SWP-Aktuell
NR. 69 DEZEMBER 2024

Wie man erfolgreich Desinformation bekämpft

Reaktive Ansätze – Potentiale und Grenzen

Aldo Kleemann

Regierungen, Medien, Wissenschaft und Gesellschaft sind sich einig, dass zur Bekämpfung von Desinformation vielfältige und koordinierte Anstrengungen notwendig sind. Uneinigkeit herrscht oft darüber, wie diese Erkenntnis in konkrete Maßnahmen umgesetzt werden soll. Besonders greifbare, unmittelbare und sichtbare Ansätze erhalten regelmäßig überproportional viel Aufmerksamkeit. Der Vorteil von Faktenchecks, von Kennzeichnung und Deplatforming ist, dass sie ein zählbares Ergebnis liefern. So lässt sich öffentlichkeitswirksam kommunizieren, dass bereits mehr als 18.000 Desinformationen widerlegt wurden oder man ein Netzwerk mit 50.000 inauthentischen Accounts abgeschaltet hat. Doch solche Zahlen sagen noch nichts über die Wirksamkeit derartiger Interventionen aus. Der empirische Forschungsstand zum Effekt reaktiver Instrumente auf die Verbreitung und Wirkung von Desinformation ist lückenhaft und mitunter widersprüchlich. Mit Blick auf Faktencheck, Kennzeichnung und Deplatforming stellen sich die Fragen: Was beinhaltet die Maßnahme, welcher Effekt ist aus wissenschaftlicher Sicht zu erwarten, welche nicht-intendierten Effekte können eintreten und wie skalierbar ist der jeweilige Ansatz?

Desinformation — also die gezielte Verbreitung von falschen oder irreführenden Informationen mit der Absicht, einzelne Menschen oder Gruppen zu manipulieren — hat das Potential, den demokratischen Diskurs zu beeinträchtigen, Radikalisierung zu befeuern und Hass und Gewalt auszulösen. Sie wird auf der ganzen Welt als Bedrohung für Demokratien angesehen. Viele politische Entscheidungsträger:innen suchen daher nach schnellen und wirksamen Wegen, Menschen davon abzuhalten, Desinformation anzunehmen und zu verbreiten.

Empirische Forschungsergebnisse zum Effekt von Desinformation auf Menschen und Gesellschaften — etwa, wie sie Überzeugungen oder Wahlverhalten beeinflusst oder Gewalt auslöst — sind jedoch rar und variieren stark. Dies erschwert die Bewertung der Bedrohung durch Desinformation und die Wahl geeigneter präventiver und reaktiver Gegenmaßnahmen. Nachfolgend liegt der Fokus auf drei der häufigsten reaktiven Ansätze: Faktenchecks, Kennzeichnung und Deplatforming. Dabei werden vier Hauptfragen beleuchtet:

Stiftung Wissenschaft und Politik
Deutsches Institut für Internationale Politik und Sicherheit

1. Was wird unter dem jeweiligen Ansatz verstanden?
2. Wie wird die Wirkung des Ansatzes von der aktuellen Forschung eingeschätzt?
3. Welche möglichen nicht-intendierten Effekte gibt es?
4. Lässt sich der Ansatz im großen Maßstab umsetzen?

Faktencheck

Was ist das? — Unter diesem Ansatz, auch als Fact-Checking oder Debunking bekannt, versteht man die Veröffentlichung korrigierender Informationen zur Widerlegung irreführender Behauptungen. Faktenchecks werden häufig im Rahmen journalistischer Arbeit durchgeführt oder von spezialisierten Institutionen betrieben. Diese können entweder dauerhaft oder temporär tätig sein, oft mit einem spezifischen thematischen Fokus, etwa zu Themen wie der Corona-Pandemie, kriegerischen Konflikten oder Wahlen. Staatlich organisierte Faktenchecks sind hingegen eher selten.

Wie wird die Wirkung eingeschätzt? — Zu der Frage, wie genau und beständig die korrigierte Information wirkt, liefert die Forschung komplexe und teils widersprüchliche Ergebnisse.

