

Bundestagswahl 2021

Eine Evaluation der Regeln gegen digitale Bedrohungen

Mauritius Dorn, Sara Bundtzen



Autor:innen

Mauritius Dorn

Mauritius Dorn ist Digital Policy Fellow beim ISD Germany. Er koordiniert digitalpolitische Empfehlungen, führt Desinformationstrainings durch und koordiniert Initiativen mit anderen Forschungsorganisationen. Er unterstützt zudem das Team des Digital Policy Lab (DPL).

Sara Bundtzen

Sara Bundtzen ist Research & Policy Associate beim ISD Germany. Sie unterstützt die digitalpolitische Arbeit des Digital Policy Lab (DPL) und untersucht die Verbreitung von Desinformation und Propaganda durch staatliche und nichtstaatliche Akteure.

Herausgeberische Verantwortung:

Huberta von Voss, Executive Director ISD Germany

Das Projekt wurde finanziell von Reset Tech unterstützt. Die inhaltliche Verantwortung liegt ausschließlich beim ISD Germany.

Copyright © Institute for Strategic Dialogue (2022).
Das Institute for Strategic Dialogue (gGmbH) ist beim
Amtsgericht Berlin-Charlottenburg registriert (HRB 207 328B).
Die Geschäftsführerin ist Huberta von Voss. Die Anschrift lautet:
Postfach 80647, 10006 Berlin. Alle Rechte vorbehalten.

Überblick

Noch nie lief eine Bundestagswahl so digital ab wie 2021. Damit standen auch die Aktivitäten auf (und von) großen Online-Plattformen wie Facebook und Twitter im Fokus des öffentlichen Diskurses. Ziel dieser Analyse ist es, die Durchsetzung der Regelwerke gegen die Verbreitung von digitalen Bedrohungen während der Bundestagswahl 2021 zu untersuchen, darunter digitale Gewalt, Falsch- und Desinformationen. Die Ergebnisse der Aufarbeitung zeigen, dass bestehende nicht-regulative wie regulative Maßnahmen teilweise unzureichend und fehlerhaft von den Plattformanbietern durchgesetzt wurden. Auch die politischen Parteien gaben zwar mehrheitlich Selbstverpflichtungen ab, jedoch waren diese in einigen Bereichen nicht ausreichend, um einen fairen digitalen Wahlkampf zu ermöglichen. Angesichts der Defizite empfiehlt der Bericht, ambitionierte Initiativen auf nationaler und europäischer Ebene zu konkretisieren und auszuweiten. Hier gilt es, demokratische und pluralistische Debatten im Netz zu gewährleisten und innovative Lösungsansätze zu fördern, damit diese tatsächlich stattfinden können. Gerade systemische Bestrebungen, basierend auf robusten Transparenzvorschriften und Sorgfaltspflichten, darunter Anforderungen zur Risikoanalyse und -minderung, sollten weiter definiert und gestärkt werden.

Inhaltsverzeichnis

Einleitung	5
Wie effektiv haben Plattformen die bestehenden Regeln durchgesetzt?	6
Digitale Gewalt	6
Netzwerkdurchsetzungsgesetz	6
Gemeinschaftsstandards	7
Falsch- und Desinformationen	9
Verhaltenskodex zur Bekämpfung von Desinformation	9
Gemeinschaftsstandards	10
Risiken algorithmischer Empfehlungssysteme	12
Verhaltenskodex und Gemeinschaftsstandards	12
Medienstaatsvertrag	13
Inwiefern haben Parteien zu einem fairen digitalen Wahlkampf beigetragen?	15
Fazit: Unzureichende und fehlerhafte Durchsetzung	17
Politischer Ausblick	18
Empfehlungen	21
Endnoten	23

Einleitung

Mit dem Ende der Ära Merkel stand die Bundestagswahl 2021 wie nie zuvor im Zentrum der Aufmerksamkeit inländischer wie ausländischer Akteur:innen.

Gleichzeitig führte die Corona-Pandemie dazu, dass der Wahlkampf mehr im digitalen Debattenraum ausgetragen wurde. Die Aktivitäten auf und von Online-Plattformen¹ stellen in diesem Kontext ein besonderes Interesse für Forschende aus Journalismus, Wissenschaft und Zivilgesellschaft dar.

In einem breit angelegten Projekt, das von der Initiative Reset finanziert wurde, analysierte das Institute for Strategic Dialogue (ISD) neben anderen Teams und Organisationen von Juli bis Dezember 2021 die Verbreitung digitaler Gewalt sowie von Falsch- und Desinformationen. Der vorliegende retrospektive Bericht greift sowohl ISD-Forschungsbefunde als auch solche anderer Forschungsorganisationen auf, um die Durchsetzung bestehender Regelwerke wie dem Netzwerkdurchsetzungsgesetz (NetzDG), dem Medienstaatsvertrag, dem Verhaltenskodex der Bundestagsparteien zur Bekämpfung von Desinformation oder den Gemeinschaftsstandards der Plattformanbieter zu untersuchen. Die Untersuchung unterscheidet dabei zwischen folgenden Bereichen:

1. Digitale Gewalt
2. Falsch- und Desinformationen
3. Risiken automatisierter Empfehlungssysteme

Dieser Bericht fokussiert sich auf die Lücken, Versäumnisse und Grenzen der von den Online-Plattformen durchgeführten Maßnahmen. Während viele Anbieter die Gefahren, die von ihren Diensten ausgehen, mittlerweile einräumen, verschärft sich die digitale Bedrohungslage besonders dort, wo sie durch unternehmerische Gewinn-Entscheidungen ermöglicht werden. Einschränkend sei angemerkt, dass sich die Analysen nur auf eine Auswahl relevanter Plattformen wie Facebook, Twitter und Telegram beziehen, wodurch eine aussagekräftige, vergleichende Analyse der Regeldurchsetzung nur begrenzt möglich war. Der limitierte Datenzugang zu Plattformen schränkte die Forschungsgemeinschaft zudem in ihren Möglichkeiten ein, die Verbreitung problematischer Inhalte auf verschiedenen Plattformen zu analysieren. Daher empfiehlt dieser Bericht abschließend unter anderem eine Ausweitung des Datenzugangs für Forscher:innen.

Ziel dieser Untersuchung ist es, eine Grundlage für die Weiterentwicklung der nationalen wie europäischen Regelungsrahmen zu schaffen, basierend auf kritischen Befunden zur Durchsetzung ausgewählter bestehender Regeln im Kontext der Bundestagswahl 2021.

Wie effektiv haben Plattformen die bestehenden Regeln durchgesetzt?

Digitale Gewalt

Digitale Gewalt ist ein Phänomen, das sich mit der Verbreitung von elektronischen Kommunikationsmitteln und Online-Plattformen ständig weiterentwickelt und unterschiedliche Formen annimmt, darunter illegale Hassrede. In Deutschland sollen Nutzer:innen vor digitaler Gewalt auf Plattformen bislang vor allem durch zwei Regelwerke geschützt werden:

- das Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (Netzwerkdurchsetzungsgesetz - NetzDG)²; und
- die Gemeinschaftsstandards der Plattformanbieter

Netzwerkdurchsetzungsgesetz

Im Kontext der zunehmenden Verbreitung von Hasskriminalität im Netz, der Erfahrungen im US-Wahlkampf und weitestgehend folgenlose Gespräche mit den Plattformanbietern über freiwillige Maßnahmen legte die Bundesregierung 2017 das NetzDG vor. Dieses zielt darauf ab, Hasskriminalität, strafbare Falschnachrichten und andere strafbare Inhalte auf den sozialen Netzwerken effektiver zu bekämpfen. Anbieter sozialer Netzwerke³ müssen demnach Beschwerden der Nutzer:innen unverzüglich zur Kenntnis nehmen und auf strafrechtliche Relevanz prüfen. „Offensichtlich rechtswidrige Inhalte“ müssen innerhalb von 24 Stunden gelöscht oder gesperrt werden, jeder rechtswidrige Inhalt innerhalb von sieben Tagen. Außerdem müssen die Anbieter halbjährlich einen Transparenzbericht vorlegen. Im Jahr 2021 wurde das NetzDG novelliert, einige relevante Änderungen treten allerdings erst ab Februar 2022 in Kraft (siehe „Politischer Ausblick“).

Durchsetzung

Mit der gesetzlichen Verpflichtung zur Löschung oder Sperrung rechtswidriger Inhalte sehen sich die Anbieter mit der Aufgabe konfrontiert, Abgrenzungen zwischen der Rechtmäßigkeit und Rechtswidrigkeit gemeldeter Beiträge innerhalb eines relativ kurzen Zeitraums vorzunehmen.

Hass und Hetze gegen Politiker:innen weit verbreitet trotz NetzDG

Analysen des ISD Germany und der Opferschutzorganisation HateAid zeigen, dass die Löschung oder Sperrung rechtswidriger Inhalte teilweise unzureichend war. Die Auswertung zeigt zudem, dass politische Kandidat:innen vielfach hasserfüllten Inhalten ausgesetzt waren, die schädlich, aber nicht rechtswidrig waren.

- **Telegram:** Eine Analyse des ISD zeigt, dass neun der jeweils 100 meistgelesenen Beiträge zu den drei Spitzenkandidat:innen (von CDU, SPD und Bündnis 90/Die Grünen) in qualitativ ausgewählten Telegram-Kanälen des rechtsextremen und verschwörungsaffinen Milieus⁴ im Zeitraum 1. Februar bis 31. Juli 2021 potenziell rechtswidrige Inhalte aufwiesen. In einem Fall wurden Nutzer:innen über Telegram sogar mobilisiert, koordiniert Hasskommentare gegen CDU-Spitzenkandidaten Armin Laschet bei Instagram zu verfassen.⁵
- **Twitter:** Auswertungen von HateAid für den Zeitraum 19. bis 26. August 2021 zeigen, dass Laschet 5.662 potenziell hasserfüllte Kommentare⁶ auf Twitter erhielt - fünf- bis neunmal häufiger als im Fall der anderen Spitzenkandidat:innen.⁷ In dem Zeitraum vom 7. August bis 7. September 2021 waren es 27.476 entsprechende Kommentare. Olaf Scholz (SPD) wurde im selben Zeitraum in 6.690, Annalena Baerbock (Bündnis 90/Die Grünen) in 5.524 Tweets im Kontext mit potenziell beleidigender und verletzender Sprache genannt. Basierend auf einer stichprobenartigen rechtlichen Überprüfung von je 100 potenziell hasserfüllten Kommentaren zu Laschet, Scholz und Baerbock wurden etwa 7 % als potenziell rechtswidrig eingestuft.⁸

Löschung erfolgt nach Gemeinschaftsstandards trotz NetzDG-Beschwerde

Analysen der Süddeutschen Zeitung (SZ) sowie der Initiative Reset und HateAid legen dar, dass die Löschung oder Sperrung von potenziell „offensichtlich rechtswidrigen Inhalten“ auf **Facebook**

teilweise nicht fristgerecht und häufig auf Basis der Gemeinschaftsstandards erfolgte, anstatt auf Basis des NetzDG-Meldeverfahrens. Die untersuchten Inhalte waren vor Beschwerde nach NetzDG von einer Juristin als rechtswidrig eingestuft worden.

