Desinformationen und Messengerdienste: Herausforderung und Lösungsansätze

Nicole Krämer, Gerrit Hornung, Carolin Jansen, Jan Philipp Kluck, Lars Rinsdorf, Tahireh Setz, Martin Steinebach, Inna Vogel und York Yannikos

Zusammenfassung

In diesem Beitrag werden Fragestellungen und Lösungsansätze zum Thema der Erkennung und Bekämpfung von Desinformation in Messengerdiensten betrachtet. Durch die Zusammenarbeit der Disziplinen Informatik, Journalistik, Medienpsychologie und Rechtswissenschaften wird der hochdynamische Gegenstand der digitalen Desinformation im bislang wenig erforschten Bereich der Messengerdienste einer multiperspektivischen Analyse unterzogen. Dabei werden von jeder Disziplin der Stand der Forschung, bisherige eigene Erkenntnisse sowie Forschungsfragen für die Zukunft dargestellt. Außerdem wird ein Überblick über disziplinübergreifende Forschungsfragen gegeben. Vertieft diskutiert werden dabei die Einflüsse datenschutzrechtlicher Anforderungen auf die Projektarbeit.

1. Einleitung

Maßnahmen gegen die Verbreitung digitaler Desinformation werden weltweit unter Hochdruck gesucht. Eine besondere Herausforderung stellt dabei die Erkennung und Bekämpfung bewusst irreführender Informationen dar, die in Messengerdiensten verbreitet werden. Problematisch sind hier sowohl die sozio-technischen Besonderheiten der massenhaften Desinformationsverbreitung als auch die Anwendbarkeit, Durchsetzbarkeit und Effektivität rechtlicher Maßnahmen. Desinformation erweist sich dabei als hochdynamischer Gegenstand sowohl bezogen auf die verwendeten Technologien als auch im Hinblick auf die Verbreitungskanäle, die Aufbereitungsformen, die rechtliche Bewertung und die politischen Kontroversen, an denen entlang sie sich besonders stark verbreitet. Als Desinformation werden falsche Informationen bezeichnet, die wissentlich mit der Intention verbreitet werden, einem Individuum, einer sozialen Gruppe, Organisation oder einem Staat Schaden zuzufügen. Wird Desinformation weiterverbreitet, entsteht aus ihr häufig eine Misinformation. Diese beschreibt

ebenfalls falsche Tatsachenbehauptungen, allerdings erfolgt die Weiterverbreitung nicht in dem Wissen, dass es sich um falsche Informationen handelt und nicht in der Absicht, Schaden zu verursachen (Wardle 2017). Diese zwei Kerndimensionen der Definition sind dabei zugleich Gegenstand wissenschaftlicher Debatten, da sowohl der Wahrheitsbegriff als auch die Intention als volatil erachtet werden können. Bildeten in den vergangenen Jahren noch soziale Netzwerke sowie quasi-journalistisch aufbereitete Web-Portale den Schwerpunkt von Desinformationsdynamiken, verlagern sie sich heute zunehmend in Messengerdienste. Denn nachdem die sozialen Netzwerke mehr und mehr reguliert wurden, zogen sich Desinformationsakteur:innen in den sicher gewähnten Hafen der Messengerdienste zurück. Gleichzeitig gewinnt Videomaterial an Bedeutung. Neben den privatheitsbezogenen Aspekten, die dadurch berührt werden, haben diese Entwicklungen eine hohe Relevanz für das Funktionieren von Demokratien. Diese komplexe Problemlage verlangt nach einem multiperspektivischen Ansatz, um sie in all ihren Dimensionen angemessen bearbeiten zu können. Interdisziplinäre Konsortien müssen Strategien und Instrumente entwickeln, um Desinformation unter den aktuellen Bedingungen zu erkennen und zu bekämpfen. Dabei erscheint es besonders vielversprechend, technische Ansätze wie etwa maschinelles Lernen (ML) mit Regulierungen zu kombinieren, die auf die Praktiken der Nutzer:innen bei der Verbreitung von Desinformation zugeschnitten sind und gleichzeitig einen angemessenen Grundrechtsschutz gewährleisten.

Desinformationsstrategien folgen der Dynamik politischer Kontroversen bzw. gesellschaftlicher Problemlagen. Nachdem seit 2015 Desinformation als Phänomen im deutschen Sprachraum sehr stark von den Themen Migration, Integration und innerer Sicherheit geprägt war (Bader u.a. 2020, S. 49), lässt sich derzeit eine Dynamik beobachten, die sich primär entlang der COVID 19-Pandemie und des Ukrainekrieges entwickelt.

Kern dieses Beitrags ist die Frage, welche Auswirkung die Verwendung von Messengerdiensten auf die Verbreitung von Desinformation hat. Entsprechend orientiert sich der Beitrag an einem gemeinsamen, interdisziplinären Vorgehen entlang folgender forschungsleitender Fragen:

- Wie verbreitet sich Desinformation von Messengerdiensten aus weiter? Wie erreicht sie andere soziale Medien? Existieren Muster, die Bekämpfungsstrategien ermöglichen?
- Welche Nutzungspraktiken stehen hinter den beobachtbaren Verbreitungsmustern? Welche Rolle spielt dabei insbesondere die Verbreitung emotionaler Inhalte?

- Welche Eigenschaften weisen Desinformationen auf, die sich besonders gut in Messenger-Netzwerken verbreiten?
- Wie lassen sich Desinformationen, die primär über Messengerdienste verbreitet und initiiert werden, bekämpfen? Welche gesellschaftlichen, juristischen und technischen Maßnahmen erscheinen als besonders erfolgversprechend?

In der jüngeren Vergangenheit wurde das Aufkommen von Desinformationen in sozialen Netzwerken ausführlich betrachtet, wobei unter anderem fehlende inhaltliche Kontrolle, die hohe Verbreitungsgeschwindigkeit und Effekte von Filterblasen als Gründe für das Anwachsen des Phänomens identifiziert wurden. Messengerdienste sind anderes gestaltet als soziale Netzwerke und können daher neue Ausprägungen von Desinformationen ermöglichen. Dabei ist allerdings ein Trend zur Hybridisierung der Medientypen auszumachen, da neben Messengerdiensten, die vor allem Individualkommunikation anbieten, immer mehr sog. Hybrid-Medien existieren, die neben Messenger-Funktionen auch andere Kommunikationsformen beinhalten, wie beispielsweise One-to-many oder Many-to-many. Je nach Typ kann hier untersucht werden, welche Dynamiken sich um welche Kommunikationsformen von Desinformation entwickeln. Gleichzeitig wird Desinformation stärker eingebettet in die Interaktion von Akteuren, die auf diesen Plattformen agieren, sodass dies ebenfalls bei der Frage zu berücksichtigen ist, welche Themenschwerpunkte, Positionen, narrative und argumentative Strukturen, affektiven Potentiale und sprachlichen Gestaltungsmuster zur Maximierung von Aufmerksamkeit genutzt werden.

Im Folgenden werden Aspekte aus den Perspektiven unterschiedlicher Disziplinen betrachtet. Dabei werden von jeder Disziplin der Stand der Forschung, bestehende eigene Erkenntnisse sowie Forschungsfragen für die Zukunft dargestellt. Anschließend wird die disziplinübergreifende Problematik des mit der Anonymisierung bzw. Pseudonymisierung personenbezogener Daten einhergehenden Informationsverlusts vertieft. Im Schlusskapitel liegt das Augenmerk auf den Ausblick des interdisziplinären Vorgehens.

2. Informatik

Die technische Perspektive des Projektes beinhaltet primär das Erfassen der Inhalte aus verschiedenen Kanälen und danach deren Analyse durch Methoden der Wiedererkennung von Inhalten (beispielsweise durch robuste Hashverfahren), der Multimedia Forensik und des Natural Language

Processings (NLP). Weiterhin ist aber auch Datenschutz von hoher Relevanz, weshalb auch Ansätze von Privacy by Design eingesetzt werden.

2.1 Stand der Forschung

Für die Datenerfassung im Projekt werden Methoden des Web-Scrapings und -Crawlings verwendet. Hierbei wird zunächst untersucht, ob die relevanten Plattformen, auf denen Daten erhoben werden sollen (bspw. Messengerdienste wie Telegram), bereits Schnittstellen zur Verfügung stellen, die eine automatisierte Datenerfassung effizient ermöglichen. Ist dies nicht der Fall, müssen speziell angepasste Crawler entwickelt werden, um die Inhalte der Plattformen zu erfassen. Dabei kann der Implementierungsaufwand eines solchen Crawlers schnell steigen, insbesondere dann, wenn durch die Plattformbetreiber Mechanismen eingesetzt werden, die eine automatisierte Datenerfassung verhindern sollen (bspw. Captchas). Ziel der Datenerfassung im Projekt ist, ein robustes und effizientes Crawling relevanter Daten zu ermöglichen und dabei "Privacy by Design" zu berücksichtigen.

Um Desinformationen in Texten automatisch zu erkennen, können die Forschungsansätze in drei grobe Bereiche aufgeteilt werden: context-, style- und knowledge-based (Potthast u.a. 2018). Bei den ersten beiden Ansätzen wird entweder der Schreibstil oder Metainformationen (wie Profildaten in sozialen Netzwerken) genutzt, um Fake News mithilfe von maschinellen Lernverfahren zu erkennen. Der wissensbasierte Ansatz (knowledge-based) zielt darauf ab, externe Quellen zu nutzen, um zu überprüfen, ob es sich um gefälschte oder echte Nachrichten handelt. Viele Forschungsarbeiten befassen sich mit der automatisierten Identifizierung von Falschmeldungen sowie deren Verbreitungswege in Sozialen Medien (Shu u.a. 2017; Vogel/Meghana 2020). Wie Fake News in Messenger-diensten wie Telegram oder WhatsApp als solche automatisiert identifiziert werden können bzw. deren Verbreitungswege wurde bis jetzt wenig erforscht. Für Telegram wird bis heute z.B. keine offizielle Unterstützung der Faktenüberprüfung angeboten.