Einige Studien legen nahe, dass die Korrektur falscher Informationen aus vertrauenswürdigen Quellen zu einem Wissenszuwachs führen kann, der Verhalten und Einstellung der Empfänger:innen nachhaltig beeinflusst. Andere Untersuchungen zeigen auf, dass die Wirkung von Faktenchecks vor allem davon abhängt, wie stark die Desinformation mit der Überzeugung der Betroffenen übereinstimmt. Ein Grund dafür ist der sogenannte »confirmation bias«, also die Tendenz, Informationen selektiv und passend zu den eigenen Überzeugungen aufzunehmen.

Selbst wenn Desinformation unmittelbar und überzeugend widerlegt wurde und dieser Akt der Korrektur von den Betroffenen erinnert wird, bedingt das Wissen um die Unwahrheit der Information nicht zwingend auch eine Einstellungs- und Verhaltensänderung. Dieses Phänomen wird als »continued influence effect« und »belief echo« bezeichnet. Betroffene können etwa um die Lügen von Kriegsparteien oder politischen Kandidat:innen wissen und diese Parteien oder Personen weiterhin unterstützen. Das bedeutet, dass im öffentlichen Diskurs eingesetzte Desinformation die Haltungen gegenüber Persönlichkeiten und Themen auch dann noch beeinflussen kann, wenn die Information unmittelbar und schlagend widerlegt wurde.

Nicht abschließend geklärt sind der Einfluss und die Wechselwirkung einer Vielzahl an Faktoren, die immer wieder in Untersuchungen zu Faktenchecks auftauchen. Dazu gehören einerseits Bildung, das Vertrauen der Betroffenen in die Wissenschaft und in das bestehende politische System, aber auch die Schnelligkeit, Art und Formulierung, in der die Richtigstellung der Desinformation präsentiert wird, und ob die korrigierende Quelle den Betroffenen bekannt ist. Einige Forschungsergebnisse deuten zudem darauf hin, dass es wirksamer ist, eine Desinformation in Form einer Erzählung zu widerlegen als schlicht die nackten Fakten darzulegen. Hier besteht ein Bedarf an weiteren tiefgehenden Studien, um die angesprochenen Widersprüche aufzulösen oder zu erklären.

Gefahr nicht-intendierter Effekte? — Für alle reaktiven Maßnahmen gilt: ein möglicher negativer Effekt einer Reaktion auf Desinformation ist, dass die Auseinandersetzung mit derselben zu deren Verbreitung und Wirkung beiträgt.

Wie skalierbar ist der Ansatz? — Faktenchecks, die als Bestandteil journalistischer Arbeit durchgeführt werden, finanzieren sich durch Unternehmensgewinne. Eigenständige Institutionen, die sich auf Faktenchecks spezialisiert haben, sind hingegen oft auf externe Unterstützung angewiesen. Die Vielzahl an Initiativen weltweit deutet darauf hin, dass dieser Ansatz grundsätzlich gut hochskalierbar ist. Allerdings soll-

ten dabei grundlegende strukturelle Nachteile nicht übersehen werden:

1. **Zeit- und Fachaufwand:** Die Korrektur von Desinformation erfordert oft deutlich mehr Zeit und Fachwissen als deren Erstellung. Der Zugang zu den für die Verifikation notwendigen Informationen ist in autoritären Regimen, Krisen- und Konfliktregionen nicht immer möglich. Eine umfassende Korrektur aller Desinformationen ist daher nicht realistisch.
2. **Effektive Verteilung:** Die Verteilungsmechanismen müssen so gestaltet sein, dass die korrigierten Informationen schnell, zielgerichtet und wiederholt diejenigen erreichen, die am empfänglichsten für die spezifische Desinformation sind. Idealerweise sollte die Korrektur gleichzeitig mit der Verbreitung der Desinformation erfolgen. Dies dürfte jedoch aufgrund der in Punkt 1 beschriebenen Herausforderungen (Zeit- und Fachaufwand) nie restlos zu bewerkstelligen sein.
3. **Fragmentierung sozialer Medien:** Die zunehmende Fragmentierung sozialer Medien erschwert sowohl die Erfassung von Desinformation als auch die Verbreitung der entsprechenden Korrekturen.

