DYNAMO

Policy Paper

Desinformation in Messenger-Diensten

Aktuelle Herausforderungen & Handlungsempfehlungen für rechtliche und gesellschaftliche Maßnahmen

Autor:innen

Tahireh Panahi, Carolin Jansen, Amancay Ancina, Katarina Bader, Jeong-Eun Choi, Gerrit Hornung, Nicole Krämer, Lars Rinsdorf, Karla Schäfer, Inna Vogel, York Yannikos und Martin Steinebach

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

Impressum

Autor*innen:
Tahireh Panahi, Carolin Jansen, Amancay Ancina, Katarina Bader, Jeong-Eun Choi, Gerrit Hornung, Nicole Krämer, Lars Rinsdorf, Karla Schäfer, Inna Vogel, York Yannikos und Martin Steinebach

https://www.dynamo.sit.fraunhofer.de/

Kontakt/Projektkoordinator:
Prof. Dr. Martin Steinebach

Telefon: +49 6151 869-349
E-Mail: martin.steinebach@sit.fraunhofer.de

Fraunhofer SIT
Nationales Forschungszentrum für angewandte Cybersicherheit ATHENE
Rheinstraße 75
64295 Darmstadt

Details zur Publikation:
DOI:10.17185/duepublico/82406

Veröffentlichende Institution: Universität Duisburg-Essen, Universitätsbibliothek, DuEPublico, Universitätsstraße 9-11, 45141 Essen, https://duepublico2.uni-due.de

1. Auflage, September 2024

Zusammenfassung

Die Verbreitung von Desinformation in digitalen Medien stellt eine wachsende Herausforderung und ernsthafte Bedrohung für den gesellschaftlichen Frieden und demokratische Prozesse dar. Insbesondere Messenger-Dienste wie Telegram und WhatsApp haben sich zu Plattformen entwickelt, auf denen falsche Informationen massenhaft verbreitet werden. Bislang mangelt es an effektiven Gegenstrategien, die auf die komplexen Dynamiken der Desinformationsverbreitung auf Messenger-Diensten zugeschnitten sind. In diesem Policy Paper präsentieren Forschende der Informatik, Rechtswissenschaften, Psychologie und Journalismus (gefördert durch das BMBF-Projekt DYNAMO) gemeinsame Handlungsempfehlungen und zeigen weiteren Forschungsbedarf auf.

Zunächst wird beschrieben, wie Messenger-Dienste zur Bildung von verfassungsfeindlichen und staatsskeptischen Gegenöffentlichkeiten genutzt werden. Bei der Analyse des aktuellen Rechtsrahmens wird festgestellt, dass bestehende Rechtsakte, wie der Digital Services Act (DSA) der EU, die spezifischen Herausforderungen von Desinformation in Messenger-Diensten nicht angemessen adressieren. Es besteht daher ein deutlicher Bedarf an ergänzenden Maßnahmen, um die Verbreitung von Desinformation in Messenger-Diensten effektiv und grundrechtskonform einzudämmen. Das Policy Paper macht konkrete Vorschläge zur Ergänzung des DSA, die auf die besondere Situation von Messenger-Diensten zugeschnitten sind.

Im Zentrum des Policy Papers steht die Analyse und Bewertung eines Ansatzes zur Prävention der Desinformationsverbreitung in Messenger-Diensten, das sog. Prebunking. Im Vergleich zu herkömmlichen Faktenchecks setzt Prebunking vor der eigentlichen Verbreitung von Desinformation an, indem entweder eine spezifische Aufklärung über einzelne Desinformationsinhalte erfolgt (Narrow-Spectrum) oder die generelle Medienkompetenz geschult wird (Broad-Spectrum). Unter Berücksichtigung aktueller psychologischer und kommunikationswissenschaftlicher Forschungsergebnisse wird das Prebunking kritisch hinsichtlich seiner technologischen und rechtlichen Umsetzbarkeit bewertet. Auch wenn Prebunking insgesamt als gut umsetzbare Strategie betrachtet werden kann und breitgefasste Ansätze eine technisch einfach umsetzbare sowie grundrechtsschonende Möglichkeit der Regulierung von Messenger-Diensten darstellen, deuten psychologische Studien darauf hin, dass inhaltlich spezifische Ansätze wirksamer sein könnten. Es bedarf daher weiterer Untersuchungen zur empirischen Wirksamkeit und grundrechtskonformen und praktischen Anwendbarkeit.

1 Einleitung

In der digitalen Welt ist seit einigen Jahren eine Zunahme der Verbreitung von Desinformation zu beobachten. Unter Desinformation werden unwahre Tatsachenbehauptungen verstanden, die mit einer schädigenden Absicht verbreitet werden. Nicht selten zielen sie darauf ab, Misstrauen und Hass gegen Personengruppen, Organisationen und Staaten zu schüren und damit gesellschaftlichen Unfrieden zu erzeugen. Daher werden effektive Gegenstrategien immer dringlicher. Dabei reicht es nicht, nur öffentliche soziale Netzwerke in die Verantwortung zu nehmen. Längst versammeln sich Desinformationsakteur*innen und ihre Anhängerschaft in Messenger-Diensten wie Telegram oder WhatsApp. Durch ihre öffentlichen und privaten (Gruppen-)Chats und Kanäle bieten Messenger-Dienste eine Möglichkeit, große staatsskeptische Gegenöffentlichkeiten zu bilden und sich auszutauschen.

Das Projekt DYNAMO setzt hier an. Wir untersuchen, wie Messenger-Dienste zur Verbreitung von Desinformation genutzt werden, wie dort Desinformation aufgespürt werden kann und welche Ansätze es gibt, Desinformation entgegenzuwirken. Dabei betrachten wir das Problem aus unterschiedlichen wissenschaftlichen Perspektiven. Die Expertisen aus den Bereichen Informatik, Recht, Psychologie und Journalismus fließen in das Projekt ein und werden dort kombiniert.

In diesem Policy Paper werden auf Basis unserer Analysen und Studien Erkenntnisse zu den Verbreitungswegen von Desinformation, Strategien ihrer Akteur*innen, der Rolle von Emotionen sowie mögliche Bekämpfungsmaßnahmen vorgestellt. Aufbauend auf diesen Befunden zeigen wir auf, welche Herausforderungen bei der gesetzlichen Regulierung bestehen und wie wichtig und zugleich schwierig die Differenzierung zwischen öffentlicher und privater Kommunikation in Messenger-Diensten ist. Dabei beziehen wir uns vor allem auf den Digital Services Act (DSA), einer EU-Verordnung, die die Risiken durch Desinformationsverbreitung eindämmen soll. Wir stellen dar, warum diese Verordnung gegen Desinformation in den hybriden Kommunikationsstrukturen von Messenger-Diensten bisher nur unzureichend wirken kann. Im Anschluss formulieren wir Vorschläge für ergänzende rechtliche Maßnahmen und schlagen weitere Bekämpfungsstrategien vor. Den Schwerpunkt legen wir auf das sogenannte Prebunking.[1] Darunter werden präventive Maßnahmen verstanden, die der Desinformationsverbreitung vorgelagert sind und entweder spezifische Aufklärung über einzelne Desinformationstypen bieten oder die generelle Medienkompetenz schulen. Abschließend bewerten wir die Eignung des Prebunkings aus unseren verschiedenen wissenschaftlichen Perspektiven und zeigen weiteren Forschungsbedarf auf.

[1] Roozenbeek, J. & van der Linden, S. (2019): Fake News Game Confers Psychological Resistance Against Online Misinformation. *Palgrave Commun 5* (1), S. 1–10. Verfügbar unter https://www.doi.org/10.1057/s41599-019-0279-9 (abgerufen am 17.07.2024) sowie Lewandowsky, S. & van der Linden, S. (2021) Countering Misinformation and Fake News Through Inoculation and Prebunking, *European Review of Social Psychology, 32* (2), S. 348–384. Verfügbar unter https://www.doi.org/10.1080/10463283.2021.1876983 (abgerufen am 19.07.2024).

2 Problemanalyse

2.1 Nutzung von Messenger-Diensten als Infrastruktur zur Vergemeinschaftung staatsskeptischer Gegenöffentlichkeiten

Messenger-Dienste sind ein ideales Umfeld für die Bildung von verfassungsfeindlichen und staatsskeptischen Gegenöffentlichkeiten[2], etwa im Hinblick auf die Verbreitung von Desinformation, die realweltliche Vernetzung und Radikalisierung desinformationsaffiner Gruppen.[3] Dies stellt – insbesondere aufgrund ihres abschottenden Charakters und häufig fehlender Gegenrede – eine Herausforderung für den gesellschaftlichen Zusammenhalt und die repräsentative Demokratie dar, da ohne gemeinsame Diskursräume nicht zwischen unterschiedlichen Interessen vermittelt oder Kompromisse erzielt werden können. Zum einen nutzen staatsskeptische Akteur*innen, die Desinformation verbreiten, Messenger-Dienste, weil einige Diensteanbieter wie Telegram die Rechtsdurchsetzung durch ihre Passivität erschweren. Zum anderen sind die technischen Kommunikationsfunktionen von Messenger-Diensten (Affordanzen) – insbesondere von Telegram – für den Aufbau einer staatsskeptischen Gegenöffentlichkeit ideal geeignet.[4] Dazu zählt insbesondere der fluide Übergang zwischen Individual-, Gruppen- und Massenkommunikation:

Individualkommunikation (one-to-one) findet statt, wenn ausschließlich zwei Nutzer*innen miteinander kommunizieren (private Unterhaltung, vergleichbar mit Kommunikation via SMS). Im Beispiel Telegram kann sich die Gruppenkommunikation auf kleine Gruppen wie private Familienchats beschränken, sie kann aber auch in weit größeren Kommunikationsräumen, in geschlossenen wie öffentlich zugänglichen Gruppen von bis zu 200.000 Mitgliedern, stattfinden (few-to-few, many-to-many). Während in Gruppen der Austausch zwischen den Mitgliedern möglich und gewünscht ist, bildet die Kanalfunktion eine weitere Dimension der spezifischen Kommunikation ab: Telegram-Kanäle erlauben es Kanalbetreibenden, eine unbegrenzte Zahl an Abonnent*innen zu erreichen (one-to-many) und dort ohne Möglichkeit der Gegenrede die eigenen Narrative zu verbreiten. Andere Messenger-Dienste wie WhatsApp und Signal verfügen über vergleichbare Funktionen, die jedoch in einigen Bereichen Unterschiede aufweisen (z.B. Begrenzung der Mitgliederzahlen in Gruppen).

Das Kuratieren von Inhalten erweist sich auf Telegram-Kanälen als zentrale Praktik jener, die sich auf Telegram in Gegenöffentlichkeiten vernetzen: Kanalbetreiber*innen mischen eigene und weitergeleitete Inhalte und setzen dazu häufig Desinformationen ein, die als zentrales Element komplexer alternativer Wirklichkeitskonstruktionen

[2] Anders: Jungherr, A., & Schroeder, R. (2021). Disinformation and the Structural Transformations of the Public Arena: Addressing the Actual Challenges to Democracy. *Social Media + Society, 7*(1), S. 1–13. Verfügbar unter https://doi.org/10.1177/2056305121988928 (abgerufen am 23.07.2024).