Bei der Erkennung von Inhalten muss zwischen merkmalsbasierten und robusten hashbasierten Verfahren unterschieden werden. Beide versuchen, Inhalte auch nach einer Veränderung wiederzuerkennen. Das Grundkonzept der Merkmalserkennung besteht darin, Merkmale aus relevanten Bildbereichen zu finden. Diese Bereiche werden extrahiert und durch einen Merkmalsdeskriptor beschrieben. Diese Beschreibung kann dann für die Re-Identifizierung verwendet werden (Hassaballah u.a. 2016). Hashba-

sierte Algorithmen werden in verschiedenen Anwendungsbereichen eingesetzt, z. B. bei der Bildsuche, der Erkennung von Duplikaten oder Beinahe-Duplikaten oder der Bildauthentifizierung (Du u.a. 2020). Robuste Hashes überstehen Veränderungen wie verlustbehaftete Komprimierung oder Skalierung. Analog sind solche Verfahren auch für Video und Ton bekannt.

2.2 Eigene Beiträge

In mehreren Forschungsarbeiten hat sich gezeigt, dass maschinelle Lernverfahren geeignet sind, um potenzielle Desinformationen in Nachrichtentexten und Sozialen Medien automatisiert zu erkennen. So werden je nach Verfahren und Datensatz Genauigkeitswerte von bis zu 90% erreicht (Vogel/Jiang 2019; Steinebach u.a. 2020; Vogel/Meghana 2020). Stilbasierte- und linguistische Merkmalsanalysen haben gezeigt, dass Falschmeldungen oft emotionaler und reißerischer verfasst sind. Der Ton ist oft negativer, es wird weniger auf Rechtschreibung geachtet sowie auf den sprachlichen Ausdruck (Vogel/Meghana 2018/2020). In sozialen Netzwerken wie Twitter unterscheiden sich die Posts dadurch, dass Emojis verwenden werden und auf andere User:innen referiert wird (mithilfe von @-Mentions), wohingegen Fake News-Spreader öfter Hashtags nutzen und URLs posten, um die Falschinformationen zu verbreiten. Ob solche Merkmale (Features), die für das Training von maschinellen Lernverfahren relevant sind, in Messengerdiensten zu finden sind, muss im Verlauf des Projekts erforscht werden. Das Problem bei stil- und metadatenbasierten maschinellen Lernverfahren ist, dass diese an neue Ereignisse und die damit einhergehenden sprachlichen Änderungen („Coronaleugner", „Rape-Fugees", „Demokratur") angepasst und neu trainiert werden müssen. Das bedeutet, dass neue händisch gelabelte Daten zur Verfügung gestellt werden müssen. Da dies einen hohen Zeit- und Personalkostenbedarf erfordert, werden wissensbasierte maschinelle Lernansätze und robuste Text- und Medienhashverfahren erforscht, um so Verbreitungswege auch beim Wechsel von Kanälen oder leichten Änderungen der Inhalte erkennen zu können (Steinebach u.a. 2013; Steinebach u.a. 2019; Steinebach u.a. 2020).

2.3 Forschungsfragen

Fragestellung 1: Sind stil- und metabasierte maschinelle Verfahren geeignet, um Falschinformationen in Messengerdiensten zu erkennen? Bei der automa-

tisierten Erkennung von Desinformationen werden stilbasierte maschinelle Lernansätze und Methoden der Computerlinguistik angewandt, um zu erforschen, ob diese Desinformationen in Messengerdiensten wie Telegram von korrekten Meldungen unterscheiden können. Dabei können Merkmale herangezogen werden wie Emotionen und Stimmungen in den Meldungen, aber auch syntaktische Merkmale wie Hashtags, URLs oder Themenschwerpunkte. Ziel ist es zu erforschen, ob die Verbreitung von Falschmeldungen in sozialen Netzwerken sich von der in Messengerdiensten wie Telegram unterscheidet.

Fragestellung 2: Gleichzeitig sollen quantitative Analysen durch einen Vergleich von Datenquellen und deren Metadaten umgesetzt werden, beispielsweise durch Wiedererkennen von Nachrichten in verschiedenen Kanälen durch robuste Text- und Medienhashverfahren, um so Verbreitungswege auch bei Wechseln von Kanälen oder leichten Änderungen der Inhalte erkennen zu können. Ziel ist es neben der Wiedererkennung von Falschmeldungen auch Mechanismen zur Gegenaufklärung zu entwickeln, beispielsweise durch zeitige Reaktionsmöglichkeiten.

Die erzielten Analyseergebnisse sollen dabei jeweils nachvollziehbar darstellbar sein.

3. Journalistik

Die Journalistik arbeitet durch qualitative und quantitative Inhaltsanalysen textimmanente Merkmale, Narrative und Strukturen von Desinformationen heraus. Darüber hinaus wird untersucht, inwieweit Themenkarrieren in ausgewählten Massenmedien und Messengerdiensten parallel verlaufen und wie Nutzer:innen von Messengerdiensten die dort verbreiteten (Falsch-)Informationen in ihre übrigen Medienrepertoires integrieren. Indem Diskursverläufe und Nutzungspraktiken medienübergreifend untersucht werden, werden technische, psychologische und rechtliche Ansatzpunkte für die Bekämpfung der Verbreitung von Desinformation identifiziert. Nicht zuletzt evaluiert die Journalistik die erarbeiteten Regulierungsansätze aus der Perspektive wichtiger Adressat:innengruppen und transferiert sie in die Praxis: Dazu führt sie einen systematischen Dialog mit Akteur:innen aus Journalismus und Zivilgesellschaft, die Desinformation bekämpfen, um über die Effektivität und Effizienz der erarbeiteten Ansätze zu diskutieren und Einsatzmöglichkeiten in der Praxis zu erörtern.

3.1 Stand der Forschung

Verbreitungsdynamiken von Desinformationen lassen sich nur entlang von Themen identifizieren, zu denen in größerem Umfang Desinformationen erstellt und verteilt werden. Dies ist stark abhängig von aktuellen, kontroversen Debatten. So haben sich Desinformationsschwerpunkte in den vergangenen Jahren von den Themen Innere Sicherheit und Migration (Bader u.a. 2020) zu den Themen Klimawandel und insbesondere COVID-19 (Lamberty/Holnburger 2021) verschoben. Im Laufe der Pandemie haben sich die Gruppierungen der „Querdenker"-Proteste (Holzer 2021; Pantenburg u.a. 2021) zu einer spektrenübergreifenden rechtsoffenen Protestszene gewandelt (BMB 2021), die bundesweit durch den Verfassungsschutz beobachtet wird (Bundesamt für Verfassungsschutz 2021). Anhänger:innen dieser Protestszene lassen sich in heterogene, häufig disparate Gruppen einteilen (Nachtwey u.a. 2020, S. 51), die Verbindungen zu sogenannten „Reichsbürgern" und Rechtsextremisten offenlegen. Charakteristisch ist eine starke Entfremdung von den Institutionen des politischen Systems, den alten Volksparteien und den etablierten Medien (ebd., S. 52). Letzteres führt zu einer zunehmenden Radikalisierung des öffentlichen Diskurses sowie zu einer Hinwendung zu alternativen Medien und ihren Plattformen wie Telegram (RND 2021). Für den deutschsprachigen Raum existieren bereits hunderte Kanäle, die dem Austausch und der Protestorganisation dienen (Dittrich/Holnburger 2021). Während für das außereuropäische Ausland bereits eine Vielzahl an Studien vorliegt, die die inhaltliche Komponente der Verbreitung von Desinformation über Telegram systematisch durchleuchtet (siehe u.a. Baumgartner u.a. 2020; Bovet/Grindrod 2020; Guhl/Davey 2020; Rogers 2020; Scheffler u.a. 2021), ist für den deutschsprachigen Raum derzeit noch wenig bekannt. Ausnahmen stellen die Analysen des Bundesverbands Mobile Beratung (BMB 2021) sowie die Untersuchungen des Centers für Monitoring, Analyse, Strategie (Lamberty/Holnburger 2021) dar. Um die Gefahren, die von über Messengerdienste verbreiteter Desinformation ausgehen (Flade/Mascolo 2021), erkennen und bekämpfen zu können, ist daher eine systematische Auseinandersetzung mit den inhaltlichen Schwerpunkten notwendig.

3.2 Eigene Beiträge

Um belastbare Schwerpunktthemen zu identifizieren, an denen entlang Desinformationsdynamiken untersucht werden, beobachtet die Journalistik systematisch den medialen Diskurs auf kontroverse Themen. Im Fokus

stehen Themen mit erkennbar populistischen Tendenzen, die Themenschwerpunkte von Websites, die Hubs in Netzwerken zur Verbreitung von Desinformationen sind, und Inhalte öffentlicher Telegram-Kanäle und -Gruppen von zentralen Akteur:innen aus dem populistischen, extremistischen und Verschwörungsmythen verbreitenden Milieu. Dieser Medienbeobachtung liegen qualitative Leitfäden zur Identifikation kontroverser Themen zugrunde. Die Ergebnisse werden in Expert:inneninterviews validiert.

Für die systematische Analyse greift die Journalistik auf bewährte Methoden quantitativer und qualitativer Inhaltsanalyse zurück. Die Methodik wird im Folgenden anhand spezifischer Forschungsfragen näher betrachtet.

3.3 Forschungsfragen

Fragestellung 1 (Themen): Entlang welcher politischer bzw. gesellschaftlicher Kontroversen entwickeln sich Desinformationsdynamiken?

