Desinformationserkennung anhand von Netzwerkanalysen – ein Instrument zur Durchsetzung der Pflichten des DSA am Beispiel von Telegram

Tahireh Panahi, Gerrit Hornung, Karla Schäfer, Jeong-Eun Choi, Martin Steinebach und Inna Vogel

Zusammenfassung

Desinformationen werden verstärkt zu Krisenzeiten wie z. B. während der Covid-19 Pandemie oder dem Angriffskrieg auf die Ukraine erstellt und verbreitet. Propagandistische Akteure aus dem Ausland, aber auch z. B. extremistische Gruppen im Inland verwenden für ihre Desinformationskampagnen soziale Medien. Telegram stellt einen fast unmoderierten Kommunikationsdienst dar, der die Möglichkeit der nahezu ungestörten Verbreitung von Desinformationen ermöglicht. Unter anderem, um der Verbreitung von falschen und irreführenden Tatsachenbehauptungen entgegenzutreten, wurde von der EU der Digital Services Act (DSA) erlassen. Für die Erfüllung der darin angeordneten risikobezogenen Pflichten wird in diesem Beitrag das Instrument der Netzwerkanalyse vorgeschlagen und anhand des Hybrid-Mediums Telegram näher erklärt. Desinformationen können durch eine Netzwerkanalyse zwar auf inhaltlicher Ebene nicht direkt erkannt werden; in einem nutzerstarken Dienst wie Telegram ist es ist aber möglich, die Verbindungen zwischen Akteuren zu ermitteln und zu charakterisieren. Die Netzwerkanalyse kann daher als ein erstes Tool zur Desinformationserkennung im Rahmen der risikobezogenen Pflichten des DSA eingesetzt werden, verursacht aber auch rechtliche und technische Herausforderungen bei der Umsetzung.[1]

1 Dieser Beitrag ist im Rahmen des BMBF-Verbundprojekts „DYNAMO – Dynamiken der Desinformation erkennen und bekämpfen“ (FKZ: 16KIS1498) entstanden.

1. Einleitung

Desinformation ist ein wachsendes Problem in der heutigen Online-Welt, das darin besteht, dass falsche oder irreführende Informationen gezielt verbreitet werden, um eine bestimmte, typischerweise illegitime Agenda zu unterstützen. Besonders der Dienst Telegram steht für die Duldung massenhafter Desinformation in der Kritik. Um Desinformation entgegenzutreten, hat die EU mit dem DSA[2] neue Vorschriften erlassen. Insbesondere die für sehr große Online-Plattformen geltenden risikobezogenen Pflichten können einen Beitrag zur langfristigen Erkennung und Bekämpfung von Desinformation leisten.

Eine offene Frage ist bislang jedoch, wie die in diesem Rahmen geforderten Risikobewertungen technisch umgesetzt werden können. Ein Instrument, das hierzu verwendet werden kann, ist die Netzwerkanalyse. Durch diese werden Interaktionen in Online-Plattformen analysiert, sodass Muster erkannt werden können, die auf eine Desinformationskampagne hindeuten. Die Netzwerkanalyse ermöglicht es damit, die Verbreitung von Desinformation besser zu verstehen und ihr entgegenzuwirken. In diesem Beitrag werden die risikobezogenen Pflichten des DSA der technischen Umsetzung durch die Netzwerkanalyse gegenübergestellt. Abschließend werden rechtliche und technische Probleme diskutiert. Zur Konkretisierung und Veranschaulichung wird das Beispiel des Dienstes Telegram herangezogen.

2. Telegram

Der Dienst Telegram ist in den letzten Jahren durch die Duldung massenhafter Desinformation negativ aufgefallen.[3] Die Betreiber verfolgen einen sehr weitgehenden „Free-Speech"-Ansatz[4], bei dem Sperrungen, Löschun-

2 Verordnung (EU) 2022/2065 des Europäischen Parlaments und des Rates vom 19. Oktober 2022 über einen Binnenmarkt für digitale Dienste und zur Änderung der Richtlinie 2000/31/EG.

3 *Borscher/Woltert*, Monitor Medienpolitik – Ort der Meinungsfreiheit und Verschwörungsplattform, 2022; *Jünger/Gärtner*, Die Verbreitung und Vernetzung problembehafteter Inhalte auf Telegram, 2021; zum Aufstieg Telegrams: *Denga*, EuR 2021, 569 (574).

4 Die Redefreiheit ist hier im Sinne eines Free-Speech-Absolutism gemeint. Dieses Verständnis einer schrankenlosen Redefreiheit ist nicht gleichbedeutend mit der Meinungsfreiheit nach Art. 5 Abs. 1 GG bzw. Art. 11 GRCh.

gen und andere Moderationsmaßnahmen nur in seltenen Ausnahmefällen durchgeführt werden.[5] Verschiedene Länder, so auch die Bundesrepublik Deutschland, versuchten in jüngerer Zeit durchzusetzen, dass der Dienst aktiver gegen Desinformation vorgeht.[6] Dies hatte zumindest teilweise Erfolg; Telegram hat seine Bereitschaft zur Kooperation mit staatlichen Behörden in bestimmten Konstellationen angepasst.[7]

Bei Telegram handelt es sich um ein sog. Hybrid-Medium, da der Dienst neben rein interpersonaler Kommunikation auch Austausch in öffentlichen Gruppen und Kanälen anbietet.[8] Gruppen können dabei eine Größe von bis zu 200.000 Mitgliedern[9] annehmen; Kanäle sind unbegrenzt in ihrer Abonnentenzahl. Während in Gruppen jedes Mitglied Nachrichten veröffentlichen kann, ist dies in Kanälen grundsätzlich den Administratoren vorbehalten; Nutzer benötigen hierzu die Genehmigung des Administrators. Ein Kanal ist daher zur einseitigen Informationsverbreitung bzw. reinen Konsumierung von Informationen gedacht.

3. Der Digital Services Act

Mit dem DSA erfolgt eine vollständige Harmonisierung der Vorschriften für Vermittlungsdienste innerhalb des EU-Binnenmarkts (vgl. ErwG 9 S. 1). Als Vermittlungsdienste gelten gem. Art. 3 lit. g DSA Dienste zur reinen Durchleitung, Caching-Dienste und Hosting-Dienste. Als Verordnung i. S. d. Art. 288 Abs. 2 AEUV gilt der DSA unmittelbar in allen Mitgliedsstaaten der EU. Der Geltungsbeginn ist in Artt. 92 f. DSA innerhalb abgestufter Fristen vorgesehen.

Ziel der Verordnung ist ein sicheres, vertrauenswürdiges Online-Umfeld zu gewährleisten, das unter anderem den gesellschaftlichen Risiken durch Desinformation entgegenwirken soll (vgl. ErwG 9 DSA). Bemerkenswerterweise wird Desinformation weder im DSA, noch in sonstigen EU-Rechtsakten definiert. Stattdessen können lediglich rechtlich nicht bindende Stellungnahmen der EU herangezogen werden. Diese definieren Desinformati-

5 Vgl. *Telegram.org*, FAQ.
6 *Zeit.de*, Brasiliens oberstes Gericht nimmt Telegram-Sperrung zurück.
7 *Balser,* Telegram sperrt 64 Kanäle.
8 *Jünger/Gärtner*, Datenanalyse von Rechtsverstoßenden Inhalten in Gruppen und Kanälen von Messengerdiensten am Beispiel Telegram, 2020.
9 Vgl. *Telegram.org*, FAQ.

on als nachweislich falsche oder irreführende Informationen , die mit dem Ziel des wirtschaftlichen Gewinns oder der vorsätzlichen Täuschung der Öffentlichkeit konzipiert, vorgelegt und verbreitet werden und öffentlichen Schaden anrichten können.[10]

Vermittlungsdiensten wird in Kapitel 3 des DSA eine breite Palette unterschiedlicher Sorgfaltspflichten auferlegt, die zum Teil auch Desinformation adressieren. Gerade der risikobasierte Ansatz der Regelungen aus Kapitel 3 Abschnitt 5 DSA erfordert den Einsatz passender technischer Analyse-Instrumente, wie die in diesem Beitrag erläuterte Netzwerkanalyse.

3.1 Anwendungsbereich

Zunächst ist darzustellen, unter welchen Bedingungen Dienste wie Telegram in den Anwendungsbereich des DSA fallen. Dieser ist gem. Art. 2 Abs. 1, Art. 3 lit. g DSA für Vermittlungsdienste eröffnet (s.o.). Telegram und ähnliche Dienste sind i. d. R. jedenfalls als Hosting-Dienste zu qualifizieren, da sie gem. Art. 3 lit. g iii) DSA Informationen im Auftrag der Nutzer speichern.

Zu beachten ist, dass es sich bei Telegram um ein sog. Hybrid-Medium handelt, das interpersonale Kommunikation, Austausch in Gruppen und öffentliche Kanäle anbietet (s.o.). Die Teile von Telegram, die die gespeicherten Informationen gem. Art. 3 lit. k DSA öffentlich verbreiten, also für eine potenziell unbegrenzte Zahl von Dritten bereitstellen, können als Online-Plattformen i. S. d. Art. 3 lit. i DSA gelten (dies ist eine Untergruppe der Hosting-Dienste). Auch Gruppen, die durch einen öffentlich geposteten Einladungslink erweitert werden, stellen gem. ErwG 14 öffentliche Kommunikation dar, da die Nutzer, die auf die Informationen zugreifen möchten, automatisch registriert oder aufgenommen werden.[11] Funktionen, die „Instant-Messaging-Dienste" darstellen, können hingegen gem. ErwG

10 *Europäische Kommission*, Mitteilung zur Bekämpfung von Desinformation im Internet (COM(2018) 236 final); *Europäische Kommission*, Aktionsplan gegen Desinformation (JOIN(2018) 36 final).