[3] Gegenöffentlichkeiten werden hier in den Blick genommen, wenn sie verfassungsfeindlich und desinformierend handeln (vgl. dazu auch Schulze, H., Hohner, J., Greipl, S,: Girgnhuber, M., Desta, I. & Rieger, D. (2022). Far-Right Conspiracy Groups on Fringe Platforms: A Longitudinal Analysis of Radicalization Dynamics on Telegram. *Convergence: The International Journal of Research into New Media Technologies, 28*(4), S. 1103–1126. Verfügbar unter https://doi.org/10.1177/13548565221104977 (abgerufen am 25.07.2024).

[4] Siehe dazu auch Schulze, H., Greipl, S., Hohner, J. & Rieger, D. (2024). Social Media and Radicalization: An Affordance Approach for Cross-Platform Comparison. *M&K 72*(2), S. 187–212.

dienen. Häufig werden hier auch divergierende Wirklichkeiten aus Qualitätsmedien und „alternativen" Medien zusammengeführt. In der Folge entstehen Verknüpfungen zwischen unterschiedlichen Ideologien und Milieus.

Derartige Akteur*innen nutzen sowohl lokal vernetzte kleine Gruppen, also dem Austausch über die Organisation realweltlicher Vernetzung dienliche Kanäle (few-to-few-Kommunikation) und große Gruppen (many-to-many-Kommunikation) sowie die Kanalfunktion (one-to-many-Kommunikation). Telegram ermöglicht es so Akteur*innen, auch auf der lokalen Ebene spezifische Gemeinschaften zu bilden. Diese (oft auch lokalen) Gruppen tragen zur Konstruktion einer (staatsskeptischen Gegen-)Öffentlichkeit bei, indem sie Themen und Diskussionen fördern, die in ihrem thematischen und/oder geografischen Kontext relevant sind.[5] Diese durch die Verbreitung von Desinformationen entstehenden staatsskeptischen Gegenöffentlichkeiten tragen somit über Messenger-Dienste sowohl im digitalen als durch die dort vollzogene Vernetzung im realen Raum zu einem zunehmenden Vertrauensverlust in staatliche Institutionen und einer zunehmenden Polarisierung der Gesellschaft bei.

Durch ihre Broadcasting-Funktion sind Kanäle per se als öffentlichkeitswirksam einzuordnen,[6] insbesondere dann, wenn sie hohe Reichweiten zu erzielen. Diese sind jedoch nicht zwingend notwendig, um öffentlich wirksam zu werden.[7] Daher reicht es nicht aus, sich nur an quantitativen Metriken zu orientieren. Denn die Herstellung von Gegenöffentlichkeiten auf Messenger-Diensten kann von weiteren Faktoren abhängen, die sich durch die Affordanzen der Plattformen ergeben: Durch Anschlusskommunikation (Reichweite) innerhalb der Messenger-Dienste sowie durch Verlinkungen auf alternative Plattformen und/oder über die Plattformen hinaus.

Um weiter zu klären, inwieweit Kommunikation über Kanäle öffentlichkeitswirksame Relevanz entfalten kann, ist zudem ein Blick auf die Motive und Praktiken von Gruppen-Admins und Kanalbetreibenden notwendig. Hier ist der Fokus auf die tatsächlich verbreiteten Inhalte zweckdienlich. Um Desinformation auf Messenger-Diensten wie Telegram zu verbreiten, werden häufig journalistische Produktionsroutinen adaptiert und strategisch eingesetzt, um öffentlichkeitswirksam zu werden.[8] Dazu zählen insbesondere emotionalisierende Darstellungen und Inhalte,

[5] In diesem Zusammenhang ist auch das sogenannte "Engagement Farming" relevant - eine Strategie, die darauf abzielt, die eigene Präsenz in sozialen Medien durch verschiedene Taktiken künstlich zu steigern, um Likes, Kommentare, Shares und andere Formen der Interaktion zu erhöhen. Dies kann sich darin äußern, dass User*innen an mehreren Diskussionen teilnehmen, ohne viel Mehrwert zu bieten, zahlreiche Nutzer*innen markieren, um Aufmerksamkeit zu erregen, oder kontroverse Inhalte verwenden, um Reaktionen zu provozieren.

[6] In den Telegram-FAQs zu Kanälen heißt es dazu im Wortlaut: „Kanäle sind eine Möglichkeit, öffentliche Mitteilungen an ein großes Publikum zu schicken, da Kanäle eine unbegrenzte Anzahl von Mitgliedern haben können."

[7] Bader, K.; Müller, K. & Rinsdorf, L. (2023): Zwischen Staatsskepsis und Verschwörungsmythen: Eine Figurationsanalyse zur kommunikativen Konstruktion von Gegenöffentlichkeiten auf Telegram. In: *M&K, 71* (3-4), S. 248–265. Verfügbar unter https://doi.org/10.5771/1615-634X-2023-3-4-248 (abgerufen am 29.07.2024) sowie Rinsdorf, L., Bader, K. & Jansen, C. (2024). Telegram als Plattform für staatsskeptische Akteur:innen. In: C. Nuernbergk, J. Haßler, J. Schützeneder, & N. Schumacher (Hrsg.), *Politischer Journalismus: Konstellationen - Muster - Dynamiken* (S. 97–108). Baden-Baden: Nomos Verlag. Im Erscheinen.

[8] Siehe dazu auch Eisenegger, M. (2021). Dritter, digitaler Strukturwandel der Öffentlichkeit als Folge der Plattformisierung (S. 31). In: M. Eisenegger, M. Prinzing, P. Ettinger & R. Blum (Hrsg.), *Mediensymposium. Digitaler Strukturwandel der Öffentlichkeit: Historische Verortung, Modelle und Konsequenzen* (1. Auflage 2021, S. 17–39). Springer Fachmedien Wiesbaden sowie Hepp, A. & Coudry, N (2023). Necessary Entanglements: Reflections on the Role of a "Materialist Phenomenology" in Researching Deep Mediatization and Datafication. *Sociologica, 17*(1), 137–153. Verfügbar unter https://doi.org/10.6092/issn.1971-8853/15793 (abgerufen am 29.07.2024).

die geeignet sind, gemeinschaftsbildende Effekte anzustoßen. Unsere Analysen[9], die sich auf den Dienst Telegram konzentrieren, zeigen neben vielen weiteren Untersuchungen[10], dass sich auf Telegram tatsächlich staatsskeptische Gegenöffentlichkeiten bilden, die sich in verschiedene Typen der Messenger-Kommunikation einordnen lassen. Dazu zählen **(1) die Vermittlung von Weltanschauungen mit monothematischer Fokussierung auf ein Schwerpunktthema**, etwa Covid-19 oder den Krieg in der Ukraine, **(2) das Pflegen von Communities durch Kultivierung verschwörungserzählerischer Perspektiven**, indem aus Insiderperspektive aktuelle Themen mittels QAnon-Sicht[11] kommuniziert werden, **(3) das Erzeugen von Glaubwürdigkeit durch Seriosität** mittels Kommunikation im journalistischen Stil, **(4) Politisches Influencing**, in dem durch einen meinungsbetonten Akteur oder eine meinungsbetonte Akteurin im Zentrum die eigene Persönlichkeit dargestellt und Nähe erzeugt wird, **(5) das Erzeugen von Aufmerksamkeit und Reichweite durch Verlinken und Verweisen**, indem Inhalte verbreitet werden und Werbung für Kanäle und Blogs gemacht wird sowie **(6) Verschwörungserzählungen für Einsteiger*innen,** in deren Kanälen verschwörungserzählerische Weltbilder entfaltet werden und Überzeugungsarbeit geleistet wird.

2.2 Rechtliche Herausforderungen bei Gegenmaßnahmen

Vor dem Hintergrund der kommunikationswissenschaftlichen Analysen erscheinen wirksame Gegenmaßnahmen dringend erforderlich, um den demokratischen Diskurs und gesellschaftlichen Zusammenhalt zu schützen. Die gesetzliche Regulierung von Bekämpfungsmaßnahmen gegen Desinformation in Messenger-Diensten ist jedoch nicht ohne weiteres möglich. Bereits in grundrechtlicher Hinsicht bestehen maßgebliche Herausforderungen. Dabei gelten für die öffentlichen und privaten Kommunikationsfunktionen von Messenger-Diensten jeweils unterschiedliche grundrechtliche Anforderungen.

2.2.1 Grund- und verfassungsrechtliche Vorgaben

Relevant kann vor allem die **Meinungsfreiheit** der Verbreiter*innen von Inhalten, die (mutmaßlich) Desinformation darstellen, sein (Art. 5 Abs. 1 S. 1 GG bzw. Art. 10 Abs. 1 EMRK, Art. 11 Abs. 1 GRCh). Unwahren Tatsachenbehauptungen, die bewusst geäußert werden oder erwiesen unwahr sind, wird grundsätzlich kein Schutz durch die Meinungsfreiheit zugebilligt (so z.B. das Bundesverfassungsgericht). Werden falsche

[9] Bader, K.; Müller, K. & Rinsdorf, L. (2023): Zwischen Staatsskepsis und Verschwörungsmythen: Eine Figurationsanalyse zur kommunikativen Konstruktion von Gegenöffentlichkeiten auf Telegram. In: *M&K, 71* (3-4), S. 248–265. Verfügbar unter https://doi.org/10.5771/1615-634X-2023-3-4-248 (abgerufen am 29.07.2024) sowie Rinsdorf, L., Bader, K. & Jansen, C. (2024). Telegram als Plattform für staatsskeptische Akteur:innen. In: C. Nuernbergk, J. Haßler, J. Schützeneder, & N. Schumacher (Hrsg.), *Politischer Journalismus: Konstellationen - Muster - Dynamiken* (S. 97–108). Baden-Baden: Nomos Verlag. Im Erscheinen.

[10] z.B. Holnburger, J. (2023, 29. März). *Chronologie einer Radikalisierung: Wie Telegram zur wichtigsten Plattform für Verschwörungsgläubige und Rechtsextremismus wurde*. Report des Centers für Monitoring, Analyse und Strategie (CeMAS). Verfügbar unter https://cemas.io/publikationen/telegram-chronologie-einer-radikalisierung/ (abgerufen am 28.06.2024).

[11] QAnon ist keine feste Organisation, sondern eher eine Idee oder Legende, die sich als lose Bewegung im Internet formiert hat und zunehmend auch offline sichtbar wird. Die sogenannten Q-Texte sind oft kryptisch und schwer verständlich, meist bestehend aus Satzfragmenten oder Fragen. Ein zentrales Thema ist der Mythos einer dunklen, geheimen Elite, die angeblich die USA durch den „Deep State" kontrolliert. Diese Aussagen enthalten oft versteckte antisemitische Andeutungen. Reale Ereignisse werden häufig als Beweise für diese Behauptungen interpretiert. In Deutschland wurden die QAnon-Theorien zunächst vor allem von Rechtsextremisten und Anhängern der Reichsbürgerbewegung verbreitet. Mit den Protesten gegen die Corona-Maßnahmen fanden sie jedoch auch in Teilen dieser neuen Bewegung Anklang.