Im Fokus des Untersuchungsinteresses stehen jene Themen, zu denen in größerem Umfang Desinformationen erstellt und verbreitet werden und die stark abhängig von aktuellen, kontroversen Debatten sind. Daher erfolgt zur Identifikation belastbarer Schwerpunktthemen eine systematische Beobachtung des medialen Diskurses auf kontroverse Themen mit erkennbar populistischen Tendenzen und der Themenschwerpunkte von Websites, die Hubs in Netzwerken zur Verbreitung von Desinformationen sind (Rathje 2021). Ziel ist es anschließend, systematisch die Inhalte von öffentlichen Telegram-Kanälen von zentralen Akteur:innen aus dem populistischen, extremistischen und Verschwörungsmythen verbreitenden Milieu zu tracken.

Fragestellung 2 (Kanäle): Mit welchen Messengerkanälen lassen sich welche Themen scannen, um möglichst breitenwirksam die Verbreitung von Desinformation erfassen zu können?

Die Journalistik automatisiert die Erfassung von Daten aus Messengerdiensten, um deren Rolle als Verbreitungswege zu untersuchen. Dabei sollen große, quasi-öffentliche Kanäle und Gruppen von Meinungsführer:innen beobachtet werden (Jalilvand/Neshati 2020). Hier wird überprüft, ob dort initial verbreitete Nachrichten ihren Weg in andere soziale Medien wie Twitter finden und welchen Einfluss die Nennung eines Videolinks in YouTube auf dessen Nutzungszahlen hat. So kann quantitativ die

Bedeutung der Messengerdienste erfasst werden. Weiterhin werden Desinformationskanäle und -gruppen, nicht nur bezogen auf eine individuelle Nachricht, sondern mit Blick auf den gesamten Kanal bzw. die gesamte Gruppe erfasst. So lassen sich statistisch belastbarere Aussagen treffen und die Fehlerraten der Erkennung reduzieren.

Fragestellung 3 (Inhalte): Welche Nutzungspraktiken stehen hinter den beobachtbaren Verbreitungsmustern? Welche Rolle spielt dabei insbesondere die Verarbeitung von emotionalen Inhalten?

Im Vordergrund stehen zunächst die Praktiken von Nutzer:innen bei der Verbreitung von Desinformation, die mit ethnographischen Methoden rekonstruiert werden: Welche Arten von Desinformation rezipieren und teilen sie auf welchen Kanälen? Des Weiteren explorieren wir, wie die Nutzung und Verbreitung von Desinformation in Medienrepertoires eingebettet werden (Lou u.a. 2021; Schwarzenegger 2022). Die Ergebnisse fließen in die rechtswissenschaftliche Analyse der Schutzbedürftigkeit unverzerrter Kommunikation in demokratischen Gesellschaften ein.

Fragestellung 4 (Verbreitung): Welche Eigenschaften haben Desinformationen, die sich besonders gut in Messenger-Netzwerken verbreiten?

Bezogen auf die Bedeutung von emotionalen Komponenten in der Interaktion mit Desinformation untersucht die Journalistik in qualitativen, quantitativen und automatisierten Inhaltsanalysen, was Desinformation kennzeichnet, die besonders starke Verbreitungsdynamiken auslöst (Knuutila u.a. 2020). Um Desinformation gezielt bekämpfen zu können, werden die Daten um weitere Eigenschaften wie Formate (z.B. Video, Podcast, Text) und Gestaltungsmerkmale (thematisch, visuell und sprachlich) angereichert. Dazu greift die Journalistik insbesondere auf eigene Vorarbeiten zurück (Bader u.a. 2020). Instrumente und Strategien zur Eindämmung von Desinformation in der öffentlichen Kommunikation berücksichtigen zugleich die Kontroversen, entlang derer sich Desinformationsdynamiken entwickeln. In einer explorativen Inhaltsanalyse wird daher nachgezeichnet, wie die Verbreitung von Desinformationen in der Bevölkerung verschränkt ist mit den Themenkarrieren der Kontroversen, an die sie anknüpfen.

Fragestellung 5 (Bekämpfen): Wie lassen sich Desinformationen, die primär über Messengerdienste verbreitet oder initiiert werden, bekämpfen?

Zur Bekämpfung von Desinformation, die primär über Messengerdienste verbreitet werden, identifiziert die Journalistik Handlungsmuster, basierend auf einer qualitativen Analyse der Praktiken der Weiterverbreitung

von Desinformation und deren Einbettung in Medien-Repertoires durch Online-Beobachtung, Desk-Research und qualitative Interviews (Buchanan 2020; Nachtwey u.a. 2020; Schwarzenegger 2022).

Fragestellung 6 (Evaluation): Wie sind die Strategien zur Eindämmung von Desinformationsdynamiken aus Sicht der Produktion und Nutzung journalistischer Inhalte und Desinformation bzw. deren Einbettung in Medienrepertoirs zu bewerten?

Die Journalistik evaluiert Strategien zur Eindämmung von Desinformationsdynamiken aus Sicht der Produktion und Nutzung journalistischer Inhalte und Desinformation bzw. deren Einbettung in Medienrepertoirs (Michailidou/Trenz 2021).

Fragestellung 7 (Dialog): Wie bewerten Praktiker:innen Effektivität und Effizienz der im Projekt erarbeiteten Ansätze zur Bekämpfung von über Messengerdienste verbreiteter Desinformation?

Im Hinblick auf die Verbreitung und Darstellung der Projektergebnisse leitet die Journalistik, auch unter Zuhilfenahme eines von der Informatik entwickelten Demonstrators und Einbeziehung der Partner:innen, einen systematischen Dialog mit Akteur:innen aus Journalismus und Zivilgesellschaft, die Desinformation bekämpfen, um über die Effektivität und Effizienz der im Projekt erarbeiteten Ansätze im Sinne eines „Member Check" zu diskutieren, deren Bekanntheit zu steigern und Einsatzmöglichkeiten in der Praxis zu erörtern.

4. Medienpsychologie

Anknüpfend an die journalistische Perspektive untersucht das medienpsychologische Teilprojekt, welche Rolle Menschen bei der Verbreitung von Falschinformationen (d.h. sowohl Des- als aus Misinformation) spielen. Anhand von Umfragen und Experimenten wird insbesondere exploriert, inwiefern das (emotionale) Erleben einer Falschinformation dazu führt, dass diese weitergeleitet wird, inwiefern die wahrgenommene Glaubwürdigkeit diesen Prozess beeinflusst und welche Motive zur Weiterleitung vorherrschen. Zusätzlich wird analysiert, welche Wirkung verschiedene Medienmerkmale auf die Wahrnehmung von Falschinformationen haben, um zu eruieren, welche technischen Faktoren einen Einfluss auf die individuelle Bereitschaft haben, Falschinformationen zu teilen.

4.1 Stand der Forschung

Es konnten bereits wertvolle Erklärungsansätze für die psychologischen Wirkmechanismen von Desinformation hervorgebracht werden. Unter anderem wurden kognitive Verzerrungen wie der Confirmation Bias (Nickerson 1998) oder Motivated Reasoning (Kunda 1990) als Erklärung dafür herangezogen, dass Menschen Falschinformationen Glauben schenken (z.B. Lazer u.a. 2018;). Demnach werden vor allem Inhalte für glaubwürdig befunden, die dem eigenen Weltbild entsprechen und die eigene (politische) Identität schützen, da sich wiedersprechende Kognitionen, also kognitive Dissonanz, negative Gefühle verursachen (Festinger 1957). Aber auch die individuelle Fähigkeit und Neigung, Inhalte kognitiv zu reflektieren, werden als zentrale Erklärung für die Wirkung von Desinformation hervorgehoben (z.B. Pennycook/Rand 2021). Weiterhin wird in der Forschung betont, dass Menschen aufgrund limitierter kognitiver Kapazitäten nur ein gewisses Maß an Informationen elaboriert verarbeiten können. Daher wird angenommen, dass sich Internetnutzer:innen häufig auf sogenannte heuristische Hinweisreize verlassen, die ihnen mentale Abkürzungen bei der Bewertung von Online-Inhalten erlauben (z.B. Metzger/Flanagin 2013; Sundar 2008). Solche Hinweisreize sind in der Regel saliente, leicht zu verarbeitende Merkmale einer Nachricht. Zum Beispiel ziehen Menschen die Reputation einer Informationsquelle heran, wenn sie entscheiden, welche Inhalte sie für glaubwürdig erachten (Reinhard/Sporer 2010) und lesen (Winter/Krämer 2014). Ferner zeigen verschiedene Studien (z.B. Kim u.a. 2019), dass Menschen dazu neigen, der sichtbaren Bewertung anderer Nutzer:innen (z.B. Likes oder Ratings) im Internet zu folgen, wenn es zur Bewertung von Online-Inhalten kommt (vgl. Bandwagon Heuristic; Sundar 2008). Aufbauend auf diese Ergebnisse werden Warnhinweise (z.B. durch Faktencheck-Initiativen) als vielversprechende Maßnahme gegen Desinformation im Internet erachtet, da diese Intervention Menschen ressourcenschonend bei der Evaluation von Informationen unterstützen kann (Pennycook/Rand 2021). Die aktuelle Befundlage zur Wirkung von Warnhinweisen ist jedoch gemischt. Zwar konnten mehrere Studien zeigen, dass Rezipient:innen Falschinformationen mit Warnhinweisen als weniger akkurat bewerten und seltener teilen (z.B. Mena 2020) - allerdings konnten solche Effekte nicht konstant verifiziert werden (z.B. Oeldorf-Hirsch u.a. 2020).