11 An der Vorschrift des DSA kritisch zu betrachten ist jedoch, dass besonders mitgliederstarke geschlossene Gruppen mangels Verbreitung an eine potenziell unbegrenzte Zahl Dritter keine Online-Plattformen darstellen können (bei Telegram bis zu 200.000 Mitglieder möglich), vgl. *Setz*, in: Bernzen u.a. (Hrsg.), Immaterialgüter und Medien im Binnenmarkt Europäisierung des Rechts und ihre Grenzen, 2022, 175 (188 f.).

14 sogar ausdrücklich keine Online-Plattformen sein.[12] Eine solche funktionsbasierte Einordnung der Dienste in den Anwendungsbereich ist gem. ErwG 13-15 DSA vorgesehen.[13]

Wenn Online-Plattformen eine durchschnittliche aktive Nutzerzahl von monatlich mindestens 45 Millionen in der Union aufweisen, können sie gem. Art. 33 Abs. 1, 4 DSA von der EU-Kommission als „sehr große Online-Plattform"[14] benannt und damit den in diesem Beitrag relevanten Risikobewertungspflichten unterworfen werden.[15] Online-Plattformen (mit Ausnahme von Mikro- und Kleinunternehmen) waren und sind gem. Art. 24 Abs. 2 DSA dazu verpflichtet, bis zum 17. Februar 2023 – danach mindestens alle sechs Monate – Informationen über die durchschnittliche monatliche Zahl ihrer aktiven Nutzer in der Union zu veröffentlichen.[16] Nach eigenen Angaben erreichte Telegram den Schwellenwert von 45 Millionen aktiven Nutzern in der EU zum Stichtag am 17. Februar 2023 nicht, sondern schätzt seine durchschnittliche Nutzerzahl auf 38 Millionen.[17] Telegram macht zudem darauf aufmerksam, dass die Zahl tatsächlich noch niedriger sein könnte, da nur einige der Funktionen als Online-Plattformen i. S. d. DSA eingestuft werden könnten (s.o.).[18]

Der Koordinator für digitale Dienste[19] und die EU-Kommission können nun gem. Art. 24 Abs. 3 DSA zusätzliche Informationen über die in jenem

12 Dazu auch: *Gielen/Uphues*, EuZW 2021, 627 (634); *Kuhlmann/Trute*, GSZ 2022, 115 (115).

13 Anders als bei vergleichbaren deutschen Regelwerken, ist der Anwendungsbereich des DSA auch für hybride Dienste wie Telegram eröffnet. Während etwa bei NetzDG und MStV der Anwendungsbereich aufgrund starrer Legaldefinitionen von Dienste-Typen nicht ohne weiteres für diese doppelfunktionalen Dienste eröffnet ist, sieht der DSA in ErwG 12 ff. ausdrücklich eine funktionsbasierte Einordnung der Dienste in den Anwendungsbereich vor, vgl. *Setz*, in: Bernzen u.a. (Hrsg.), Immaterialgüter und Medien im Binnenmarkt. Europäisierung des Rechts und ihre Grenzen, 2022, 175 (189); so auch: *Gielen/Uphues*, EuZW 2021, 627 (635); *Eisenreich*, RDi 2021, 289 (289 f.); *Kalbhenn*, ZUM 2022, 266 (272); zur Einordnung von Gruppen: *Spindler*, GRUR 2021, 653 (654).

14 Gebräuchlich ist auch die Abkürzung des englischen Begriffs „VLOP" („Very large Online-Plattform"); s. dazu auch *Kuß/Lehmann*, DB 2021, 605 (607).

15 Die genaueren Modalitäten der Berechnung werden in ErwG 76 f. DSA beschrieben.

16 *Europäische Kommission*, Presseerklärung vom 17.02.23.

17 *Telegram.org*, FAQ.

18 *Telegram.org*, FAQ.

19 Gem. Art. 49 Abs. 3 DSA haben die Mitgliedstaaten bis zum 17.2.2024 jeweils eine Behörde als Koordinator für digitale Dienste zu benennen. Die Behörden müssen die Anforderungen des Art. 50 DSA erfüllen (unter anderem Unabhängigkeit, Unparteilichkeit, Transparenz).

Absatz genannte Berechnung sowie Erläuterungen und Begründungen in Bezug auf die verwendeten Daten verlangen. Zu prüfen wäre dabei, ob Telegram bei dieser Schätzung bereits die eher „großzügigen" Berechnungsvorgaben des ErwG 77 DSA angewendet hat. Danach gelten bereits alle diejenigen als aktive Nutzer, die den Dienst in dem bestimmten Zeitraum mindestens einmal in Anspruch nehmen, z. B. indem sie Informationen bereitstellen oder diesen auch nur ausgesetzt sind, wobei eine Registrierung beim Dienst nicht erforderlich ist und auch eine einmalige Nutzung ausreicht. Zu berücksichtigen sind nach ErwG 77 alle Online-Schnittstellen wie Websites oder Anwendungen. Nutzer müssen zwar nicht aktiv mit der Information interagieren. Eine nur gelegentliche indirekte Nutzung des Dienstes durch Nutzer anderer Anbieter von Vermittlungsdiensten reicht jedoch nicht aus.

Gem. Art. 24 Abs. 3 DSA dürfen bei der Übermittlung zusätzlicher Informationen an die genannten Behörden keine personenbezogenen Daten enthalten sein. Dies ist bei einem Hybrid-Medium (s.o.) wie Telegram problematisch, da der Anbieter berechnen muss, wie viele Nutzer die Funktionen, die als Online-Plattform gelten, nutzen, und welche nur die Messenger-Funktionen, die nicht unter den Anwendungsbereich des DSA fallen (s.o.).[20] Bei einer solch nutzer-spezifischen Unterscheidung fallen regelmäßig personenbezogene Daten an; z.B. kann eine Gruppen- oder Kanalzugehörigkeit Daten über weltanschauliche und politische Einstellungen enthalten. Inwiefern die Daten in anonymisierter oder pseudonymisierter Form übermittelt werden können, ohne die Aussagekraft der Berechnung zu gefährden, bleibt fraglich.

Schließlich ist festzuhalten, dass selbst wenn Telegram aktuell die Schwelle noch nicht erreicht, eine Neuberechnung alle sechs Monate zu erfolgen hat und zumindest damit gerechnet werden muss, dass die Zahlen mittelfristig ansteigen. Selbst wenn dies nicht der Fall sein sollte, könnte die im Folgenden erläuterte Netzwerkanalyse für den Anbieter sinnvoll und nützlich sein, z. B. für freiwillige desinformationsbezogene Untersuchungen oder zu Forschungszwecken.

20 *Jung*, DÖV 2023, 141 (147).

3.2 Risikobezogene Pflichten

Der DSA enthält unterschiedliche Pflichten, die sich zum Teil repressiv oder präventiv gegen die Verbreitung von Desinformation richten.[21] Zu nennen sind etwa Vorgaben zur Gestaltung und Durchführung von AGB (Art. 14 DSA), Transparenzpflichten (vgl. Artt. 15, 24, 38, 39, 42 DSA), Melde- und Abhilfeverfahren (Art. 16 ff. DSA) und die Deplatforming-Pflicht (Art. 21 Abs. 1 DSA).[22] Neu für das Recht der Vermittlungsdienste sind eine Reihe risikobezogener Pflichten aus Artt. 34 ff. DSA, die indes nur für „sehr große Online-Plattformen" (s.o.) gelten.[23] Die weiteren Ausführungen beschränken sich auf diese risikobezogenen Pflichten, da die in diesem Beitrag vorgestellte Netzwerkanalyse vor allem als Instrument zu ihrer technischen Umsetzung eingesetzt werden kann.

Risikobasierter Ansatz

Der zu Grunde liegende risikobasierte Ansatz wird in ErwG 75f. beschrieben.[24] Danach kommt sehr großen Online-Plattformen auf Grund ihrer Reichweite eine bedeutende Rolle z. B. für die öffentliche Debatte zu, was gesellschaftliche Risiken bewirken kann. Gem. ErwG 79 sind auch die Art und Weise, in der solche sehr großen Online-Plattformen genutzt werden können, unter anderem für die Online-Sicherheit, die öffentliche Meinungsbildung und den öffentlichen Diskurs von Bedeutung. Die Reichweite wird gerade bei der Verbreitung von Desinformation als kritischer Faktor betrachtet.[25]

21 Berberich und Seip sahen Desinformation im DSA-E nicht ausreichend berücksichtigt: *Berberich/Seip*, GRUR-Prax 2021, 4 (7). Die finale Verordnung enthält keine wesentlichen Änderungen im Bereich der Desinformation.

22 Zur Effektivität gegen Desinformation *Setz*, in: Bernzen u.a. (Hrsg.), Immaterialgüter und Medien im Binnenmarkt. Europäisierung des Rechts und ihre Grenzen, 2022, 175 (190 ff.).

23 Im nationalen Recht enthält § 2 NetzDG bislang Berichtspflichten, die sich zumindest teilweise mit den Risikobewertungspflichten des DSA überschneiden, jedoch auf rechtswidrige Inhalte beschränkt sind.

24 Zum risikobasierten Ansatz: *Achleitner*, MR-Int 2022, 114, (116); *Rau et al.*, Rechtsextreme Online-Kommunikation in Krisenzeiten, 2022, S. 8.

25 *Achleitner*, MR-Int 2022, 114 (116).

Risikobewertung, Art. 34 DSA

Nach Art. 34 Abs. 1 S. 1 DSA sind sehr große Online-Plattformen[26] dazu verpflichtet, bestimmte systemische Risiken in der Union, die sich aus der Konzeption, dem Betrieb und der Nutzung ihrer Dienste sowie eingesetzter algorithmischer Systeme ergeben, sorgfältig zu ermitteln, zu analysieren und zu bewerten.[27] ErwG 84 stellt klar, dass Anbieter dabei auch besonders darauf achten sollten, wie ihre Dienste zur Verbreitung oder Verstärkung von Desinformation genutzt werden. Die Risikobewertung erfolgt gem. Art. 34 Abs. 1 S. 2 DSA mindestens einmal jährlich, jedenfalls aber vor jeder Einführung risiko-relevanter Funktionen.