Tatsachenbehauptungen jedoch mit subjektiven Werturteilen vermengt, werden sie trotzdem durch die Meinungsfreiheit geschützt. Grundsätzlich kann jede staatliche Maßnahme, die eine Meinungsäußerung verbietet oder behindert, einen Eingriff in die Meinungsfreiheit darstellen, so etwa gesetzliche Verpflichtungen zum Löschen und Sperren (mutmaßlicher) Desinformation auf Messenger-Diensten. Aber auch die staatliche Überwachung von Kommunikation in Messenger-Diensten kann in die Meinungsfreiheit eingreifen, da die Meinungen nicht mehr unbefangen geäußert werden können. Zudem schützt die Meinungsfreiheit auch davor, gezwungenermaßen eine fremde Meinung als eigene verbreiten zu müssen. Dies kann z.B. bei der (automatisierten) Kennzeichnung von Desinformation der Fall sein.

Empfänger*innen (mutmaßlicher) Desinformation können sich auf das Grundrecht auf **Informationsfreiheit** berufen (Art. 5 Abs. 1 S. 1 GG, Art. 11 Abs. 1 GRCh und Art. 10 Abs. 1 S. 2 EMRK). Dieses gewährleistet das Recht, sich aus allgemein zugänglichen Quellen ungehindert zu unterrichten. Öffentliche Messenger-Funktionen können in den Schutzbereich dieses Grundrechts fallen. Sogar Informationsquellen (z.B. Kanäle oder Gruppen), die erwiesenermaßen überwiegend Desinformation verbreiten,[12] fallen grundsätzlich unter den weiten Schutzbereich der Informationsfreiheit. Jede Verhinderung oder wesentliche Erschwerung des Zugangs zu Informationsquellen greift in dieses Grundrecht ein. Geschützt wird auch vor aufgedrängten Informationen, wozu dann nicht der desinformationsverbreitende Kanal, sondern die verpflichtende Kennzeichnung von Desinformation zählen kann.

Zudem spielt für Messenger-Dienste das **Fernmeldegeheimnis** bzw. Grundrecht auf **private Kommunikation** eine große Rolle (Art. 10 Abs. 1 S. 1 GG; Art. 7 GRCh, Art. 8 EMRK). Es schützt die Vertraulichkeit von Kommunikationsinhalten sowie die Umstände der Kommunikation. Jedes Auslesen, Filtern und Bewerten privater Kommunikation greift in das Fernmeldegeheimnis ein. Durch dieses Grundrecht sind somit zwar private Chats geschützt, nicht hingegen öffentliche Kanäle und Gruppen.

Das Recht auf **informationelle Selbstbestimmung** bzw. Grundrecht auf **Datenschutz** umfasst wiederum das Recht, selbst über die Erhebung und Verarbeitung der persönlichen Daten zu bestimmen (Art. 2 Abs. 1 i.V.m. Art. 1 Abs. 1 GG Art. 8 Abs. 1 GRCh und Art. 8 Abs. 1 EMRK). Maßnahmen gegen Desinformation in Messenger-Diensten können auf verschiedene Weise in das Grundrecht eingreifen, z.B. indem Inhaltsdaten oder Meta-Daten gespeichert, weitergegeben, verändert oder auf andere Weise verarbeitet werden.

Auf der anderen Seite sind die Grundrechte der Anbieter von Messenger-Diensten zu achten, insbesondere die **Berufsfreiheit** bzw. **unternehmerische Freiheit** (Art. 12 Abs. 1 GG; Art.16 GRCh). Darunter fällt der Schutz der unternehmerischen Betätigung, wozu u.a. Verfügungen über technische Ressourcen zählen. Auch die Vertragsfreiheit, z.B. die Gestaltung der Allgemeinen Geschäftsbedingungen der Messenger-Dienste werden grundsätzlich geschützt.

[12] z.B. Holnburger, J. (2023, 29. März). *Chronologie einer Radikalisierung: Wie Telegram zur wichtigsten Plattform für Verschwörungsgläubige und Rechtsextremismus wurde.* Report des Centers für Monitoring, Analyse und Strategie (CeMAS). Verfügbar unter https://cemas.io/publikationen/telegram-chronologie-einer-radikalisierung/ (abgerufen am 28.06.2024).

Neben grundrechtlichen Problemen sind in verfassungsrechtlicher Hinsicht auch die Auswirkungen auf die Demokratie zu berücksichtigen. Desinformation kann den freien gesellschaftlichen Meinungsbildungsprozess beeinträchtigen, der als wichtige Grundlage der Demokratie gilt. Jedoch können umgekehrt auch Maßnahmen gegen Desinformation den freien Meinungsbildungsprozess und damit die Demokratie beeinträchtigen. Praktische Herausforderungen der Rechtsanwendung bestehen in der **Beweisbarkeit** einer schädigenden Absicht.

Darüber hinaus müssen mögliche negative Wechselwirkungen der Regulierung berücksichtigt werden. Problematisch kann es sein, wenn Nutzer*innen sich bei einer strengeren gesetzlichen Regulierung von Diensten auf weniger regulierte Bereiche der Dienste oder kooperationsunwilligere Dienste zurückziehen.[13] Ein solcher **Verdrängungseffekt** zeigte sich etwa, als sich Verbreiter*innen von rechtswidrigen oder AGB-widrigen Inhalten von sozialen Medien wie Facebook auf Messenger-Dienste wie Telegram zurückzogen.[14]

Mittlerweile gibt es einige konkrete gesetzliche Maßnahmen, die zur Bekämpfung von Desinformation beitragen sollen (s. 2.2.2) und an diesen grund- und verfassungsrechtlichen Anforderungen zu messen sind. Ob durch diese Vorschriften möglicherweise verursachte **Grundrechtseingriffe gerechtfertigt** werden können, hängt von der Ausgestaltung der konkreten Maßnahmen und von der Abwägung mit jeweils relevanten Grundrechten ab.

2.2.2 Die neuen EU-Rechtsakte

Die Bekämpfung von Desinformation ist in den letzten Jahren Gegenstand neuer Rechtsakte geworden. Relevant ist z.B. der gestärkte **EU-Verhaltenskodex gegen Desinformation**, eine Ko-Regulierung,[15] die von den Unterzeichner*innen in erster Linie freiwillig einzuhalten ist. Daneben werden einzelne Teilbereiche des Problems in der Verordnung über die Transparenz und das Targeting **politischer Werbung** und im Gesetz über **Künstliche Intelligenz** geregelt. Das wohl strengste Vorgehen der EU gegen Desinformation besteht in **restriktiven Maßnahmen** (Sanktionen) gegen mehrere russische Staatsmedien, die aufgrund ihrer anhaltenden Desinformationstätigkeit zeitweise mit einem Sende- und Verbreitungsverbot belegt wurden.

Vor allem aber der **Digital Services Act (DSA)** soll als zentrale Verordnung zur Plattformregulierung in der EU gegen Desinformation wirken.

[13] Panahi, T. & Zurawski, P. (2023), Messenger & Co: Das Unsichtbare regulieren?, in: Kemmesies, U. et al. (Hrsg.), MOTRA-Monitor 2022, Wiesbaden 2023, 410-429.

[14] Jünger, J.& Gärtner, C. (2020). *Datenanalyse von rechtsverstoßenden Inhalten in Gruppen und Kanälen von Messengerdiensten am Beispiel Telegram.* Düsseldorf: Landesanstalt für Medien NRW, S. 6.

[15] Bei einer Ko-Regulierung gibt der Gesetzgeber Ziele vor, deren Verwirklichung wird jedoch nichtstaatlichen Akteuren überlassen. Der gestärkte EU-Verhaltenskodex gegen Desinformation basiert auf Leitlinien der EU-Kommission, die von Wirtschaftsunternehmen, aber auch Fact-Checking Organisationen umgesetzt wurden. Der aktuelle Desinformationskodex wurde am 16. Juni 2022 veröffentlicht und bisher von 44 Parteien unterzeichnet, darunter auch von Meta und TikTok.

3 Desinformationsbekämpfung im Digital Services Act

Der DSA bezweckt, den Rechtsrahmen für Vermittlungsdienste EU-weit zu vereinheitlichen und zu einem sicheren, berechenbaren und vertrauenswürdigen Online-Umfeld beizutragen. Im Folgenden wollen wir die Anwendbarkeit des DSA auf Messenger-Dienste sowie seine Effektivität gegen Desinformation vertiefen. Dabei werden wir die aktuelle Rechtslage nach dem DSA und damit einhergehende Probleme erläutern und sodann Lösungsvorschläge unterbreiten.

3.1 Anwendbarkeit auf Messenger-Dienste

Eine wichtige Frage ist, ob Messenger-Dienste überhaupt in den Anwendungsbereich des DSAs fallen. Der DSA enthält verschiedene Vorschriften, die sich an unterschiedliche Typen von Vermittlungsdiensten richten (Art. 3 g) DSA).

3.1.1 Online-Plattformen

Vorschriften, die für die Bekämpfung von Desinformation relevant sein können, gelten in erster Linie für **Online-Plattformen** (z.B. Art. 16, 23, 34, 35). Ob eine Online-Plattform vorliegt, bestimmt sich danach, ob Informationen im Auftrag der Nutzer*innen gespeichert und **öffentlich verbreitet** werden (Art. 3 i) DSA). Der DSA bestimmt, dass eine öffentliche Verbreitung nur dann vorliegt, wenn Informationen für eine **potenziell unbegrenzte Zahl von Dritten** bereitgestellt werden (Art. 3 k) DSA).

In den Erwägungsgründen[16] des DSA wird ausdrücklich klargestellt, dass "Instant Messaging-Dienste" nicht in den Anwendungsbereich der Begriffsbestimmung für Online-Plattformen fallen, da sie "für die interpersonelle Kommunikation zwischen einer endlichen Zahl von Personen verwendet werden, die vom Absender der Kommunikation bestimmt wird" (Erwägungsgrund 14). Damit fallen jedoch nicht alle Messenger-Dienste aus dem Anwendungsbereich der Vorschriften für Online-Plattformen heraus. Denn die Erwägungsgründe sehen eine **funktionsbasierte Adressatenregelung** für Online-Plattformen vor (Erwägungsgrund 15). Danach soll der Grundsatz gelten, dass nach den einzelnen Diensten eines Anbieters unterschieden werden soll, ob diese unter Vorschriften der Verordnung fallen oder nicht (z.B. Einzelchats, Gruppen, Kanäle). Dies ermöglicht es folglich, dass ein Messenger-Dienst zumindest teilweise als Online-Plattform gilt. Zum Beispiel könnte ein Telegram-Kanal ohne Zugangsbeschränkungen mit 500.000 Abonnent*innen als Online-Plattform gelten, ein Chat zwischen zwei Personen oder eine geschlossene Gruppe mit 30 Mitgliedern jedoch nicht.

Problematisch bleibt jedoch der Bereich der **besonders mitgliedsstarken, aber geschlossenen Gruppen**. Denn hier liegt keine „potenziell unbegrenzte Zahl" von Personen vor.

[16] Erwägungsgründe zählen nicht zum rechtlich verbindlichen Gesetzestext, dienen aber als Gesetzesbegründung und können zur Auslegung des Gesetzes herangezogen werden.