4.2 Eigene Beiträge

Auch im Zuge eigener Forschung konnte die medienpsychologische Perspektive dazu beitragen, die Wirkmechanismen von Desinformation besser zu verstehen. Es wurde zum Beispiel untersucht, welche Charakteristika einen falschen Online-Nachrichtenartikel glaubwürdig bzw. unglaubwürdig erscheinen lassen (Schaewitz u.a. 2020). Es zeigte sich, dass Nachrichtenfaktoren wie zum Beispiel inhaltliche Widersprüche oder Sensationalismus weniger wichtig für die Glaubwürdigkeitsbewertung eines Online-Artikels sind. Allerdings erwies sich das generelle Bedürfnis nach kognitiver Beschäftigung (Need for Cognition) als wichtiger Prädiktor für die Bewertung der Nachrichtenkorrektheit und -glaubwürdigkeit. Auch waren Menschen eher dazu geneigt, die Falschinformation zu glauben und weiterzuleiten, wenn der Inhalt des Nachrichtenartikels ihre eigene Meinung stützte.

In einer weiteren Studie wurde gezeigt, dass Menschen bei der Evaluierung eines falschen Online-Nachrichtenartikels eher auf die Glaubwürdigkeitsbewertung anderer zurückgreifen, wenn diese in Form eines konkreten Kommentars präsentiert wird und nicht in Form eines numerischen Ratings (Kluck u.a. 2019). Darüber hinaus demonstrierten Schaewitz und Krämer (2020), dass detailreichere Korrekturen von Falschinformationen das Erinnern von zugehörigen Fakten begünstigen. Der Zeitpunkt der Korrektur wies allerdings einen widersprüchlichen Effekt auf: Wenn die detailliertere Korrektur mit der Falschinformation zusammen präsentiert wurde und nicht im Nachhinein, konnten sich Individuen zwar besser die richtigen Fakten merken, der Glaube an die Kernaussage der Falschinformation verstärkte sich allerdings. Daraus kann erschlossen werden, dass Interventionen genau geplant und orchestriert werden müssen.

4.3 Forschungsfragen

Zusammengefasst hat die bisherige medienpsychologische Forschung wichtige Mechanismen identifizieren können, die erklären, warum Menschen Falschinformationen glauben. Gleichzeitig erweist sich Desinformation als ein sehr dynamischer Forschungsgegenstand. Insbesondere die Covid-19 Pandemie hat aufgezeigt, dass sich solche Inhalte über eine Vielzahl von Kanälen verbreiten und in sehr unterschiedlicher Art und Weise manifestieren (z.B. Hansson u.a. 2021). Daher ist weitere Forschung von Nöten, um zu verstehen, welchen Einfluss diese Dynamiken auf der Rezeptionsebene haben. Um zu helfen, die übergreifenden Forschungsfragen

des Projekts zu beantworten, wird das medienpsychologische Teilprojekt die folgenden Fragestellungen fokussieren:

Fragestellung 1: Welche Falschinformationen *nehmen Menschen wahr, die über Messengerdienste verbreitet werden und in welchem Maße werden diese Inhalte weitergeleitet?*

Vor allem die Mechanismen, die dazu führen, dass Falschinformationen weitergeleitet wird, sind noch vergleichsweise wenig erforscht. Das ist insofern kritisch, als die intuitive Annahme, dass Menschen vornehmlich Inhalte teilen, die sie auch für glaubwürdig halten, nicht zuzutreffen scheint (Pennycook u.a. 2021, 2020). Wenngleich zum Teilen von Online-Inhalten bereits Forschung existiert (Kümpel u.a. 2015), konzentrieren sich Untersuchungen meist auf das Weiterleiten von Nachrichtenartikeln in sozialen Medien wie Twitter oder Facebook. Allerdings haben insbesondere Messenger-Anwendungen wie WhatsApp einen starken Einfluss auf die Verbreitung von falschen Inhalten (Resende u.a. 2019)). Da bisher wenig darüber bekannt ist, in welchem Maße Falschinformationen über Messengerdienste tatsächlich wahrgenommen und weitergeleitet werden, wird in einem ersten Schritt des Teilprojekts exploriert, welche und wie viele dieser Inhalte von Menschen gesehen und mit anderen geteilt werden.

Fragestellung 2: Welche spezifischen psychologischen Mechanismen begünstigen das Weiterleiten von Falschinformationen *über Messengerdienste?*

Da vor allem Menschen und nicht etwa automatisierte Programme Hauptkatalysator für die Verbreitung von Desinformation sind (Vosoughi u.a. 2018), ist es unerlässlich, die psychologischen Mechanismen zu identifizieren, die dazu führen, dass irreführende Inhalte über weniger einsehbare Kanäle wie Messengerdienste geteilt werden und schließlich ihren Weg in andere Netzwerke finden. In diesem Zusammenhang wird Erregung als wichtige Einflussvariable erachtet, da vor allem emotionale Inhalte geteilt werden (z.B. Weismueller u.a. 2022). Zudem konnte herausgefunden werden, dass neben nachrichtenbezogenen Motiven zum Teilen einer Information soziale Motive eine wichtige Rolle spielen (Chen u.a. 2015; Lee/Ma 2012). Bisher fehlen allerdings Erkenntnisse, wie sich diese Befunde auf Messengerdienste übertragen lassen. Auch gibt es bisher kaum Forschung, die untersucht hat, wie das emotionale Erleben bei der Konfrontation mit (falschen) Informationen die unterschiedlichen Motive des Weiterleitens beeinflusst.

Fragestellung 3: Wie beeinflussen Medienmerkmale das Weiterleitungsverhalten der Rezipient:innen?

Da Messengerdienste anders funktionieren als soziale Netzwerke wie Facebook oder Twitter, ist zu erwarten, dass es Unterschiede bei den Inhalten gibt, die frequentiert weitergeleitet werden. Auch konzentrierten sich Untersuchungen zu sozialen Netzwerken primär auf die Wirkung von (falschen) Nachrichtenartikel (Pennycook/Rand 2021). In Messengerdiensten scheinen jedoch insbesondere visuelle Stimuli wie Bilder oder Videos geglaubt (Sundar u.a. 2021) und geteilt zu werden (Resende u.a. 2019). Ein weiterer wichtiger Aspekt ist, dass Messengerdienste stärker auf private interpersonelle Interaktion ausgelegt sind als andere soziale Netzwerkplattformen. Daher könnte hier vor allem die verstärkte Interaktion mit Freund:innen und Familienmitgliedern dazu führen, dass Desinformation in einer „natürlichen" Weise verbreitet wird (Buchanan/Benson 2019).

Fragestellung 4: Welche technischen Interventionsmaßnahmen gegen die Verbreitung von Falschinformationen *in Messengerdiensten können aus medienpsychologischer Perspektive abgeleitet werden und wie effizient sind diese Maßnahmen?*

Die Beantwortung der ersten drei Forschungsfragen mündet schließlich innerhalb des interdisziplinären Ansatzes von DYNAMO in dem Beitrag, Instrumente zur Bekämpfung von Desinformation zu entwickeln. In dem medienpsychologischen Projekt wird dann vor allem evaluiert, welche Strategien auf der Rezeptionsebene wirksam sind.

5. Rechtswissenschaften

Messengerdienste stellen aus juristischer Perspektive eine besondere Herausforderung dar. Seitdem soziale Netzwerke mit Plattformcharakter mehr und mehr reguliert wurden, ziehen sich Desinformationsakteure zunehmend in Messengerdienste zurück. Indes ist unklar, ob und inwieweit bestehende rechtliche Vorgaben auf diese anwendbar sind, zumal hinsichtlich öffentlicher und privater Kommunikationsfunktionen unterschiedliche rechtliche Anforderungen zu beachten sind.

5.1 Stand der Forschung inkl. eigener Beiträge

Die rechtswissenschaftliche Forschung zur digitalen Desinformation behandelt bisher primär öffentlich sichtbare Kommunikationsräume sozialer Netzwerke. Die für Messengerdienste typischen geschlossenen Gruppen, private One-to-one-Kommunikation und Grenzbereiche zwischen öffentlicher und privater Kommunikation z.B. in sehr großen geschlossenen Gruppen wurden hingegen bislang wenig untersucht.

Bereits die Definition der Desinformation wird kontrovers diskutiert. Viele Autoren (Steinebach u.a. 2020 S. 149, Holznagel 2020 S. 18, Feldmann 2021 S. 35, Gräfe 2020 S. 39) orientieren sich an der aus dem Kontext der Rechtsprechung des Bundesverfassungsgerichts zur Meinungsfreiheit stammenden Formulierung der „bewusst unwahren Tatsachenbehauptung" (BVerfG NJW 1976, 1677). Jedoch wird diese ständige Rechtsprechung im Zuge der Desinformationsdebatte zunehmend in Frage gestellt. Umstritten sind vor allem der zu Grunde liegende Wahrheitsbegriff und das Erfordernis einer Täuschungsabsicht.

Welcher Wahrheitsbegriff in Gesetzesformulierungen und der Rechtspraxis implementiert wird, ist eine bedeutsame Frage, da im epistemologischen und soziologischen Diskurs größtenteils eine Abkehr von einem objektivistischen Wahrheitsbegriff auszumachen ist (vgl. Pörksen 2015 S. 4 ff.;Kleeberg/Suter 2014 S. 217). Kritik an einem objektivistischen Verständnis findet sich auch in der rechtswissenschaftlichen Literatur (Schmalenbach 2005 S. 749; Theile 2012 S. 666). Im Rahmen der Desinformationsforschung wird zum Teil gefordert , diese Begriffswandlung in der Rechtswissenschaft und -praxis zu übernehmen. So werden zahlreiche rechtliche Probleme aufgezeigt, die durch das Kriterium der „Unwahrheit" verursacht werden (Dreyer u.a. 2021 S. 13). Flint wendet dagegen ein, dass es nicht zielführend sei, bei der Rechtsanwendung zunächst die Realität zu hinterfragen und philosophische Überlegungen neu zu durchdenken (Flint 2021 S. 40).