Die in Art. 34 Abs. 1 S. 3 lit. a – d DSA aufgezählten systemischen Risiken stehen zum Teil in unmittelbarem, zum Teil in mittelbarem Bezug zu Desinformation. In Art. 34 Abs. 1 S. 3 lit. a DSA werden Risiken genannt, die durch die Verbreitung rechtswidriger Inhalte entstehen. Manche mitgliedsstaatlichen Regelungen sind so ausgestaltet, dass bestimmte Formen von Desinformation als rechtswidrig gelten (z. B. in Deutschland der Tatbestand der Volksverhetzung nach § 130 StGB, jeweils in der Variante des „Verleumdens" oder „Leugnens"). Dabei ist allerdings darauf zu achten, dass nach dem DSA nur als systemisch einzustufende Risiken zu bewerten sind, also solche, die eine gewisse qualitative und/oder quantitative Erheblichkeitsschwelle erreichen.[28] Daneben bietet der DSA eine Reihe anderer Maßnahmen, die nicht systemische Risiken, sondern einzelne Inhalte und Akteure betreffen und im Rahmen dieses Beitrags nicht vertieft werden können (z.B. die Pflicht zur Vorhaltung eines Melde- und Abhilfeverfahrens nach Art. 16 DSA und die Aussetzungspflicht nach Art. 23 DSA).

In Art. 34 Abs. 1 S. 3 lit. c DSA werden alle nachteiligen Auswirkungen auf die gesellschaftliche Debatte und Wahlprozesse genannt. Diese können grundsätzlich durch Desinformation unmittelbar beeinträchtigt werden, indem der zu Grunde liegende freie Willensbildungsprozess verzerrt wird. Problematisch ist, dass es sich bei dem Begriff „gesellschaftliche Debatte" um eine sehr weite Formulierung handelt, die potenziell jegliche Themen und Debattenformen beinhalten kann. Um ausufernde und unbestimmte Pflichten für die Anbieter zu vermeiden, müssen einschränkende inhalt-

26 Diese Pflichten betreffen ebenso „sehr große Online-Suchmaschinen".
27 Zum Sinn und Zweck der risikobezogenen Pflichten auch: *Kalbhenn*, ZUM 2022, 266 (273).
28 *Janal*, ZEuP 2021, 227 (266).

liche und quantitative Kriterien für die Schwelle eines systemischen Risikos hinzugezogen werden.[29] ErwG 79 nennt beispielhaft die Anzahl der betroffenen Personen, die Unumkehrbarkeit und den Wiederherstellungsaufwand. Zu bewerten ist danach sowohl die Schwere als auch die Wahrscheinlichkeit dieser Risiken.

Auch die in Art. 34 Abs. 1 S. 3 lit. b DSA genannten Risiken für die Ausübung der Grundrechte können in Bezug auf Desinformation relevant sein. Hervorgehoben werden etwa Meinungs- und Informationsfreiheit, Medienfreiheit und -pluralismus aus Art. 11 Abs. 1 bzw. 2 GRCh. Durch Desinformation kann der Prozess der individuellen Meinungsbildung verzerrt werden. Umgekehrt können aber auch Maßnahmen die von Online-Plattformen gegen Desinformation eingesetzt werden, die Grundrechte der Nutzer beeinträchtigen und damit ein systemisches Risiko darstellen. Die Risikobewertung nach Art. 34 Abs. 1 S. 3 lit. b DSA wird also i. d. R. einen Abwägungsvorgang zwischen konkurrierenden Grundrechten beinhalten.

In 34 Abs. 1 S. 3 lit. d DSA werden schließlich Risiken für einige Schutzgüter aufgezählt, die jedenfalls mittelbar durch Desinformation beeinträchtigt werden können. Dies wird anhand der Desinformationskampagnen der letzten Jahre besonders deutlich: Während der Covid-19-Pandemie wurde auf Grund der massenhaften Verbreitung von Desinformation über die Existenz und den Ursprung des Corona-Virus und über Schutzmaßnahmen von einer „Infodemie" gesprochen.[30] Hiermit korrespondieren die in 34 Abs. 1 S. 3 lit. d DSA genannte Risiken für die öffentliche Gesundheit und das körperliche Wohlbefinden, die sich gem. ErwG 83 auch aus koordinierten Desinformationskampagnen ergeben können. Auch die in 34 Abs. 1 S. 3 lit. d DSA gelistete geschlechtsspezifische Gewalt ist in Fällen von Gewalt befördernder misogyner Desinformation öffentlichkeitsrelevant geworden.[31]

Bei der Risikobewertung müssen gem. Art. 34 Abs. 2 lit. a–e DSA die technischen Funktionen der Dienste berücksichtigt werden, wie z. B. Empfehlungs-, Moderations-, Werbeauswahl- und Anzeigesysteme sowie

29 Zur Flexibilität des Risikobegriffs: *Achleitner*, MR-Int 2022, 114 (117).

30 *UNRIC.org*, UN und Partner fordern Länder auf „Infodemie" zu bekämpfen.

31 Z. B. ergab eine Umfrage der Interparlamentarischen Union von 2018 unter 123 weiblichen Abgeordneten oder Mitarbeiterinnen in nationalen Parlamenten in der EU, dass fast 47 Prozent im Laufe ihrer Karriere schon einmal in sozialen Medien Vergewaltigung oder sonstige Gewalt angedroht wurden, vgl. *Klimpel*, Wie Politikerinnen im Netz diskreditiert werden; *Ipu.org*, Sexism, harassment and violence against women in parliaments in Europe.

andere relevante algorithmische Systeme[32] der Anbieter. In die Bewertung einzuschließen sind zudem voluntative Elemente, wie die AGB und die datenbezogene Praxis der Anbieter. Auch durch Nutzerverhalten bedingte Faktoren, wie die vorsätzliche Manipulation, unauthentische Nutzung und automatisierte Ausnutzung der Dienste sowie die ihren AGB widersprechende Verbreitung und Verstärkung von Inhalten sind zu analysieren (Art. 34 Abs. 2 S. 2 DSA). Als Beispiel nennt ErwG 84 Risiken, die sich aus automatisierten oder teilautomatisierten Verhaltensweisen ergeben, die zu Desinformationskampagnen beitragen.

Risikominderung

Art. 35 Abs. 1 S. 1 DSA verpflichtet Anbieter i. S. d. Art. 33 DSA dazu, verhältnismäßige und wirksame Risikominderungsmaßnahmen zu ergreifen, die auf die gemäß Art. 34 ermittelten Risiken zugeschnitten sind. Dabei müssen die Auswirkungen solcher Maßnahmen auf die Grundrechte besonders berücksichtigt werden. Art. 35 Abs. 1 S. 2 lit. a–k DSA nennt exemplarisch Risikominderungsmaßnahmen, die von der Anpassung technischer Funktionen über AGB bis zu Moderationsmaßnahmen reichen. Die Kennzeichnungsmaßnahme gegen Deepfakes in lit. k dient der Bekämpfung visueller Desinformation.[33] Abs. 2 enthält Berichtspflichten und in Abs. 3 wird der EU-Kommission und den Koordinatoren für digitale Dienste die Befugnis übertragen, unter bestimmten Voraussetzungen gemeinsame Leitlinien für die Risikominderung für besondere Risiken herauszugeben. Zurecht werden diese Maßnahmen als zu vage kritisiert.[34]

Krisenreaktionsmechanismus

Gem. Art. 36 Abs. 1 DSA kann die EU-Kommission in einem Krisenfall auf Empfehlung des europäischen Gremiums für digitale Dienste[35] einen Beschluss erlassen, durch den Anbieter sehr großer Online-Plattformen

32 Zur Funktionsweise von Algorithmen und Regulierungserfordernissen vgl. *Kühling*, JZ 2021, 529 (531).

33 Eine Kennzeichnungspflicht für Deep Fakes ist auch im Art. 52 Abs. 3 KI-VO-E geplant. Dieser richtet sich jedoch an Anbieter bestimmter KI-Systemen und ihrer Nutzer. Art. 35 DSA adressiert hingegen nur "sehr große Online-Plattformen".

34 *Flamme*, MMR 2021, 770 (774).

35 Das europäische Gremium für digitale Dienste (Art. 61 DSA) setzt sich zusammen aus den Koordinatoren für digitale Dienste; die Kommission hat den Vorsitz (Art. 62 DSA).

(s.o.) aufgefordert werden, eine oder mehrere Maßnahmen nach Art. 36 Abs. 1 lit. a–c DSA zu ergreifen. Als Krise gilt nach Abs. 2 das Auftreten außergewöhnlicher Umstände, die zu einer schwerwiegenden Bedrohung der öffentlichen Sicherheit oder öffentlichen Gesundheit in der Union oder in wesentlichen Teilen der Union führen. Zu den Maßnahmen gehören nach Art. 36 Abs. 1 DSA die Bewertung, ob und wie der Betrieb und die Nutzung des Dienstes erheblich zu einer schwerwiegenden Bedrohung beitragen oder beitragen werden (lit. a), das Ergreifen gezielter, wirksamer und verhältnismäßiger Maßnahmen (lit. b) und die Berichterstattung an die Kommission (lit. c).[36] Auch für die Erfüllung dieser Pflichten können analytische Instrumente wie die Netzwerkanalyse eingesetzt werden.

Zwischenfazit

Zusammenfassend ist festzustellen, dass durch die risikobezogenen Pflichten der Artt. 34 ff. DSA der Grundgedanke der Bekämpfung systemischer Risiken fortentwickelt wird, indem diese explizit benannt und die Plattformarchitektur und AGB der Plattformen ausdrücklich in die Bewertung miteinbezogen werden müssen. Angesichts koordinierter Desinformationskampagnen, die auf die Ausnutzung der Plattformfunktionen angelegt sind, ist dieses Vorgehen für eine effektive Desinformationsbekämpfung sinnvoll.