Jedoch wäre es sachfremd, dass das Merkmal der „Öffentlichkeit" nicht erfüllt sein soll, wenn z.B. bei Telegram Gruppen mit bis zu 200.000 Nutzer*innen möglich sind.[17]

3.1.2 Sehr große Online-Plattformen

Ein weiteres Problem besteht bei der Einstufung von sog. **sehr großen Online-Plattformen**, für welche besondere Pflichten gelten, die zum Teil den Umgang mit Desinformation betreffen können (Art. 34, 35 DSA). Eine sehr große Online-Plattform liegt vor, wenn der Dienst über 45 Millionen aktive Nutzer*innen in der EU aufweist und von der EU-Kommission als solche benannt wurde (Art. 33 Abs. 1 DSA). Die Ermittlung der erforderlichen Nutzer*innenzahl ist gerade bei hybriden Messenger-Diensten problematisch. Unklar ist bislang, ob nur solche Nutzer*innen gezählt werden sollen, die den Dienst für die öffentliche Verbreitung von Informationen nutzen (Art. 3 i DSA). Dann wäre für jede*n Nutzer*in einzeln zu prüfen, ob der Dienst öffentlich oder nur privat genutzt wird. Dies ginge mit der Erhebung massenhaft personenbezogener Daten einher, wofür in dieser Konstellation bislang keine Rechtsgrundlage gegeben ist.[18] Zudem unterliegen die Nutzungsweisen erheblichen Schwankungen.[19]

3.1.3 Lösungsvorschlag: Kriterien für mitgliedsstarke Gruppen

Wir schlagen eine Anpassung der Kriterien für die Einstufung von sehr mitgliedsstarken geschlossenen Gruppen vor, die speziell auf die hybride Kommunikationsstruktur von Messenger-Diensten ausgerichtet sind. Diese kann durch einen delegierten Rechtsakt der EU-Kommission erfolgen (Art. 33 Abs. 3, 24 Abs. 2 DSA). Um eine Beeinträchtigung des Fernmeldegeheimnisses bzw. private Kommunikation zu vermeiden (s. Kap. 2.2.1.), sind restriktive und eindeutige Kriterien zu wählen. In Betracht kommt etwa die Festlegung eines **hohen nummerischen Schwellenwerts** (z.B. 10.000 Mitglieder). Zusätzlich sollten die **Beitrittsmöglichkeiten** zu einer Gruppe oder einem Kanal berücksichtigt werden. Ist der Beitritt öffentlich möglich, z.B. für jede*n Nutzer*in über einen öffentlich zugänglichen Einladungslink, liegt ein Indiz für eine öffentliche Kommunikation vor.

3.1.4 Lösungsvorschlag: Einstufung sehr großer Online-Plattformen

Zudem schlagen wir die Festlegung von Berechnungsmethoden sehr großer Online-Plattformen vor. Da bislang unklar ist, wann ein Messenger-Dienst als sehr große Online-Plattform eingestuft werden kann, bedarf es hier der Konkretisierung, z.B. mittels eines delegierten Rechtsakts der EU-Kommission (Art. 33 Abs. 3 DSA). Die Problematik

[17] Telegram-Gruppen erlauben etwa bis zu 200.000 Mitglieder, sogenannte Giga-Gruppen arbeiten gänzlich ohne Beschränkung der Mitgliederanzahl, vgl. Telegram.org (o.J.): *Fragen und Antworten.* Verfügbar unter https://telegram.org/faq/de#f-was-ist-der-unterschied-zwischen-gruppen-und-kanalen (abgerufen am 12.03.2024) sowie https://core.telegram.org/api/channel (abgerufen am 28.06.2024).

[18] Auch die Pflicht, Datenzugang für bestimmte Institutionen und Forscher*innen zu gewähren, gilt wiederum erst dann, wenn klar ist, dass es sich um eine sehr große Online-Plattform handelt (Art. 40 DSA).

[19] Panahi, T., Hornung, G., Schäfer, K., Choi, J.-E., Steinebach, M. & Vogel, I. (2023), Desinformationserkennung anhand von Netzwerkanalysen – ein Instrument zur Durchsetzung der Pflichten des DSA am Beispiel von Telegram, in: Friedewald, M., Roßnagel, A. Neuburger, R., Bieker, F. & Hornung, G. (Hrsg.), *Daten-Fairness in einer globalisierten Welt*, (S. 343–370). Baden-Baden. Verfügbar unter https://www.nomos-elibrary.de/10.5771/9783748938743-343/desinformationserkennung-anhand-von-netzwerkanalysen-ein-instrument-zur-durchsetzung-der-pflichten-des-dsa-am-beispiel-von-telegram?page=1 (abgerufen am 18.07.2024).

besteht insbesondere darin, dass unklar ist, inwiefern auch die Nutzer*innen in die Berechnung mit einbezogen werden sollen, die zwar einen Messenger-Dienst, aber gerade nicht seine öffentlichen Kommunikationsfunktionen nutzen. Denn es kommt für die Einordnung einer großen Online-Plattform gerade auf die öffentliche Nutzung des Dienstes an (s. Kap. 3.1.1).
Eine Möglichkeit wäre es, Kriterien für die Berechnung der Nutzer*innenzahlen zu formulieren, bei denen die Nutzer*innen, die die öffentlichen Funktionen nicht nutzen, möglichst nicht einbezogen werden. Um diese zu ermitteln, müssten jedoch zahlreiche personenbezogenen Daten über das Nutzungsverhalten (automatisiert) verarbeitet werden, was nicht nur praktisch, sondern auch datenschutzrechtlich herausfordernd wäre (s.o.). Um Verstöße gegen das Datenschutzrecht zu vermeiden, sollte stattdessen gerade mit Blick auf die extensiven Berechnungsmethoden, die in Erwägungsgrund 77 DSA anheimgestellt werden, klargestellt werden, dass diese nicht ohne Weiteres auf Messenger-Dienste übertragbar sind. Demgegenüber wären regelmäßige auf datenschutzrechtliche Einwilligungserklärungen basierende Nutzer*innenumfragen zwar datenschutzfreundlicher, jedoch kein verlässliches Instrument, wenn eine Überprüfung der ermittelten Angaben fehlt.

Eine andere, einfachere Möglichkeit wäre es, Art. 33 Abs. 1 DSA im Rahmen eines delegierten Rechtsakts so auszulegen, dass jegliche Nutzer*innen, die sich für den Dienst registriert haben, als relevante Nutzer*innen im Rahmen der Berechnung herangezogen werden müssen.[20] Gegenüber der oben genannten Differenzierung nach Nutzer*innenverhalten, würde die **Einbeziehung aller registrierter Nutzer*innen** in die Berechnung zu eindeutigeren sowie konstanteren Ergebnissen führen und wäre wohl aus Perspektive der Unternehmen am ressourcenschonendsten. Zwar widerspricht ein solches Vorgehen auf den ersten Blick dem Wortlaut der Vorschrift, welcher sich lediglich auf Online-Plattformen und damit auf öffentliche Kommunikation bezieht. Jedoch steht bei den meisten hybriden Messenger-Diensten allen Nutzer*innen jedenfalls die Möglichkeit offen, die öffentlichen Funktionen zu nutzen, sodass diese daher auch bei der Berechnung einbezogen werden sollten. Dies deckt sich auch mit der gesetzlichen Definition der Nutzer*in nach Art. 2 lit. b) DSA, welche zwar auch auf die öffentliche Kommunikation abstellt, aber nicht nur. Denn durch den Begriff „insbesondere" werden gerade auch solche Nutzer*innen erfasst, welche die öffentliche Kommunikation nicht in Anspruch nehmen. Schließlich würde diese Herangehensweise grundsätzlich am wenigsten in die Grundrechte auf Datenschutz und private Kommunikation eingreifen, da hierfür in der Regel keine oder nur wenige personenbezogenen Daten verarbeitet werden müssten, denn eine solche Berechnung könnte durch eine anonymisierte Zählung der Registrierungen erfolgen.

3.2 Effektivität gegen Desinformation

Der DSA soll – jedenfalls den Erwägungsgründen nach – gegen Desinformation wirken (Erwägungsgrund 9). Trotz dieser Zielrichtung enthält der DSA allerdings keine Definition für Desinformation und erwähnt den Begriff der Desinformation in keiner

[20] Damit ist nicht gemeint, dass die Vorschriften der Art. 34 ff. DSA dann auch auf die privaten Kommunikationsfunktionen anwendbar sein sollen. Der Vorschlag bezieht sich nur auf die Einstufung eines Dienstes als sehr große Online-Plattform.

Vorschrift. Dennoch gibt es einige Regelungen, die in einem begrenzten Umfang (mittelbar) gegen Desinformation wirken können.

Der DSA enthält einige repressive Vorschriften, die die **Moderation rechtswidriger Inhalte** betreffen (z.B. Melde- und Abhilfeverfahren, Art. 16 DSA, Aussetzungspflicht, Art. 23 Abs. 1 DSA). Desinformation ist an sich nicht rechtswidrig. Jedoch können bestimmte Formen von Desinformation nach EU- und mitgliedsstaatlichem Recht rechtswidrig sein. Das deutsche Strafrecht enthält z.B. einige Straftatbestände, die unter Umständen Desinformation erfassen und daher auch im Rahmen der Vorschriften des DSA relevant sind (z.B. Verleumdung gemäß § 187 StGB, Volksverhetzung gemäß § 130 StGB, jeweils in der Variante des „Verleugnens").

Die im DSA geregelten Pflichten zu **Risikobewertung und Risikominderung** richten sich hingegen nicht nur gegen rechtswidrige Inhalte, sondern auch gegen „systemische Risiken" (Art. 34, 35 DSA), wozu auch Desinformation zählt (Erwägungsgrund 84). Mindestens einmal jährlich müssen sehr große Online-Plattformen und Suchmaschinen alle systemischen Risiken ermitteln, analysieren und bewerten, die sich aus der Konzeption oder dem Betrieb ihrer Dienste und damit verbundenen (algorithmischen) Systemen, oder der Nutzung ihrer Dienste ergeben. Um diese Risiken zu mindern, müssen sie, wenn nötig, ihre technischen Funktionen, Algorithmen und Allgemeinen Geschäftsbedingungen überarbeiten.

Zudem enthält der DSA präventive Maßnahmen, die mittelbar gegen die Verbreitung von Desinformation wirken können. Dazu zählen etwa die **Transparenz** der algorithmischen Empfehlungssysteme und der Online-Werbung (Art. 26, 27 DSA). Zudem müssen Allgemeine Geschäftsbedingungen transparent ausgestaltet werden und etwa Angaben über die Art und Weise der Moderation von Inhalten enthalten (Art. 14 Abs. 4 DSA).

Eine der größten Schwachstellen des DSA ist, dass er viele **unbestimmte Formulierungen** enthält, die der Konkretisierung bedürfen. Dies zeigt sich zum Beispiel in den interpretationsoffenen Formulierungen der Art. 34, 35 DSA (zum Beispiel "gesellschaftliche Debatte"). Diese unbestimmten Rechtsbegriffe bergen einerseits die Gefahr, dass sie zu weit interpretiert werden und Messenger-Dienste unangemessen in die private Kommunikation eingreifen, um diese Pflichten zu erfüllen. Andererseits können die vagen Formulierungen auch zu eng interpretiert werden und damit Raum für Passivität bei der Risikominderung bieten.

3.2.1 Lösungsvorschläge zur Effektivität des DSA gegen Desinformation

Da der DSA nur wenige Vorschriften enthält, die (mittelbar) gegen Desinformation wirken können, und diese Vorschriften zahlreiche unbestimmte Formulierungen enthalten, braucht es Anpassungen und Konkretisierungen des DSA, um seine Effektivität gegen Desinformation in Messenger-Diensten zu steigern. Diese sollten auf die hybride Struktur von Messenger-Diensten zugeschnitten sein.