Eine weitere Kontroverse besteht bei der Frage der Intentionalität für das Phänomen der Desinformation in Abgrenzung zu weiteren Formen der sog. Information Disorder. So stellen einige zur Unterscheidung von der unabsichtlichen Misinformation bei der Desinformation auf eine Täuschungsabsicht ab (Steinebach u.a. 2020 S. 149, Ferreau 2021 S. 204). Teilweise wird sogar gefordert, Wahrhaftigkeit als immanentes Attribut von Information zu behandeln (Lipowicz/Szpor 2021 S. 348). Dreyer u.a. kritisieren hieran, dass durch diesen Ansatz, viele Verbreitungsformen falscher Informationen, z.B. die leichtfertige Weiterleitung, welche als Misinformation zu qualifizieren ist (s.o.), nicht hinreichend berücksichtigt würden,

obwohl diese wesentlich zur Verbreitung und damit zum Schadenspotenzial beitrügen (Dreyer u.a. 2021 S. 11). Zu beachten ist auch die schwierige Beweisbarkeit dieses Aspekts.

Um geeignete Maßnahmen gegen die Verbreitung von Desinformation zu finden, wird anhand unterschiedlicher Ansätze der Rechtsrahmen abgesteckt. Steinebach u.a. identifizierten Schutzgüter, die durch Desinformation betroffen sein können, namentlich die demokratische Willensbildung und die Meinungs- und Informationsfreiheit (Steinebach u.a. 2020 S. 150 ff.). Betont wird in diesem Zusammenhang die Einschätzungsprärogative des Gesetzgebers, Regelungen zu erlassen, die Risiken vorbeugen und zur Aufrechterhaltung einer funktionierenden Kommunikationsordnung beitragen (Steinebach u.a. S. 165 m.w.N.). Ferreau bejaht obendrein eine grundsätzliche Gewährleistungspflicht des Gesetzgebers für den Meinungsbildungsprozess insgesamt, welche ihn verpflichte, den Prozess vor Verfälschungen und Verzerrungen zu bewahren (Ferreau 2021 S. 205). Dagegen wählen Dreyer u.a. einen risikobasierten Ansatz, bei dem durch Desinformation verursachte abstrakte Gefahren auf Rezipientenseite im Vordergrund stehen (Dreyer u.a. 2021 S. 14). Zur Feststellung hinreichend evidenter Risiken seien die Ergebnisse der empirischen Wirkungsforschung zu Grunde zu legen, welche jedoch in vielen Bereichen noch fehlten (Dreyer u.a. 2021 S. 15 ff.). Risiken, die zu einem regulatorischen Handlungsbedarf führen, werden jedenfalls für die positive Informationsfreiheit sowie die Meinungsvielfalt nach Art. 5 Abs. 1 GG gesehen, sofern durch die künstliche Schaffung von Relevanz und/oder Reichweite die Sichtbarkeit und der Zugang zu anderen Informationen oder Ansichten faktisch ausgeschlossen würden (ebd. S. 21). Gefahren bestünden auch für die Wahlfreiheit gem. Art. 38 Abs. 1 GG, jedoch nur in unmittelbarer zeitlicher Nähe zum Wahlakt, da hier die diskursive Selbstregulierung nicht ohne weiteres möglich sei (ebd. S. 26; vgl. auch Lipowicz/Szpor 2021 S. 383). Keine ausreichende Risikoevidenz wird hingegen für die Verfassungsgüter der individuellen Autonomie, der Freiheitlichkeit der öffentlichen Meinungsbildung sowie der kommunikativen Chancengerechtigkeit gesehen (Dreyer u.a. 2021 S. 17 ff.).

Als Maßnahmen gegen die Verbreitung werden verschiedene Ansätze diskutiert. Auf gesetzlicher Ebene wird u.a. die Ergänzung des Volksverhetzungstatbestands des § 130 StGB (Mafi-Gudarzi 2019 S. 68) und die Vorgabe grober Leitplanken für die Hausregeln der Intermediäre zur Moderation desinformierender Inhalte vorgeschlagen (Kühling 2021 S. 467). Der BGH hat kürzlich entschieden, dass Anbieter sozialer Netzwerke ihren Nutzer:innen grundsätzlich objektive, überprüfbare Kommunikationsstandards vorgeben dürfen, die über die gesetzlichen Vorgaben hinausgehen.

Diese dürfen auch Löschungen von Inhalten und Sperrungen von Profilen beinhalten, wobei bestimmte Verfahrensrechte einzuräumen sind (BGH, Urteil vom 29.07.2021 - III ZR 179/20; BGH, Urteil vom 29.07.2021 - III ZR 192/20). Diese Verschränkung staatlicher und anbietereigener Regelungen, die sogenannte Hybrid Governance, wird auch von Dreyer u.a. befürwortet (Dreyer u.a. 2021 S. 46). Plattformen könnten dadurch abstrakte, auf Desinformation bezogene Infrastrukturmaßnahmen vorgegeben werden, die sie z.B. dazu verpflichten könnten, Missbrauchsszenarien zu identifizieren (Ebd. S. 65). In Bereichen, die auf Grund ihrer technischen Gestaltung für die Nutzer:innen und den Staat eine Black Box darstellen – etwa Bots – wird die Bedeutung der freiwilligen Selbstregulierung der sozialen Netzwerke betont (Steinebach u.a. 2020 S. 185). Weitere Vorschläge reichen von harten Content Moderation Maßnahmen (Mafi-Gudarzi 2019 S. 68), über die Diskursstärkung (z.B. durch rational Nudges, Enghofer 2021 S. 71), bis hin zur Einbeziehung unabhängiger Fact-Checking-Initiativen (Lipowicz/Szpor 2021 S. 383).

Indes werden nur vereinzelt Vorschläge unterbreitet, um Desinformation in Messengerdiensten zu bekämpfen. Im Rahmen der medienwissenschaftlichen Analyse des Mediums Telegram wird etwa vorgeschlagen, die Auffindbarkeit von desinformierenden Inhalten zu erschweren (Jünger/Gärtner 2020 S. 33).

5.2 Forschungsfragen

Im Forschungsprojekt DYNAMO soll der rechtswissenschaftliche Diskurs über digitale Desinformation mit dem Schwerpunkt auf Messengerdiensten fortgeführt werden. In der rechtlichswissenschaftlichen Analyse sollen rechtliche Auswirkungen sozio-technischer Besonderheiten dieses prävalenten Kommunikationsmittels untersucht werden und rechtliche Gestaltungsvorschläge, die Forschungsergebnisse der Partnerdisziplinen berücksichtigen, entwickelt werden.

Die Befassung mit Grundlagenfragen darf zu Beginn des Projekts nicht ausbleiben: Prämisse einer fundierten Desinformationsforschung ist insbesondere die kritische Auseinandersetzung mit dem rechtswissenschaftlichen Wahrheitsbegriff und der Frage nach der Implementierung eines Täuschungsvorsatzes in die Definition der Desinformation. Die Frage nach dem Wahrheitsbegriff ist u.a. für den Schutzumfang der Kommunikationsgrundrechte sowie für die freie gesellschaftliche Willensbildung relevant. Schließlich gilt es eine allgemeine staatliche Deutungshoheit über wahr und falsch zu vermeiden und zugleich ein faktisches Wahrheitsmonopol

privater Unternehmen auszuschließen. Bestehende Gesetzesformulierungen und Rechtspraktiken sind anhand geeigneter und für die Rechtswissenschaften praktikabler Wahrheitskriterien zu evaluieren. Diese wären auch bei der Formulierung von Normvorschlägen zu beachten.

Weiterhin scheint die Einbeziehung eines Täuschungsvorsatzes in die Desinformationsdefinition und die Abgrenzung zur Misinformaton (s.o.) einen effektiveren Grundrechtsschutz für die Mediennutzer:innen zu gewährleisten, die unabsichtlich zur Verbreitung falscher Informationen beitragen. Die Differenzierung zwischen bewusst unwahren und unabsichtlich unzutreffenden Tatsachenbehauptungen bildet ein kommunikationsverfassungsrechtliches Schutzniveaugefälle ab – sei es auf Ebene des Schutzbereichs (BVerfGE 61, 1 (8); 90, 241 (254); 90, 1 (15)) oder der Rechtfertigung (Wendt 2021 Art. 5 Rn. 29). Eine solche Differenzierung sollte – trotz der schwierigen Beweisbarkeit – mithin auch bei der Entwicklung einfachgesetzlicher Normen mit Sanktionscharakter berücksichtigt werden, während eine Differenzierung bei gefahrabwehrrechtlichen Normen nicht erforderlich wäre.

Sodann soll der Rechtsrahmen für die konkret zu untersuchenden Dienstetypen identifiziert werden. Dabei ist zwischen „reinen" Messengerdiensten, die in erster Linie Individualkommunikation anbieten und sogenannten Hybrid-Medien, welche daneben auch öffentlich sichtbare Kommunikationsfunktionen anbieten, zu unterscheiden (Vgl. Jünger/Gärtner 2021 S. 31), sodass die Grenzen zwischen den Geschäftsmodellen der Messengerdienste und sozialen Netzwerke verschwimmen (Jünger/Gärtner 2020 S. 6, Sunyaev u.a. 2021 S. 77). Um grundrechtsschonende, aber effektive Regularien zu entwickeln, sind bei der Identifikation des Rechtsrahmens beide Kommunikationsmodi differenziert zu betrachten.

