-19 disinformation](#)“, Kurzdossier 1, (Paris, UNESCO, 2020), und „[Disinfodemic: dissecting responses to COVID-19 disinformation](#)“, Kurzbrief 2, (Paris, UNESCO, 2020).
- 24 Center for Countering Digital Hate, *Pandemic Profiteers: The Business of Anti-Vaxx* (2021).

- 25 Michael A Gisondi et al., „[A deadly infodemic: social media and the power of COVID-19 misinformation](#)“, *Journal of Medical Internet Research*, Vol. 24, Nr. 2 (Februar 2022).
- 26 [A/77/288](#).
- 27 [A/77/287](#), Ziff. 6.
- 28 2018 erklärte eine vom Menschenrechtsrat beauftragte unabhängige internationale Ermittlungsmission Facebook zur „Plattform, über die in Myanmar am meisten Hetze verbreitet wird“ ([A/HRC/42/50](#), Ziff. 72).
- 29 Siehe United Nations News, „[Hate speech: a growing, international threat](#)“, 28. Januar 2023, und „[Digital technology, social media fueling hate speech like never before, warns UN expert](#)“, 20. Oktober 2022
- 30 Siehe John Cook, „Deconstructing climate science denial“, in *Research Handbook in Communicating Climate Change*, David C. Holmes und Lucy M. Richardson, Hrsg. (Cheltenham, United Kingdom, Edward Elgar, 2020). Laut Cook enthält Abraham et al. (2014) eine Zusammenfassung klimawandelskeptischer wissenschaftlicher Arbeiten, etwa Artikel über eine vermeintliche Abkühlung laut Satellitenmessungen oder Schätzungen einer geringen Klimasensitivität, die in der wissenschaftlichen Literatur klar widerlegt wurden. Auch Benestad et al. (2016) unternahmen den Versuch, in nonkonformistischen Publikationen veröffentlichte Ergebnisse zu reproduzieren, und stellten dabei unzählige Fehler fest, etwa unsachgemäße statistische Methoden, falsche Dichotomien und Schlussfolgerungen, die auf einem falschen Verständnis physikalischer Prinzipien basieren.
- 31 Global Witness, „[The climate divide: how Facebook’s algorithm amplifies climate disinformation](#)“, 28. März 2022.
- 32 Analyse der Hauptabteilung Globale Kommunikation der Vereinten Nationen anhand von Daten von Talkwalker.
- 33 Jeffrey A. Hicke et al., „North America“, in Intergovernmental Panel on Climate Change, *Climate Change 2022: Impacts, Adaptation and Vulnerability*, Beitrag der Arbeitsgruppe II des Sechsten IPCC-Sachstandsberichts (Cambridge, United Kingdom, Cambridge University Press, 2022).
- 34 Mei Li, Gregory Trencher und Jusen Asuka, „[The clean energy claims of BP, Chevron, ExxonMobil and Shell: a mismatch between discourse, actions and investments](#)“, PLOS ONE, Ausgabe 17, Nr. 2 (Februar 2022).
- 35 Robert J. Brulle und Carter Werthman, „[The role of public relations firms in climate change politics](#)“, *Climatic Change*, Vol. 169, Nr. 1–2 (November 2021). Laut des Global Disinformation Index, einer gemeinnützigen Aufsichtsbehörde, haben Werbetreibende in der Technologiebranche 2021 36,7 Millionen USD an 98 Websites gezahlt, auf denen Desinformation zum Klimawandel auf Englisch veröffentlicht wurde. Im November 2022 deckte die Aktionsgruppe Center for Countering Digital Hate in einem Bericht auf, dass allein auf Google nahezu die Hälfte der 23,7 Millionen USD, die Öl- und Gasunternehmen in den vorangegangenen zwei Jahren für suchgebundene Werbung ausgegeben hatten, auf Suchbegriffe zum Thema ökologische Nachhaltigkeit abzielte. Bei Untersuchungen fand InfluenceMap im Jahr 2020 25.147 irreführende Anzeigen von 25 Organisationen aus dem Öl- und Gassektor auf Facebook Plattformen in den USA, wobei sich die Gesamtausgaben auf 9.597.376 USD beliefen. Angesichts des Ausmaßes des Problems waren die Reaktionen bisher unzureichend.
- 36 Siehe Resolution [76/227](#) der Generalversammlung, siehe Resolution [49/21](#) des Menschenrechtsrats und European Union External Action, „[Tackling disinformation, foreign information manipulation and interference](#)“, 27. Oktober 2021.
- 37 Lucina Di Meco, „[Monetizing misogyny: gendered disinformation and the undermining of women’s rights and democracy globally](#)“, #ShePersisted, Februar 2023.