Zum einen könnte die EU-Kommission die gesetzlich vorgesehene Möglichkeit nutzen, spezielle Leitlinien herauszugeben, die **Best-Practice-Beispiele für Messenger-Dienste** enthalten und der Orientierung dienen können (Art. 35 Abs. 3 DSA). In Anbetracht der vagen Formulierungen der Art. 34, 35 DSA ist für die Anbieter von Messenger-Diensten klarzustellen, welche (personenbezogenen) Daten sie zur Umsetzung dieser Pflichten mindestens erheben müssen und darüber hinaus noch dürfen. Dies sollte für die unterschiedlichen Kommunikationsfunktionen von Messenger-Diensten gesondert dargestellt werden. Da sich die Pflichten aus Art. 34, 35 DSA dem Wortlaut nach nur auf sehr große Online-Plattformen beziehen, könnten Messenger-Dienste zweckwidrig ihre hybride Struktur bei der Risikobewertung außer Acht lassen, obwohl z.B. auch die bloße Möglichkeit der Nutzung privater Funktionen auf die Nutzung öffentlicher Funktionen Einfluss nehmen kann. Daher sollte im Rahmen der Leitlinien klargestellt werden, dass Messenger-Dienste im Rahmen der Risikobewertung die Bedeutung ihrer unterschiedlichen Kommunikationsfunktionen im Hinblick auf die zu untersuchenden systemischen Risiken zu analysieren haben. Sie sollten etwa dazu verpflichtet sein, zu untersuchen, wie sich die möglichen Mitgliederzahlen und Beitrittsmöglichkeiten auf die Nutzung des Dienstes auswirken könnten und welche technischen Funktionen die Möglichkeiten begünstigen könnten, verfassungsfeindliche Gegenöffentlichkeiten zu begründen.

Zum anderen könnte der nationale Gesetzgeber aktiv werden. Wie gezeigt, knüpfen viele der zentralen Vorschriften des DSA an rechtswidrigen Inhalten an und nicht an Desinformation (Art. 16, 23 DSA). Der deutsche Gesetzgeber könnte weitere Formen der Desinformation als rechtswidrig erklären, auf welche diese Vorschriften des DSA dann anwendbar wären. Dabei dürfen die Grundrechte – insbesondere die Meinungsfreiheit – nicht unverhältnismäßig eingeschränkt werden.[21] Da dieser Vorschlag jedoch die Analyse des insbesondere strafrechtlichen Rechtsrahmens und damit vieler einzelner Vorschriften erfordern würde, gehen wir an dieser Stelle nicht näher darauf ein.

3.2.2 Zwischenfazit

Die rechtliche Analyse zeigt, dass öffentliche Kommunikationsfunktionen von Messenger-Diensten bereits in den Anwendungsbereich des DSA fallen können (s. 3.1.1). Die Kommunikation in geschlossenen Chatgruppen bleibt in einem großen Umfang jedoch außen vor, obwohl auch diese zur Bildung verfassungsfeindlicher Gegenöffentlichkeiten und zur Radikalisierung beitragen können. Daher braucht es weitere Lösungen, die auf die hybriden Kommunikationsfunktionen von Messenger-Diensten zugeschnitten sind. Zudem ist durch den DSA vor allem eine risikobasierte Regulierung gegeben, die eher reaktiv und repressiv gegen Desinformation wirkt und keine, die langfristig und direkt zur Medienkompetenz der einzelnen Nutzerin bzw. des einzelnen Nutzers beiträgt. Auch hier sehen wir Reformpotenzial.

[21] Viel mehr könnten die Modalitäten einzelner Tatbestände dahingehend überprüft werden, ob sie zu der in Messenger-Diensten vorgefundenen soziotechnischen Realitäten passen oder zu reformieren sind. Gerade die in Straftatbeständen häufig vorausgesetzte Eignung zur Störung des "öffentlichen Friedens" sowie andere Erheblichkeitsschwellen sollten auf dieser Basis überprüft werden.

4 Prebunking als präventive Interventionsmaßnahme

Eine denkbare Maßnahme, die hinsichtlich der Desinformationsbekämpfung in aller Munde ist und diese Lücke schließen könnte, ist das sogenannte **Prebunking**. So nannte der ehemalige EU-Außenbeauftragte Josep Borell Prebunking etwa als mögliche Maßnahme zur Bekämpfung von Desinformation im Super-Wahljahr 2024.[22] Im Folgenden wollen wir vertiefen, was unter Prebunking zu verstehen ist und inwiefern es für die Regulierung von Desinformation in Messenger-Diensten fruchtbar gemacht werden kann.

Prebunking ist eine **präventive Maßnahme**, die darauf abzielt, Individuen vor dem Einfluss von Falschinformationen zu schützen. Im Gegensatz zu Debunking, bei dem Falschinformationen nachträglich korrigiert werden, soll Prebunking bereits vor der Konfrontation mit einer Falschnachricht ansetzen und diese so vorzeitig widerlegen.

Die Idee des Prebunkings lässt sich auf die Inokulationstheorie von McGuire zurückführen.[23, 24] Nach dieser Theorie können Menschen ähnlich einer Impfung gegen Krankheiten gegen Überzeugungsversuche mit Falschinformationen „immunisiert" werden. Eine vorherige Konfrontation mit einer abgeschwächten Version der Falschinformation mit gleichzeitiger argumentativer Widerlegung soll „mentale Antikörper" aktivieren, ohne dabei eine Einstellungsveränderung herbeizuführen.[25] Das bedeutet, dass peripher vorhandene Überzeugungen durch die Mikrodosen von Falschnachrichten „attackiert" werden und dazu führen, dass sich Rezipient*innen sowohl mit eigenen Überzeugungen als auch mit den „falschen Mikrodosen" aktiv auseinandersetzen müssen. Das Ziel des Prebunkings ist es daher, die **Widerstandsfähigkeit** einer Person gegenüber Desinformationskampagnen zu erhöhen, indem den Tatsachen entsprechende Überzeugungen gefestigt werden.

4.1 Vorstellung des Prebunking-Ansatzes

Der Begriff „Prebunking" wird in verschiedenen Kontexten verwendet und kann unterschiedliche Maßnahmen beschreiben. Prebunking-Maßnahmen variieren dabei in ihrem Fokus und ihrem Inhalt und können spezifisch (narrow) oder breit (broad) formuliert sein.[25]

4.1.1 Narrow-Spectrum-Prebunking

In der spezifischeren Variante des Prebunkings (**Narrow-Spectrum**) werden Ausschnitte oder dieselbe Nachricht wie die nachfolgende Falschinformation verwendet. Nur wenn

[22] Heise.de vom 24.01.2024, verfügbar unter: https://www.heise.de/news/EU-Aussenbeauftragter-Vergiftete-Informationen-untergraben-die-Demokratie-9606550.html (abgerufen am 27.08.2024).

[23] McGuire, W. J. (1961). The Effectiveness of Supportive and Refutational Defenses in Immunizing and Restoring Beliefs Against Persuasion. *Sociometry, 24(2),* S. 184. Verfügbar unter https://doi.org/10.2307/2786067 (abgerufen am 29.07.2024).

[24] McGuire, W. J. (1964). Inducing Resistance to Persuasion: Some Contemporary Approaches. In L. Berkowitz (Ed.), *Advances in experimental social psychology, 1*, S. 191–229. New York, NY: Academic Press.

[25] Lewandowsky, S. & Van Der Linden, S. (2021). Countering Misinformation and Fake News Through inoculation and Prebunking. *European Review of Social Psychology, 32*(2), S. 348–384. Verfügbar unter https://doi.org/10.1080/10463283.2021.1876983 (abgerufen am 29.07.2024).

die Prebunking-Nachricht spezifisch formuliert ist, ist sie tatsächlich mit der klassischen Inokulation vergleichbar. Eine Inokulation enthält zwei zentrale Komponenten:[26] Zunächst werden Rezipient*innen vor einstellungsgefährdenden Falschnachrichten gewarnt. Anschließend wird ihnen eine „abgeschwächte (Mikro-)Dosis" der Falschnachricht präsentiert. Die Präsentation der abgeschwächten Version der Falschnachricht wird **Refutational Preemption** oder **Prebunking** genannt. Diese zweigeteilte Struktur der Inokulation soll Widerstandsmechanismen wie das Empfinden von Bedrohung und eine Gegenargumentation aktivieren.

4.1.2 Broad-Spectrum-Prebunking

In der breit gefassten Form des Prebunkings (**Broad-Spectrum**) werden Rezipient*innen über gängige **Manipulationstechniken und -strategien** aufgeklärt.[27] Ein Beispiel für eine gängige Manipulationsstrategie ist die Berufung auf vermeintliche Forscher*innen, Expert*innen oder Institutionen, die die Glaubwürdigkeit von geteilten Falschinformationen erhöhen sollen, in Wahrheit aber weder die wissenschaftlichen Kenntnisse noch die Expertise besitzen.[28] Ein weiteres Beispiel ist die Erklärung manipulativer Rhetorik, wie die Nutzung emotionaler Sprache oder der Sündenbock-Argumentation, bei der Schuldzuweisungen an eine Gruppe oder einzelne Personen erfolgen ohne tatsächliche Problemlösungen zu suchen.[29] Broad-Spectrum Prebunking konfrontiert Rezipient*innen also nicht mit der „Mikro-Dosis" einer Falschnachricht, sondern soll aufklären und die kritische Denkfähigkeit fördern. Mit der Vermittlung gängiger Strategien vor der Konfrontation mit Desinformationen soll die Erkennung von Manipulationsversuchen erleichtert und die Resistenz gegenüber Desinformationskampagnen aufgebaut werden. Auch wenn dieser neuere Prebunking-Ansatz nicht auf einen spezifischen Inhalt fokussiert, kann er dabei helfen, den Glauben in Falschinformationen zu verringern.[28] In Kombination mit einer Vorwarnung gegen spezifische Desinformationen wird Broad-Spectrum-Prebunking ebenfalls als Inokulation verstanden. [29]

Im Folgenden verwenden wir den Begriff des Prebunkings als Oberbegriff für Narrow-Spectrum und Broad-Spectrum-Methoden.

4.2 Intradisziplinäre Bewertung des Prebunkings

4.2.1 Psychologische Bewertung

Frühere Forschungsansätze zeigen, dass klassische Inokulationen zu einer umfassenden kognitiven Verarbeitung (Prozess des Denkens und Verstehens) des gesehenen

[26] Compton, J. A., & Pfau, M. (2005). Inoculation Theory of Resistance to Influence at Maturity: Recent Progress In Theory Development and Application and Suggestions for Future Research. *Communication Yearbook, 29*(1), S. 97–145. Verfügbar unter https://doi.org/10.1080/23808985.2005.11679045 (abgerufen am 29.07.2024).

[27] Lewandowsky, S. & Van Der Linden, S. (2021). Countering Misinformation and Fake News through Inoculation and Prebunking. *European Review of Social Psychology, 32*(2), S. 348–384. Verfügbar unter https://doi.org/10.1080/10463283.2021.1876983 (abgerufen am 29.07.2024).

[28] Cook, J. (2020). Deconstructing Climate Science Denial. In D. C. Holmes & L. M. Richardson (Eds.), *Research Handbook on Communicating Climate Change* (S. 62–78). Edward Elgar Publishing. Verfügbar unter https://doi.org/10.4337/9781789900408.00014 (abgerufen am 29.07.2024).