Die Einbindung von KI zur Erkennung von Desinformation, zur Recherche und zur Erstellung und Verteilung korrektiver Kurznachrichten ist denkbar und wird schon erprobt. Für alle hier betrachteten Ansätze gilt: Die Qualität der Trainingsdatensätze ist entscheidend, um die Gefahr von »algorithm biases« zu reduzieren. Ansätze zur Reduzierung dieses Risikos und zur Bereitstellung ausgewogener KI-Trainingsdatensätze und Überprüfungsmethoden gibt es bereits. Die Vielfalt der Desinformationsformate — Text, Bild, Bewegtbild und Ton — wird allerdings weiterhin eine Herausforderung für eine automatisierte Erkennung bleiben.

Kennzeichnung

Was ist das? — Unter diesem Ansatz, auch als Labeling bekannt, versteht man die Anfügung von (Warn-)Hinweisen oder relevantem Kontext bei einzelnen Beiträgen oder Quellen. Beispielsweise werden Accounts in sozialen Medien, Beiträge oder Links als journalistisch qualitativ hochwertig oder als staatlich finanzierte Quellen gekennzeichnet. Zu den üblichen Kennzeichnungen gehören auch Hinweise darauf, dass zu einem bestimmten Sachverhalt noch nicht alle Informationen vorliegen oder dass es dazu widersprüchliche Informationen gibt. Zudem wird angegeben, ob Informationen möglicherweise veraltet sind. Auch das Labeling von Bildern, Videos oder Tonaufnahmen — also ob sie authentisch sind oder mit Hilfe von KI erzeugt wurden — fällt unter diesen Ansatz. Das Ziel der Kennzeichnung ist, Einfluss darauf zu nehmen, wie Rezipient:innen die Richtigkeit und Wichtigkeit der dargebotenen Informationen einschätzen. Kennzeichnungen können, müssen aber nicht mit einem Faktencheck einhergehen.

Wie wird die Wirkung eingeschätzt? — Zahlreiche Forschungsergebnisse belegen, dass die Kennzeichnung von strittigen und falschen Inhalten die Wahrscheinlichkeit verringern kann, dass diese Informationen geglaubt oder geteilt werden. Dafür muss die Kennzeichnung jedoch unmittelbar ins Auge springen und der enthaltene Text kurz und vor allem explizit sein. Vage Hinweise, wie etwa die Feststellung, dass ein Sachverhalt unter Beobachtung stehe, genügen vor allem in sozialen Medien nicht, um die Verbreitung effektiv einzudämmen.

Neuere Studien haben gezeigt, dass eine einfache Aufforderung zum Nachdenken über die Exaktheit eines Beitrags oft ausreicht, um das unreflektierte Teilen von Inhalten und damit die Verbreitung von Desinformation einzuschränken. Die Forscher:innen gehen davon aus, dass die meisten Menschen nicht bewusst Unwahrheiten verbreiten wollen, in sozialen Medien aber andere Faktoren — etwa Belustigung

oder der Wunsch nach Bestätigung — das Nutzungsverhalten erheblich beeinflussen. Die Aufforderung, vor dem Teilen über die Korrektheit und den Wahrheitsgehalt eines Beitrags nachzudenken, kann dieses Handlungsschema aufbrechen.

Zusätzlich zur Kennzeichnung einzelner Beiträge kann auch die allgemeine journalistische Qualität einer Quelle bewertet werden. Diesen Ansatz verfolgt beispielsweise das Portal Newsguard, das anhand grundlegender journalistischer Kriterien, wie Faktentreue und Transparenz, eine Bewertung auf einer Skala von 0 bis 100 vergibt. Eine Studie zum Mediennutzungsverhalten ergab, dass Newsguards Kennzeichnung bei den stärksten und häufigsten Konsument:innen von Desinformation zu einem qualitativ höherwertigen Nachrichtenkonsum führt.