Im DSA selbst ist nicht vorgegeben, wie die erläuterten Pflichten technisch umzusetzen sind. Im Folgenden soll daher die Netzwerkanalyse als ein mögliches Instrument zur Erfüllung der risikobasierten Pflichten aus Artt. 34 ff. DSA dargestellt werden. Derartige Instrumente können nicht nur für die Plattformanbieter, sondern auch für zugelassene Forscher i. S. d. Art. 40 Abs. 4 DSA nützlich sein, denen ein Datenzugang zur Erforschung systemischer Risiken ermöglicht wird.[37]

4. Die Netzwerkanalyse

Durch die Analyse von Interaktionen auf großen Online-Plattformen können Muster identifiziert werden, die auf eine Desinformationskampagne hindeuten können. Dabei kann, wie im Folgenden beschrieben, zwischen der Analyse der über einzelne Akteure hinausgehenden Verbreitungswege und der Betrachtung einzelner Akteure durch selbstdefinierte Netzwerk-

36 Vgl. dazu *Kuhlmann/Trute*, GSZ 2022, 115 (122).
37 Vgl. *Löber*, ZD-Aktuell 2022, 014290.

strukturen unterschieden werden, wobei unter Akteuren in diesem Kontext neben Nutzern auch Gruppen und Kanäle verstanden werden.

4.1 Analyse der Verbreitungswege

Die Verbreitungswege von Nachrichten zwischen Akteuren werden im Folgenden zunächst anhand von Metadaten, anschließend anhand des Nachrichtentextes diskutiert.

Metadatenbasiert

Durch Metadaten können Verbreitungswege durch sogenannte Nachrichtenkaskaden, d. h. die direkte Darstellung der Nachrichtenverbreitung, ermittelt werden. Eine Nachrichtenkaskade bezeichnet einen Baum oder eine baumähnliche Struktur, welche die Verbreitung eines bestimmten Nachrichtenartikels in einem Online-Netzwerk erfasst. Der Wurzelknoten repräsentiert den Nutzer, der einen Nachrichtenartikel zuerst teilt. Die anderen Knoten in der Kaskade repräsentieren Nutzer, die den Artikel nach der Veröffentlichung durch ihre übergeordneten Knoten, mit denen sie über Kanten verbunden sind, weiterverbreiten. Eine Nachrichtenkaskade kann durch die Anzahl der Schritte (d. h. hop-based), welche die Nachricht durchlaufen hat, oder durch die Zeitpunkte, zu denen sie veröffentlicht wurde (d. h. time-based), dargestellt werden.[38] Der Inhalt der Nachricht wird dabei nicht betrachtet.

Telegramkanäle und -gruppen bestehen zum Teil aus weitergeleiteten Nachrichten. Die Abbildung der Verbreitungsstruktur dieser Nachrichten kann dabei helfen, die Quelle einer Desinformation zu identifizieren und zu verstehen, wie sie durch das Netzwerk verbreitet wurde. Die Verbreitungswege einer Nachricht können dabei mit Kanälen oder Gruppen als Knoten untersucht werden. Je nach Online-Plattform unterscheiden sich die Metadaten, die über weitergeleitete Nachrichten vorhanden sind. In Telegram ist von den weitergeleiteten Nachrichten nur der Ursprungskanal/-gruppe oder der originäre Nutzer bekannt. Durch die Angabe des Zeitpunkts der Veröffentlichung kann damit der Zeitverlauf der Weiterleitungen ermittelt werden, genannt time-based Nachrichtenkaskade. Die Zwischenkanäle/-gruppen, von denen die weitergeleitete Nachricht wiederum weitergeleitet wurde (hops), sind bei Telegram nicht bekannt. Verbin-

38 *Zhou/Zafarani,* ACM Comput. Surv. 2020, 21 (22f.).

dungen können daher immer nur zwischen einem Kanal/einer Gruppe und einem Ursprungskanal/einer Ursprungsgruppe hergestellt werden. Dies führt nicht zu den aus der Literatur bekannten Kaskaden[39] mit mehreren übergeordneten Knoten (siehe Abb.1a), sondern zu Kaskaden mit jeweils zwei Knoten (siehe Abb. 1b) oder zu einer langen Kaskade mit allen Knoten (unter der Annahme, dass die Gruppen/Kanäle nach dem Zeitpunkt der Veröffentlichung sortiert werden können).

Anders verhält es sich z. B. bei Twitter. Hier ist immer nur die Kennung des letzten Postings der Nachricht bekannt, so dass der Weg der Nachricht zurückverfolgt werden kann. Dies würde bei Twitter zu einer hop-based Nachrichtenkaskade führen (siehe Abb.1a).

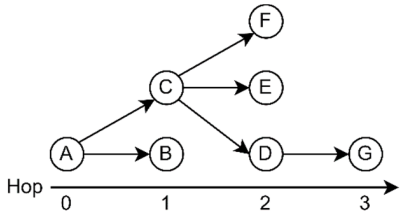

Abb. 1a: Nachrichtenkaskaden bei Twitter (hop-based), nach Zhou/Zafarani[40]

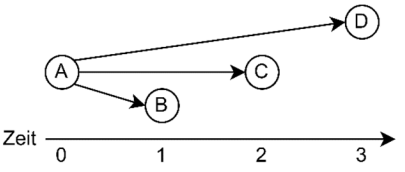

Abb. 1b: Nachrichtenkaskaden bei Telegram (time-based), eigene Darstellung

Durch die Nachverfolgung der Verbreitungswege können time-based Merkmale wie „Lifetime", „Real-time heat" und „Overall heat" ermittelt werden. Nach Zhou/Zafarani[41] geben diese drei Kennzahlen Hinweise darauf, ob es sich bei der vorliegenden Nachricht um Desinformation handeln könnte:

- *Lifetime* gibt die längste Kette an, in dem sich die Nachricht verbreitet hat (Abb. 1a, 4).
- *Real-time heat* (zum Zeitpunkt t) ist die Anzahl der Nutzer, die die Nachricht jeweils zum Zeitpunkt t gepostet/weitergeleitet haben und
- *Overall heat* die Gesamtzahl der Nutzer, die die Nachricht weitergeleitet/gepostet haben.

Ein weiteres Merkmal ist die Zeit, die Kaskaden von Desinformationen benötigen, um eine beliebige Tiefe und Größe zu erreichen. Die Tiefe gibt

39 *Zhou/Zafarani,* ACM Comput. Surv. 2020, 21 (22).
40 *Zhou/Zafarani,* ACM Comput. Surv. 2020, 22.
41 *Zhou/Zafarani,* ACM Comput. Surv. 2020, 22.

die maximale Anzahl an Schritten an, die die Nachricht innerhalb der Kaskade durchlaufen hat (Abb. 1a, Tiefe: 3). Die Größe wird durch die Gesamtzahl der Knoten bestimmt (Abb. 1a, Größe: 7). Die Zeit für das Erreichen einer beliebigen Größe und Tiefe ist laut Vosoughi u.a.[42] bei Desinformationen kürzer als bei Kaskaden mit echten Nachrichten. Die Tiefe einer Nachrichtenkaskade ist bei Telegram nicht bekannt, da mit den verfügbaren Informationen keine Verzweigungen identifiziert werden können, aber die Tiefe ist der Lifetime sehr ähnlich, die stattdessen als Merkmal verwendet werden könnte. Die Größe, also die Anzahl der Knoten innerhalb der Nachrichtenkaskade (Overall heat), ist auch bei Telegramdaten bekannt.

Für die Klassifizierung der Nachricht als Desinformation oder sonstige Information wird bei der Untersuchung der Verbreitungswege stellvertretend die Kaskade zur Klassifikation herangezogen. Dazu können Methoden des traditionellen maschinellen Lernens wie Support Vector Machine (SVM)[43], Entscheidungsbäume[44] oder Naive Bayes[45] verwendet werden. Auch (tiefe) neuronale Netze werden z. B. für Twitterdaten vorgeschlagen.[46] Dabei wird die Struktur der Nachrichtenkaskade in neuronalen Netzen nachgebildet, und es werden Werte für die Blattknoten ermittelt. Dies ist bei Telegram aufgrund der bereits erwähnten mangelnden Nachvollziehbarkeit der Verzweigungen nicht möglich. Auch Kontoeinstellungen wie die Anonymisierung des Nutzers in Telegram in den Privatsphäre-Einstellungen, die eine Nachverfolgung von geposteten Nachrichten unmöglich macht, erschweren diese Analyse. Aufgrund dieser Schwierigkeiten sollten

42 *Vosoughi* u.a., science 2018, 1147.

43 *Castillo* u.a., Information credibility on twitter in: Sadagopan u.a. (Hrsg.), Proceedings of the 20th international conference on World wide web, 2011, 680; *Kwon* u.a., Prominent Features of Rumor Propaganda in Online Social Media, in: 2013 IEEE 13th international conference on data mining 2013, 1108; *Wu* u.a., False rumors detection on Sina Weibo by propagation structures, in: 2015 IEEE 31st international conference on data engineering 2015, 653 (654).

44 *Castillo* u.a., Information credibility on twitter in: Sadagopan u.a. (Hrsg.), Proceedings of the 20th international conference on World wide web, 2011, 680; *Kwon* u.a., Prominent Features of Rumor Propaganda in Online Social Media, in: 2013 IEEE 13th international conference on data mining 2013, 1108.

45 *Castillo* u.a., Information credibility on twitter in: Sadagopan u.a. (Hrsg.), Proceedings of the 20th international conference on World wide web, 2011, 680.

46 *Zhou/Zafarani,* ACM Comput. Surv. 2020, 22 (23).

inhaltliche Analysen wie die „Semantic Similarity“[47] neben der Metadaten-Analyse in Betracht gezogen werden.