- Ein investigatives Projektteam der SZ fand heraus, dass zwölf von 69 (17,4 %) Facebook-Kommentaren, die es nach NetzDG im August 2021 gemeldet hatte, trotz potenzieller Rechtswidrigkeit - unter anderem Beleidigungs- und Volksverhetzungstatbestände - auch eine Woche nach Meldezeitpunkt immer noch nicht gelöscht oder gesperrt waren. Außerdem wurden die Kommentare in den meisten Fällen mit der Begründung gelöscht, dass sie die Gemeinschaftsstandards verletzen würden. Die Kommentare richteten sich gegen eine Politikerin von Bündnis 90/Die Grünen.⁹
- Ein ähnliches Experiment von HateAid und Reset zeigt, dass von den 100 potenziell „offensichtlich rechtswidrigen“ Kommentaren, die nach NetzDG im Vorfeld der Bundestagswahl 2021 gemeldet wurden, Facebook in gar 33 Fällen befand, dass kein Verstoß vorläge. Fast alle übrigen Kommentare wurden mit Verweis auf die eigenen Gemeinschaftsstandards gelöscht. Insbesondere Inhalte, die sich gegen hochrangige Politiker:innen richteten, /werden in den seltensten Fällen entfernt.¹⁰

Selbst wenn Anbieter alle gemeldeten potenziell rechtswidrigen Inhalte fristgerecht nach NetzDG oder Gemeinschaftsstandards löschen oder sperren würden, würde dies nicht bedeuten, dass Hasskriminalität von den Plattformen verschwindet. **Im Experiment von Reset und HateAid waren die 100 selektierten Inhalte zum Meldezeitpunkt beispielsweise im Schnitt bereits seit 409 Tagen online.** Aufgrund des im NetzDG vorgesehenen „Notice and take down“-Verfahrens muss eine Prüfung durch die Anbieter jedoch erst nach Meldung durch die Nutzer:innen erfolgen. Da die Anbieter keine Nachforschungspflicht trifft, ist entscheidend, dass Nutzer:innen den Anbietern rechtswidrige Inhalte auf eine nutzerfreundliche Art übermitteln können, damit diese davon Kenntnis erlangen. Anbieter sollten ihrer Verantwortung beim Umgang mit Beschwerden zu rechtswidrigen Inhalten gerecht werden und insofern

mehr Transparenz über die Melde- und Reaktionswege schaffen. Indem Inhalte trotz NetzDG-Beschwerde auf Grundlage der Gemeinschaftsstandards entfernt werden, umgehen die Anbieter häufig die Erfassung von Beschwerden in den Transparenzberichten sowie die im NetzDG festgeschriebene Begründungspflicht. Ursächlich ist ein zweistufiges Meldeverfahren, bei dem gemeldete Inhalte zunächst auf eine Verletzung der Gemeinschaftsstandards hin und erst bei Nichtverletzung nach NetzDG überprüft werden - eine Problematik, die bereits im Evaluationsbericht der Bundesregierung Erwähnung fand.¹¹ Im jüngsten Transparenzbericht von Facebook für den Zeitraum 1. Januar bis 30. Juni 2021 wurden 77.671 NetzDG-Beschwerden dokumentiert, in denen insgesamt 67.028 Inhalte genannt wurden¹² - ein massiver Anstieg um 1.844 %.¹³ Dieser ist jedoch vermutlich eher auf ein neues allgemein zugängliches Meldeformular zurückzuführen als auf eine gestiegene Bereitschaft der Plattformen, demokratieschädigende Inhalte zeitnah und nachhaltig zu entfernen.

Während das NetzDG die Anbieter dazu verpflichtet, rechtswidrige Inhalte nach Beschwerde der Nutzer:innen zu löschen oder zu sperren, obliegt die Bekämpfung legaler, aber dennoch hasserfüllter und problematischer Inhalte wiederum den Gemeinschaftsstandards.

Gemeinschaftsstandards

Die meisten Plattformen verfügen mittlerweile über umfassende Hausregeln - sogenannte Gemeinschaftsstandards (engl.: Community Standards)- die festlegen, welche Inhalte und welches Verhalten erlaubt oder nicht erlaubt sind, insbesondere hinsichtlich der Formen digitaler Gewalt und Hassrede. Die Gemeinschaftsstandards sind in der Regel in den Nutzungsbedingungen (engl.: Terms of Service) der Anbieter dargelegt. Nach einem Urteil des Bundesgerichtshofs im Juli 2021 dürfen sie dabei auch legale Inhalte verbieten.¹⁴ Allerdings müssen die Anbieter die betroffenen Nutzer:innen über ihre Entscheidungen angemessen informieren. Hinsichtlich digitaler Gewalt wird in den Gemeinschaftsstandards gemeinhin zwischen Hassrede, Gewalt und Anstiftung zu Gewalt sowie Belästigung und Mobbing unterschieden, wobei Hassrede bereits in vielen

Fällen abdeckt, was auch nach dem NetzDG potenziell rechtswidrig ist. Dennoch gibt es Bereiche, in denen die Gemeinschaftsstandards hinter dem gesetzlichen Rahmen zurückbleiben. Die Standards enthalten in den meisten Fällen auch Regeln für gefährliche Gruppen und Personen, die aber in diesem Bericht nicht berücksichtigt werden.

Durchsetzung

Bei der Erkennung und Entfernung unzulässiger Inhalte setzen die Plattformanbieter sowohl auf die Meldung durch Nutzer:innen, als auch verstärkt auf automatisierte Erkennungssysteme. Facebook selbst gibt an, von April bis Juni 2021 weltweit 31,5 Millionen Inhalte von Facebook entfernt zu haben, weil diese gegen die Richtlinien zu Hassrede verstoßen hätten (oder verstießen). Knapp 98 % von ihnen sollen dabei sogar eigenständig vor einer Meldung durch Nutzer:innen entfernt worden sein.¹⁵

Unzureichende und selektive Durchsetzung der Gemeinschaftsstandards

Analysen der SZ, HateAid und Reset sowie des ISD offenbaren Defizite bei der Entfernung von Inhalten, die offensichtlich gegen die Gemeinschaftsstandards der Anbieter verstoßen.

- **Facebook:** In einer Untersuchung der SZ vom 1. September 2021 waren 11 von 70 (15,7 %) Facebook-Kommentaren, die wegen offensichtlicher Verletzung der Gemeinschaftsstandards gemeldet worden waren, auch eine Woche nach Meldung noch weiterhin zugänglich. In 56 Fällen antwortete Facebook innerhalb einer Woche.¹⁶ In einem ähnlichen Experiment von HateAid und Reset wurden nur 50 % der aufgrund einer offensichtlichen Verletzung der Gemeinschaftsstandards - unter anderem der Hassrede-Richtlinie - gemeldeten Kommentare auf öffentlichen Seiten der AfD gelöscht, darunter zahlreiche Kommentare, die sich gegen Kandidat:innen richteten.¹⁷
- **Twitter:** Eine gemeinsame Analyse von ISD Germany und HateAid kommt zu dem Ergebnis, dass im Zeitraum 1. bis 30. September 2021 10.592 potenziell hassgefüllte Tweets¹⁸ veröffentlicht

wurden, die sich gegen 60 Bundestagskandidat:innen richten. Dabei finden sich Verschwörungsmythen, Fremdenfeindlichkeit, Rassismus, Sexismus und LGBTQ-Feindlichkeit in einer Stichprobe der meistgeteilten Tweets. Die untersuchten Kandidat:innen repräsentieren unterschiedliche Identitäten im Hinblick auf Gender, Sexualität, ethnische Zugehörigkeit, Migrationshintergrund und Parteizugehörigkeit.¹⁹

Dass die Durchsetzung der Gemeinschaftsstandards unzureichend und teilweise selektiv erfolgt, zeigt, dass eine weitgehend unbeaufsichtigte und intransparente Durchsetzung der Gemeinschaftsstandards das Problem der digitalen Gewalt nicht effektiv bewältigt. Die Veröffentlichung Facebook-interner Dokumente, Studien und Memos belegt zudem, dass Facebook Länder kategorisierte und solche in der höchsten Prioritätsstufe, wie die Vereinigten Staaten, bei der Bekämpfung von Hassrede bevorzugte. 84 % des Zeitbudgets für die Löschung und Eindämmung von Desinformation sollten demnach bewusst auf den US-Markt entfallen.²⁰ Bisher widersprechen Anbieter wie Facebook jedoch solchen Erkenntnissen und verweisen auf die Wirksamkeit ihrer meist automatisierten Erkennungs- und Entfernungssysteme.

Künstliche Intelligenz doch nicht so wirksam wie erhofft?

Auch im Vorfeld der Bundestagswahl 2021 verwiesen die Plattformanbieter abermals auf ihre technologischen Ressourcen. So gab YouTube an, 94 % der Inhalte mit Richtlinienverstößen im Netzwerk durch automatische Meldungen zu erkennen.²¹ Laut Johannes Baldauf, Public Policy Manager bei Facebook Deutschland, sei Facebook sogar „Vorreiter im Einsatz künstlicher Intelligenz, um unzulässige Inhalte zu erkennen und schneller zu erkennen.“²²

Die Enthüllungen der Whistleblowerin Frances Haugen stellen diese Äußerungen in Frage. So schrieb ein leitender Facebook-Ingenieur bereits Mitte 2019, dass **Facebooks** automatisierte Systeme Beiträge von der Plattform nahmen, die nur 2 % der Aufrufe richtlinienwidriger Hassrede generierten. Im März 2021 kam ein anderes Team zu einer

ähnlichen Schlussfolgerung und schätzte, dass die automatisierten Systeme nur Beiträge entfernten, die 3 bis 5 % der Aufrufe von Hassrede ausmachten, und 0,6 % aller Inhalte, die gegen die Richtlinien gegen Gewalt und Aufwiegelung verstießen. Facebook entgegnete, dass sich diese Prozentsätze nur auf Beiträge bezögen, die mithilfe von Künstlicher Intelligenz entfernt wurden und keine anderen Maßnahmen berücksichtigten wie zum Beispiel ein niedrigeres Ranking im NewsFeed.²³ Während automatisierte Moderation von Inhalten skalierbar und schnell arbeitet, steigen gleichzeitig die Risiken der Beschränkungen der Meinungsfreiheit. Förderung menschlicher Moderator:innen mit regionalen und nationalen Kontextwissen für umsichtige Einzelfallentscheidungen bleibt ein zentraler Aspekt von Regulierungsbemühungen.

Falsch- und Desinformationen

Nach wie vor ist der Begriff „Desinformation“ nicht eindeutig definiert, was die Erarbeitung effektiver Regelwerke erschwert. Desinformationsexperte James Pamment von der Universität Lund schlägt die Verwendung folgender Begrifflichkeiten vor: Falschinformation, Desinformation, Einflussoperation und ausländische Beeinflussung. Falschinformation meint dabei eine falsche Information, die eine Einzelperson ohne Absicht der Irreführung verbreitet. Desinformation bezieht sich wiederum auf einen nachweislich falschen Inhalt, der zum wirtschaftlichen Vorteil oder zur absichtlichen Täuschung der Öffentlichkeit erstellt, präsentiert und verbreitet wird, und der der Öffentlichkeit Schaden zufügen kann.²⁴ In Deutschland sollen Nutzer:innen vor Falsch- und Desinformationen auf den Plattformen vor allem durch den europäischen Verhaltenskodex zur Bekämpfung von Desinformationen und (daraus resultierende) Gemeinschaftsstandards geschützt werden. Aber auch das NetzDG bietet Schutz vor ehrverletzenden Delikten wie beispielsweise Verleumdung.

Verhaltenskodex zur Bekämpfung von Desinformation

Im Jahr 2018 haben unter anderem Facebook, Google, Twitter und TikTok den Verhaltenskodex zur Bekämpfung von Desinformation (engl.: Code of Practice on Disinformation - CoPD)²⁵ unterzeichnet,

den die Europäische Kommission initiiert hatte. Basierend auf der Definition von Desinformation der Kommission - diese Definition verwendet auch James Pamment - umfasst der CoPD freiwillige Standards und Verpflichtungen, die die Verbreitung von Desinformationen in folgenden Bereichen eindämmen sollen: Stärkung der Forschungsgemeinschaft, Kontrolle der Anzeigenschaltung, politische und themenbezogene Werbung, Integrität der Dienste und Stärkung der Verbraucher:innen. Inzwischen sind weitere Anbieter dem Kodex beigetreten. Nachdem die Kommission im Mai 2021 Leitlinien für die Stärkung des Kodex herausgegeben hatte, planen die unterzeichnenden Unternehmen und Organisationen die Vorlage einer überarbeiteten Version bis Ende 2021 (siehe Abschnitt „Politischer Ausblick“).

Durchsetzung

Bereits im Kontext der Europawahl 2019 evaluierte das ISD den CoPD. Dabei konnte nachgewiesen werden, dass der Kodex zwar zu Verbesserungen insbesondere bei der Transparenz politischer Werbung geführt hatte. Gleichzeitig wurde aber auch deutlich, dass aufgrund des selbstregulativen Ansatzes, bei dem keine Konsequenzen bei Nichteinhaltung drohen, kaum substantielle Veränderungen erreicht werden konnten.²⁶ Nachfolgende Erkenntnisse unterstreichen die bereits festgestellten Mängel. Die Kontrolle der Anzeigenschaltung wurde in der Analyse nicht berücksichtigt. Die Bestimmungen zu Algorithmen werden im Abschnitt „Algorithmische Empfehlungssysteme“ genauer untersucht.