Gefahr nicht-intendierter Effekte? — Die Forschung zur Kennzeichnung zeigt vor allem zwei mögliche nicht-intendierte Konsequenzen auf: den »implizierten Wahrheitseffekt« sowie den »implizierten Unwahrheitseffekt«. Einige Studien kamen zu dem Ergebnis, dass das Vorhandensein von Warnhinweisen bei einzelnen Beiträgen die Neigung der Proband:innen erhöht, die ungekennzeichneten Beiträge allgemein als glaubwürdiger oder unglaubwürdiger einzustufen. Warum bestimmte Gruppen tendenziell mehr von dem einen oder anderen Effekt betroffen sind, kann bisher nicht abschließend beantwortet werden. Die umfassende Kennzeichnung aller Beiträge und Quellen als geprüft (wahr oder unwahr) oder ungeprüft stellt angesichts der schieren Masse an Inhalten in sozialen Medien keinen praktikablen Lösungsansatz dar.

Eine Untersuchung von Twitter-Beiträgen aus 2020 — unter anderem von Donald Trump — hat gezeigt, dass Kennzeichnung die Verbreitung von Desinformation nicht immer eindämmt. Seine als »umstritten« gekennzeichneten Beiträge wurden deutlich öfter geteilt als die ungekennzeichneten. Hier bedarf es weiterer Forschung, vor allem weil sehr prominente Einzelpersonen oftmals eine überproportionale Wirkung bei der Verbreitung von Desinformation entfalten können.

Wie skalierbar ist der Ansatz? — Die schon jetzt umfangreich praktizierte Kennzeichnung von Informationen in sozialen Medien zeigt, dass dieser Ansatz im großen Stil implementiert werden kann. Im Vergleich zu weniger transparenten Methoden, wie der bloßen Entfernung oder der Verringerung der Sichtbarkeit von Beiträgen, wird die Kennzeichnung von vielen Nutzer:innen zudem besser akzeptiert. Ähnlich dem Faktencheck sind jedoch einige strukturelle Nachteile zu berücksichtigen:

1. **Individuelle Kennzeichnung:** Eine individuelle Prüfung und explizite Kennzeichnung — durch Menschen — ist effektiver, erfordert jedoch Zeit und Ressourcen, was die Umsetzbarkeit eines solchen Ansatzes einschränkt.
2. **Automatische Kennzeichnung:** Eine auf Algorithmen basierende automatische Kennzeichnung ist verhältnismäßig unpräzise und birgt die Gefahr von »algorithm biases«, ermöglicht jedoch eine schnellere und breitere Anwendung.
3. **Fragmentierung sozialer Medien:** Die wachsende Zahl sozialer Medienplattformen mit unterschiedlicher Ausrichtung, Ressourcenausstattung und öffentlicher Sichtbarkeit erschwert die Etablierung eines einheitlichen Kennzeichnungsprozesses und schränkt die Skalierbarkeit des Ansatzes ein.

Die Einbindung von KI hat das Potential, bei der Bewältigung großer Datenmengen im Prozess präziser Kennzeichnungen zu unterstützen.

Deplatforming

Was ist das? — Unter diesem Ansatz wird hier die Entfernung von inauthentischen Netzwerken von sozialen Plattformen zusammengefasst, die eine authentische Identität oder einen authentischen Zweck in der Absicht vortäuschen, Desinformation zu verbreiten. Die Abschaltung von Seiten und

Accounts des sogenannten Doppelgänger-Netzwerks ist ein aktuelles Beispiel für eine derartige Maßnahme.

Das Deplatforming wird von den sozialen Netzwerken selbst, von Dienstleistern sowie von staatlichen oder zivilgesellschaftlichen Akteuren initiiert. Die Abschaltung erfolgt aufgrund unterschiedlicher gesetzlicher Vorgaben, öffentlichen Drucks oder Verstößen gegen Plattformrichtlinien. Das Vorgehen ist nicht umfassend und einheitlich geregelt, daher kann es sich von Plattform zu Plattform oder innerhalb derselben Plattform von Land zu Land unterscheiden.

Wie wird die Wirkung eingeschätzt? — Deplatforming wird aktuell regelmäßig in den Medien thematisiert. Die Anzahl der öffentlichkeitswirksamen Abschaltungen ist aber kein Beleg für deren Wirksamkeit. Die zu beobachtende Zunahme derartiger Berichte könnte ebenso auf ein gesteigertes öffentliches Interesse oder eine wachsende Zahl von Desinformationskampagnen zurückzuführen sein.