Semantische Analyse

Um die Verbreitungswege identischer und ähnlicher Inhalte zu analysieren, reicht es nicht aus, nur die Weiterleitungsfunktion von Inhalten in Telegram zu betrachten, da hier nur der Ursprungskanal oder die Ursprungsgruppe erkenntlich ist. Um gleiche Inhalte (Bilder, Videos und Texte) automatisiert erkennen zu können existieren unterschiedliche Ansätze. Auf der Textebene kann die semantische Ähnlichkeitsanalyse (engl. Semantic Similarity) in Betracht gezogen werden. Semantic Similarity ist eine Aufgabe im Bereich der Verarbeitung natürlicher Sprache („Natural Language Processing“, kurz NLP), bei der die Ähnlichkeit zwischen Texten oder Dokumenten anhand einer definierten Metrik bewertet wird. Die aktuell gängigste Methode besteht darin, ein maschinelles Lernmodell (z. B. einen Transformer) zu verwenden, welches Sätze zunächst in eine Vektordarstellung überführt.[48] In der Regel liegen dabei inhaltlich ähnliche Sätze im Vektorraum näher beieinander. Dann wird eine Ähnlichkeitsmetrik (z. B. Cosinus-Ähnlichkeit) verwendet, um zu berechnen, ob die Sätze sich inhaltlich ähnlich sind oder nicht. Es geht folglich darum, festzustellen, ob zwei oder mehr Textstücke die gleiche Bedeutung haben oder nicht.

Mithilfe der Ähnlichkeitsanalyse lässt sich nicht nur analysieren, wie dieselben Inhalte innerhalb Telegrams weitergeleitet werden, sondern auch wie identische oder ähnliche Inhalte generell plattformübergreifend verbreitet werden, d. h. über die Grenze einer Online-Plattform hinweg (Interoperabilität zwischen Online-Plattformen). Es können folglich unbekannte Verbreitungswege zwischen Kanälen/Gruppen innerhalb und außerhalb Telegrams identifiziert werden. Diese Kanäle/Gruppen teilen dieselben Inhalte, sind jedoch laut ihrer Metadaten nicht untereinander vernetzt.

Nachdem die identischen oder ähnlichen Nachrichten identifiziert wurden (durch Metadaten oder Semantic Similarity), können diese und die Akteure (Nutzer, Gruppe, Kanal) auf ihre Charakteristika (Medientyp, Thema etc.) untersucht werden~~)~~.

47 *Chandrasekaran/Mago*, ACM Comput. Surv. 2021, 1.
48 *Chandrasekaran/Mago*, ACM Comput. Surv. 2021, 1.

4.2 Analyse einzelner Akteure

Einzelne Akteure (Nutzer, Gruppe, Kanal) können durch selbstdefinierte Netzwerkstrukturen abgebildet und analysiert werden.[49]

Anders als bei den Nachrichtenkaskaden können diese Netzwerke unterschiedliche (selbstdefinierte) Strukturen annehmen. Unterschieden wird zwischen homogenen, heterogenen oder hierarchischen Netzwerken.[50] Homogene Netzwerke (Bsp. siehe Abb. 2) bestehen aus einer Art, heterogene Netzwerke aus verschiedenen Arten an Knoten und Kanten. Hierarchische Netzwerke bilden verschiedene Arten von Knoten und Kanten in Mengen-Teilmengen-Beziehungen (d. h. Hierarchien) ab. In hierarchischen Netzwerken stellen beispielsweise die Kanten die Zugehörigkeit einer Nachricht (Knoten) zu einem übergeordneten Ereignis (Knoten) dar.

Beispielsweise können Ansichten der Nutzer zu einem Thema innerhalb eines Kanals/einer Gruppe durch homogene Netzwerke dargestellt werden. Dieses Netzwerk wird als Stance Net bezeichnet und bildet auf den Knoten die unterschiedlichen Ansichten der Nutzer zu einem Thema und auf den Kanten die Beziehung dieser (Unterstützend (+) oder Gegensätzlich (-)) ab (Abb. 2), wobei die Themen durch Methoden des NLP wie Topic Modeling identifiziert werden können. Topic Modeling, auch bekannt als Textkategorisierung, ist eine Technik der Textanalyse, mit der vorherrschende Themen eines Textkorpus identifiziert werden können.[51] Hierfür wird z. B. ein Algorithmus verwendet, der die Nachrichten auf Basis ihrer semantischen Ähnlichkeit zu Themen gruppiert, wobei jedes Thema mithilfe von Schlagwörtern extrahiert bzw. beschrieben werden kann.[52]

Desinformationen provozieren oft kontroverse Ansichten, wobei verneinende und hinterfragende Haltungen einen Hinweis darauf geben können, ob Nachrichten als falsch zu klassifizieren sind.[53] Informationen für die Bildung dieser Netzwerke können aus den Metadaten wie „gefällt mir"-Angaben oder dem Nachrichtentext mittels NLP, z. B. durch die Stimmungsanalyse (eng. Sentiment Analysis) oder der Analyse des Standpunktes einer

49 *Zhou/Zafarani,* ACM Comput. Surv. 2020, 24 (25f).
50 *Zhou/Zafarani,* ACM Comput. Surv. 2020, 24 (25f).
51 *Churchill/Singh,* ACM Comput. Surv. 2022, 1 (2f).
52 *Grootendorst*, arXiv preprint 2022, 1.
53 *Shu* u.a., Emerging research challenges and opportunities in computational social network analysis and mining 2019, 7; *Jin* u.a., News Verification by Exploiting Conflicting Social Viewpoints in Microblogs, in: Proceedings of the AAAI conference on artificial intelligence 2016, 2972 (2973f).

Nachricht (eng. Stance Detection) ermittelt werden. Hierbei kann ein „Profil" mit den Eigenschaften des jeweiligen Nutzers erstellt oder aber es können Standpunkte repräsentativ für eine Gruppe von Nutzern (einer/einem Telegram Gruppe/Kanal) gebildet werden. Werden Standpunkte auf Basis von Nutzerprofilen erstellt, kann auch hier die bereits erwähnte Anonymisierungsfunktion von Telegram den Autor einer Nachricht unkenntlich machen. Weiter zu beachten ist, dass hierfür in Telegram nur Daten aus Gruppen verwendet werden können und nicht aus Kanälen, da in diesen nur die Administratoren Nachrichten veröffentlichen können. In Telegram existieren keine Folgebeziehungen zwischen Nutzern („Follower"), wie dies bei Twitter der Fall ist. Analysen auf Basis von „Freundschaften" und den daraus resultierenden Einflüssen aufeinander sind daher, zumindest auf Nutzerebene, nicht möglich (z. B.: Spreader Net, nach Zhou/Zafrani[54]).

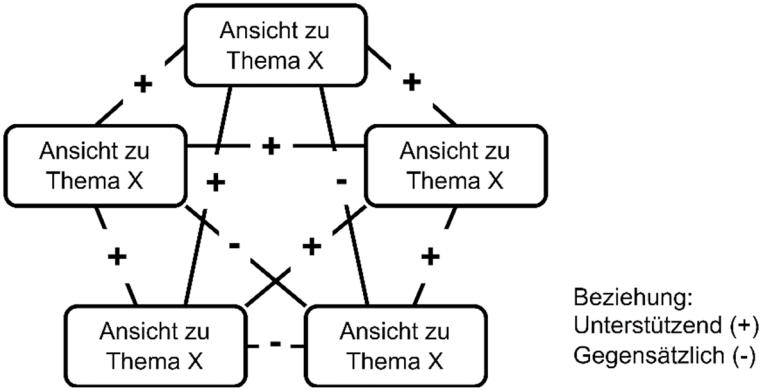

Abb. 2: Stance Net (Homogenes Netzwerk), eigene Darstellung nach Zhou/Zafarani[55] und Jin u.a.[56]

Mithilfe der Netzwerkanalyse können zwar keine Desinformationen als solche identifiziert werden. Allerdings können verdächtige Themen erkannt werden, die sich sehr schnell über die Kanäle/Gruppen hinweg oder auch plattformübergreifend verbreiten. Auf diese Weise ist es möglich die Relevanz eines Themas und entsprechende Kanäle/Gruppen mit einem hohen

54 *Zhou/Zafrani*, ACM SIGKDD explorations newsletter 2019, 48 (49f).
55 *Zhou/Zafarani,* ACM Comput. Surv. 2020, 25.
56 *Jin* u.a., News Verification by Exploiting Conflicting Social Viewpoints in Microblogs, in: Proceedings of the AAAI conference on artificial intelligence 2016, 2972 (2973f).

Einfluss zu erkennen, um zeitig darauf reagieren zu können (z. B. durch eine entsprechende Kennzeichnung durch Fact-Checking-Organisationen).

5. Diskussion

Die neuen risikobezogenen Pflichten aus Artt. 34 ff. DSA sind ein im Ausgangspunkt begrüßenswerter Versuch des europäischen Gesetzgebers, netzbasierte Desinformation, die oftmals keinen Straftatbestand erfüllt, regulatorisch einzuhegen. Allerdings zieht diese Regulierung auch rechtliche Bedenken nach sich. Allgemein ist problematisch, dass keinerlei Qualitätsvorgaben zur technischen Umsetzung gemacht werden. Zu befürchten ist, dass von den Plattformen Verfahren eingesetzt werden, die ein verzerrtes Bild wiedergeben. Hinzu kommen spezifische rechtliche und technische Herausforderungen konkreter Analyseinstrumente, die die Anbieter einsetzen könnten. Diese werden im Folgenden für das Beispiel der Netzwerkanalyse näher betrachtet werden.

5.1 Rechtliche Probleme

Aus rechtlicher Perspektive wird grundsätzlich kritisiert, dass die Vorschriften der Artt. 34 ff. DSA insgesamt sehr vage formuliert sind.[57] Dies führ zu einigen Folgeproblemen.

Überwachungspflichten der Anbieter

Zu befürchten ist zunächst, dass durch die neuen risikobezogenen Pflichten aus Artt. 34 ff. DSA de facto eine Überwachungspflicht für sehr große Online-Plattformen begründet wird. In Art. 8 DSA wird zwar statuiert, dass die Verordnung Vermittlungsdiensten gerade keine allgemeine Verpflichtung zur Überwachung auferlegt. ErwG 30 spezifiziert, dass eine solche Pflicht weder de jure noch de facto bestehen soll, wobei Ausnahmen gem. ErwG 30 lediglich für bestimmte Fälle gestattet sind. Zur Erfüllung der Risikobewertung nach Art. 34 DSA ist es jedoch faktisch erforderlich, dass die Dienste jegliche Inhalte auf ihre potenzielles Risikorelevanz prüfen und feststellen, ob die Schwelle zum „systemischen Risiko" überschritten

57 *Buri/Van Hoboken*, The Digital Services Act (DSA) proposal: a critical overview, 2021, S.33.

wurde. Wird die Pflicht nach ErwG 30 so ausgelegt, dass keine allgemeine Überwachungspflicht besteht, bleibt zweifelhaft, ob die Risikobewertung effektiv zur Pflichterfüllung erfolgen kann.