- 38 Siehe Andrew Puddephatt, „[Social media and elections](#)“, Cuadernos de Discusión de Comunicación e Información, Nr. 14 (Montevideo, UNESCO, 2019); und Julie Posetti et al., „[The chilling: global trends in online violence against women journalists](#)“, Thesenpapier (UNESCO, 2021).
- 39 EU Disinfo Lab, „[The role of “media” in producing and spreading disinformation campaigns](#)“, 13. Oktober 2021.
- 40 Siehe United Nations News, „[Social media poses ‘existential threat’ to traditional, trustworthy news: UNESCO](#)“, 10. März 2022; und Anya Schiffrin et al., „[Finding the funds for journalism to thrive: policy options to support media viability](#)“, World Trends in Freedom of Expression and Media Development (Paris, UNESCO, 2022).
- 41 European Commission, „[Shaping Europe’s digital future: the 2022 Code of Practice on Disinformation](#)“, 4. Juli 2022.
- 42 Siehe die Bemerkungen der Vizepräsidentin der Europäischen Kommission und Kommissarin für Werte und Transparenz, Věra Jourová, in: European Commission, „[Code of Practice on Disinformation: new Transparency Centre provides insights and data on online disinformation for the first time](#)“, Presseerklärung, 9. Februar 2023. Auf Englisch abrufbar unter: https://ec.europa.eu/commission/presscorner/detail/en/mex_23_723.
- 43 Auf Deutsch abrufbar unter: <https://www.globalcompact.de/mediathek/publikationen/publikation/un-leitprinzipien-fuer-wirtschaft-und-menschenrechte-1>.
- 44 Center for Countering Digital Hate and Human Rights Campaign, „[Digital hate: social media’s role in amplifying dangerous lies about LGBTQ+ people](#)“, 10. August 2022.
- 45 Siehe Andrew Puddephatt, „[Letting the sun shine in: transparency and accountability in the digital age](#)“, World Trends in Freedom of Expression and Media Development (Paris, UNESCO, 2021).
- 46 Whose Knowledge?, Oxford Internet Institute and The Centre for Internet and Society, *State of the Internet’s Languages Report (2022)*.
- 47 [A/HRC/38/35](#).

- 48 Billy Perrigo, „[Inside Facebook’s African sweatshop](#)“, *Time*, 17. Februar 2022.
- 49 [A/HRC/42/50](#).
- 50 Zu den herausragenden Beispielen zählen Maria Ressa, *How to Stand up to a Dictator* (New York, HarperCollins, 2022). Deutsche Ausgabe: *How to Stand up to a Dictator – Der Kampf um unsere Zukunft* (Köln, Quadriga, 2022); und Max Fisher, *The Chaos Machine* (New York, Little, Brown and Company, 2022).
- 51 Siehe UNICEF, „[Young reporters fact-checking COVID-19 information](#)“.
- 52 Siehe <https://shareverified.com/>.
- 53 Der Global Disinformation Index, ein nicht gewinnorientiertes Unternehmen, verfolgt Werbeanzeigen, die im Umfeld von Desinformation platziert werden. Dieser redaktionellen Praxis sind auch die Vereinten Nationen bereits zum Opfer gefallen: Der Global Disinformation Index entdeckte Werbeanzeigen von UNICEF neben Textbeiträgen, die Impfskepsis verbreiteten, und Anzeigen des Hohen Flüchtlingskommissariats der Vereinten Nationen in unmittelbarer Nähe von flüchtlingsfeindlichen Inhalten.
- 54 Conscious Advertising Network, manifestos. Auf Englisch abrufbar unter: www.consciousadnetwork.com/the-manifestos/.
- 55 UNESCO, *Journalism is a Public Good: World Trends in Freedom of Expression and Media Development – Global Report 2021/2022* (Paris, 2022).
- 56 ITU, „[Facts and figures 2021: 2.9 billion people still offline](#)“, 29. November 2021. Der [Globale Digitalpakt](#), der von den Mitgliedstaaten beim Zukunftsgipfel im Jahr 2024 beraten werden soll, wird gemeinsame Grundsätze für eine offene, freie und sichere digitale Zukunft für alle enthalten (siehe www.un.org/techenvoy/global-digital-compact).
- 57 Krystal Hu, „[ChatGPT sets record for fastest-growing user base – analyst note](#)“, Reuters, 2. Februar 2023.
- 58 Siehe Center for Countering Digital Hate, „[Misinformation on Bard, Google’s new AI chat](#)“, 5. April 2023; und Tiffany Hsu and Stuart A. Thompson, „[Disinformation researchers raise alarms about A.I. chatbots](#)“, *The New York Times*, 13. Februar 2023.
- 59 Der Entwurf der Richtlinien ist auf Englisch abrufbar unter www.unesco.org/en/internet-conference.