Stärkung der Forschungsgemeinschaft

- **Facebook:** Obwohl sich Facebook als Unterzeichner des CoPD dazu bekannt hat, gewissenhafte Forschung zu Desinformation und politischer Werbung nicht zu verbieten oder zu entmutigen - ja sogar derartige Forschungszwecke proaktiv zu fördern - wurde das Projekt Ad Observer der New York University (NYU) im Kontext der Bundestagswahl 2021 unterbunden. Die Forscher:innen wollten analysieren, wem Facebook wann welche politischen Anzeigen anzeigt, um damit Erkenntnisse über die Verbreitung von Falschinformationen zu sammeln. Facebook begründete die Sperrung von Konten und

Datenzugang mit einer Anordnung der Federal Trade Commission (FTC) zum Schutz der Privatsphäre der Nutzer:innen. Die FTC dementierte diese Sichtweise jedoch.²⁷ Die Organisation AlgorithmWatch sah sich im Juli 2021 ebenfalls gezwungen, ein Datenspende-Projekt zu dem Newsfeed-Algorithmus von Instagram zu beenden, um nicht zu riskieren, von Facebook verklagt zu werden. Facebook argumentierte, das Projekt verstoße gegen die Nutzungsbedingungen, die die automatisierte Erfassung von Daten verbieten und drohte „formellere Schritte“ einzuleiten. Allerdings hatten die Datenspende:innen nur auf ihren eigenen Feed zugegriffen und die Inhalte freiwillig und zu Forschungszwecken über ein Datenspende Add-On mitgeteilt.²⁸

Politische Werbung und themenbezogene Werbung

- **Facebook:** Eine Recherche des ZDF Magazins Royale - gemeinsam mit der Initiative Who Targets Me - analysierte von April bis September 2021 mittels der Datenspenden von mehr als 17.000 Nutzer:innen 134.000 maßgeschneiderte Werbeanzeigen. Die Analyse weist unter Verwendung einer alternativen Werbebibliothek nach, dass tausende der geschalteten politischen Werbeanzeigen in Facebooks Werbebibliothek fehlten. Darunter befanden sich insbesondere auch Werbeanzeigen von Parteien, der Bundesregierung sowie von Ministerien. Die Projektverantwortlichen warfen in der Folge Facebook grobe Fehleinschätzungen bei der Einstufung von Werbeanzeigen vor.²⁹ Eine Analyse des Recherchezentrums CORRECTIV zeigt zudem, dass bei der Sperrung politischer Werbung teilweise selektiv und nicht konsequent vorgegangen wurde.³⁰

Integrität der Dienste

- **TikTok:** Als Unterzeichner des CoPD erkennt TikTok an, wie wichtig es ist, die Bemühungen zur Schließung von Fake-Accounts zu intensivieren. Eine Untersuchung der Internetfirma Mozilla legt jedoch dar, dass während der Bundestagswahl 2021 gefälschte Konten mit politisch-propagandistischen Inhalten tausende Nutzer:innen erreichten. So gab es unter anderem einen gefälschten TikTok Account (@derbundestag) mit 14,1K Followern und 130,8K Likes, der AfD-Politiker:innen prominent herausstellte.³¹

Stärkung der Verbraucher:innen

- **TikTok:** Dieselbe Analyse von Mozilla zeigt, dass TikToks automatisierte Info-Banner zur Bundestagswahl 2021 teilweise nicht wirkungsvoll waren. Zwar hatte der Anbieter ab dem 20. Juli 2021 über bestimmten Videos ein Banner eingeblendet, das mit einem Link zu einer Informationsseite versehen war. In der Praxis wurden die Info-Banner jedoch nur für einzelne Schlagwörter und Hashtags wie #SPD oder #AfD automatisiert eingeblendet, die auch nicht-wahlbezogene Inhalte einbezogen, andere relevante politische Inhalte wiederum nicht. So waren beispielsweise die allermeisten Videos der Accounts der Parteien AfD und die LINKE im Vorfeld der Wahl nicht gekennzeichnet, da sie keine der definierten Schlagwörter oder Hashtags aufwiesen. Dies unterminiert die Bemühungen, vertrauenswürdige Informationsquellen anzubieten.³²

Die Fallbeispiele zeigen, dass Maßnahmen wie Facebooks Werbebibliothek oder der Info-Banner von TikTok in der Praxis oft mangelhaft umgesetzt wurden. Bei der Stärkung der Forschungsgemeinschaft erweisen sich wiederum Bedenken hinsichtlich des Schutzes der Privatsphäre als größere Herausforderung. **Bekennnisse zum Vorgehen gegen Fake-Accounts wurden stellenweise nicht erfüllt.** Da der CoPD auf inhaltlicher Ebene kaum Regeln beinhaltet, wird der Umgang mit Falschinformationen vor allem durch die Gemeinschaftsstandards geregelt.

Gemeinschaftsstandards

Unter Falschinformationen fassen die Plattformanbieter in ihren Gemeinschaftsstandards gemeinhin Meldungen, die von externen Faktenprüfer:innen als falsch eingestuft wurden. **Die meisten als falsch eingestuften Inhalte werden nicht von den Plattformen entfernt, sondern lediglich weniger stark verbreitet, indem sie weiter unten im News Feed angezeigt werden.** Dennoch existieren Ausnahmen. Beispielsweise entfernen Facebook und Instagram Falschinformationen, die gegen Gemeinschaftsstandards verstoßen - darunter Inhalte über den Wahlprozess, die Menschen von der Stimmabgabe abhalten könnten.³³ Auf Twitter besteht für falsche oder irreführende Inhalte gegen

die Integrität staatsbürgerlicher Prozesse ebenfalls die Möglichkeit zur Entfernung.³⁴ Auf YouTube sind zudem auch Falschinformationen zur Wählbarkeit von Kandidat:innen nicht erlaubt.³⁵ TikTok untersagt zusätzlich sogar bestimmte verschwörungsideologische Inhalte³⁶, während Facebook keine Inhalte von Politiker:innen prüfte. Anders als in der Vorwoche der US-Wahl 2020, sahen die meisten anderen Anbieter keine Sonderrechte vor und moderierten sämtliche Inhalte auf Basis ihrer Standards. Laut Facebook erhalten durch Politiker:innen geteilte und zuvor als falsch eingestufte Inhalte eine geringere Verbreitung im News Feed.

Durchsetzung

Im Umgang mit Falschinformationen setzen die Anbieter sozialer Netzwerke vor allem auf die Zusammenarbeit mit externen Faktenprüfer:innen. So arbeitet Facebook in Deutschland beispielsweise seit einigen Jahren mit den Nachrichtenagenturen AFP, DPA und dem Recherchezentrum CORRECTIV zusammen.³⁷ TikTok hingegen startete seine Zusammenarbeit mit externen Faktenprüfer:innen erst wenige Wochen vor der Bundestagswahl 2021.³⁸ Nachdem Faktenprüfer:innen einen Inhalt identifiziert und geprüft haben, wird dieser in der Regel mit einem Hinweis bzw. Label versehen und seine Verbreitung eingeschränkt.

Gemeinschaftsstandards bieten kaum Schutz vor Falschinformationen

- **Telegram:** Eine Untersuchung des ISD Germany zeigt, dass die jeweils 100 meistgelesenen Beiträge zu den drei Spitzenkandidat:innen aus qualitativ ausgewählten Telegram-Kanälen des rechtsextremen und verschwörungsideologischen Milieus³⁹ im Zeitraum 1. Februar bis 31. Juli 2021 zu einem hohen Anteil Falschinformationen oder Anspielungen auf Verschwörungserzählungen enthielten: 43 der Posts über Baerbock (Bündnis 90/Die Grünen), 26 der Posts über Laschet (CDU) und 17 der Posts über Scholz (SPD).⁴⁰ Das Bedrohungspotenzial, das von Verschwörungsideolog:innen auf Telegram ausgeht, wird auch in einer Analyse des Center für Monitoring, Analyse und Strategie (CeMAS) erkenntlich. Die analysierten Telegramkanäle ermöglichten es unter anderem zentralen Akteur:innen

des verschwörungsideologischen Milieus, Desinformationen mit einer großen Öffentlichkeit zu teilen. Dabei waren die Manipulation von Umfragen, Handlungsanweisungen zu „Shitstorming“ gegenüber Behörden und Institutionen sowie die Erstellung eigener Umfragen im Kontrast zu repräsentativen Erhebungen wiederkehrende Elemente der Kommunikation.⁴¹

- **Facebook:** Forscher:innen des ISD fanden zudem heraus, dass insgesamt 21 der jeweils 100 meistgeteilten öffentlichen Beiträge zu den drei Spitzenkandidat:innen aus 58.858 Facebook-Beiträgen vom 1. Februar bis 1. Juli 2021 Falschinformationen oder Anspielungen auf Verschwörungserzählungen enthielten. Dabei war Annalena Baerbock mit 18 Beiträgen im Vergleich zu Olaf Scholz und Armin Laschet deutlich häufiger Ziel von Verschwörungsmythen, Des- und Falschinformationen.⁴² Eine Studie des Sustainable Computing Lab (SCL) und der Wirtschaftsuniversität Wien (WU) zeigt außerdem, dass etwa 6,7 % aller Facebook-Beiträge vom 15. bis 31. März 2021 mit Schlagwort- beziehungsweise Hashtag-Bezug zur Bundestagswahl 2021 potenziell rechtswidrige Inhalte, Desinformation oder Verletzungen des Wahlrechts darstellten. Dabei waren mehr als 90 % der Inhalte Desinformationen, etwa 35 % verstießen gegen die Gemeinschaftsstandards und 4 % waren wahrscheinlich rechtswidrig im Sinne der Tatbestände des Strafgesetzbuchs.⁴³

- **Twitter:** Für Twitter analysiert dieselbe Studie von SCL und WU, dass etwa 5,6 % aller wahlbezogenen Tweets vom 15. bis 31. März 2021 problematisch waren. Von den als problematisch identifizierten Inhalten weisen alle Desinformationen auf, 14,5 % verstießen gleichzeitig gegen die Gemeinschaftsstandards.⁴⁴ Im Zuge der Landtagswahl in Sachsen-Anhalt am 6. Juni 2021 beobachtete das Team von ISD Germany deutschsprachige Beiträge in den sozialen Medien, vor allem auf Twitter und Facebook. So stieg beispielsweise die Anzahl von Beiträgen mit dem Hashtag #Wahlbetrug, der in den Tagen vor der Wahl erst wenige Male genutzt worden war, nach Verkündung der ersten Hochrechnungen deutlich: der Hashtag wurde innerhalb von 24 Stunden 4.860-mal und im Zeitraum 1. Juni bis 14. Juni 2021 6.480-mal

verwendet. Neben Twitter wurden Desinformationen über die Wahl auch über Facebook, Telegram oder Bitchute, sowie über die Blogs und Newsletter rechter Influencer wie Roland Tichy, Boris Reitschuster und Oliver Janich verbreitet.⁴⁵

Nachweislich falsche Inhalte ohne Warnhinweise auf Facebook ersichtlich

- Forscher:innen der Bewegung Avaaz fanden heraus, dass mehr als 50 % der 137 Facebook-Beiträge vom 1. Mai bis 31. Juli 2021, die von unabhängigen Faktenprüfer:innen als falsch eingestufte Narrative über Annalena Baerbock enthielten, ohne Warnhinweise ersichtlich blieben. Dabei waren nahezu identische Beiträge manchmal mit und manchmal ohne Warnhinweis versehen - ein Indiz für unzureichende Erkennungssysteme, bei denen auch KI zum Einsatz kommt.⁴⁶
- Eine Untersuchung des ISD legt zudem dar, dass Beiträge vom 1. Januar 2020 bis 13. Juni 2021, die die Allianz der Weltärzte (WDA) oder ihre Mitglieder erwähnten und bereits als falsch eingestufte, auf Deutsch verfasste Narrative beinhalteten, seltener mit einem Warnhinweis versehen waren als Beiträge auf Englisch. Dabei unterschieden sich die Beiträge oft nur marginal von bereits mit einem Hinweis versehenen Beiträgen. Darüber hinaus fanden die Forscher:innen heraus, dass auch mit einem Warnhinweis versehene Beiträge immer noch zehn- oder hunderttausende Aufrufe erhielten.⁴⁷

Die Fallbeispiele veranschaulichen, dass Falschinformationen im Kontext der Bundestagswahl 2021 auf einigen Plattformen stark verbreitet waren. Besonders problematisch ist der Befund, dass nachweislich falsche Inhalte teilweise nur selektiv mit Warnhinweisen versehen waren. Damit rücken auch die Bedingungen und Regeln für algorithmische Empfehlungssysteme, die die Reichweite problematischer Inhalte im Rahmen der Kuratierung weniger prominent darstellen sollen, als zusätzliche Komponente in den Fokus dieses Berichts.