[29] Roozenbeek, J., Van der Linden, S., Goldberg, B., Rathje, S. & Lewandowsky, S. (2022). Psychological Inoculation Improves Resilience Against Misinformation on Social Media. *Science Advances, 8*(34), eabo6254. Verfügbar unter https://doi.org/10.1126/sciadv.abo6254 (abgerufen am 29.07.2024).

Inhalts führen können[30] und dass die Motivation, sich mit der Falschinformation zu beschäftigen, und der Widerstand gegen sie durch ausgelöste Bedrohungsgefühle erhöht werden können.[31] Aktuellere Studienergebnisse weisen darauf hin, dass Inokulationen das Nachdenken über ein Nachrichtenthema anregen.[32] Darüber hinaus belegen mehrere Feldstudien auch die Effektivität von breit gefassten Prebunking-Nachrichten auf der Plattform YouTube (z.B. kurze Videos, die gängige Manipulationstechniken von Desinformationskampagnen erklären).[33] Dies deutet darauf hin, dass insbesondere **Broad-Spectrum-Prebunking** bei der Erkennung gängiger Manipulationstechniken und der Förderung einer kritischen Haltung gegenüber Falschinformationen helfen kann. Eine Meta-Analyse aus dem Jahr 2023 zeigt zudem, dass **Narrow-Spectrum-Prebunking** die **Glaubwürdigkeit von Falschinformationen verringern** und positive Effekte auf das Teilen wahrer Informationen haben kann.[34] Die Wirkung von Inokulationen auf das Teilen von Falschinformationen wird allerdings nicht eindeutig belegt; es zeigt sich, dass Inokulationen nur die Weiterleitung von gesundheitsbezogenen Falschinformationen reduziert, nicht aber die von anderen Falschinformationen. In einer unserer aktuellen Studien im Rahmen des DYNAMO-Projekts zeigt sich, dass Prebunking die Glaubwürdigkeit von Desinformation nur dann signifikant reduziert, wenn sie **sehr spezifisch (narrow-spectrum)** erfolgt. Weitere Forschung ist daher notwendig, um die Wirksamkeit von Prebunking-Maßnahmen umfassend zu untersuchen und nachhaltig zu verbessern.

Um ein Urteil über Prebunking als Maßnahme gegen Desinformationen bilden zu können, müssen sowohl die Chancen als auch die Risiken betrachtet werden. Der größte Vorteil liegt darin, dass Falschinformationen widerlegt werden können, bevor sie verarbeitet werden. Dies könnte den „**Continued Influence Effect**" verhindern, der beschreibt, wie einmal verarbeitete Desinformation plausibel mit dem Weltwissen verbunden werden und auch nach einer Korrektur nachhaltig das Verhalten, das Denken und die Einstellung einer Person beeinflussen können.[35] Eine breit gefasste Prebunking-Nachricht, die über gängige Manipulationstechniken informiert, könnte zudem zur Erweiterung der **digitalen Medien- und Informationskompetenz** beitragen. Rezipient*innen könnten so dabei unterstützt werden, neue Informationen kritisch zu hinterfragen und zu prüfen. Allerdings muss beachtet werden, dass die Dauer der Wirkung von Prebunking-Nachrichten begrenzt ist und nachlässt, wenn keine weiteren Nachrichten

[30] Pfau, M., Tusing, K. J., Lee, W., Godbold, L. C., Koerner, A. F., Penaloza, L., Hong, Y. H. & Yang, V. S. H. (1997). Nuances in Inoculation: The Role of Inoculation Approach, Ego-Involvement, and Message Processing Disposition in Resistance. *Communication Quarterly, 45*(4), S. 461–481. Verfügbar unter https://doi.org/10.1080/01463379709370077 (abgerufen am 29.07.2024).

[31] Compton, J. & Pfau, M. (2004). Use of Inoculation to Foster Resistance to Credit Card Marketing Targeting College Students. Journal of Applied *Communication Research, 32*(4), S. 343–364. Verfügbar unter https://doi.org/10.1080/0090988042000276014 (abgerufen am 29.07.2024).

[32] Compton, J., Van Der Linden, S., Cook, J. & Basol, M. (2021). Inoculation Theory in the Post-truth Era: Extant Findings and New Frontiers for Contested Science, Misinformation, and Conspiracy Theories. Social and Personality *Psychology Compass, 15*(6), e12602. Verfügbar unter https://doi.org/10.1111/spc3.12602 (abgerufen am 29.07.2024).

[33] Roozenbeek, J., Van der Linden, S., Goldberg, B., Rathje, S. & Lewandowsky, S. (2022). Psychological Inoculation Improves Resilience Against Misinformation on Social Media. Science Advances, 8(34), eabo6254. Verfügbar unter https://doi.org/10.1126/sciadv.abo6254 (abgerufen am 29.07.2024).

[34] Lu, C., Hu, B., Li, Q., Bi, C. & Ju, X. (2023). Psychological Inoculation for Credibility Assessment, Sharing Intention, and Discernment of Misinformation: Systematic Review and Meta-Analysis. *Journal of Medical Internet Research, 25*, e49255. Verfügbar unter https://doi.org/10.1080/10463283.2021.1876983/10.2196/49255 (abgerufen am 29.07.2024).

[35] Johnson, H. M., & Seifert, C. M. (1994). Sources of the Continued Influence Effect: When Misinformation in Memory Affects Later Inferences. Journal of Experimental Psychology: *Learning, Memory, and Cognition, 20*(6), S. 1420. Verfügbar unter https://doi.org/ 10.1037/0278-7393.20.6.1420 (abgerufen am 29.07.2024).

folgen.[36] Zudem könnten Prebunking-Nachrichten abgelehnt oder falsch interpretiert werden, wenn sie nicht der eigenen Überzeugung entsprechen, was dem psychologischen Mechanismus des „Confirmation Bias" entspricht.[37] Die Auswirkungen von Prebunking-Maßnahmen auf die Glaubwürdigkeit **wahrer** Informationen ist noch **nicht vollständig geklärt.**[38] Hier besteht etwa die Gefahr, dass Vertrauen in (auch seriöse und korrekte) Nachrichten nachhaltig untergraben wird.

Prebunking-Maßnahmen können folglich die Resistenz gegenüber Desinformation erhöhen und die Glaubwürdigkeit von falschen Informationen reduzieren. Der Einfluss auf die Weiterleitung von Desinformationen ist jedoch nicht eindeutig belegt. Dabei ist gerade dies maßgeblich für die Bekämpfung der Desinformation. Trotz eines geringeren Glaubens an die Falschnachricht könnten Menschen diese dennoch weiterleiten. Unklar bleibt, ob Prebunking-Maßnahmen das Vertrauen in wahre Nachrichten mindern und welche Rolle persönliche Faktoren und Einstellungen spielen. Gleichwohl ist Prebunking aus psychologischer Perspektive insgesamt eine geeignete Methode. Jedoch ist weitere Forschung ist notwendig, um die psychologischen Mechanismen der unterschiedlichen Prebunking-Maßnahmen zu verstehen und die Effektivität auf das Weiterleitungsverhalten zu verbessern.

4.2.2 Kommunikationswissenschaftliche Bewertung

Aus kommunikationswissenschaftlicher Perspektive scheinen Prebunking-Maßnahmen sinnvoll zu sein, um die Verbreitung von Desinformationen zu bekämpfen. Trotz der oben bereits angesprochenen ersten Bestätigungen, dass Prebunking hilfreich sein kann[39], müssen entsprechende Maßnahmen umsichtig geplant werden. Es existieren bereits eine Reihe von Erkenntnissen, die für die **Ausgestaltung von Prebunking-Maßnahmen** genutzt werden können. Sie sollten beispielsweise an typische Formen der Aufbereitung von Inhalten angepasst werden, die wir in unserer empirischen Forschung (s. Kap. 2.1) herausarbeiten konnten. Broad-Spectrum-Maßnahmen könnten z.B. über besonders emotionalisierte Schreibweisen, aber auch die für die Vermittlung von Desinformation häufig genutzte Fokussierung auf ein Schwerpunktthema, die Vermittlung alternativer Wirklichkeitskonstruktionen und das Angebot abgeschotteter Communitys aufklären. Diese Erkenntnisse könnten aber auch zur Erkennung von Desinformationskampagnen dienen, auf die dann mit einer Narrow-Spectrum-Maßnahme reagiert werden kann. Hier muss ebenfalls berücksichtigt werden, dass mit der Kuratierung (s. Kap. 2.1) eine Praxis existiert, die auf die Verbreitung von Desinformation hinweisen kann:

[36] Maertens, R., Roozenbeek, J., Basol, M., & Van Der Linden, S. (2021). Long-Term Effectiveness of Inoculation Against Misinformation: Three Longitudinal Experiments. *Journal of Experimental Psychology: Applied, 27*(1), 1–16. Verfügbar unter https://doi.org/10.1037/xap0000315 (abgerufen am 29.07.2024).

[37] Nickerson, R. S. (1998). Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology, 2*, S. 175–220. Verfügbar unter https://doi.org/10.1037/1089-2680.2.2.175 (abgerufen am 29.07.2024).

[38] Studienergebnisse deuten einerseits darauf hin, dass sowohl wahre als auch falsche Informationen als weniger glaubwürdig eingeschätzt werden könnten (siehe Modirrousta-Galian, A., & Higham, P. A. (2023). Gamified inoculation interventions do not improve discrimination between true and fake news: Reanalyzing existing research with receiver operating characteristic analysis. *Journal of Experimental Psychology: General, 152*(9), S. 2411–2437. Verfügbar unter https://doi.org/10.1037/xge0001395, abgerufen am 29.07.2024), während andere Studien dies nicht bestätigen (siehe Lu, C., Hu, B., Li, Q., Bi, C. & Ju, X. (2023). Psychological Inoculation for Credibility Assessment, Sharing Intention, and Discernment of Misinformation: Systematic Review and Meta-Analysis. *Journal of Medical Internet Research, 25*, e49255. Verfügbar unter https://doi.org/10.1080/10463283.2021.1876983/10.2196/49255 (abgerufen am 29.07.2024).

[39] Roozenbeek, J., van der Linden, S., Goldberg, B., Rathje, S. & Lewandowsky, S. (2022). Psychological Inoculation Improves Resilience Against Misinformation on Social Media. *Science Advances, 8*(34), eabo6254. Verfügbar unter https://doi.org/10.1126/sciadv.abo6254 (abgerufen am 29.07.2024).

Wirklichkeit wird auf desinformierenden Telegram-Kanälen nicht nur durch eigens verfasste Posts konstruiert, sondern auch zu einem bedeutenden Teil durch Auswahl und Kommentierung.