Verfassungsrechtlich sind insbesondere die Grundrechte auf Meinungsfreiheit und Informationsfreiheit aus Art. 5 Abs. 1 S. 1, 1. bzw. 2. Hs. GG sowie auf Berufsfreiheit der Dienste-Anbieter aus Art. 12 Abs. 1 GG einschlägig - auf europäischer Ebene Art. 11 Abs. 1 S. 1 bzw. 2; Art. 15 Abs. 1 GRCh. Sofern eine Übermittlung an individuelle Kommunikationsempfänger:innen vorliegt (Jarass/Pieroth 2020 Art. 10 Rn. 6), ist auch das Fernmeldegeheimnis nach Art. 10 Abs. 1 GG zu beachten. Darüber schützt das Grundrecht auf informationelle Selbstbestimmung aus Art. 1 Abs. 1 i.V.m. Art. 2 Abs. 1 GG bzw. Art. 7 und 8 GRCh die Preisgabe und Verwendung personenbezogener Daten auch unabhängig von ihrer Übermittlung an Individualempfänger:innen oder an die Öffentlichkeit und unabhängig davon, ob der Kommunikationsvorgang bereits abgeschlossen ist (Durner 2021 Art. 10 Rn 78). In Bezug auf massenhaft verbreitete Desinformation

kann schließlich die freie öffentliche Willensbildung als Grundvoraussetzung des Demokratieprinzips aus Art. 20 Abs. 2 GG betroffen sein.

Auf einfachgesetzlicher Ebene sind insbesondere der Medienstaatsvertrag (MStV), das Netzwerkdurchsetzungsgesetz (NetzDG), das Telemediengesetz (TMG), das Telekommunikationsgesetz (TKG) und das Telekommunikation-Telemedien-Datenschutz-Gesetz (TTDSG) sowie europarechtliche Regelungen wie der Verhaltenskodex gegen Desinformation und der Digital Service Act (DSA) relevant. Wie sich in der Debatte um eine mögliche Sperrung des Anbieters Telegram zeigt, ist die Anwendbarkeit des NetzDG und des MStV auf Dienste mit Messenger-Funktionen umstritten (Tuchtfeld 2021). Nach § 1 Abs. 1 S. 3 Alt. 1 NetzDG ist Individualkommunikation ausdrücklich aus dem Anwendungsbereich des NetzDGs ausgeschlossen. Hierbei ist fraglich, was unter Individualkommunikation zu verstehen ist und ob das Gesetz zumindest auf öffentliche Kommunikationsfunktionen der Hybrid-Medien (s.o.) anwendbar ist. Zudem könnte es sich bei Hybrid-Medien um Medienintermediäre i.S.d § 2 Abs. 2 Nr. 16 MStV handeln. Diese wären nach § 93 Abs. 1 MStV verpflichtet, ihre Kriterien über den Zugang und den Verbleib von Informationen (Nr. 1) sowie bestimmte Funktionsweisen ihrer Algorithmen (Nr. 2) transparent zu machen. Diesbezüglich wird u.a. die Frage aufgeworfen, ob die Transparenzpflicht ausreichend konkret formuliert wurde, um hinreichend informative Erklärungen der Intermediäre zu erhalten (Gahntz u.a. 2021 S. 13).

Zu beachten ist, dass durch die zunehmende staatliche Regulierung und Rechtsdurchsetzung auf sozialen Netzwerken und die teilweise strengere Lösch- und Sperrpraxis der Plattformen nach eigenen Community Standards (Vgl. zum virtuellen Hausreich: BGH, Urteil vom 29.07.2021 - III ZR 179/20; BGH, Urteil vom 29.07.2021 - III ZR 192/20; Eydlin 2021), eine Rückzugswelle der Desinformationsverbreiter weg von sozialen Netzwerken und hin zu Messengerdiensten beobachtet werden kann (Jünger/Gärtner 2020 S. 4, 7, 33). Daher sind regulierungsbedingte Wechselwirkungen zwischen öffentlichen sozialen Netzwerken,Messengerdiensten und ggf. neuen Medientypen bei der Entwicklung systemisch-wirkungsvoller Regulierungsvorschläge zu berücksichtigen.

Weiterhin solltenaktuelle psychologische Erkenntnisse beachtet werden, die verschiedene kognitive Verzerrungen als mitursächlich für die Verbreitung von Desinformation sehen (s.o.). Unter Einbeziehung der in der Projektlaufzeit gewonnenen medienpsychologischen Erkenntnisse der Projektpartner, sollte die rechtswissenschaftliche Forschung insbesondere Maßnahmen eruieren, die die rationale Auseinandersetzung der Nutzenden mit kontroversen Themen fördern, etwa Nudges to reason (Enghofer 2021 S. 71). Gerade bezüglich der Sphäre des Dark Social, also Kommuni-

kationsbereiche sozialer Medien, die für die Öffentlichkeit unsichtbar stattfinden, ist auf Grund des Fernmeldegeheimnisses weder eine durchgängige effektive staatliche Aufsicht, noch eine reine Selbstkontrolle der Medien möglich (z.B. Content Moderation ohne vorherige Meldung durch andere Nutzer:innen). Hier könnten Nutzer:innen durch eine verpflichtende diskursfördernde Technikgestaltung befähigt werden, Infragestellungen und Gegendarstellungen auszuüben.

Als Ergebnis des rechtswissenschaftlichen Teilprojekts kommt schließlich ein Regulierungsvorschlag für ein Gesetz oder Community Standards in Betracht. Dabei sind die Grundrechte der Nutzenden und Dienste-Anbietenden in einen optimalen Ausgleich zu bringen.

5.3 Datenschutzrecht: Anonymisierung und Informationsverlust

Eine wichtige disziplinübergreifende Forschungsfrage stellt das Thema Anonymisierung personenbezogener Daten und der damit einhergehende Informationsverlust dar. Bei der Betrachtung des Spannungsfelds zwischen Privatheit und Datenqualität spielen technische Möglichkeiten ebenso eine Rolle wie rechtliche Rahmenbedingungen. Aber auch die Bedarfe der anderen Disziplinen, die die zu erhebenden Daten im Rahmen ihrer empirischen Forschung verarbeiten werden, sind zu berücksichtigen. Schließlich können nur diese beurteilen, inwiefern ein Verlust der Datenqualität und damit der Verwertbarkeit durch eine Anonymisierung oder Pseudonymisierung zu erwarten ist.

Gesetzliche Anonymisierungs- und Pseudonymisierungserfordernisse sind sowohl für die Entwicklung praktischer Lösungsansätze als Forschungsergebnis als auch für die Datenerhebung im Rahmen des Forschungsprojekts selbst zu beachten. Im Hinblick auf praktische Lösungen gilt nach Art. 5 Abs. 1 lit. c, Art. 25 Abs. 1 und Art. 32 Abs. 1 DS-GVO der Grundsatz, dass Daten zu pseudonymisieren sind, wenn dies nach dem Verwendungszweck möglich ist und in Beziehung zum angestrebten Schutzzweck keinen unverhältnismäßigen Aufwand erfordert. Aus dem in Art. 5 c) DS-GVO normierten Zweck der Datenminimierung ergibt sich, dass anonyme oder anonymisierte Daten demgegenüber grundsätzlich vorranging zu verwenden sind (Heberlein 2018 Art. 5 DS-GVO Rn. 22) Für Messengerdienste ist überdies § 19 Abs. 2 TTDSG relevant, der die Ermöglichung einer grundsätzlich anonymen bzw. pseudonymen Nutzung von Telemedien vorsieht. Demgegenüber wird die Datenverarbeitung zu wissenschaftlichen Zwecken in der DS-GVO grundsätzlich privilegiert behandelt. Art. 89 Abs. 1 DS-GVO begrenzt die Privilegierung wiederum, indem

technisch-organisatorische Maßnahmen, wie die Pseudonymisierung (S. 3) als Garantien zum Schutz der betroffenen Personen gefordert werden. Indes sind weitergehende Maßnahmen, also auch die Anonymisierung, durch Art. 89 Abs. 1 DS-GVO nicht ausgeschlossen, sondern nach Maßgabe der Datenminimierung stets zu prüfen und grundsätzlich vorranging anzuwenden (Caspar 2019 Art. 89 DS-GVO Rn 51 f.). Gem. § 27 Abs. 3 BDSG ist eine Anonymisierung für besondere Kategorien personenbezogener Daten nach Art. 9 Abs. 1 DS-GVO sogar ausdrücklich erforderlich, sobald dies nach dem Forschungszweck möglich ist, es sei denn, berechtigte Interessen der betroffenen Person stehen dem entgegen. § 27 Abs. 3 BDSG ist im Kontext der Desinformationsforschung besonders einschlägig, da die zu erhebenden Daten häufig Bezüge zu politischen Meinungen und Weltanschauungen aufweisen, die besondere Kategorien i.S.d. Art. 9 Abs. 1 DS-GVO darstellen. Die Pflicht zur Anonymisierung entsteht nicht erst bei Abschluss des Forschungsprojektes, sondern bereits dann, wenn die personenbezogenen Daten für den weiteren Verlauf des Forschungsprojektes nicht mehr in personenbezogener Form erforderlich sind (Pauly 2021 § 27 BDSG Rn. 18). Daneben sind die Landesdatenschutzgesetze zu beachten, die ebenfalls Regelungen zur Datenverarbeitung zu wissenschaftlichen Forschungszwecken enthalten, z.B. Art. 25 BayDSG, § 11 HmbDSG, § 24 HDSIG.

Erwägungsgrund Nr. 26 S. 5 zur DS-GVO bestimmt, dass die Grundsätze des Datenschutzes nicht für anonyme Informationen gelten. Unter Anonymisierung ist die Auflösung der Beziehung zwischen den Daten und der betroffenen Person zu verstehen (Winter u.a. 2020 S. 26), wobei eine faktische Anonymisierung ausreicht (Gierschmann 2021 S. 483). Diese meint den Fall, dass die Re-Identifikation nur mit unverhältnismäßig hohem Aufwand zu erreichen ist. Während die h.M. Anonymisierung mittlerweile als eine zu rechtfertigende Datenverarbeitung anerkennt (Thüsing/Rombey 2021 S. 548; Gierschmann 2021 S. 483), wird die Frage, auf welche Rechtsgrundlage diese gestützt werden kann, sehr unterschiedlich bewertet (Hornung/Wagner 2020 S. 223; Stürmer 2020 S. 630; Gierschmann 2021 S. 483).