Risiken algorithmischer Empfehlungssysteme

Beiträge und Kommentare auf großen Online-Plattformen werden durch algorithmische Empfehlungssysteme priorisiert und dargestellt. Sie treffen eine grundlegende Auswahl großer Informationsmengen und sortieren diese nach Kriterien so, dass möglichst relevante Informationen zuerst angezeigt werden.⁴⁸ Damit kommt den Empfehlungssystemen eine elementare Rolle bei der Verbreitung und Verstärkung problematischer Inhalte auf den Plattformen zu. Bisher wurden für Deutschland kaum Regeln eingeführt, die Empfehlungssysteme regulieren. Dennoch enthalten der CoPD und die Gemeinschaftsstandards sowie der Medienstaatsvertrag bereits einzelne Bestimmungen zu Priorisierung und Transparenz.

Verhaltenskodex und Gemeinschaftsstandards

Die Unterzeichner des CoPD⁴⁹ verpflichten sich dazu, in Technologien zu investieren, um relevante, authentische und maßgebliche Informationen - wo möglich - in der Suche, in Feeds oder anderen automatisch geordneten Verbreitungschanälen zu priorisieren. Darüber hinaus geben die meisten Anbieter großer Plattformen in ihren Gemeinschaftsstandards an, die Verbreitung problematischer Inhalte wie entlarvte Falschinformationen zu verringern.⁵⁰ YouTube verweist sogar darauf, dass 2019 Änderungen an den Empfehlungssystemen eingeführt wurden, um die Verbreitung problematischer Falschinformationen und grenzwertiger Inhalte zu begrenzen.⁵¹ Im CoPD legen sich die Unterzeichner außerdem darauf fest, Forscher:innen und Faktenprüfer:innen generelle Informationen über ihre Algorithmen bereitzustellen.

Durchsetzung

Bei der Weiterentwicklung algorithmischer Empfehlungssysteme setzen die Anbieter unter anderem aufs Feedback der Nutzer:innen.⁵² In den letzten Jahren haben einige Anbieter außerdem einen Austausch mit externen Expert:innen angestoßen. Facebook hat sich dabei zum Ziel gesetzt, dass jedes eingeführte System Schutz vor Manipulation und Gaming bietet, persönliche Voreingenommenheit

vermeiden und die Stimmen der Minderheiten schützen müsse.⁵³ Ein Blick auf die Analysen im Kontext der Bundestagswahl 2021 zeigt jedoch, dass gewisse Stimmen im Feed durch die Empfehlungssysteme bevorzugt werden.

- **Facebook:** Eine Untersuchung des Projektes Citizen Browser der Organisation The Markup in Zusammenarbeit mit der Süddeutschen Zeitung (SZ) belegt auf Basis der Daten eines politisch diversen Panels von 473 deutschsprachigen Nutzer:innen, dass Beiträge AfD-bezogener Seiten vom 20. Juli bis 16. September 2021 dreimal so häufig in den NewsFeeds der Nutzer:innen vorzufinden waren, wie die einer anderen Partei.⁵⁴ Eine Studie von HOPE NOT HATE zeigt außerdem, dass Beiträge der offiziellen AfD-Seite vom 1. Mai bis 30. Juni 2020 im Schnitt rund 15-mal häufiger als die der CDU-Seite und 5-mal häufiger als die der Seite von Bündnis 90/Die Grünen geteilt wurden.⁵⁵ Eine ethnografische Analyse des ISD legt darüber hinaus dar, dass insbesondere AfD-bezogene Seiten und Gruppen Zweifel am Wahlergebnis in Deutschland auf unterschiedlichen Plattformen säten.⁵⁶ Dies legt nahe, dass problematische Inhalte priorisiert worden sein könnten.

- **Instagram:** Eine andere Untersuchung von SZ und AlgorithmWatch auf Grundlage von mehr als 6.000 Beiträgen von mehr als 100 Accounts von Politiker:innen und Parteien belegt außerdem, dass der Empfehlungsalgorithmus von Instagram bestimmte Themen und Parteien - insbesondere die AfD - weiter oben im Feed platzierte. Dabei blieben die Entscheidungsparameter letztlich unklar - Facebook verwies in einer Antwort auf eine Anfrage der Forscher:innen einzig darauf, dass inhaltliche Kriterien keine Rolle bei der Platzierung von Beiträgen spielten.⁵⁷

Jede Verzerrung der Rangfolge und Darstellung von Inhalten im Feed wirkt sich auf tausende wahlberechtigte Nutzer:innen und ihre Meinungsbildung aus. Eine Untersuchung der Hintergründe von Verzerrungen ist jedoch einzig durch eine verlässliche Transparenzverpflichtung und weiterführende Informationen zur Funktionsweise algorithmischer Empfehlungssystemen möglich - ein Aspekt, den auch der Medienstaatsvertrag aufgreift.

Die Enthüllungen der Whistleblowerin Frances Haugen geben Einblicke, die bislang verwehrt blieben. So führten **Facebooks** Änderungen des NewsFeed Algorithmus im Jahr 2018 dazu, dass Verlage und sogar politische Parteien verstärkt auf sensationelle und empörende Inhalte setzten, da dies zu mehr Engagement auf Facebook führte. Hintergrund war eine neue Formel, mit der gemessen wurde, wie viel sinnvolle, soziale Interaktion (engl.: meaningful interaction) ein Beitrag auslöst. Um Sichtbarkeit zu gewinnen, mussten Beiträge daher möglichst viel „hochwertiges Engagement“ erhalten (wie beispielsweise Reaktionen oder Antworten auf Kommentare). Nach einem internen Punktesystem war ein „Gefällt mir“ einen Punkt, eine Reaktion, ein erneutes Teilen ohne Text oder eine Antwort auf eine Einladung fünf Punkte und ein substantieller Kommentar, eine Nachricht oder ein erneutes Teilen 30 Punkte wert. Zusätzliche Multiplikatoren wurden hinzugefügt, je nachdem, ob die Interaktion zwischen Mitgliedern einer Gruppe, Freunden oder Fremden stattfand.⁵⁸ Trotz genannter Beschwerden wurden die Probleme erst ab August 2021 in einigen europäischen Ländern im Rahmen einer Testphase angegangen.⁵⁹

Medienstaatsvertrag

Um Meinungsvielfalt und Chancengleichheit auch im Netz zu ermöglichen, verabschiedeten die 16 Landesparlamente im Jahr 2020 den Staatsvertrag zur Modernisierung der Medienordnung in Deutschland (Medienstaatsvertrag - MStV) als Ergebnis eines mehrjährigen Prozesses. Dieser löste den bisherigen Rundfunkstaatsvertrag ab und führte unter anderem umfassende Pflichten für sogenannte Medienintermediäre (MI) wie Facebook oder Google ein, die im Rahmen einer spezifischen MI-Satzung der Länder weiter konkretisiert wurden. Soweit Medienintermediäre Einfluss auf die öffentliche Meinungsbildung haben, sind sie nach MStV zur Offenlegung der zentralen Kriterien einer Aggregation, Selektion und Präsentation von Inhalten und ihre Gewichtung einschließlich der Funktionsweise ihrer Algorithmen verpflichtet. Außerdem dürfen sie keine einzelnen journalistisch-redaktionellen Inhalte systematisch und ohne Rechtfertigung benachteiligen (Diskriminierungsfreiheit). Eine Diskriminierung liegt konkret dann vor, wenn ohne sachlich gerechtfertigten

Grund von genannten Kriterien zugunsten oder zulasten eines bestimmten Angebots systematisch abgewichen wird oder Kriterien Angebote unmittelbar oder mittelbar unbillig systematisch behindern. Zudem müssen sich alle Telemedien mit journalistisch-redaktionellen Angeboten an anerkannte journalistische Grundsätze halten, was jedoch nicht Teil dieser Untersuchung ist. Schließlich enthält der MStV auch eine Kennzeichnungspflicht für automatisiert erstellte Inhalte oder Mitteilungen, die im nächsten Kapitel angesprochen werden. Die Prüfung von Verstößen obliegt den Landesmedienanstalten, die Anzeigen erhalten und prüfen können⁶⁰ – Befugnisse, die insbesondere von der Europäischen Kommission kritisch gesehen werden.⁶¹

Durchsetzung

Obwohl die konkretisierenden Regelungen der MStV in den Bundesländern erst zum 1. Januar 2022 in Kraft treten sollen, sind Medienintermediäre bereits seit Inkrafttreten des MStV in dessen Anwendungsbereich einbezogen. Auf dieser gesetzlichen Grundlage beanstandeten die Landesmedienanstalten bereits Verstöße gegen die Diskriminierungsfreiheit.

So sprach die Kommission für Zulassung und Aufsicht (ZAK) im Juni 2021 in einem von der Medienanstalt Hamburg/Schleswig-Holstein (MA HSH) gegen **Google** eingeleiteten Verfahren eine Beanstandung wegen Verletzung der Diskriminierungsfreiheit aus. Gegenstand des Verfahrens war eine Kooperation von Google und dem Bundesministerium für Gesundheit (BMG), die eine priorisierte Darstellung von Inhalten des nationalen Gesundheitsportals www.gesund.bund.de in der deutschen Version der Google Suchmaschine vom 10. November 2020 bis zum 10. Februar 2021 umfasste. In dem Kontext wurde einer Beschwerde des betroffenen Betreibers von www.apotheken-umschau.de stattgegeben. Die ZAK stellte fest, dass durch diese Kooperation eine unbillige Behinderung anderer Anbieter journalistisch-redaktioneller Inhalte stattgefunden habe.⁶² Die Kooperation zwischen Google und dem BMG wurde inzwischen beendet.