Im Frühjahr 2024 durchgeführte **Interviews** mit insgesamt neun **Expert*innen**, die in verschiedenen Fachdisziplinen auf dem Feld der Desinformationsbekämpfung tätig sind (Medienrecht, Politikberatung, Wissenschaft, staatliche Akteur*innen), bestätigen den für die Bekämpfung der Desinformationsverbreitung positiven Effekt. Bezüglich der Umsetzbarkeit seien Prebunking-Maßnahmen aus Sicht jener Expert*innen, die Interventionsmaßnahmen in ihrem Arbeitskontext anwenden und an die Zivilgesellschaft herantragen, sehr anschlussfähig. Zugleich werden aber auch Zweifel geäußert, wer am ehesten die Zielgruppe von Prebunking-Maßnahmen sein sollte und wie diese erreicht werden könne. Daran knüpft auch das Argument an, dass zumindest Broad-Spectrum-Prebunking-Maßnahmen erst auf lange Sicht und nicht kurzfristig wirken könnten. Hervorgehoben werden muss zudem, dass die Zielgruppe von Bekämpfungsmaßnahmen, insbesondere jene, die in staatskeptischen Gegenöffentlichkeiten verortet sind, teils nur schwer erreichbar ist und dies sowohl inhaltlich als auch technisch. Schließlich bestehen zudem erhebliche Zweifel an der Kooperationsbereitschaft von Diensteanbietern (allen voran der Messenger-Dienst Telegram).

4.2.3 Technische Bewertung

Aus technischer Sicht ist es wichtig, bei der Untersuchung von Prebunking-Maßnahmen zwischen dem Broad- und Narrow-Spectrum zu differenzieren. Gängige Manipulationstechniken und -strategien zu präsentieren (Broad-Spectrum), ist technisch leicht umsetzbar. **Pop-up-Mitteilungen, Informationsfeeds oder Kanäle** können Rezipient*innen vordefinierte Informationen technisch einfach umsetzbar näherbringen. Auch **spielerische Medienkompetenztrainings** (Gamification) können hier entwickelt werden.

Beim spezifischen „Impfen" im Narrow-Spectrum Bereich müssen hingegen erst Informationen über die jeweilige Desinformation gesammelt und analysiert werden. Soll dies über Nachrichten mit Themenschwerpunkten des spezifischen Kanals bzw. der spezifischen Gruppe geschehen, müssen die Messenger-Dienste **kontinuierlich überwacht** werden.[40] Die klassische Inokulation (zweigeteilte Struktur) wäre technisch also nur schwer umsetzbar. Einfacher wäre es, generell über aktuelle Desinformationskampagnen zu informieren, hierfür wäre auch keine fortlaufende Überwachung notwendig. Um beispielsweise Fakten über eine spezifische Nachricht in einem Messenger-Dienst ermitteln zu können, müsste eine **Datenbank aufgebaut** werden, mit der der Inhalt dieser Nachricht verglichen werden kann. Diese Datenbank müsste fortlaufend um weitere und aktuelle Informationen erweitert werden. Auch wenn der Aufbau der Datenbank erfolgreich wäre, wäre es aus technischer Sicht immer noch sehr anspruchsvoll, Daten aus der Datenbank und den Messenger-Diensten systematisch zu extrahieren. In diesem Kontext stellt sich besonders die Frage, auf welche Nachrichten Prebunking-

[40] Panahi, T., Hornung, G., Schäfer, K., Choi, J. E., Steinebach, M., & Vogel, I. (2023). Desinformationserkennung anhand von Netzwerkanalysen – ein Instrument zur Durchsetzung der Pflichten des DSA am Beispiel von Telegram. In *Daten-Fairness in einer globalisierten Welt* (Vol. 2, S. 343–370). Nomos-elibrary, 28, S. 343–370. Verfügbar unter https://www.nomos-elibrary.de/10.5771/9783748938743-343.pdf (abgerufen am 29.07.2024).

Maßnahmen angewendet werden sollen, denn es wäre ineffizient, alle Nachrichten auf mögliche Falschinformationen zu überprüfen.

Außerdem muss sorgfältig abgewogen werden, wie die Prebunking-Nachrichten ausgestaltet sein sollen. Sollen sie so nah wie möglich an die Falschinformation angepasst oder sollten eher allgemeine, themenbezogene Fakten dargestellt werden? Je nach Design würde die Schwierigkeit der technischen Implementierung unterschiedlich ausfallen. Wird beispielsweise eine Prebunking-Nachricht mit ähnlichem Inhalt wie die Falschnachricht bevorzugt, muss diese Ähnlichkeit messbar gemacht und entschieden werden, wie hoch diese Ähnlichkeit sein soll. Eine weitere Herausforderung stellt die **Reaktionszeit** dar, die abhängig von der erstellten Datenbank und dem Design ist und umso höher ausfällt, je stärker das Prebunking auf die aktuellen Desinformationen zugeschnitten sein soll.

4.2.4 Rechtliche Bewertung

Prebunking-Maßnahmen sind **bislang nicht gesetzlich verankert**. Zwar können Prebunking-Maßnahmen als mögliche Risikominderungsmaßnahme im Rahmen des DSA eingesetzt werden (Art. 35 DSA), allerdings wird Prebunking in dem Katalog an beispielhaften Maßnahmen nicht genannt, sodass kaum argumentiert werden kann, dass Anbieter hierzu verpflichtet sind. Auf der Ebene der als rechtlich unverbindlichen Selbstverpflichtungen im Rahmen **des EU-Verhaltenskodex gegen Desinformation** lässt sich immerhin eine Maßnahme so verstehen, dass Prebunking-Maßnahmen als Instrument eingeschlossen werden. Im Sinne der Broad-Spectrum-Methode verpflichten sich die unterzeichnenden Parteien nämlich dazu, Funktionen oder Initiativen zu entwickeln und umzusetzen, die die Nutzer*innen dazu befähigen, kritisch über die erhaltenen Informationen nachzudenken und ihnen helfen, zu überprüfen, ob diese korrekt sind (Commitment Nr. 25).

Jedoch scheint eine über eine bloße Selbstverpflichtung hinausgehende, **gesetzliche Verpflichtung** von Diensteanbietern zu Prebunking-Maßnahmen ein vielversprechender Ansatz zu sein. Aus rechtlicher Sicht können Prebunking-Maßnahmen **grundrechtsschonender** sein als viele andere Maßnahmen (s. Kap. 2.2.1.). Wie bereits dargestellt, können insbesondere das Löschen und Sperren von Inhalten oder Konten die Grundrechte auf Meinungs- und Informationsfreiheit stark beeinträchtigen. Auch die Kennzeichnung von Desinformation (Debunking, Flagging) kann in die Meinungsfreiheit eingreifen (s. 2.2.1). Viele der möglichen Maßnahmen würden zudem ein Filtern von privaten Nachrichten erfordern, was die Grundrechte auf private Kommunikation und auf Datenschutz beeinträchtigen würde.

Allerdings ist es nicht ausgeschlossen, dass auch Prebunking-Maßnahmen zur Beeinträchtigung dieser Grundrechte führen können. Grundsätzlich gilt: Eine Prebunking-Maßnahme ist grundrechtsverträglicher, je meinungsneutraler sie gestaltet ist, je mehr selbstbestimmtes Verhalten der Nutzer*innen möglich bleibt, je weniger externe Interventionen in den Kommunikationsverlauf stattfinden und je weniger personenbezogene Daten verarbeitet werden.

Prebunking als präventive Interventionsmaßnahme

Am grundrechtsverträglichsten scheinen daher Prebunking-Maßnahmen zu sein, welche allgemein die Medienkompetenz der Nutzer*innen durch die Technikgestaltung der Messenger-Dienste selbst fördern und ohne Echtzeit-Interventionen auskommen. Aus grundrechtlicher Sicht setzt eine Intervention idealerweise nicht bei einer konkreten Tatsachenbehauptung an, sondern bietet allgemein zugängliche, abstrakt gefasste Angebote zur Erweiterung der Medienkompetenz. Anstatt der Filterung oder Kennzeichnung von Inhalten, die z.B. bei bestimmten Schlagwörtern zum Zweck der Vorwarnung erfolgen, sollten Prebunking-Maßnahmen den Nutzer*innen **allgemein by design** zur Verfügung stehen (z.B. durch hervorgehobene Informationsfeeds/Kanäle/Stories, spielerische Medienkompetenz-Trainings). Zu bedenken ist auch, dass eine Kennzeichnung einzelner Inhalte mit dem Schutz der sog. negativen Meinungsfreiheit konfligieren würde (s. Kap. 2.2.1.). Dies kann durch die Broad-Spectrum-Methode verhindert werden, die **Meinungsneutralität gewährleisten** kann und im Hinblick auf die Meinungsfreiheit als milderes Mittel gegenüber der Narrow-Spectrum-Methode vorzuziehen ist. Der Vorschlag für eine Gesetzesformulierung könnte inhaltlich auf Commitment Nr. 25 des Verhaltenskodex gegen Desinformation basieren, da es hier schon einen festgeschriebenen, aber rechtlich unverbindlichen Konsens vieler Unternehmen, Fact-Checking-Organisationen und der EU-Kommission zu Broad-Spectrum-Prebunking-Maßnahmen gibt.

Prebunking-Maßnahmen würden gerade für die privaten, bislang wenig regulierten Kommunikationsfunktionen der Messenger-Dienste (One-to-one, Few-to-few) eine angemessene Lösung zur Bekämpfung von Desinformation darstellen, denn viele der möglichen Prebunking-Maßnahmen, vor allem im Broad-Spectrum, kommen ohne Eingriffe in das Fernmeldegeheimnis bzw. Grundrecht auf private Kommunikation aus. Prebunking-Maßnahmen passen auch zur hybriden Struktur vieler Dienste, die öffentliche und private Kommunikationsfunktionen in einer gemeinsamen technischen Anwendungsfläche verknüpfen und zur Netzwerkbildung beitragen. Eine Verpflichtung zu Prebunking-Maßnahmen wäre eine Möglichkeit, den **gesamten Messenger-Dienst zu regulieren** und somit Desinformation in gesamten Nutzer*innen-Netzwerken einzudämmen.[41]

Zudem sollten Diensteanbieter zur **grundrechtskonformen Technikgestaltung** der eingesetzten Prebunking-Maßnahmen verpflichtet werden. Eine Überwachung und ein Filtern von privaten Nachrichten sollte explizit ausgeschlossen werden.

Dienste-Anbieter sollten darüber hinaus auch zu **transparenten Allgemeinen Geschäftsbedingungen** verpflichtet werden, die etwa Auskunft über die eingesetzten Prebunking-Mechanismen geben. Art. 14 Abs. 1 DSA ist hierfür noch nicht ausreichend, da dieser nur die Transparenz von Beschränkungen des Dienstes fordert, wozu Prebunking nicht gezählt werden kann.

Grundsätzlich sollte eine möglichst **technikneutrale Gesetzesformulierung** gewählt werden, um die Kreativität und damit vielfältige Innovationen der Diensteanbieter bei der Entwicklung von Prebunking-Maßnahmen zu fördern. Das heißt, dass keine

[41] Sollte der Vorschlag bei einer Gesetzesnovelle des DSA berücksichtigt werden, müsste die neue Vorschrift konsequenterweise nicht in den Katalog der Pflichten für Online-Plattformen fallen, sondern in die allgemeinen Vorschriften für Hosting-Dienste, damit auch die privaten Kommunikationsfunktionen von Messenger-Diensten umfasst wären.

konkreten Prebunking-Maßnahmen gesetzlich vorgeschrieben werden sollten, sondern durch offene Formulierungen auch neue Technologien die gesetzlichen Voraussetzungen erfüllen könnten. Dadurch würde das Gesetz auch weniger in das Grundrecht auf Berufsfreiheit bzw. unternehmerische Freiheit der Messenger-Diensteanbieter aus Art. 12 Abs. 1 GG, Art. 16 GRCh eingreifen.