Eine ähnliche Widersprüchlichkeit ist hinsichtlich der Haftungsprivilegierung der Hosting-Dienste aus Art. 6 Abs. 1 DSA zu befürchten. Danach haften Anbieter von Hosting-Diensten nicht, sofern sie keine tatsächliche Kenntnis von rechtswidrigen Inhalten oder diese begründenden Umständen haben oder zügig tätig werden, um diese Inhalte zu sperren oder zu entfernen. Diese grundsätzlich geltende Privilegierung droht aber dadurch ausgehöhlt zu werden, dass sehr großen Online-Plattformen bei Verletzung der Risikobewertungspflichten, die eine aktive Ermittlung erfordern, Sanktionen nach Artt. 74 und 76 DSA drohen. Sollte die sog. „gute Samariter Privilegierung" aus Art. 7 DSA[58] so ausgelegt werden, dass Dienste bei freiwilligen Risikobewertungsmaßnahmen nicht sanktioniert werden, droht die Risikobewertungspflicht wiederum an Effektivität einzubüßen.[59]

Fehlende datenschutzrechtliche Verarbeitungsgrundlage

Zudem zeichnen sich durch die risikobezogenen Pflichten aus Artt. 34 ff. DSA datenschutzrechtliche Konfliktlagen ab. Problematisch ist in diesem Zusammenhang, dass die Ermittlung und Bewertung systemischer Risiken eine breite Datengrundlage erfordert, wobei aus funktionaler Sicht auch zahlreiche personenbezogene Daten i. S. d. Art. 4 Nr. 1 DSGVO verarbeitet werden müssten. Grundsätzlich lässt der DSA gem. Art. 2 Abs. 4 lit. g DSA Unionsvorschriften zum Schutz personenbezogener Daten, z. B. die DSGVO, unberührt. Daher ist etwa neben der Risikobewertung nach Artt. 34 ff. DSA bei der Einführung risikorelevanter neuen Funktionen ggf. auch eine Datenschutz-Folgenabschätzung nach Art. 35 DSGVO durchzuführen.

Besonders zweifelhaft ist, ob die Artt. 34 ff. DSA geeignete Rechtsgrundlagen i. S. d. Art. 6 Abs. 1 lit. c, Abs. 3 lit. a DSGVO darstellen, um die Analyse personenbezogener Daten im Rahmen des Risikomanagements zu legitimieren. Dafür müsste die Datenverarbeitung durch die gesetzliche

58 Dieser stellt klar, dass die Haftungsprivilegierung auch bei freiwilligen Untersuchungen der Dienste-Anbieter unter bestimmten Voraussetzungen gilt. So wird Vermittlungsdiensten einen Anreiz zum Ergreifen solcher Maßnahmen gesetzt.

59 So auch: *Holznagel*, CR 2021, 123 (132); Weitere grundrechtliche und medienrechtliche Probleme thematisieren *Berberich/Seip*, GRUR-Prax 2021, 4 (6).

Rechtsgrundlage festgelegt werden und zur Erfüllung einer rechtlichen Verpflichtung erforderlich sein.[60]

Bereits aus Art. 8 Abs. 2, Art. 52 Abs. 1 GRCh ergibt sich, dass Rechtsgrundlagen klar und präzise ausgestaltet und ihre Anwendung vorhersehbar sein müssen. Deklaratorisch stellt Art. 6 Abs. 3 S. 2 und 4 DSGVO klar, dass die Rechtsgrundlage die Zwecke der dazu erforderlichen Verarbeitung festlegen muss.[61] Die rechtliche Verpflichtung muss sich unmittelbar auf die Datenverarbeitung beziehen, und der Zweck muss spezifisch bestimmt sein. Allein der Umstand, dass ein Verantwortlicher, um eine rechtliche Verpflichtung erfüllen zu können, auch personenbezogene Daten verarbeiten muss, reicht nicht aus.[62] In den Vorschriften der Artt. 34 ff. DSA wird indes an keiner Stelle ausdrücklich erwähnt, dass personenbezogene Daten verarbeitet werden dürfen. Somit muss den Artt. 34 ff. DSA eine Konturlosigkeit attestiert werden, die verschiedenste Datenverarbeitungen möglich erscheinen lässt (z. B. den Abgleich personenbezogener Daten verschiedener Nutzer, die dauerhafte Speicherung, ggf. sogar das Anlegen umfassender Nutzerprofile). Dies ist mit grundrechtlichen Bestimmtheitsanforderungen nicht vereinbar.

Weiterhin besteht zwar das nach Art. 6 Abs. 3 S. 4 DSGVO geforderte im öffentlichen Interesse liegende Ziel, indem Artt. 34 ff. DSA die Schaffung eines sicheren und vertrauenswürdigen Online-Umfelds bezwecken (vgl. ErwG 75 DSA). Dieses legitime Ziel ist fraglos gewichtig, jedoch kann es unter Verhältnismäßigkeitsgesichtspunkten nicht sämtliche, heute ggf. noch gar nicht absehbare für Artt. 34 ff. DSA sinnvolle Datenverarbeitungen legitimieren.

Dies wird besonders deutlich, wenn es um sensible, also besondere Kategorien personenbezogener Daten geht, deren Verarbeitung Art. 9 Abs. 1 DSGVO grundsätzlich untersagt. Von den Ausnahmen in Art. 9 Abs. 2 DSGVO kommt lit. g in Betracht, da die Bekämpfung von Desinformation zu den „Gründen eines erheblichen öffentlichen Interesses" zählt. Die Norm verlangt jedoch, dass die entsprechende Rechtsgrundlage „angemessene und spezifische Maßnahmen zur Wahrung der Grundrechte und Interessen der betroffenen Personen" regelt. Diese enthält der DSA nicht.

60 *Roßnagel*, in: Simitis u.a. (Hrsg.), Datenschutzrecht, Art. 6 DS-GVO, Rn. 52.

61 *Frenzel* in: Paal/Pauly (Hrsg.), DS-GVO/BDSG, Art. 6 DS-GVO, Rn. 45; *Heberlein* in: Ehmann/Selmayr (Hrsg.), DS-GVO, Art. 6 DS-GVO, Rn. 15 m.w.N.

62 *Albers/*Veit in: Wolff/Brink (Hrsg.), BeckOK Datenschutzrecht, Art. 6 DS-GVO, Rn. 48; *Roßnagel*, in: Simitis u.a. (Hrsg.), Datenschutzrecht, Art. 6 DS-GVO, Rn. 29; a.A. *Frenzel* in: Paal/Pauly (Hrsg.), DS-GVO/BDSG, Art. 6 DS-GVO, Rn. 41.

Besonderheiten der Netzwerkanalyse

Auch die Anwendung der Netzwerkanalyse birgt Gefahren für den Schutz personenbezogener Daten, da diese Personen zugeordnet werden, um zu Diagrammen verarbeitet zu werden. Dies ist unter anderem deswegen problematisch, da bei der Analyse der Nachrichteninhalte regelmäßig Daten verarbeitet werden, die in den Katalog des Art. 9 Abs. 1 DSGVO fallen. Beschränkt sich die Analyse allerdings auf Metadaten, bei denen dies nicht der Fall ist, ist es möglich, andere Rechtsgrundlagen heranzuziehen, insbesondere berechtigte Interessen der Anbieter, wenn diese gegenüber Grundrechten und Grundfreiheiten der betroffenen Personen überwiegen (Art. 6 Abs. 1 UAbs. 1 DSGVO lit. f).

Greift Art. 9 Abs. 1 DSGVO, kommt von den Ausnahmen des Art. 9 Abs. 2 DSGVO zwar nicht lit. g (s.o.), ggf. aber lit. e in Betracht, wenn die Daten offensichtlich durch die betroffene Person veröffentlicht wurden.[63] Indes wird gerade beim Anwendungsbeispiel Telegram deutlich, dass die Einstufung als „veröffentlicht" von den jeweils genutzten Kommunikationsfunktionen (Kanal, Gruppe) als auch anderen Faktoren, wie der Gruppengröße, abhängt und im Einzelfall beurteilt werden muss.[64] Ohnehin veröffentlichen Nutzer auf den Plattformen vielfach auch Daten Dritter, die diese dann nicht selbst öffentlich gemacht haben; insgesamt bleibt die Verarbeitung öffentlicher Massendaten damit ein teilweise ungelöstes Problem,[65] für das Rechtsgrundlagen geschaffen werden sollten. Losgelöst von der grundsätzlichen Frage der Verarbeitungsbefugnis müssen bei Anwendung der Netzwerkanalyse Anonymisierung, Pseudonymisierung oder andere Maßnahmen des Datenschutzes durch Technikgestaltung (Art. 25 DSGVO) vorgenommen werden.[66] Zu bedenken ist allerdings, dass eine Anonymisierung oder Pseudonymisierung bei Inhaltsdaten nur schwer zu erreichen ist, da z. B. Texte durch eine einfache Web-Suche Personen zugeordnet und prominente Personen anhand ihres Schreibstils identifiziert werden können.

63 Schutzlos gestellt ist die betroffene Person durch eine Veröffentlichung jedoch nicht. Vorausgesetzt ist immerhin das Vorliegen ein Erlaubnistatbestands nach Art. 6 Abs. 1 UAbs. 1 DSGVO, vgl. *Petri*, in: *Simitis/Hornung/Spiecker gen. Döhmann*, DSGVO, Art. 9, Rn 57; *Hornung/Gilga*, CR 2020, 367 (374).

64 *Petri*, in: *Simitis/Hornung/Spiecker gen. Döhmann*, DSGVO, Art. 9, Rn 58.