Im Kontext der Bundestagswahl 2021 haben Anbieter wie **Google** und **Facebook** außerdem bestimmte

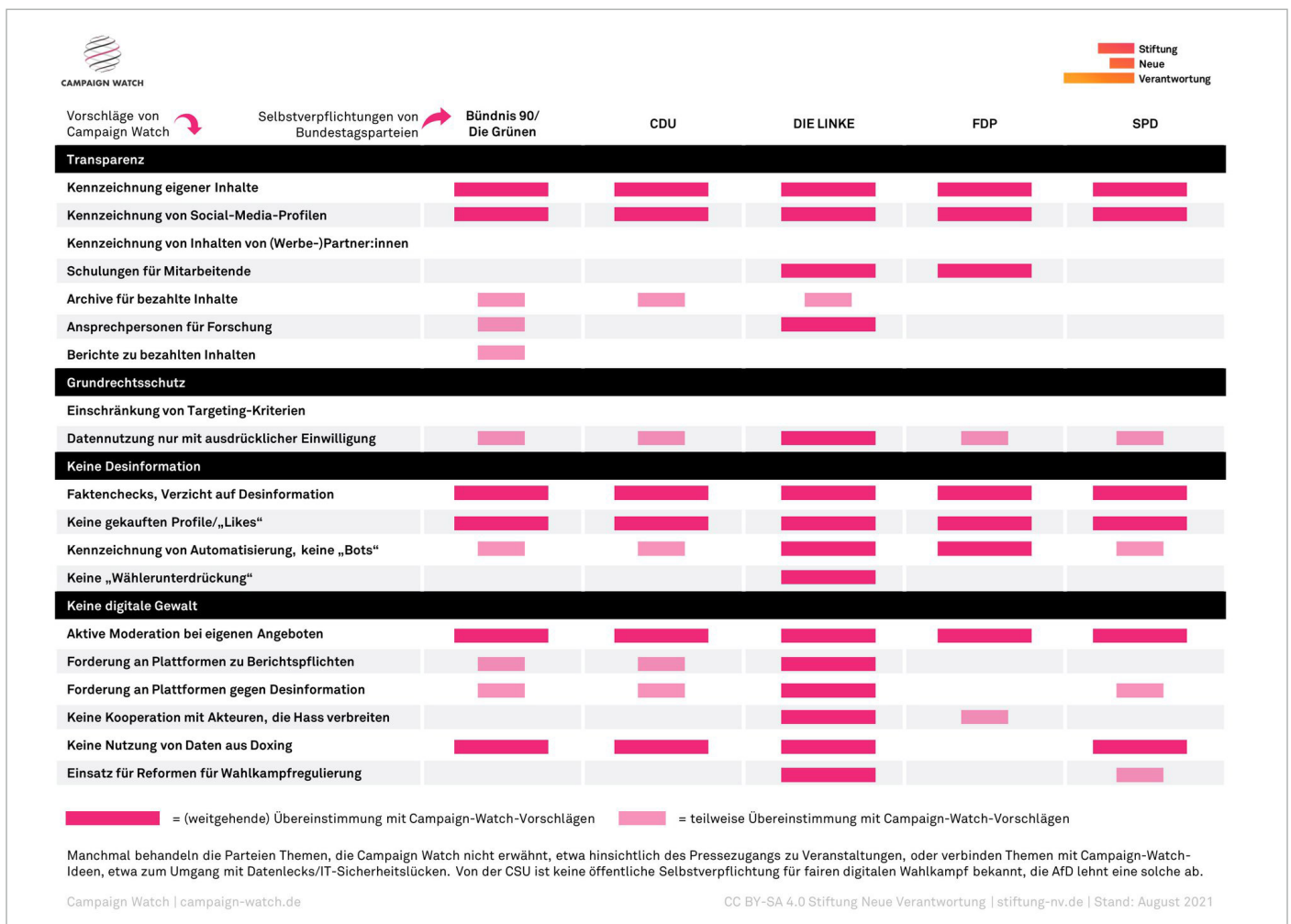
Informationsangebote zur Wahl prioritär behandelt. Auf Facebook News - ein gesonderter Bereich für journalistische Inhalte - wurde ein spezieller Themenbereich zur Bundestagswahl aufbereitet und Beiträge von Medienpartner:innen in gesonderten Kategorien (unter anderem „Die Wahlkampfthemen“) aufgegriffen. Zudem bestand am Wahltag und am Folgetag der gesamte obere Bereich der Facebook News Startseite aus kuratierten Artikeln zur Wahl.⁶³ Ob für dieses Angebot die Regeln für MI gelten, ist jedoch unwahrscheinlich - vielmehr handelt es sich bei Facebook News womöglich um eine medienrechtliche Qualifizierung als Medienplattform, welche zwar einen diskriminierungsfreien und chancengleichen Zugang für Inhalteanbieter gewährleisten müssen, nicht jedoch dem Kriterium der Diskriminierungsfreiheit unterliegen.⁶⁴ Anders sieht es hingegen in Fällen aus, in denen bestimmte Inhalte priorisiert im allgemeinen NewsFeed oder den Suchergebnissen dargestellt werden.

Dieser Befund zeigt, dass gewisse Spannungsverhältnisse unter den Regelwerken gegen digitale Bedrohungen im Netz existieren. Denn während der CoPD eine Priorisierung relevanter, authentischer und maßgeblicher Informationen festschreibt, beinhaltet der MStV Regeln zur Diskriminierungsfreiheit für die Angebote von Medienintermediären.

Inwiefern haben Parteien zu einem fairen digitalen Wahlkampf beigetragen?

Abseits der Regelwerke für Plattformanbieter gaben sämtliche im Bundestag vertretenen Parteien außer der AfD im Kontext der Bundestagswahl 2021 freiwillige Selbstverpflichtungen für den digitalen Wahlkampf ab. Vorausgegangen war die Initiative Campaign Watch, an der sich neben dem ISD Germany auch 21 weitere zivilgesellschaftliche Organisationen als Partner beteiligten. Um eine „Firewall der Demokratie“ zu bauen, legte die Initiative Anfang Juni 2021 einen Leitfaden für digitale Demokratie mit konkreten Handlungsempfehlungen vor und verschickte ihn an die politischen Parteien.⁶⁵ Dieser umfasste Bestimmungen - basierend auf den Ansichten und Ideen der beteiligten Partner - in folgenden Bereichen: Transparenz, Grundrechtsschutz, Desinformation und digitale Gewalt. Dazu gehörten etwa die Kennzeichnung eigener Inhalte, die Einschränkung von Targeting-Kriterien auf Ort und Alter in der Zielgruppenansprache und die Kennzeichnung von Inhalten, Profilen und Seiten, die automatisiert erstellt und betrieben wurden (im Einklang mit der Kennzeichnungspflicht für automatisiert erstellte Inhalte oder Mitteilungen im Sinne des MStV).

Eine Auswertung des Leibniz-Instituts für Medienforschung | Hans-Bredow-Institut (HBI) im August 2021 anhand eines Fragenkatalogs, der auf den Kriterien des Leitfadens für digitale Demokratie basiert, dokumentiert eine hohe Übereinstimmung der im Bundestag vertretenen Parteien - außer der AfD - mit den Zielen und Verpflichtungen des Leitfadens für digitale Demokratie. So ernannten alle Parteien Verantwortliche für den (digitalen) Wahlkampf. CDU, SPD, Bündnis 90/Die Grünen, FDP und die LINKE bekannten sich außerdem dazu, ihre Werbung und andere Inhalte als der Partei zugehörig zu kennzeichnen und Inhalte vor Verbreitung auf ihre Richtigkeit hin zu überprüfen. Auch gaben einige Parteien an, Mitarbeiter:innenschulungen durchzuführen. Die meisten Parteien verzichteten außerdem auf durch Doxing gewonnene Informationen. Dennoch wurde auch Nachbesserungsbedarf festgestellt. So veröffentlichte nur Bündnis 90/Die Grünen detaillierte Berichte über eingesetzte digitale Werbemaßnahmen. Andere Parteien verwiesen in diesem Zusammenhang lediglich auf die Werbebibliotheken von Plattformen. Eine Einschränkung der Targeting-Kriterien auf die Parameter „Ort“ und „Alter“ - eine zentrale Vorgabe von Campaign Watch - wurde zudem von keiner der befragten Parteien übernommen.⁶⁶



Die Abbildung der Stiftung Neue Verantwortung (SNV) visualisiert die Umsetzung pro Vorgabe des Leitfadens. Dabei wird zwischen (weitgehender) und teilweiser Übereinstimmung unterschieden.

Während die Auswertungen von HBI und SNV auf den Selbstverpflichtungen der Parteien und einer Befragung von ihnen basieren, wurde die Durchsetzung der Selbstverpflichtungen und das digitale Verhalten der Parteien in der Praxis bisher kaum analysiert.

Für **Twitter** veranschaulicht eine weitere Analyse des HBI interessanterweise, dass Tweets von Bundestagsabgeordneten einzelner im Bundestag vertretener Parteien überdurchschnittlich häufig verschwanden. Während der Tweetschwund der meisten Bundestagsfraktionen zwischen April und

September 2021 relativ konstant rund um den Durchschnitt aller Fraktionen liegt, bleibt die Quote der AfD meist deutlich oberhalb des Durchschnitts. Außerdem wurden die Accounts von AfD-Abgeordneten deutlich öfter privat gestellt als die der anderen. Für eine Untersuchung der Gründe würde es jedoch einer Inhaltsanalyse bedürfen, die bereits gelöschte Daten enthält - ein möglicher Verstoß gegen die Gemeinschaftsstandards von Twitter. Bisherige Studien (siehe Abschnitt „Falschinformationen“) bestätigen jedoch, dass insbesondere von AfD-zugehörigen Seiten Falschinformationen und Verschwörungserzählungen verbreitet wurden.⁶⁷ Damit liegt nahe, dass der Tweetschwund durch eine Umstellung auf „Privat“ mit der Verbreitung problematischer Inhalte in Verbindung stehen könnte. Weitere Analysen sind hier jedoch unerlässlich.

Fazit: Unzureichende und fehlerhafte Durchsetzung

Dieser Bericht konnte auf Basis einzelner kritischer Fallstudien Defizite bei der Durchsetzung von Regeln gegen digitale Bedrohungen im Netz im Kontext der Bundestagswahl 2021 identifizieren. **So konnten digitale Gewalt sowie Falsch- und Desinformationen trotz NetzDG, CoPD und Gemeinschaftsstandards nicht ausreichend eingedämmt werden.**

Rechtswidrige Inhalte wurden teilweise zu spät oder gar nicht entfernt oder trotz NetzDG-Beschwerden auf Grundlage der Gemeinschaftsstandards bearbeitet. Dies dürfte unter anderem Verzerrungen in den Transparenzberichten nach NetzDG zur Folge haben. Falschinformationen waren überdies im Kontext der Bundestagswahl 2021 weit verbreitet und vielfach trotz bereits erfolgter Faktenprüfung für die aufgegriffene Erzählung ohne Warnhinweis weiterhin ersichtlich. Diese Befunde legen außerdem den Verdacht nahe, dass automatisierte Systeme zur Erkennung problematischer Inhalte noch nicht so weit entwickelt sind, wie es die Anbieter öffentlich verkünden. Darüber hinaus trugen automatisierte Empfehlungssysteme dazu bei, dass Inhalte bestimmter politischer Akteur:innen - insbesondere AfD-zugehörige Seiten - im Feed bevorzugt wurden - ein Befund, der Aussagen der Anbieter zur Einschränkung der Verbreitung problematischer Inhalte widerspricht. Denn vor allem AfD-zugehörige Seiten säten Zweifel an dem Wahlprozess sowie Wahlergebnis und verbreiteten Verschwörungserzählungen. Zudem waren auch Info-Banner, die vertrauenswürdige Informationen prioritär

darstellen, mitunter mangelhaft umgesetzt. Gleichzeitig stellen diese möglicherweise ein Spannungsverhältnis zur im MStV vorgesehenen Diskriminierungsfreiheit dar. Überdies gab es Belege dafür, dass Facebooks Werbebibliothek unvollständig war. Bei politischen Anzeigen selbst verzichteten alle politische Parteien auf eine weitgehende Einschränkung des Microtargeting (der zielgerichteten Ansprache von Nutzer:innen basierend auf persönlichen Daten zu deren Onlineverhalten). Trotzdem bleibt positiv hervorzuheben, dass alle im Bundestag vertretenen Parteien bis auf die AfD erstmals Selbstverpflichtungen für den digitalen Wahlkampf abgaben.

Die in diesem Bericht aufgegriffenen Forschungsbefunde im Kontext der Bundestagswahl 2021 können eine Grundlage für die Weiterentwicklung der Regelwerke für Plattformanbieter und politische Parteien im In- und Ausland bieten. Trotzdem unterliegt der Bericht der Limitation, dass vorliegende Forschungsarbeiten große Online-Plattformen wie Facebook bevorzugt untersuchten und dabei andere Anbieter weniger berücksichtigt blieben. Der weiterhin begrenzte Zugang zu Plattformdaten - gerade im Bereich der Reichweite problematischer Inhalte - sowie die unterschiedliche Beschaffenheit der Plattformen erschwerte zudem eine systematische und vergleichende Analyse der Durchsetzung bestehender Regelwerke. Eine Gesamtevaluation bestehender staatlicher und selbstregulativer Regelwerke ist daher nicht möglich.