Die in Art. 4 Nr. 5 DS-GVO legal definierte Pseudonymisierung meint die Funktionstrennung von Zuordnungsinformation und Daten (Schleipfer 2020 S. 285). Sie stellt den „Kernpfeiler" der technisch-organisatorischen Maßnahmen des Datenschutzes nach der DS-GVO dar (Roßnagel 2018 S. 243). Grundsätzlich ist zwischen zwei Arten von pseudonymen Daten zu unterscheiden. Während die erste – in Art. 4 Nr. 5 DS-GVO nicht geregelte – eine anonymisierende Wirkung hat, dient die zweite lediglich zur Minderung der Risiken der Datenverarbeitung für die Grundrechte be-

troffener Personen (ebd. S. 246). Jedoch können anonyme Phasen bei Erscheinen geeigneter Kontextinformation in einen Personenbezug umschlagen (Schleipfer 2020 S. 285). Diesbezüglich ist im Rahmen einer Risikoprognose die Wahrscheinlichkeit der faktischen Durchführbarkeit der Bestimmbarkeit einer Person zu ermitteln (Roßnagel 2018 S. 244), wobei auch auf den Verarbeitungskontext abzustellen ist (Gierschmann 2021 S. 484 m.w.N.). Auch wenn die Wahrscheinlichkeit der De-Identifizierung nur schwer quantifizierbar ist (ebd. S. 483), genügt eine rein hypothetische Einschätzung nicht (Roßnagel 2018 S. 244). Zudem kann aus Erwägungsgrund 26 S. 2 zur DS-GVO abgeleitet werden, dass pseudonymisierte Daten dann als personenbezogen gelten sollen, wenn sie durch Heranziehung von Zusatzinformationen einer natürlichen Person zugeordnet werden könnten. Das ist jedenfalls dann der Fall, wenn es wahrscheinlich ist, dass der Datenverarbeiter in den Besitz der Zuordnungsregel kommen könnte (Roßnagel 2018 S. 245). Für Datenverarbeiter und für Dritte, die nicht über die Zuordnungsregel und/oder über andere Möglichkeiten der Kenntniserlangung verfügen, sind die pseudonymisierten Daten anonym (Winter u.a. 2020 S. 26).

Eine Anonymisierung bzw. anonymisierende Pseudonymisierung kann allerdings dazu führen, dass Daten nicht mehr aussagekräftig genug und deshalb nicht verwertbar sind. Um festzustellen, wie Anonymität garantiert und gleichzeitig Informationsverlust minimiert werden kann (Winter u.a. 2019 S. 489), sollte vor der Anonymisierung konkret geprüft werden, welche Daten für die weitere Verwendung des anonymisierten Datensatzes besonders wichtig sind und möglichst aussagekräftig erhalten bleiben sollten (Gierschmann 2021 S. 484 m.w.N.). Zwar sollten nach dem oben genannten möglichst wenige personenbezogene Daten verarbeitet werden. Für die empirische Forschung der Partnerdisziplinen ist indes die Verarbeitung einiger Datentypen unerlässlich. Für die Medienpsychologie ist vor allem das Verhalten der Nutzer:innen von Bedeutung. Von besonderem Interesse wären neben den reinen Inhaltsdaten auch Daten über Reaktionen der Nutzer:innen auf desinformierende Inhalte wie z.B. Weiterleitungen oder das Verlassen eines Kanals. Aus kommunikationswissenschaftlicher Sicht stellen zudem z.B. die Anzahl der Abonnent:innen, die Anzahl veröffentlichter Postings im Untersuchungszeitraum und die Anzahl der Aufrufe / Views eines Postings eines Kanals wichtige publizistische Relevanzfaktoren dar. Diese Daten sind für die weitere Forschung von Bedeutung, da sie systematische Aussagen über die Breitenwirksamkeit der Urheber:innen desinformierender Inhalte zulassen. Sollte es in diesen Bereichen zu Personenbezügen kommen, könnte eine risikomindernde Pseudonymisierung (Roßnagel 2018 S. 246) eine Erleichterung der Datenverarbeitung

darstellen, insbesondere hinsichtlich der Erfordernisse nach Art. 89 Abs. 1 Satz 1 und 3 DS-GVO (s.o.).

Problematischer ist hingegen die Anonymisierung bzw. Pseudonymisierung von Inhaltsdaten. Zum einen ist zu prüfen, ob anonymisierte bzw. pseudonymisierte Kommunikationsinhalte über den Messengerdienst hinaus im Internet veröffentlicht wurden, da so der Personenbezug einfach hergestellt werden könnte. Zum anderen könnten – vor allem prominente – Desinformationsverbreiter anhand allgemein bekannter Kontextinformationen z.B. über ihren Sprachstil identifiziert werden.

Der Datenschutz kann bereits durch die Technikgestaltung erfolgen (Privacy by Design). Während des Sammelns von Daten aus dem Internet können Crawler so angepasst werden, dass Daten bei der Erfassung automatisch anonymisiert oder pseudonymisiert werden (z.B. E-Mail-Adressen, Telefonnummern etc.) (Kamocki/Witt 2020). Um den Datencontent zu schützen, können verschiedene computerlinguistische Verfahren angewandt werden. Das kann die Eigennamenerkennung (engl. Named-Entity-Recognition, kurz „NER“), Coreference Resolution (dt. Koreferenzauflösung) oder auch die Schreibstilverschleierung (engl. Authorship Obfuscation) sein. NER-Verfahren identifizieren Eigennamen im Text wie z.B. Personen, Orte oder nummerische Daten, die folglich gelöscht oder "geschwärzt" werden können. Das kann allerdings zum Verlust der Lesbarkeit führen. Die Aufgabe der Coreference Resolution besteht darin alle Ausdrücke zu finden, die sich auf dieselbe Entität (z.B. „George Bush“, „Microsoft“) in einem Text beziehen. Das können Pronomen (z.B. „er“, „seine“) und andere referierende Ausdrücke (z.B. „der Politiker“, „der ehemalige US-Präsident“) sein (Kenton u.a. 2018). Die erkannten Referenzen könnten beispielsweise genutzt werden, um Personennamen durch ihre generischen Substitute zu ersetzen. Dadurch könnte sowohl der Datenschutz als auch die Erhaltung der Lesbarkeit gewährleistet werden. Um auch die Anonymität des Verfassers zu wahren, können Authorship Obfuscation-Verfahren eingesetzt werden, die den Schreibstil im Text so verändern, dass dieser weder von Menschen noch von modernen Verfahren zur Überprüfung der Autorschaft dem ursprünglichen Autor zugeordnet werden kann. Zu beachten ist, dass nicht jede Methode, die technisch unter den Begriff der „Anonymisierung“ fällt, auch eine Anonymisierung im datenschutzrechtlichen Sinne darstellt (Gierschmann 2021 S. 485 m.w.N.). Dies ist im Einzelfall zu prüfen.

6. Ausblick

In DYNAMO werden verschiedene Facetten der Desinformation in Messengerdiensten interdisziplinär erforscht. Die Untersuchung der Unterschiede zwischen sozialen Plattformen mit Netzwerkcharakter und Messengerdiensten stellen einen wichtigen gemeinsamen Forschungsstrang dar. Zu analysieren ist etwa, ob Unterschiede hinsichtlich der Bedeutung von visuellen Inhalten, wie Bilder oder Videos, und der Interaktion mit Freund:innen und Familienmitgliedern bestehen. Zudem ist zu klären, ob und inwieweit bestehende rechtliche Vorgaben auf Messengerdienste anwendbar sind. Neben der Betrachtung der Unterschiede ist auch zu berücksichtigen, dass die Grenzen zwischen den Geschäftsmodellen der sozialen Plattformen und Messengerdienste zunehmend verschwimmen. Daher gilt es zu untersuchen, wie der Verbreitung von Desinformation angesichts dieser Hybridisierung der Medientypen effektiv und grundrechtsschonend entgegengewirkt werden kann.

Weiterhin ermöglicht der multiperspektivische Ansatz, Erkenntnisse über die Verbreitung von Desinformation innerhalb von Messengerdiensten und im Zusammenspiel zu anderen Medien zu gewinnen. Für alle Disziplinen ist es von Bedeutung zu untersuchen, welche Rolle öffentlich gepostete Einladungslinks, Weiterleitungen von Nachrichten und Verlinkungen von Gruppen und Kanälen spielen. Die Psychologie erforscht dabei die Motive der Weiterleitung sowie die dahinterstehenden Kognitionen und Emotionen. Konsekutiv muss erörtert werden, ob und wie diesen Faktoren durch Technikgestaltung und rechtliche Vorgaben z.B. durch Begrenzung der Weiterleitungsmöglichkeiten, entgegengewirkt werden kann, wobei der Grundrechtsschutz unbedingt zu wahren ist. Auch Kennzeichnungspflichten und die Zusammenarbeit mit unabhängigen Fact-Checking-Organisationen sind für den Bereich der Messengerdienste zu untersuchen.

Aktuell werden viele Desinformationen über den Ukrainekrieg verbreitet. Umso relevanter wird die übergeordnete Frage nach der Beeinträchtigung der freien demokratischen Willensbildung durch Desinformation. Erkenntnisse über emotionale Reaktionen der Mediennutzer:innen und die wahrgenommene Glaubwürdigkeit von Quellen können zur Feststellung von Risiken für die freie demokratische Willensbildung ausgewertet werden. Auch die Auswirkungen solcher Desinformationen auf die Medienberichterstattung sind von Bedeutung. Mit der EU-Verordnung 2022/350 vom 1.3.2022 erfolgte als Ad-Hoc-Reaktion der EU ein Verbot zweier russischer Staatsmedien (RT und Sputnik). Durch automatisierte Verfahren und journalistische Analysen könnte beobachtet werden, ob die

durch die Verordnung verbotenen Medien, trotz des in der Verordnung geregelten Umgehungsverbots, Messengerdienste zur Verbreitung ihrer Inhalte nutzen.