65 Näher *Hornung/Gilga*, CR 2020, 367ff.

66 Weitere grundrechtliche und medienrechtliche Probleme thematisieren *Berberich/Seip*, GRUR-Prax 2021, 4 (6).

5.2 Technische Herausforderungen

Für die Erkennung von Desinformation müssen die Faktizität (enthält die Nachricht eine nicht faktengemäße Aussage?) und die Absicht (zielt sie darauf ab, Menschen in die Irre zu führen oder Menschen zu schaden?) analysiert werden.[67] Faktizität kann über sogenanntes Fact-Checking untersucht werden. Dieses ist jedoch von der Domäne abhängig und in Echtzeit momentan aufgrund der Schnelligkeit der Verbreitung von Nachrichten nicht anwendbar. Auch die Irreführungs- oder Schädigungsabsicht ist kaum zu ermitteln und nachzuweisen. Der gute Glaube des Verbreiters würde eine solche Absicht an sich entfallen lassen, ist aber ebenso schwer zu ermitteln. Ein Ansatz besteht darin, die Irreführungs- oder Schädigungsabsicht durch das Verhalten der Nutzer zu ermitteln, z. B. durch das Weiterleiten oder Veröffentlichen von Nachrichten.[68] Shu u.a.[69] klassifizieren z. B. auf der Basis von Weiterleitungen Nutzer, die wiederholt Desinformationen weiterleiten, als eher geneigt Desinformation zu glauben. Dieses beobachtbare Verhalten in sozialen Medien spiegelt jedoch nicht zwangsläufig den Glauben oder die Absicht des Nutzers wider.[70] Sowohl für die Analyse der Faktizität als auch der Absicht und damit für die Erkennung von Desinformation ist ein tieferes Verständnis des Inhalts, der zugrundeliegenden Logik und des menschlichen Verhaltens erforderlich.

In der Literatur lassen sich Methoden zur Desinformationserkennung in wissensbasierte (Überprüfung, ob Wissen im Nachrichtentext mit Fakten übereinstimmt), stilbasierte (wie sind Nachrichten geschrieben, Bsp. Analyse der Emotionen/Intentionen), propagationsbasierte (Untersuchung der Verbreitungswege) und quellbasierte Ansätze (Untersuchung der Nachrichtenquelle, Bsp. Nutzer) unterscheiden.[71] Mittels der Netzwerkanalyse wurden hier vor allem propagationsbasierte (die Verbreitungswege von Weiterleitungen) und quellbasierte Ansätze (Analyse der Akteure) dargestellt. Die Netzwerkanalyse bietet die Möglichkeit, Akteure und ihre Verbindungen untereinander zu beschreiben, stellt aber keine alleinige Methode zur automatisierten Desinformationserkennung dar. Durch eine Analyse der Verbreitungswege können Akteure mit meinungsbildendem Einfluss, d. h.

67 *Zhou/Zafarani,* ACM Comput. Surv. 2020, 3 (4f.); *Baptista/Gradim*, Encyclopedia 2022, 640; *Lazer* u.a., Science 2018, 1094.
68 *Pennycook* u.a., Nature 2021, 590.
69 *Shu* u.a., IEEE MIPR 2018, 3.
70 *Ellison* u.a., Journal of Computer-Mediated Communication 2020, 402 (403f.).
71 *Zhou/Zafarani,* ACM Comput. Surv. 2020, 7 (8f.).

Ursprünge von vielen Weiterleitungen, identifiziert werden. Wurde eine Vorauswahl an Akteuren vorgenommen, kann deren Inhalten näher untersucht werden, z. B. durch wissensbasierte und stilbasierte Methoden oder selbstdefinierten Netzwerkstrukturen.

6. Fazit und Ausblick

Die rechtliche Analyse der risikobezogenen Pflichten des DSA zeigt, dass diese grundsätzlich zu einem sicheren und vertrauenswürdigen Online-Umfeld beitragen können. Durch die jährlich zu erfüllenden Pflichten können systemische Risiken langfristig evaluiert werden, was sukzessiv ermöglicht, die bestehende Regulierung – auch hinsichtlich einer effektiveren Desinformationsbekämpfung – adäquat anzupassen.

Zugleich offenbart der DSA eine Fülle offener Rechtsfragen. Beim Anwendungsbereich (Online-Plattform vs.„sehr große Online-Plattform“) zeigt das Beispiel Telegram, dass die Berechnungsmaßstäbe für Hybrid-Medien weiterer Konkretisierung bedürfen. Der DSA enthält überdies etliche unbestimmte Gesetzesformulierungen und verursacht neue medien- und datenschutzrechtliche Probleme.

Als ein Instrument der Risikoabschätzung ermöglicht die Netzwerkanalyse, Kanäle/Gruppen und ihre Verbindungen untereinander zu bestimmen und zu beschreiben. Angesichts der großen Datenmengen kann eine Vorauswahl relevanter Akteure für eine weitergehende Analyse hilfreich sein.

Hierdurch können z. B. die folgenden Fragen untersucht werden:

- Von welchem Kanal/welcher Gruppe werden besonders viele Nachrichten verbreitet?
- Welche Art von Nachrichten (Medientyp,[72] Thema etc.) wird viral verbreitet?
- Liegt der Ursprung der Nachrichten in Telegram oder bei anderen Plattformen?
- Existieren erste Merkmale, die auf Desinformation hindeuten (Eigenschaften der Nachrichtenkaskade)?
- Welche Standpunkte werden zu bestimmten Themen in den Kanälen/Gruppen vertreten?

72 Dies kann z. B. ein Bild, Video oder auch eine Nachricht mit oder ohne URL sein.

- Welche Nutzer sind besonders aktiv und welche Inhalte verbreiten sie?
- Welche Nutzer sind wie häufig in mehreren Kanälen/Gruppen aktiv?

Antworten auf diese Fragen können ein wichtiger Bestandteil einer Analyse der einzelnen in Art. 34 Abs. 1 S. 3 DSA genannten systemischen Risiken sehr großer Online-Plattformen sein, indem sie beispielsweise als solche erkennen lassen, wie sich bestimmte rechtswidrige Inhalte verbreiten (S. 3 lit. a) oder die Grundlage für weitergehende, vertiefte Risikobewertungsmaßnahmen bilden.

Bei der Risikobewertung nach Art. 34 DSA muss auch auf die Verbreitung von Desinformation eingegangen werden. Hierbei kann die Netzwerkanalyse als erste Übersicht dienen und Akteure mit großer Reichweite (Einfluss) identifizieren. Zu diesen einflussreichen Akteuren können dann nähere Untersuchungen auf Desinformation (z. B. durch NLP, Fact-Checking) durchgeführt werden. Die Netzwerkanalyse kann damit als erstes, aber nicht abschließendes Instrument zur Erkennung und Bekämpfung von Desinformation verwendet werden.[73]

Angesichts der Bedeutung, die der europäische Gesetzgeber dem Risikomanagement im DSA beimisst, bleibt abschließend allerdings auf das erläuterte regulatorische Defizit hinzuweisen. Datenschutzrechtlich steht die Verarbeitung personenbezogener Daten zur Risikobewertung und Risikominderung auf tönernen Füßen, da der DSA keine Befugnis enthält, die Umfang und Grenzen der zulässigen Verarbeitung erkennen lässt. Diese sowohl für die Anbieter als auch für die betroffenen Nutzer schwer erträgliche Situation sollte der Gesetzgeber beheben.

Literatur

Achleitner, Ranjana Andrea (2022): Der Digital Services Act als risikobasierte Regulierung. *Medien und Recht International (MR-Int)*, 18(4), S. 114-121.

Albers, Marion und Veit, Raoul-Darius (2022): Artikel 6 Datenschutz-Grundverordnung. In: Wolff, Amadeus und Brink, Stefan (Hrsg.): *Beck'scher Online-Kommentar zum Datenschutzrecht*. München: C.H.Beck.

Balser, Markus (11. Feb. 2022): *Telegram sperrt 64 Kanäle*. URL: https://www.sueddeutsche.de/politik/telegram-kanaele-sperrung-1.5527255.

73 Neben der Netzwerkanalyse werden auch andere Verfahren vorgeschlagen, um Artt. 34 ff. DSA zu erfüllen; vgl. z. B. den Szenarien-basierten Ansatz von *Meßmer/Degeling*, Auditing Recommender Systems – Putting the DSA into practice with a risk-scenario-based approach, 2023; *Achleitner*, MR-Int 2022, 114 (116), schlägt die Heranziehung des internationalen Standards ISO 31000 vor.

Baptista, João P. und Gradim, Anabela (2022): A working definition of fake news. *Encyclopedia,* 2(1). doi: 10.3390/encyclopedia2010043.

Berberich, Matthias und Seip, Fabian (2021): Der Entwurf des Digital Services Act. *Praxis im Immaterialgüter- und Wettbewerbsrecht (GRUR-Prax)*, 13(1), S. 4- 7.

Borschert, Nils und Wolter, Daphne (April 2022): *Ort der Meinungsfreiheit und Verschwörungsplattform*. Berlin: Konrad-Adenauer-Stiftung.

Buri, Ilaria; van Hoboken, Joris (2021): *The Digital Services Act (DSA) proposal: a critical overview*, URL: https://dsa-observatory.eu/wp-content/uploads/2021/11/Buri-Van-Hoboken-DSA-discussion-paper-Version-28_10_21.pdf.

Castillo, Carlos; Mendoza Marcelo und Poblete Barbara (2011): Information credibility on twitter. *Proceedings of the 20th international conference on World wide web,* S675-684. doi: 10.1145/1963405.1963500.

Chandrasekaran, Dhivya und Mago, Vijay (2021): Evolution of semantic similarity—a survey. *ACM Computing Surveys*, 54(2), S. 1-37. doi: 10.1145/3440755.

Churchill, Rob und Singh, Lisa (2022): The evolution of topic modeling. *ACM Computing Surveys*, 54(10s), S.1-35. doi: 10.1145/3507900.