Politischer Ausblick

Im Dezember 2020 veröffentlichte die Europäische Kommission einen Vorschlag für ein Gesetz über digitale Dienste („Digital Services Act“, DSA), das künftig einheitliche horizontale Regeln für Anbieter von Vermittlungsdiensten, einschließlich der Online-Plattformen, einführen und einer Zersplitterung im europäischen Rechtsraum entgegenwirken soll.⁶⁸

Der Vorschlag der Kommission verpflichtet Online-Plattformen zu Melde- und Abhilfeverfahren (sog. „Notice-and-action-Verfahren“) für illegale Inhalte oder Inhalte, die gegen die Gemeinschaftsstandards verstoßen. Die vorgeschlagenen Haftungsregeln enthalten zudem Haftungsausschlüsse für Anbieter, die auf Eigeninitiative freiwillige Tätigkeiten zur Erkennung, Feststellung und Entfernung illegaler Inhalte durchführen (sog. „Good Samaritan Clause“). Zudem sieht der Vorschlag einen Beschwerde- und Rechtsbehelfsmechanismus, Vorgaben für „Trusted Flaggers“ (also akkreditierte Personen und Organisationen, die rechtswidrige Inhalte melden) sowie Transparenzpflichten hinsichtlich der Moderation von Inhalten vor. Transparenzberichte sollen wie beim NetzDG Daten über Beschwerden und Löschungen offenlegen. Dabei soll zwischen Meldungen auf gesetzlicher Grundlage oder aufgrund der Gemeinschaftsstandards unterschieden werden. Offengelegt werden soll auch, in welchen Fällen Anbieter eigenständig Inhalte gelöscht oder gesperrt haben. Im Bereich der Online-Werbung sollen ebenfalls Transparenzpflichten eingeführt werden. Nutzer:innen sollen Informationen darüber erhalten, anhand welcher grundlegenden Kriterien bestimmt wird, welche Werbung ihnen angezeigt wird.

Sehr große Online-Plattformen, deren Dienste von mehr als 45 Millionen Menschen in der Europäischen Union genutzt werden, sollen zusätzliche Sorgfaltspflichten auferlegt werden. Dazu gehören Erklärungen zu automatisierten Empfehlungssystemen und dazugehörige Einstellungsmöglichkeiten für Nutzer:innen, Transparenz-Datenbanken für Online-Werbung, jährliche Audits sowie Risikobewertungen hinsichtlich der möglichen Auswirkungen auf die Grundrechte der Bürger:innen. Die Plattformen sollen demnach die von ihren Systemen ausgehenden Risiken bewerten, um geeignete Risikomanagementinstrumente zu entwickeln, die die

Integrität ihrer Dienste vor manipulativen Techniken schützen. Zu möglichen Instrumenten gehören etwa die Beendigung von Werbeeinnahmen für bestimmte Inhalte, die Anpassung der algorithmischen Empfehlungssysteme oder der Online-Schnittstellen. Systemische Risiken umfassen hier Risiken, die durch die Verbreitung illegaler Inhalte entstehen (darunter illegale Hassrede), die Auswirkungen auf die Ausübung der Grundrechte haben (einschließlich der Freiheit der Meinungsäußerung), oder die absichtliche und oftmals koordinierte Manipulation der Plattform, welche absehbare Auswirkungen auf Gesundheit, den gesellschaftlichen Diskurs, Wahlprozesse, die öffentliche Sicherheit sowie den Schutz Minderjähriger bedeuten können.

Anfang 2022 sollen die Verhandlungen im Trilog-Verfahren zwischen Rat, Parlament und Kommission über den endgültigen Gesetzestext beginnen, mit dem Ziel, dass die Regelungen Ende 2022 umgesetzt werden. Mehrfach wurde bereits von deutscher Seite die Komplementarität des NetzDG und des DSA betont, wodurch nach Inkrafttreten des DSA schlicht ein Änderungs- und Anpassungsbedarf geprüft werden müsse.⁶⁹ Tatsächlich beinhaltet der DSA Regelungen, die sich teilweise mit denen des NetzDG und des MStV überschneiden, wodurch eine Implementierung in Deutschland ohne Gesetzesnovellierungen kaum möglich sein wird.

Der DSA steht im Einklang mit den Zielen des Europäischen Aktionsplan für Demokratie (EDAP)⁷⁰ zur Förderung freier und fairer Wahlen, Stärkung der Medienfreiheit sowie Bekämpfung von Desinformation. Im Rahmen des Aktionsplans hat die Kommission im November 2021 zusätzlich Rechtsvorschriften für politische Werbung, Wahlrecht und Parteienfinanzierung vorgelegt.⁷¹ Mit dem Vorschlag sollen unter anderem Techniken für Targeting und Amplifikation von politischer Werbung verboten werden, bei denen sensible personenbezogene Daten wie ethnische Herkunft, religiöse Überzeugungen oder sexuelle Orientierung verwendet oder abgeleitet werden. Solche Techniken sollen jedoch noch nach „ausdrücklicher Zustimmung“ der betroffenen Person zulässig sein. Erstmals soll es zudem strenge Transparenzanforderungen zu verwendeten Kriterien sowie Instrumenten oder Methoden beim Einsatz

politischer Werbung geben. Die neuen Vorschriften sollen bis zum Frühjahr 2023, das heißt ein Jahr vor den Wahlen zum Europäischen Parlament, in Kraft treten und von den Mitgliedstaaten bis dahin auch vollständig umgesetzt werden.

Ergänzend zu dem Vorschlag lenkt die Kommission die Bemühungen um eine Überarbeitung des bereits vorgestellten Verhaltenskodex zur Bekämpfung von Desinformation (CoPD), mit dem die Anforderungen an Online-Plattformen verschärft und strenge Überwachungs- und Aufsichtsvorschriften eingeführt werden sollen. Im Mai 2021 wurden die Leitlinien für die Stärkung des Verhaltenskodex für den Bereich der Desinformation herausgegeben⁷², um die in der Bewertung des Kodex durch die Kommission⁷³ festgestellten Unzulänglichkeiten anzugehen. Zu den gefundenen Mängeln gehören u.a. uneinheitliche und unvollständige Anwendung des Kodex auf verschiedenen Plattformen und in den Mitgliedstaaten, Lücken bei der Erfüllung der Verpflichtungen, sowie Fehlen eines geeigneten Überwachungsmechanismus. In den Leitlinien wird gefordert, die Kommerzialisierung von Desinformation einzuschränken und Maßnahmen gegen manipulative Techniken zu bekräftigen. Zudem ist eine transparentere Gestaltung von Empfehlungssystemen zur Begrenzung viraler Verbreitung von Desinformation vorgesehen. Die Unterzeichner:innen sollen dabei gewährleisten, dass den Nutzer:innen die ihnen angezeigten Inhalte transparent dargestellt werden. Mit dem Kodex soll auch die Zusammenarbeit mit Faktenprüfer:innen ausgeweitet werden und in allen EU-Ländern und -Sprachen verfügbar sein. Die Unterzeichner:innen sollen zudem ein Transparenzzentrum einrichten, in dessen Rahmen sie ihre zentralen Leistungsindikatoren, relevante Daten und Parameter angeben sollen. Ein überarbeiteter Kodex soll bis Ende März 2022 vorgelegt werden.

Neben den europäischen Initiativen existieren auf nationaler Ebene bereits einige gesetzliche Bestimmungen, die von den Anbietern sozialer Netzwerke ab Februar 2022 umzusetzen sind. So sieht das Gesetz zur Bekämpfung des Rechtsextremismus und der Hasskriminalität eine Verpflichtung zur Übermittlung bestimmter nach NetzDG gemeldeter Inhalte sowie Daten der inhalteverfassenden Person

an eine Zentralstelle beim Bundeskriminalamt (BKA) vor - eine Maßnahme, gegen die einige Anbieter derzeit gerichtlich vorgehen. Das Gesetz zur Änderung des Netzwerkdurchsetzungsgesetzes (NetzDG-ÄndG) führt ähnlich zum DSA außerdem eine Verpflichtung zu Auskünften für wissenschaftliche Forschung ein. Dazu gehören auch Informationen zur Wirkweise von Verfahren automatisierter Inhalte-Erkennung und zur Verbreitung von gemeldeten, entfernten oder gesperrten Inhalten.

Am 24. November 2021 veröffentlichten SPD, Bündnis 90/Die Grünen und FDP überdies den gemeinsamen Koalitionsvertrag „Mehr Fortschritt wagen. Bündnis für Freiheit, Gerechtigkeit und Nachhaltigkeit“, der in den Folgewochen von allen drei Parteien angenommen wurde. Der Vertrag skizziert die Vorhaben der neuen Regierungskoalition für die 20. Legislaturperiode von 2021 bis 2025, wobei diese nur relativ oberflächlich beschrieben werden. Damit unterliegt die konkrete Ausgestaltung der Vorhaben auch den verantwortlichen Bundesministerien.

Anders als vorherige Koalitionsverträge beinhaltet der vorgelegte Vertrag ambitionierte Vorhaben zum Umgang mit digitaler Gewalt und Falschinformationen. Dazu gehört u. a. ein Gesetz gegen digitale Gewalt, das rechtliche Hürden für Betroffene, wie Lücken bei Auskunftsrechten, abbauen und umfassende Beratungsangebote aufsetzen soll. Damit wollen die Koalitionsparteien rechtliche Rahmenbedingungen für elektronische Verfahren zur Anzeigenerstattung und für private Verfahren schaffen und richterlich angeordnete Accountsperrungen ermöglichen. Außerdem will sich die neue Bundesregierung beim DSA für die Wahrung der Kommunikationsfreiheiten, starke Nutzerrechte, klare Meldeverfahren, den Zugang zu Daten sehr großer Plattformen für Forschungszwecke, die Überprüfbarkeit ihrer algorithmischen Systeme und klare Regelungen gegen Desinformationen einsetzen. Zudem soll eine grundlegende Überarbeitung des nationalen Rechtsrahmens (u. a. Telemediengesetz und NetzDG) auf Basis der europäischen Vorgaben erfolgen - ein Bekenntnis, das mit Blick auf Überschneidungen zwischen DSA und NetzDG sehr sinnvoll erscheint. Darüber hinaus soll der Aufbau von Plattformräten vorangebracht werden, während allgemeine Überwachungspflichten, Maßnahmen zum Scannen

privater Kommunikation und eine Identifizierungspflicht abgelehnt werden und anonyme wie pseudonyme Online-Nutzung gewahrt bleiben sollen. Schließlich soll auch die Zivilgesellschaft besser in digitalpolitische Vorhaben eingebunden und unterstützt werden, insbesondere in den Bereichen Diversität und Civic Tech. Im außenpolitischen Bereich will die Regierungskoalition zudem die liberalen Demokratien Europas dazu befähigen, Desinformation, Fake-News, Kampagnen, Propaganda sowie Manipulationen aus dem In- und Ausland besser abwehren zu können.⁷⁴

Empfehlungen

Sowohl die aktuellen Initiativen auf europäischer Ebene wie der DSA und der EDAP oder die Überarbeitung des CoPD, als auch die nationalen Initiativen wie das NetzDG-ÄndG oder der Koalitionsvertrag der neuen Bundesregierung sehen Maßnahmen vor, die die in diesem Bericht identifizierten Regelungslücken in der Bekämpfung von digitaler Gewalt und Falschinformationen teilweise schließen könnten. Dazu gehören insbesondere eine Verpflichtung der Anbieter sehr großer Online-Plattformen, Forscher:innen Zugang zu Daten für Forschungszwecke einzuräumen sowie geeignete Maßnahmen zur Minderung sog. systemischer Risiken zu entwickeln. Außerdem sind Regelungen zur Vorlage von Transparenzberichten, in denen Angaben zu Beschwerden und Löschungen auf gesetzlicher Grundlage und auf Basis der Gemeinschaftsstandards gemacht werden müssen, vielversprechend.