Einige Untersuchungen deuten darauf hin, dass kontinuierliche Abschaltungen und der damit verbundene stetige Anpassungsdruck zulasten der Qualität der einzelnen Desinformationsbeiträge (z. B. mehr Übersetzungsfehler, schlichterer Content) gehen und deren Reichweite reduzieren können, selbst wenn die Betreiber immer wieder neue Accounts oder Seiten erstellen. Auch wenn es gelingt, derartige Kampagnen klar zu attribuieren, wird eine begrenzte Wirkung attestiert, da manche Desinformationsakteure dann lieber Wege der Einflussnahme wählen, bei denen sie die Urheberschaft einfacher abstreiten können.

Die Bedeutung von inauthentischen Netzwerken in sozialen Medien für die Verbreitung von Desinformation und auf die Einstellung von Nutzer:innen ist umstritten. Eine Studie zu russischen Einflussnahmeversuchen während der US-Wahlen 2016 etwa konnte keine definitive Wirkung auf die Einstellung und das Verhalten der Twitter-Nutzer:innen nachweisen. Auch andere Untersuchungen bewerten den Impact der russischen Internet Research Agency, der sogenannten »russischen Trollfabrik«, nur als gering. Demgegenüber wird in einigen Analysen die Bedeutung authentischer/menschlicher Akteure bei der Verbreitung von Desinformation deutlich höher eingeschätzt. Die Regulierung derartiger authentischer Akteure, wie etwa der Reichsbürgerbewegung oder QAnon, ist ein besonders sensibler Bereich. Hier besteht das Risiko von Zensurvorwürfen, unabhängig davon, ob die Abschaltungen aufgrund von Verstößen gegen die Plattformrichtlinien oder auf staatliche Anordnung erfolgen.

Gefahr nicht-intendierter Effekte? — Das Deplatformen von einzelnen Accounts kann, zumindest kurzfristig, die Sichtbarkeit und damit Reichweite der Betreiber bzw. der von ihnen verbreiteten Beiträge erhöhen. Zudem haben einige Studien nachgewiesen, dass das Deplatforming zwar die Möglichkeiten zur Verbreitung von Desinformation vor allem auf großen und verhältnismäßig gut regulierten Plattformen einschränkt, gleichzeitig jedoch Migrationsbewegungen zu weniger regulierten Plattformen begünstigen und Radikalisierung beschleunigen kann. Interessanterweise haben Untersuchungen auf unregulierten Plattformen gezeigt, dass Desinformationen dort, entgegen der allgemeinen Annahme, nicht flächendeckend, sondern überwiegend innerhalb eines kleinen, aktiven Nutzerkreises geteilt werden.

Wie skalierbar ist der Ansatz? — Schwierig ist vor allem die Abschätzung der Kosten für das Deplatforming, denn die Maßnahme wird seitens der sozialen Netzwerke aktuell zumeist unter allgemeinen Posten wie Sicherheitsausgaben oder Faktenchecks subsumiert. Der Ansatz ist nur bedingt skalierbar, Grund dafür sind neben der komplizierten Kostenkalkulation folgende Herausforderungen:

1. **Auswahl der sozialen Netzwerke**: Es gibt keine einheitlichen Kriterien, nach denen Plattformen priorisiert werden könnten. Unterschiede in Größe, regio-

naler und internationaler Reichweite und potentieller politischer Wirkung machen es schwer, eine kohärente Strategie zur Bekämpfung inauthentischer Netzwerke zu entwickeln. Ein Fokus auf Plattformen wie META oder X erfordert Abwägungen, die nicht immer eindeutig messbar sind.

2. **Standardisierte Meldungen:** Die Notwendigkeit der Standardisierung von Meldungen, vor allem bei plattformübergreifenden Desinformationskampagnen, ist erkannt. Der Digital Services Act beispielsweise verweist hier auf den Code of Practice on Disinformation und den Europäischen Aktionsplan für Demokratie als Grundlage. Allerdings werden kleinere Plattformen noch gar nicht oder nur im Rahmen einer freiwilligen Teilnahme erfasst.
3. **Zusammenarbeit:** Die Natur der inauthentischen Netzwerke und deren mögliches plattformübergreifendes Wirken wird es regelmäßig erfordern, dass die Plattformen untereinander und diese mit staatlichen und zivilgesellschaftlichen Institutionen kooperieren. Allerdings unterscheiden sich diese drei Akteure deutlich, was ihre Expertise in Sachen Desinformation und ihre personelle und finanzielle Ausstattung angeht.
4. **Frühwarnfunktion:** Eine Abschaltung ist nicht immer sinnvoll. Die Beobachtung eines solchen Netzwerks kann auch als Frühwarnmechanismus dienen, der Aufschluss über aufkommende Desinformationskampagnen und das dahinterstehende Akteursnetzwerk gibt.