Schließlich sollten die Messenger-Diensteanbieter zur **(externen) Evaluation** ihrer Fortschritte verpflichtet sein. Zwar wären Daten aus den privaten Kommunikationsfunktionen von Messenger-Diensten aufgrund des Grundrechts auf informationelle Selbstbestimmung und des Fernmeldegeheimnisses voraussichtlich nur über freiwillige Datenspenden und freiwillige experimentelle Settings erhältlich, jedoch können Diensteanbieter verpflichtet werden, über die Entwicklung ihrer technischen Funktionen Bericht zu erstatten.

Eine Verpflichtung zu Prebunking-Maßnahmen im Sinne der Broad-Spectrum-Methode könnte grundsätzlich auf nationaler Ebene gesetzlich verankert werden. Zu prüfen wäre dann allerdings, ob der DSA für die öffentlichen Kommunikationsfunktionen der Messenger-Dienste abschließende Regelungen trifft, die eine Sperrwirkung entfalten und wie viel Raum für nationale Bestimmungen bleibt. Aufgrund des Ziels des DSA die Plattformregulierung im Binnenmarkt der EU zu vereinheitlichen (vgl. Art. 1 S. 1 DSA, EG 9 DSA), ist grundsätzlich von einer weitreichenden Sperrwirkung auszugehen. Denkbar wäre aber auch, eine solche **gesetzliche Neuerung auf EU-Ebene** anzuregen. Begründet werden könnte dies – wie für die Regelungen des DSA – grundsätzlich durch die Binnenmarktkompetenz der EU aus Art. 114 AEUV, die ggf. erfordern kann, dass die Vorschriften für Vermittlungsdienste, zu denen Messenger-Dienste auch zählen, im Binnenmarkt weiter harmonisiert werden.

4.2.5 Zusammenfassende Bewertung

Die Eignung von Narrow-Spectrum- und Broad-Spectrum-Maßnahmen wird von den verschiedenen Fachdisziplinen unterschiedlich bewertet.

Aus psychologischer Perspektive ist Prebunking insgesamt eine geeignete Methode, deren Wirkung jedoch von der Art der Anwendung und den individuellen Faktoren der Rezipient*innen abhängt. Prebunking-Maßnahmen können die Resistenz gegenüber Desinformation erhöhen und die Glaubwürdigkeit von falschen Informationen reduzieren. Jedoch ist weitere Forschung notwendig, z.B. über etwaige negative Auswirkungen auf die Glaubwürdigkeit von akkuraten Fakten oder den Einfluss auf die Weiterleitung von Falschinformationen. Ebenso muss die relative Wirksamkeit von Narrow- und Broad-Spectrum-Methoden weiter erforscht werden, da sowohl einzelne Studien als auch die theoretischen Grundlagen nahelegen, dass Narrow-Spectrum-Prebunking wirkungsvoller ist.

Unsere kommunikationswissenschaftlichen Expert*innen-Interviews zeigen, dass Prebunking-Maßnahmen grundsätzlich als anschlussfähig betrachtet werden, wobei zugleich Zweifel an der Bestimmung und Erreichbarkeit der Zielgruppen, der kurzfristigen Wirksamkeit und der Kooperationsbereitschaft von Messenger-Diensten bestehen.

Aus Sicht der Informatik sind Broad-Spectrum-Maßnahmen technisch leichter umzusetzen. Demgegenüber führen Narrow-Spectrum-Maßnahmen zu der Herausforderung Datenbanken aufzubauen und Methoden zu entwickeln, die eine kurzfristige und repräsentative Extraktion von Nachrichten erlaubt. Probleme bestehen zudem in der Messbarkeit der Desinformationsverbreitung.

Aus rechtlicher Sicht würde eine gesetzliche Verpflichtung zu Broad-Spectrum-Maßnahmen eine grundrechtsschonende und minimalinvasive Maßnahme darstellen. In Kommunikationsräumen von Messenger-Diensten, die aufgrund des Fernmeldegeheimnisses bzw. dem Schutz der privaten Kommunikation nicht für Überwachungsmaßnahmen zugänglich sind und sein sollten, stellt eine Verpflichtung der Diensteanbieter zu Prebunking-Maßnahmen im Vergleich zu anderen Maßnahmen (z.B. das Löschen, Filtern oder Kennzeichnen von Chatnachrichten) folglich ein relativ milderes Mittel dar.

Sollte eine Verpflichtung zum Prebunking tatsächlich gesetzlich verankert oder von den Diensteanbietern freiwillig umgesetzt werden, sollten sowohl die empirischen Erkenntnisse als auch rechtlichen Wertungen berücksichtigt werden. Narrow- und Broad-Spectrum-Methoden sollten aber weiter erforscht werden, um möglichst wirksame und grundrechtskonforme Prebunking-Maßnahmen zu entwickeln. Zwei Punkte müssen dabei durch zukünftige Forschung im Blick behalten werden: Auch wenn die meisten hier diskutierten Perspektiven eher eine Broad-Spectrum-Prebunking-Lösung empfehlen, ist nicht endgültig empirisch geklärt, ob Narrow-Spectrum-Prebunking nicht deutlich bessere Effekte erzielt. Damit zusammenhängend muss im Rahmen ethischer Reflektionen diskutiert werden, inwieweit die Steigerung der Medienkompetenz im Rahmen von Broad-Spectrum-Maßnahmen nicht wiederum die Verantwortung in hohem Maße bei einzelnen Nutzenden verortet, die weitergebildet werden müssen. Es muss daher eine Balance gefunden werden zwischen technischen Möglichkeiten, grundrechtsbezogenen Abwägungen und dem Schutz der Bevölkerung, der wiederum für diese ressourcenschonend (d.h. zum Beispiel ohne mental load durch Kompetenzaneignung) wirkt.

5 Fazit

Staatsskeptische und verfassungsfeindliche Gegenöffentlichkeiten bilden sich auf Messenger-Diensten, indem über Kanäle und Gruppen Desinformationen verbreitet werden. Hier sind Gegenmaßnahmen erforderlich, die auf rechtlicher, technischer und zivilgesellschaftlicher Ebene nachjustiert werden müssen.

In diesem Policy Paper haben wir gezeigt, dass bereits einige rechtliche Maßnahmen existieren, die die Verbreitung von Desinformation einzudämmen versuchen. Die Vorschriften sind jedoch nicht ausreichend auf die hybride Struktur von Messenger-Diensten ausgelegt, sodass es hier Konkretisierungsbedarf gibt. Zudem bleiben die privaten Kommunikationsfunktionen (one-to-one, few-to-few) weitgehend unreguliert. Daher besteht aus rechtlicher und kommunikationswissenschaftlicher Perspektive sowie aus medienpsychologischer und technischer Sicht zugleich Ergänzungsbedarf zu bestehenden Maßnahmen.

Wir kommen zu dem Schluss, dass der Fokus nicht nur auf repressiven Maßnahmen, sondern vor allem auf **Ursachenbekämpfung und Prävention** liegen sollte, da die Ursachen der Desinformationsverbreitung auf Messenger-Diensten **besser eruiert** werden müssen. Um die Entstehung und das Inszenieren von staatsskeptischen Gegenöffentlichkeiten, die **den öffentlichen Diskurs nicht primär pluralisieren, sondern gefährden,** zu verhindern, ist es notwendig, die dahinterliegenden Motive besser zu verstehen. Dies kann nur dann gelingen, wenn der Dialog zu jenen nicht abreißt, die erhebliche Zweifel an etablierten Medien und politischen Institutionen hegen, sich von demokratischen Systemen abwenden und sich zunehmend und teils ausschließlich über Messenger-Dienste informieren. Dazu müssen die kommunikativen **Vorteile der persönlichen Ebene in den sozialen Dialog** mitgenommen werden: um das Entstehen von, Verbleiben in und die Dynamik von staatsskeptischen Gegenöffentlichkeiten, die sich auf Messenger-Diensten vernetzen, besser zu verstehen, gilt es, auf Augenhöhe miteinander ins Gespräch zu kommen und nicht nur „über" die in den Gegenöffentlichkeiten Agierenden zu sprechen.

Des Weiteren müssen es Akteur*innen aus Journalismus und Politik künftig stärker **vermeiden, den argumentativen Strategien desinformierender Akteur*innen zu verfallen** und dies nicht nur, aber insbesondere auch dann, wenn sie selbst Messenger-Dienste in der eigenen Kommunikation einsetzen. In unserem Informationsökosystem, das Journalismus, Politik, Plattformen und individuelle Nutzer*innen umfasst, sind politische und journalistische Akteur*innen besonders bedeutsam bei der Verbreitung und Bekämpfung von Desinformation.

Die Politik trägt dabei eine doppelte Verantwortung: Sie muss sowohl die **Voraussetzungen schaffen, um Desinformationen zu bekämpfen**, als auch sicherstellen, dass sie **selbst keine Desinformation verbreitet**. Besonders während Wahlkämpfen besteht das Risiko, dass Informationen gezielt verwendet werden, um politische Ziele zu fördern. Dies kann nur verhindert werden, indem in **Journalismus und Politik Narrative**

Fazit

überdacht und neue etabliert werden, um populistische Strategien, wie sie auch von desinformierenden Akteur*innen genutzt werden, aus dem demokratischen Diskurs fernzuhalten. Dies ist besonders in Krisensituationen sowie vor (anstehenden) Wahlen nötig und kann mit mehr systemischen Lösungen gelingen, die kurzfristig greifen und umgesetzt werden können.

Mit den vorgeschlagenen Prebunking-Maßnahmen empfehlen wir daher für Politik, Forschung und Akteur*innen der Desinformationsbekämpfung ein präventives Maßnahmenbündel. Vor dem Hintergrund, dass nicht alle Segmente der Zivilgesellschaft (bzw. nur teilweise) erreicht werden können, wird durch breitgefasste Prebunking-Maßnahmen die Förderung von Medienkompetenz sinnvoll erweitert. Medienkompetenz by Design bedeutet in diesem Zusammenhang, dass Maßnahmen Akteur*innen innerhalb und außerhalb staatsskeptischer Gegenöffentlichkeiten gleichermaßen erreichen und so den konstruktiven Diskurs sowie den gesellschaftlichen Zusammenhalt fördern, um der weiteren Verbreitung von Desinformation auch in fluiden Kommunikationsformen effektiv entgegen treten zu können.

Um die Umsetzbarkeit und Effektivität der wissenschaftlich erforschten Handlungsempfehlungen beurteilen und bei Bedarf schärfen zu können, ist zudem die Förderung weiterer wissenschaftliche Forschung notwendig. Diese sollte untersuchen, ob das Broad-Spectrum- oder das Narrow-Spectrum-Prebunking gegen die Weiterleitung von Desinformationen in Messenger-Diensten wirksamer ist. Dabei gilt es, die empirische Eignung, technische Möglichkeiten und grundrechtliche Abwägungen gleichermaßen zu berücksichtigen und eine sinnvolle Verantwortungsteilung zwischen Diensteanbietern, zivilgesellschaftlichen Akteur*innen und Einzelpersonen zu finden.

DYNAMO

Ein gemeinsames Projekt von

Fraunhofer SIT

HOCHSCHULE DER MEDIEN

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

UNIKASSEL VERSITÄT

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

DOI: 10.17185/duepublico/82406
URN: urn:nbn:de:hbz:465-20240917-161748-9