Letztlich wird es bei den Vorschlägen jedoch vor allem auf die Details in der Ausgestaltung sowie auf eine effektivere Durchsetzung ankommen. Folgende Empfehlungen an Bundesregierung und Parteien ergeben sich hieraus:

Sorgfaltspflichten für Desinformationen und Grundrechtsverletzungen konkretisieren

Wenn sich illegale Inhalte, Falsch- und Desinformationen bereits millionenfach verbreitet haben, ist der Schaden für die Integrität unserer Demokratien kaum noch abzuwenden. Die Politik muss daher europäisch abgestimmte Verpflichtungen für Anbieter sehr großer bzw. einflussreicher Plattformen erlassen, um die von ihren Systemen ausgehenden Risiken wirksam zu minimieren. Konkret sollten Anbieter in die Verantwortung genommen werden, neben der Verbreitung illegaler Inhalte auch Falsch- und Desinformationen sowie andere digitale Bedrohungen, insbesondere vor dem Hintergrund geschlechtsspezifischer und intersektionaler Diskriminierung, durch wirksame und klar definierte Risikominderungsmaßnahmen einzuschränken. Zu den möglichen Maßnahmen gehören speziell die Offenlegung und gegebenenfalls die Anpassung der Prozesse der Moderation von Inhalten sowie algorithmischer Empfehlungssysteme. Die Entscheidung über Anpassungen sollte dabei

von der Plattform ausführlich begründet werden müssen, damit der Einsatz wirksamer Maßnahmen gewährleistet ist. Anbieter sollten außerdem verpflichtet werden, ihre Gemeinschaftsstandards im Rahmen strenger Transparenzanforderungen sorgfältig, objektiv und effektiv umzusetzen. Kontextsensible Regulierungsansätze sollten die schwierige Abwägung zwischen der Meinungsfreiheit und anderen Grundrechten vornehmen.

„Dunkle“ soziale Medien in die Pflicht nehmen, Vorbildunternehmen einbinden

Illegale und problematische Inhalte und Aktivitäten finden sich vermehrt auf kleineren, aber in einigen Ländern besonders einflussreichen Plattformen, einschließlich Plattformen wie Telegram - auch „dunkle“ soziale Medien unter Wissenschaftler:innen genannt. Diese sollten insbesondere im Wahlkontext den weitreichenden Sorgfaltspflichten wie zum Beispiel Risikominderungsmaßnahmen und Datenzugang für Forschungszwecke auf Erlass der Regulierungsbehörden hin unterworfen werden können. Kleine und mittlere Anbieter, deren Systeme voraussichtlich keine großen Auswirkungen auf den demokratischen Diskurs haben, sollten hiervon nicht betroffen sein. Außerdem sollten solche Anbieter, die bereits durch funktionierende Maßnahmen insbesondere im Bereich der Moderation von Inhalten auf sich aufmerksam gemacht haben, in die Diskussion über Sorgfaltspflichten für kleine und mittlere Plattformen verstärkt einbezogen werden. Dadurch könnten perspektivisch Selbstverpflichtungen für „alternative“ Plattformen aufgestellt werden, die über den gesetzlichen Rahmen hinausgehen.

Breiten Datenzugang für Journalismus, Wissenschaft, Zivilgesellschaft ermöglichen

Plattformanbieter sollten rechtlich dazu verpflichtet werden, eine technische Schnittstelle zur Bereitstellung von interoperablen, maschinenlesbaren und lizenzfreien Daten für vertrauenswürdige Forscher:innen aus Journalismus, Wissenschaft und Zivilgesellschaft zur Untersuchung systemischer Risiken einzurichten. Dabei muss klargestellt werden, dass auch Kommentare unter Beiträgen unter Einhaltung des Datenschutzes über die Schnittstelle zugänglich sind.

Gleichzeitig muss der Zugriff auf die Daten für nicht vertrauenswürdige Akteure gesperrt bleiben, um einen Missbrauch etwa für Desinformationskampagnen zu verhindern.

Empfehlungssysteme transparent und verbraucherfreundlich gestalten

Automatisierte Empfehlungssysteme, die Inhalte auf Plattformen sortieren, tragen immer wieder zur massenhaften Verbreitung problematischer Inhalte bei. Um diesen Umstand angemessen beurteilen zu können, müssen die Kriterien für die Funktionsweise der Algorithmen offengelegt werden - sowohl durch ungehinderte Einblicke für Forscher:innen als auch durch klar verständliche Informationsangebote für Nutzer:innen und die Kennzeichnung von Beiträgen. Nutzer:innen sollten standardmäßig außerdem die Möglichkeit haben, die wichtigsten von den Empfehlungssystemen verwendeten Parameter zu ändern oder zu beeinflussen. Zu den Parametern, die mindestens offengelegt werden sollten, gehören die Verwendung personenbezogener Daten sowie die Verfolgung von Verhaltensdaten. Empfehlungssysteme sollten überdies nicht auf Profiling beruhen und andere, sensible Datenkategorien anwenden. Nutzer:innen sollten solche auf Profiling beruhende Systeme stattdessen aktiv auswählen.

Parteiübergreifende Selbstverpflichtung vorlegen und Ressourcen ausweiten

Politische Parteien sollten ihre bereits beschlossenen Leitlinien für einen fairen digitalen Wahlkampf auf ihre Online-Aktivitäten außerhalb des Wahlkampfes ausweiten, effektiv durchsetzen und Verstöße sanktionieren. Dazu gehören die Evaluierung der Anwendung ihrer Leitlinien im Kontext der Bundestagswahl 2021 sowie der Erfahrungsaustausch mit anderen politischen Parteien und der Zivilgesellschaft. Darüber hinaus sollten die Parteien eine parteiübergreifende Selbstverpflichtung auf Basis des Leitfadens für digitale Demokratie der Initiative CampaignWatch anstreben. Sofern sich dies als nicht realisierbar erweist, sollten Reformen des Parteienrechts in Erwägung gezogen werden. Europäische Initiativen auf diesem Gebiet sollten unterstützt werden. Um Desinformationskampagnen

entgegenzuwirken, sollten die Parteien außerdem ihre Medienbeobachtung- und -analyse ausbauen, klare Bewertungsschemata einführen und Prozesse für eine schnelle Entscheidungsfindung definieren.

Datenschutzrechte bei Online-Werbung sichern

Die Plattformen sollten ihre Online-Schnittstellen für Werbung so gestalten und organisieren, dass Nutzer:innen ihre Datenschutzrechte gemäß der Datenschutzgrundverordnung (DSGVO) für jede einzelne angezeigte Anzeige einfach und effizient ausüben können. Darüber hinaus sollte es eine Kennzeichnungspflicht für alle Formen bezahlter politischer Werbung und ein Verbot der Verwendung sensibler personenbezogener Daten - mindestens im Fall einer fehlenden ausdrücklichen Zustimmung der Betroffenen - geben. Eine ambitionierte Ausgestaltung vorliegender Initiativen auf europäischer Ebene sollte unterstützt werden.

Endnoten

- 1 Unter Plattformen fasst das ISD Internetseiten oder Apps, über die Nutzer:innen miteinander in Austausch treten können. Besonders relevant sind hierbei die sozialen Medien oder sozialen Netzwerke sowie Messenger-Dienste.
- 2 Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (Netzwerkdurchsetzungsgesetz - NetzDG), https://www.bmju.de/SharedDocs/Gesetzgebungsverfahren/Dokumente/BGBl_NetzDG.pdf;jsessionid=99BEA8F221B4F1B3F85CDCECAAC05D73.1_cid297?__blob=publicationFile&v=2
- 3 Telemediendiensteanbieter, die mit Gewinnerzielungsabsicht Plattformen im Internet betreiben, die dazu bestimmt sind, dass Nutzer:innen beliebige Inhalte mit anderen Nutzern:innen teilen oder der Öffentlichkeit zugänglich machen (soziale Netzwerke). Plattformen mit journalistisch-redaktionell gestalteten Angeboten, die vom Diensteanbieter selbst verantwortet werden, gelten nicht als soziale Netzwerke im Sinne des NetzDG.
- 4 Eine qualitative Auswahl basierte auf einer Seed-Liste mit 279 relevanten Kanälen und Gruppen.
- 5 Smirnova, J., & Winter, H. (2021). An Overview of the Electoral Fraud Narratives Used to Undermine the 2021 German Federal Election. Institute for Strategic Dialogue. https://www.isdglobal.org/digital_dispatches/an-overview-of-the-electoral-fraud-narratives-used-to-undermine-the-2021-german-federal-election/
- 6 KI-Auswertung, Erwähnungen im Kontext von Sprache, die gemäß Perspective API mit einer Wahrscheinlichkeit von mindestens 85% beleidigend ist.
- 7 HateAid. (2021a). #1 Hass als Berufsrisiko – Digitale Gewalt im Wahlkampf. HateAid. <https://hateaid.org/wp-content/uploads/2021/09/Hass-als-Berufsrisiko-Digitale-Gewalt-im-Wahlkampf.pdf>
- 8 HateAid. (2021b). #2 Hass als Berufsrisiko – Digitale Gewalt im Wahlkampf. HateAid. <https://hateaid.org/wp-content/uploads/2021/09/Report-2-Hass-als-Berufsrisiko.pdf>
- 9 Baumann, S. (2021). Löschen nach Gutsherrenart. Süddeutsche Zeitung. <https://www.sueddeutsche.de/projekte/artikel/digital/hass-im-netz-wie-facebook-mit-hasskommentaren-umgeht-e994275/>
- 10 HateAid und Reset. (2021). Facebook versagt beim Schutz der Bundestagswahl. HateAid. https://hateaid.org/wp-content/uploads/2021/09/Reset_Facebook_Bundestagswahl_DE.pdf
- 11 Bundesministerium der Justiz und für Verbraucherschutz. (2020). https://www.bmju.de/SharedDocs/Downloads/DE/News/PM/090920_Evaluierungsbericht_NetzDG.pdf;jsessionid=65500B5EE1E0E4A6417F34CB5E94B1F0.1_cid289?__blob=publicationFile&v=3. BMJV
- 12 Facebook. (2021a). NetzDG Transparenzbericht. Facebook. <https://about.fb.com/de/wp-content/uploads/sites/10/2021/07/Facebook-NetzDG-Transparenzbericht-Juli-2021.pdf>
- 13 Krempf, S. (2021). NetzDG: Beschwerden bei Facebook um 1844 Prozent gestiegen. Heise online. <https://www.heise.de/news/NetzDG-Beschwerden-bei-Facebook-um-1844-Prozent-gestiegen-6152292.html>
- 14 Bundesgerichtshof, BGH. Urteile vom 29. Juli 2021 – III ZR 179/20 und III ZR 192/20. <https://juris.bundesgerichtshof.de/cgi-bin/rechtsprechung/document.py?Gericht=bgh&Art=en&nr=121741&pos=0&anz=1>
- 15 Facebook. (2021b). Hassrede. Hassrede | Transparency Center. <https://about.fb.com/de/wp-content/uploads/sites/10/2021/07/Facebook-NetzDG-Transparenzbericht-Juli-2021.pdf>
- 16 Baumann, 2021
- 17 HateAid & Reset, 2021
- 18 KI-Auswertung, Erwähnungen der Kandidat:innen in Tweets, die gemäß Perspective API mit einer Wahrscheinlichkeit von mindestens 85% Beleidigungen und identitätsbezogene Angriffe enthielten.
- 19 HateAid und Institute for Strategic Dialogue (2022). Hass als Berufsrisiko: Das ganze Ausmaß digitaler Gewalt im Bundestagswahlkampf - Eine gemeinsame Investigation von HateAid und ISD. Unveröffentlichter Forschungsbericht, Institute for Strategic Dialogue.
- 20 Lima, C. (2021). A whistleblower's power: Key takeaways from the Facebook Papers. <https://www.washingtonpost.com/technology/2021/10/25/what-are-the-facebook-papers/>
- 21 YouTube. (2021). YouTube schafft mehr Transparenz. YouTube Official Blog. <https://blog.youtube/intl/de-de/news-and-events/youtube-schafft-mehr-transparenz/>
- 22 Facebook. (2021c). Unsere Vorbereitungen auf die Wahlen in Deutschland. Newsroom. <https://about.fb.com/de/news/2021/05/unsere-vorbereitungen-auf-die-wahlen-in-deutschland/#entfernen-unzulassiger-inhalte>
- 23 Seetharaman, D., Horwitz, J., & Scheck, J. (2021). Facebook Says AI Will Clean Up the Platform. Its Own Engineers Have Doubts. Wall Street Journal. <https://www.wsj.com/articles/facebook-ai-enforce-rules-engineers-doubtful-artificial-intelligence-11634338184>
- 24 Pamment, J. (2020). Future Threats, Future Solutions | #2The EU's Role in Fighting Disinformation: Crafting A Disinformation Framework. Carnegie Endowment. <https://carnegieendowment.org/2020/09/24/eu-s-role-in-fighting-disinformation-crafting-disinformation-framework-pub-82720>