KI kommt bei der Erkennung inauthentischer Accounts und Netzwerke bereits zum Einsatz.

Abschließende Überlegungen und Handlungsempfehlungen

Komplexität und Unsicherheit — Im Kampf gegen Desinformation müssen wir die Komplexität und Unsicherheit des Problems als Teil der Herausforderung akzeptieren. Die — zum Teil widersprüchlichen — Forschungsergebnisse machen deutlich, dass unser Wissen über das komplexe Zusammenspiel zwischen kognitiven und sozialen Dynamiken noch immer begrenzt ist, ebenso wie das Verständnis der Rolle von Emotionen. Aktuell gibt es keine reaktive Maßnahme, die alle Verbreitungswege der Desinformation abdeckt und gleichermaßen umfassend erforscht, wirksam und leicht skalierbar ist. Dennoch sind Nichtstun und Schweigen keine Optionen. Die ausgewerteten Studien bieten Handlungsempfehlungen, um die knappen Ressourcen zur Bekämpfung von Desinformation gezielt und koordiniert einzusetzen.

Forschung — Obwohl die bisherige Forschung wertvolle Einblicke bietet, handelt es sich häufig um kleine Studien unter Laborbedingungen. Zudem fokussiert sich die absolute Mehrheit der Studien zur Bekämpfung von Desinformation auf den Globalen Norden. Und selbst innerhalb dieser Gruppe fokussiert sich wiederum ein erheblicher Teil der Analysen auf die USA. Wegen des spezifischen Zweiparteiensystems und des seit Jahren aufgeladenen politischen Umfelds lassen sich die so gewonnenen Erkenntnisse nicht ohne Weiteres in wirksame Strategien zur Bekämpfung von Desinformation im Rest der Welt übertragen. Daher besteht ein genereller Bedarf an mehr und breiter angelegter evidenzbasierter Forschung, welche die Auswirkungen von Desinformation auf Einstellungs- und Verhaltensänderungen über längere Zeiträume, in unterschiedlichen Regionen und unter Berücksichtigung unterschiedlicher Verbreitungswege untersucht. Zudem gibt es Forschungslücken bei nicht textbasierten Formen der Desinformation. Dazu zählen Memes, Deepfakes und subtilere Arten der Irreführung wie das Paltering — die Verwendung wahrheitsgemäßer Aussagen, um einen irreführenden Eindruck zu vermitteln. Für eine derartige Forschung ist der Zugang zu Daten auf *allen* Plattformen notwendig, nicht nur zu denen der großen, die derzeit vom Digital Services Act erfasst sind. Hier besteht politischer Regelungs-

bedarf, um technische, ethische und regulatorische Hindernisse für eine im öffentlichen Interesse liegende Forschung abzubauen. Gleichzeitig müssen die Privatsphäre der Nutzer:innen geschützt und Rechtssicherheit für Plattformbetreiber geschaffen werden.

Einsatz von KI — Algorithmen haben das Potential, bei der Erkennung und Kennzeichnung von Desinformation zu helfen und deren Verbreitung einzuschränken. So können beispielsweise »Natural Language Processing«-Tools (NLP) durch Sentimentanalyse und Entitätserkennung verdächtige Textmuster identifizieren. Maschinelle Lernklassifikatoren können dabei unterstützen, diese Muster als wahr oder falsch zu kategorisieren, während graphenbasierte Techniken helfen, die Informationsverbreitung besser zu verstehen und enthaltene Desinformation zu erkennen.