Denga, Michael (2021): Plattformregulierung durch europäische Werte: Zur Bindung von Meinungsplattformen an EU-Grundrechte. *Europarecht (EuR)*, 56(5), S. 569-595.

Eisenreich, Georg (2021): Digital Services Act – ein wirksames Instrument gegen Hass und Hetze im Netz. *Recht Digital (RDi)*, 1(6), S. 289-293.

Ellison, Nicole B.; Triệu, Penny; Schoenebeck, Sarita; Brewer, Robin und Israni, Aarti (2020): Why we don't click: Interrogating the relationship between viewing and clicking in social media contexts by exploring the "non-click". *Journal of Computer-Mediated Communication*, 25(6), S.402-426. doi: 10.1093/jcmc/zmaa013.

Europäische Kommission (2018): Mitteilung der Kommission an das Europäische Parlament, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen – Bekämpfung von Desinformationen im Internet: ein europäisches Konzept. COM(2018) 236 final. Brüssel: Europäische Kommission.

Europäische Kommission (2023): Presseerklärung vom 17.02.2023 – Digital Services Act: Commission starts collecting platform's user numbers and consults on its monitoring and investigatory procedures. URL: https://digital-strategy.ec.europa.eu/en/news/digital-services-act-commission-starts-collecting-platforms-user-numbers-and-consults-its

Flamme, Florian (2021): Schutz der Meinungsvielfalt im digitalen Raum. Transparenzpflichten für Intermediäre im nationalen und europäischen Vergleich, *Multimedia und Recht (MMR)*, 24(10), S. 770-774.

Frenzel, Michael (2021): Artikel 6 Datenschutz-Grundverordnung. In: Paal, Boris und Pauly, Daniel A. (Hrsg.), *Kommentar DS-GVO/BDSG*. München: C.H.Beck.

Gielen, Nico und Uphues, Steffen (2021): Digital Markets Act und Digital Services Act – Regulierung vonn Markt- und Meinungsmacht durch die Europäische Union. *Europäische Zeitschrift für Wirtschaftsrecht (EuZW)*, 32(14), S. 627-637.

Grootendorst, Maarten (2022): BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint*. doi: 10.48550/ARXIV.2203.05794.

Heberlein, Horst (2018): Artikel 6 Datenschutz-Grundverordnung. In: Ehmann, Eugen und Selmayr, Martin (Hrsg.), *Kommentar DS-GVO*. München: C.H.Beck.

Holznagel, Daniel (2021): Chapter II des Vorschlags der EU-Kommission für einen Digital Services Act Versteckte Weichenstellungen und ausstehende Reparaturen bei den Regelungen zu Privilegierung, Haftung & Herkunftslandprinzip für Provider und Online-Plattformen, *Computer und Recht (CR),* 37(2), S. 123-132.

Hornung, Gerrit und Gilga, Carolin (2020): Einmal öffentlich – für immer schutzlos? Die Zulässigkeit der Verarbeitung öffentlicher personenbezogener Daten, *Computer und Recht (CR),* 36(6), S. 367-379.

Inter-Parliamentary Union (Oktober 2018): Sexism, harassment and violence against women in parliaments in Europe. Genf: Inter-Parliamentary Union.

Janal, Ruth (2021) Haftung und Verantwortung im Entwurf des Digital Services Acts, *Zeitschrift für Europäisches Privatrecht (ZEuP*), 29(2), S. 227-271.

Jin, Zhiwei; Cao, Juan; Zhang, Yongdong und Luo, Jiebo (2016): News verification by exploiting conflicting social viewpoints in microblogs. *Proceedings of the AAAI conference on artificial intelligence,* 30(1). doi: 10.1609/aaai.v30i1.10382.

Jung, Laura (2023) Schutz der Demokratie durch inhaltsneutrale Regulierung digitaler Medien, *Die Öffentliche Verwaltung (DÖV)*, 76(4), S. 141-150.

Jünger, Jakob und Gärtner, Chantal (November 2020): Datenanalyse von Rechtsverstoßenden Inhalten in Gruppen und Kanälen von Messengerdiensten am Beispiel Telegram. Durchgeführt von der Universität Greifswald. Düsseldorf: Landesanstalt für Medien NRW.

Jünger, Jakob und Gärtner, Chantal (März 2021): Die Verbreitung und Vernetzung Problembehafteter Inhalte auf Telegram. Düsseldorf: Landesanstalt für Medien NRW.

Kalbhenn, Jan (2022): Medien- und wettbewerbsrechtliche Regulierung von Messenger-Diensten. *Zeitschrift für Urheber- und Medienrecht (ZUM)*, 66(4), S. 266-277.

Klimpel, Lena (1. Mai 2021): Wie Politikerinnen im Netz diskreditiert werden. URL: https://www.tagesschau.de/faktenfinder/geschlechtsspezifische-desinformation-101.html.

Kühling, Jürgen (2021): Die Verantwortung der Medienintermediäre für die demokratische Diskursvielfalt. Algorithmenregulierung für Facebook, Twitter & Co.?, *Juristen Zeitung (JZ)*, 76(6), S. 529-530.

Kuhlmann, Simone/Trute, Hans-Heinrich (2022): Die Regulierung von Desinformationen und rechtswidrigen Inhalten nach dem neuen Digital Services Act, *Zeitschrift für das gesamte Sicherheitsrecht* (GSZ), 5(3), 115-123.

Kuß, Christian und Lehmann, Daniel (2021): Digital Services Act – Entwurf eines einheitlichen Rechtsrahmens für die EU-Digitalwirtschaft. *Der Betrieb (DB)*, 74(12), S. 605-610.

Kwon, Sejeong; Cha, Meeyoung; Jung, Kyomin; Chen, Wei und Wang, Yajun (2013): Prominent features of rumor propagation in online social media. *2013 IEEE 13th international conference on data mining,* S.1103-1108. doi: 10.1109/ICDM.2013.61.

Lazer, David; Baum, Matthew; Benkler, Yochai; Berinsky, Adam; Greenhill, Kelly; Menczer, Filippo; Metzger, Miriam; Nyhan, Brendan; Pennycook, Gordon; Rothschild, David; Schudson, Michael; Sloman, Steven; Sunstein, Cass; Thorson, Emily; Watts, Duncan und Zittrain, Jonathan (2018): The science of fake news. *Science*, 359, S. 1094-1096. doi:10.1126/science.aao2998.

Löber, Lena (2022): Der Forschungsdatenzugang nach dem neuen Art. 40 DSA, *Zeitschrift für Datenschutz Aktuell (ZD-Aktuell)*, 01420.

Meßmer, Anna-Katharina und Degeling, Martin (2023): *Auditing Recommender Systems – Putting the DSA into practice with a risk-scenario-based approach*. Berlin: Stiftung Neue Verantwortung.

Pennycook, Gordon; Epstein, Ziv; Mosleh, Mohsen; Arechar, Antonio A., Eckles, Dean und Rand, David G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, 592(7855), S.590-595. doi: 10.1038/s41586-021-03344-2.

Rau, Jan; Kero, Sandra; Hofmann, Vincent; Dinar, Christina; Heldt, Amélie Pia (2022): *Rechtsextreme Online-Kommunikation in Krisenzeiten. Herausforderungen und Interventionsmöglichkeiten aus Sicht der Rechtsextremismus- und Platform-Governance-Forschung*. Hamburg: Leibniz-Institut für Medienforschung/Hans-Bredow-Institut.

Regionales Informationszentrum der Vereinten Nationen (UNRIC): *UN und Partner fordern Länder auf "Infodemie" zu bekämpfen*. URL: https://unric.org/de/24092020-infodemie/.

Roßnagel, Alexander (2019): Artikel 6 Datenschutz-Grundverordnung. In: Simitis, Spiros; Hornung, Gerrit und Spiecker genannt Döhmann, Indra (Hrsg.), *Kommentar Datenschutzrecht*. Baden-Baden: Nomos.

Setz, Tahireh (2022): Desinformation in Messenger-Diensten und Hybrid-Medien – Sind NetzDG und MStV geeignete Blaupausen für die EU? In: Bernzen, Anna K; Grisse Oliveira, Karina und Kaesling, Katharina (Hrsg.): *Immaterialgüter und Medien im Binnenmarkt*. Baden-Baden: Nomos.

Shu, Kai; Wang, Suhang und Liu, Huan (2018): Understanding user profiles on social media for fake news detection. *2018 IEEE conference on multimedia information processing and retrieval (MIPR)*, S.430-435. doi: 10.1109/MIPR.2018.00092

Shu, Kai; Bernard, Russell H. und Liu, Huan (2019): Studying fake news via network analysis: detection and mitigation. *Emerging research challenges and opportunities in computational social network analysis and mining*, S.43-65. doi: 10.1007/978-3-319-94105-9_3.

Spindler, Gerald (2021): Der Vorschlag für ein neues Haftungsregime für Internetprovider – EU-Digital Services Act. *Gewerblicher Rechtsschutz und Urheberrecht (GRUR)*, 123(5), S. 653-662.

Telegram: Fragen und Antworten (FAQ). URL: https://telegram.org/faq/de.

Vosoughi, Soroush; Roy, Deb und Aral, Sinan (2018): The spread of true and false news online. *science*, 359(6380), S.1146-1151. doi: 10.1126/science.aap9559.

Wu, Ke; Yang, Song und Zhu, Kenny Q. (2015): False rumors detection on sina weibo by propagation structures. *2015 IEEE 31st international conference on data engineering*, S. 651-662. doi: 10.1109/ICDE.2015.7113322.

Zeit Online (21. Mär. 2022): *Brasiliens oberstes Gericht nimmt Telegram-Sperrung zurück*. URL: https://www.zeit.de/politik/ausland/2022-03/telegram-messenger-sperrung-brasilien-aufhebung

Zhou, Xinyi und Zafarani, Reza (2019): Network-based fake news detection: A pattern-driven approach. *ACM SIGKDD explorations newsletter*, 21(2), S.48-60. doi:10.1145/3373464.3373473.

Zhou, Xinyi und Zafarani, Reza (2020): A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys*, 53(5), S.1-40. doi:10.1145/3395046.