- 25 Europäische Kommission. (2021b). Code of Practice on Disinformation. Policies. <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation>
- 26 Colliver, C. (2020). 2019 EU Elections. Cracking the Code: An Evaluation of the EU Code of Practice on Disinformation. Institute for Strategic Dialogue. https://www.isdglobal.org/wp-content/uploads/2020/06/isd_Cracking-the-Code.pdf
- 27 Edelson, L., & McCoy, D. (2021). We Research Misinformation on Facebook. It Just Disabled Our Accounts. The New York Times. <https://www.nytimes.com/2021/08/10/opinion/facebook-misinformation.html>
- 28 Kayser, N. (2021). Nach Drohungen von Facebook: AlgorithmWatch sieht sich gezwungen, Instagram-Forschungsprojekt einzustellen. AlgorithmWatch. <https://algorithmwatch.org/de/instagram-forschung-von-facebook-gestoppt/>
- 29 ZDF Magazin Royale. (2021). Lost Ads Library. Target Leaks. <https://targetleaks.de/lost-ads-library>
- 30 Echtermann, A. (2021). Warum Politiker mit Desinformation auf Facebook werben können. Correctiv. <https://correctiv.org/faktencheck/hintergrund/2021/05/19/facebook-verbietet-falschinformationen-in-werbeanzeigen-warum-ist-die-plattform-dann-voll-damit/>
- 31 Bösch, M., & Ricks, B. (2021). Leere Versprechungen: TikTok und die Bundestagswahl. Mozilla. https://assets.mofoprod.net/network/documents/TikTok_und_die_Bundestagswahl.pdf
- 32 Bösch, M., & Ricks, B. (2021)
- 33 Facebook, 2021c
- 34 Twitter. (2021). Twitter Richtlinie zur Integrität staatsbürgerlicher Prozesse und zu Wahlbetrug. Twitter Help Center. <https://help.twitter.com/de/rules-and-policies/election-integrity-policy>
- 35 Frank, S. (2021). YouTube – wie wir die Bundestagswahl 2021 begleiten. YouTube Official Blog. <https://blog.youtube/intl/de-de/news-and-events/youtube-wie-wir-die-bundestagswahl-2021-begleiten/>
- 36 TikTok. (2020). Community-Richtlinien. TikTok. Integrität und Authentizität. <https://www.tiktok.com/community-guidelines?lang=de#37>
- 37 Facebook. (2021d). Gemeinsam gegen Falschinformationen: Facebook, Instagram und WhatsApp starten Initiativen mit starken Partnern im Vorfeld der Bundestagswahl. Meta. <https://about.fb.com/de/news/2021/09/gemeinsam-gegen-falschinformationen-facebook-instagram-und-whatsapp-starten-initiativen-mit-starken-partnern-im-vorfeld-der-bundestagswahl/>
- 38 Bösch, M., & Ricks, B. (2021). Leere Versprechungen: TikTok und die Bundestagswahl. Mozilla. https://assets.mofoprod.net/network/documents/TikTok_und_die_Bundestagswahl.pdf
- 39 Eine qualitative Auswahl basierte auf einer Seed-Liste mit 279 relevanten Kanälen und Gruppen.
- 40 Smirnova et al., 2021
- 41 CeMAS – Center für Monitoring, Analyse und Strategie. (2021). Die Bundestagswahl 2021 Welche Rolle Verschwörungsideologien in der Demokratie spielen. CeMAS. <https://cemas.io/publikationen/die-bundestagswahl-2021-welche-rolle-verschwörungsideologien-in-der-demokratie-spielen/die-bundestagswahl-2021-welche-rolle-verschwörungsideologien-in-der-demokratie-spielen.pdf>
- 42 Smirnova et al., 2021
- 43 Kübler, J., Sekwenz, M.-T., Rachinger, F., König, A., Gsenger, R., Pírková, E., Wagner, B., Kettemann, M. C., Krennerich, M., & Ferro, C. (2021). The 2021 German Federal Election on Social Media: An Analysis of Systemic Electoral Risks Created by Twitter and Facebook Based on the Proposed EU Digital Services Act. Sustainable Computing Lab und Wirtschaftsuniversität Wien. https://www.sustainablecomputing.eu/wp-content/uploads/2021/10/DE_Elections_Report_Final_17.pdf
- 44 Kübler, J., Sekwenz, M.-T., Rachinger, F., König, A., Gsenger, R., Pírková, E., Wagner, B., Kettemann, M. C., Krennerich, M., & Ferro, C. (2021)
- 45 Mathelemuse, N., & Smirnova, J. (2021). Desinformationskampagnen gegen die Wahl: Befunde aus Sachsen-Anhalt. Institute for Strategic Dialogue. <https://www.isdglobal.org/wp-content/uploads/2021/06/Bericht-Landtagswahlen-Sachsen-Anhalt-Final.pdf>
- 46 Avaaz. (2021). Deutschlands Desinformations-Dilemma 2021. Avaaz. https://avaazimages.avaaz.org/bundestagswahl_2021_final.pdf
- 47 Gallagher, A., Hart, M., & O'Connor, C. (2021). Ill Advice: A Case Study in Facebook's Failure to Tackle COVID-19 Disinformation. Institute for Strategic Dialogue. https://www.isdglobal.org/wp-content/uploads/2021/10/Ill-Advice_v3.pdf
- 48 Zweig, K. A., Deussen, O., & Krafft, T. D. (2017). Algorithmen und Meinungsbildung. Eine grundlegende Einführung. Informatik-Spektrum, 40(4), 318-326. <https://doi.org/10.1007/s00287-017-1050-5>
- 49 Europäische Kommission. (2021a). Notification 2021/159/D. Netzpolitik. <https://cdn.netzpolitik.org/wp-upload/2021/07/Statutes-of-the-State-media-authorities-on-the-Regulation-of-media.pdf>
- 50 Facebook. (2021e). Verringerung der Verbreitung problematischer Inhalte. Transparency Center. <https://transparency.fb.com/de-de/enforcement/taking-action/lowering-distribution-of-problematic-content/>
- 51 Frank, 2021
- 52 Facebook, 2021e

- 53 Rosen, G., & Lyons, T. (2019). Remove, Reduce, Inform: New Steps to Manage Problematic Content | Meta. Meta. <https://about.fb.com/news/2019/04/remove-reduce-inform-new-steps/#reduce>
- 54 Waller, A., & Lecher, C. (2021). Germany's Far-Right Political Party, the AfD, Is Dominating Facebook This Election – The Markup. <https://themarkup.org/citizen-browser/2021/09/22/germanys-far-right-political-party-the-afd-is-dominating-facebook-this-election>
- 55 HOPE NOT HATE. (2021). Alternative for Germany and far-right networks on Telegram and Facebook. HOPE NOT HATE. <https://hopenothate.org.uk/wp-content/uploads/2021/07/AfD-and-Far-Right-Networks-on-Telegram-and-Facebook.pdf>
- 56 Smirnova, J., Winter, H., Mathelemuse, N., Dorn, M., & Schwertheim, H. (2021). Digitale Gewalt und Desinformation gegen Spitzenkandidat:innen vor der Bundestagswahl 2021. ISD Germany. https://www.isdglobal.org/wp-content/uploads/2021/09/Digitale-Gewalt-und-Desinformation_v5.pdf
- 57 Ebitsch, S., Hametner, M., Hunger, F., Kruse, B., Unterhitzberger, S., Weinmann, L., & Witzemberger, B. (2021). Wahlkampf: Wie der Instagram-Algorithmus wirkt. Süddeutsche Zeitung. <https://www.sueddeutsche.de/projekte/artikel/politik/wahlkampf-wie-der-instagram-algorithmus-wirkt-e545670/>
- 58 Peretti, J. (2021). Facebook Tried to Make Its Platform a Healthier Place. It Got Angrier Instead. Wall Street Journal. <https://www.wsj.com/articles/facebook-algorithm-change-zuckerberg-11631654215>
- 59 Gupta, A. (2021). Reducing Political Content in News Feed. Meta. <https://about.fb.com/news/2021/02/reducing-political-content-in-news-feed/>
- 60 Staatsvertrag zur Modernisierung der Medienordnung in Deutschland (Medienstaatsvertrag - MStV), <https://www.rlp.de/fileadmin/rlp-stk/pdf-Dateien/Medienpolitik/Medienstaatsvertrag.pdf>.
- 61 Europäische Kommission. (2021b). Code of Practice on Disinformation. Policies. <https://digital-strategy.ec.europa.eu/en/policies/code-practice-disinformation>
- 62 Kommission für Zulassung und Aufsicht (ZAK). (2021a). Neue Vorschriften zur Diskriminierungsfreiheit: ZAK entscheidet die ersten Fälle. Die Medienanstalten. <https://www.die-medienanstalten.de/service/pressemitteilungen/meldung/neue-vorschriften-zur-diskriminierungsfreiheit-zak-entscheidet-die-ersten-faelle>
- 63 Facebook, 2021c
- 64 Kommission für Zulassung und Aufsicht (ZAK). (2021b). ZAK qualifiziert Google News Showcase als Medienplattform und Benutzeroberfläche. Die Medienanstalten. https://www.die-medienanstalten.de/service/pressemitteilungen/meldung?tx_news_pi1%5Bnews%5D=4986&cHash=0f76c08f30a3eeb7de5b6ff126917e04
- 65 Campaign Watch. (2021). Leitfaden für Digitale Demokratie. Campaign Watch. <https://campaign-watch.de/>
- 66 Barthelmes, M., Gold, V., Hinrichs, L., Hoffmann, V., & Kettemann, M. C. (2021). Wahl-Watching. Hans-Bredow-Institut. <https://leibniz-hbi.de/de/blog/wahl-watching>
- 67 Münch, F. V. (2021). Hans-Bredow-Institut für Medienforschung. Leibniz Institut für Medienforschung | Hans-Bredow-Institut. <https://leibniz-hbi.de/de/blog/es-war-einmal-ein-tweet-wie-transparent-soll-wahlkampf-sein>
- 68 Europäische Kommission. (2020a). Vorschlag für eine Verordnung des europäischen Parlaments und des Rates über einen Binnenmarkt für digitale Dienste (Gesetz über digitale Dienste) und zur Änderung der Richtlinie 2000/31/EG. 15 Dezember 2020. IMMC. <https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:52020PC0825&from=de>
- 69 Noyan, O. (2021). Deutschland fordert Nachschärfungen beim Digital Services Act. EURACTIV.de. <https://www.euractiv.de/section/innovation/news/deutschland-fordert-nachschaerfungen-beim-digital-services-act/>
- 70 Europäische Kommission. (2020b.). Mitteilung - Europäischer Aktionsplan für Demokratie. IMMC. <https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:52020DC0790&from=EN>
- 71 Europäische Kommission. (2020c). Assessment of the Code of Practice on Disinformation – Achievements and areas for further improvement. Europäische Kommission. <https://digital-strategy.ec.europa.eu/en/library/assessment-code-practice-disinformation-achievements-and-areas-further-improvement>
- 72 Europäische Kommission. (2021c). Kommission legt Leitlinien zur Stärkung des Verhaltenskodex für den Bereich der Desinformation vor. Presseraum. https://ec.europa.eu/commission/presscorner/detail/de/ip_21_2585
- 73 Europäische Kommission. (2021d). Demokratie in Europa: Kommission legt neue Vorschriften für politische Werbung, Wahlrecht und Parteienfinanzierung fest. Presseraum. https://ec.europa.eu/commission/presscorner/detail/de/ip_21_6118
- 74 SPD, Bündnis 90/Die Grünen und FDP. (2021). Mehr Fortschritt wagen. Bündnis für Freiheit, Gerechtigkeit und Nachhaltigkeit. SPD.DE, GRUENE.DE & FDP.DE. https://www.spd.de/fileadmin/Dokumente/Koalitionsvertrag/Koalitionsvertrag_2021-2025.pdf



Beirut | Berlin | London | Paris | Washington DC

Copyright © Institute for Strategic Dialogue (2022).
Das Institute for Strategic Dialogue (gGmbH) ist beim
Amtsgericht Berlin-Charlottenburg registriert (HRB 207 328B).
Die Geschäftsführerin ist Huberta von Voss. Die Anschrift lautet:
Postfach 80647, 10006 Berlin. Alle Rechte vorbehalten.

www.isdglobal.org