Es ist somit wichtig, dass KI nicht nur als Mittel zur Erstellung und Verbreitung von Desinformation eingesetzt wird. Bei der KI-gestützten Desinformationserkennung kommt es allerdings entscheidend darauf an, die Grenzen der Technologie zu verstehen und mit entsprechenden Trainingsdatensätzen und Überprüfungsmethoden faire und zuverlässige Ansätze zu entwickeln.

Staatliche Institutionen und politische Entscheidungsträger:innen — Die betrachteten reaktiven Maßnahmen sehen aus unterschiedlichen Gründen nur eine begrenzte Rolle des Staates vor. Vor allem bei aktuellen und sehr umstrittenen Themen könnten sich staatliche Akteure mit Eingriffen in Form von Faktenchecks, Kennzeichnungen und Deplatforming schnell dem Vorwurf der Zensur aussetzen. Ohnehin sind zumindest deutsche Ministerien und Behörden aufgrund der üblichen Freigabe- und Abstimmungsverfahren nur sehr begrenzt in der Lage, selbst schnell und im großen Maßstab auf Desinformation zu reagieren. Staatliche Institutionen und politische Entscheidungsträger:innen sollten eine proaktive, transparente und faktenbasierte Kommunikation praktizieren und fördern sowie folgende Punkte beachten:

- **Unterscheidung von Wirkung und Reichweite** — Bei der Auswahl reaktiver Maßnahmen ist zu berücksichtigen, dass Reichweite, das Ausmaß der in Gang gesetzten Interaktion (z. B. durch Lesen, Weiterleiten, Likes, andere Formen der Kommentierung), die Anzahl der korrigierten Desinformationen und die Größe der abgeschalteten Netzwerke nicht mit der tatsächlichen Wirkung in Bezug auf Einstellungs- und Verhaltensänderungen gleichgesetzt werden können. Wo immer möglich, ist eine Wirkungskontrolle einzufordern, die Einstellungs- und Verhaltensänderungen erfasst.
- **Wert reaktiver Maßnahmen kennen** — Ungeachtet der Kritik an der Zuverlässigkeit der Wirkungsmessung schützen reaktive Maßnahmen eine Vielzahl potentieller Zielgruppen. Darunter sind beispielsweise Menschen, die zu einem Thema noch unschlüssig sind oder bisher noch nichts davon gehört haben.
- **Frühzeitige Identifikation von Zielen** — Mutmaßliche Ziele von Desinformationskampagnen, wie Referenden und Wahlen, müssen frühzeitig identifiziert werden. Vorbeugend sind transparente und faktenbasierte Informationsquellen anzubieten sowie Sensibilisierungsmaßnahmen, sprich Aufklärungskampagnen durchzuführen, die Wähler:innen bewusst machen, dass sie möglicherweise Ziel von Desinformationsangriffen werden. So wird der Informationsraum nicht den Desinformationsakteuren überlassen, und reaktive Maßnahmen wie Faktenchecks in Verbindung mit Kennzeichnungen können beschleunigt werden.
- **Internationaler Austausch** — Der Austausch mit regelmäßig von Desinformation betroffenen Staaten wie Taiwan oder Moldau sollte gepflegt werden. Dies dient dazu, das Vorgehen von Desinformationsakteuren in unterschiedlichen kulturellen und technischen Kontexten besser zu verstehen und sich über erfolgreiche

Maßnahmen und deren notwendige Voraussetzungen auszutauschen.

- **Forschungsförderung** — Die Forschung zur Wirkung von Desinformation, zu Gegenmaßnahmen, präventiven Ansätzen und zum Einsatz von KI zur Erstellung und Erkennung von Desinformation sollte gefördert werden. Gleichzeitig sollte der Zugang der Forschung zu qualitativ hochwertigen Daten ermöglicht werden.

SWP
Stiftung Wissenschaft und Politik
Deutsches Institut für Internationale Politik und Sicherheit

Ludwigkirchplatz 3—4
10719 Berlin
Telefon +49 30 880 07-0
Fax +49 30 880 07-100
www.swp-berlin.org
swp@swp-berlin.org

ISSN (Print) 1611-6364
ISSN (Online) 2747-5018
DOI: 10.18449/2024A69

Oberstleutnant i.G. Aldo Kleemann ist Gastwissenschaftler in der Forschungsgruppe Sicherheitspolitik.