Es muss untersucht werden, wie der Staat bzw. die EU dauerhaft auf die Beeinträchtigung der freien demokratischen Willensbildung durch die absichtliche Verzerrung von Fakten reagieren kann, ohne elementare Freiheitsrechte über Gebühr zu beeinträchtigen. Relevant ist auch die Frage, inwiefern eine Gewährleistungspflicht des Gesetzgebers für den Meinungsbildungsprozess besteht. Schließlich ist in Bezug auf den Schutz der Demokratie zu erforschen, ob Messengerdienste nicht auch eine konstruktive Rolle einnehmen können, indem sie z.B. Korrekturen durch Nutzer:innen fördern.

Literatur

Bader, Katarina; Jansen, Carolin und Rinsdorf, Lars (2020): Jenseits der Fakten: Deutschsprachige Fake News aus Sicht der Journalistik. In: Steinebach, Martin; Bader, Katarina; Rinsdorf, Lars; Krämer, Nicole und Roßnagel, Alexander (Hrsg.): *Desinformation aufdecken und bekämpfen. Interdisziplinäre Ansätze gegen Desinformationskampagnen und für Meinungspluralität.* Baden-Baden: Nomos, S. 33-76. URL: https://doi.org/10.5771/9783748904816-33 (besucht am 27.01.2022).

Baumgartner, Jason; Zannettou, Savvas; Squire, Megan und Blackburn, Jeremy (2020): The Pushshift Telegram Dataset. In: PKP Publishing Services Network (Hrsg.): *Proceedings of the International AAAI Conference on Web and Social Media*. URL: https://ojs.aaai.org/index.php/ICWSM/article/view/7348 (besucht am 27.01.2022).

Bovet, Alexandre und Grindod, Peter (2020): *The Activity of the Far Right on Telegram*. Hg. v. University of Oxford. Mathematical Institute. Oxford, UK. URL: https://www.researchgate.net/profile/Peter-Grindrod/publication/346968575_The_Activity_of_the_Far_Right_on_Telegram_v211/links/5fd5be47a6fdccdcb8c07326/The-Activity-of-the-Far-Right-on-Telegram-v211.pdf (besucht am 27.01.2022).

Brauneck, Jens (2020): EU-Desinformationsbekämpfung durch Google, Facebook u.a. unter Androhung von Gesetzen, EU-Außenpolitik durch Gegenpropaganda in Drittstaaten?, *Europarecht* 1(2020), S. 89-111.

Buchanan, Tom (2020): Why do people spread false information online? The effects of message and viewer characteristics on self-reported likelihood of sharing social media disinformation. In: *PloS one*, 15(10), S. 1-33. DOI: https://doi.org/10.1371/journal.pone.0239666 (besucht am 27.01.2022).

Buchanan, Tom und Benson, Vladlena (2019): Spreading Disinformation on Facebook: Do Trust in Message Source, Risk Propensity, or Personality Affect the Organic Reach of "Fake News"? *Social Media and Society*, 5(4). doi:10.1177/2056305119888654.

Bundesamt für Verfassungsschutz (2020): *Neuer Phänomenbereich „Verfassungsschutzrelevante Deligitimierung des Staates"*. URL: https://www.verfassungsschutz.de/SharedDocs/kurzmeldungen/DE/2021/2021-04-29-querdenker.html (besucht am 27.01.2022).

Bundesverband Mobile Beratung (2021): *Policy Paper: Auswirkungen von Verschwörungsmythen und rechtsoffenen Corona-Protesten auf die demokratische Zivilgesellschaft*. URL: https://www.bundesverband-mobile-beratung.de/wp-content/uploads/2021/12/2021-12-14_BMB_Policy-Paper_Corona-Proteste.pdf (besucht am 27.01.2022).

Chen, Xinran; Sin, Sei-ching Joanna; Theng, Yin-leng und Lee, Chei Sian (2015): Why Students Share Misinformation on Social Media: Motivation, Gender, and Study-level Differences. *The Journal of Academic Librarianship*, 41(5), S. 583–592. doi:10.1016/j.acalib.2015.07.003.

Clifford, Bennett (2018): Trucks, Knives, Bombs, Whatever: Exploring Pro-Islamic State Instructional Material on Telegram. *CTCSentinel*, 11(5), URL: https://ctc.usma.edu/trucks-knives-bombs-whatever-exploring-pro-islamic-state-instructional-material-telegram/ (besucht am 27.01.2022).

Dittrich, Miro und Holnburger, Josef (2021): *Nur einen Klick vom Rechtsterror entfernt*. URL: https://cemas.io/blog/naidoo-telegram/ (besucht am 27.01.2022).

Dreyer, Stephan; Stanciu, Elena; Potthast, Keno Christian und Schulz, Wolfgang (2021): Desinformation: Risiken, Regulierungslücken und adäquate Gegenmassnahmen: Wissenschaftliches Gutachten im Auftrag der Landesanstalt für Medien NRW. Düsseldorf: Landesanstalt für Medien NRW.

Du, Ling; Ho, Anthony und Cong, Rumin (2020): Perceptual hashing for image authentication: A survey. *Signal Processing: Image Communication, 81*.

Durner, Wolfgang (2021): Art. 10 Brief-, Post- und Fernmeldegeheimnis, in: Dürig, Günter /Herzog, Roman /Scholz, Rupert (Hrsg.): *Grundgesetz-Kommentar* München: C.H.Beck

Ehmann, Eugen/Sellmayr, Martin (Hrsg.) (2018): *DS-GVO Kommentar*. 2.Aufl. München: C.H.Beck.

Enghofer, Sebastian (2021): Nudging als Strategie gegen Fake News. *Die POLIZEI*, 2(112), S. 64-72.

Eydlin, Alexander (16.09.2021): Facebook löscht Konten und Gruppen der Querdenken-Bewegung. URL: https://www.zeit.de/digital/internet/2021-09/facebook-loescht-konten-und-gruppen-der-querdenken-bewegung (besucht am 21.01.2021).

Feldmann, Thorsten (2021): Juristische Instrumente gegen Internet-Hass. *Kommunikation & Recht (K&R)*, Beilage 1 zu 24 (6), S. 34-37.

Ferreau, Frederik (2021): Desinformation als Herausforderung für die Medienregulierung: *Zeitschrift für das gesamte Medienrecht (AfP)*, 52(3), S. 204-210.

Festinger, Leon (1957): An introduction to the theory of dissonance. In: *A theory of cognitive dissonance.* doi:10.1037/10318-001.

Flade, Florian und Maccolo, Georg (09. Dez. 2021): *Kaum zu fassen*. URL: https://www.tagesschau.de/investigativ/ndr-wdr/telegram-105.html (besucht am 27.01.2022).

Flint, Jessica (2021): *Fake News im Wahlkampf: Eine Untersuchung der rechtlichen Problemstellung der Desinformation in sozialen Netzwerken am Beispiel von Facebook*. Baden-Baden: Nomos.

Gahntz, Maximilian; Neumann, Katja T.J.; Otte, Philipp C.; Sältz, Bendix J; Steinbach, Katrin (2021): Breaking the News? Politische Öffentlichkeit und die Regulierung von Medienintermediären. Bonn: Friedrich-Ebert-Stiftung.

Geminn, Christian Ludwig (2014): *Rechtsverträglicher Einsatz von Sicherheitsmaßnahmen im öffentlichen Verkehr*. Wiesbaden: Springer Vieweg.

Gierschmann, Sibylle (2021): Gestaltungsmöglichkeiten durch systematisches und risikobasiertes Vorgehen – Was ist schon anonym? Planung und Bewertung der Risiken der Anonymisierung. *Zeitschrift für Datenschutz (ZD)*, S. 482-486.

Gräfe, Hans-Christian (2020), Desinformation im Spiegel des Rechts: Verfassungsrecht und juristische Handhabbarkeit von Falschinformation im gesellschaftlichen und im Unternehmens-Kontext. *Comply. Fachmagazin für Compliance-Verantwortliche*, *5* (4), S. 38-41.

Guhl, Jakob und Davey, Jacob (2020): *A Safe Space to Hate: White Supremacist Mobilisation on Telegram*. Hg. v. Institute for Strategic Dialogue (ISD). London, UK. URL: https://www.isdglobal.org/isd-publications/a-safe-space-to-hate-white-supremacist-mobilisation-on-telegram/ (besucht am 27.01.2022).

Hansson, Sten; Orru, Kati; Torpan, Sten; Bäck, Asta; Kazemekaityte, Austeja; Meyer, Sunniva Frislid; Ludvigsen, Johanna; Savadori, Lucia; Galvagni, Alessandro und Pigrée, Ala (2021): COVID-19 information disorder: six types of harmful information during the pandemic in Europe. *Journal of Risk Research*, 24(3–4), S. 380–393. doi:10.1080/13669877.2020.1871058.

Hassaballah, Mahmoud; Abdelmgeid, Aly Amin, und Alshazly, Hmmam A. (2016): Image features detection, description and matching. In *Image Feature Detectors and Descriptors*, S. 11-45. Springer, Cham.

Heereman, Wendy und Selzer, Annika (2019): Löschung rechtskonformer Nutzerinhalte durch Soziale-Netzwerkplattformen. Ein Überblick am Beispiel von Facebook. *Computer und Recht,* 4(2019), S. 271-276. DOI: https://doi.org/10.9785/cr-2019-350421 (besucht am 27.01.2022).

Hohlfeld, Ralf; Bauerfeind, Franziska; Braglia, Ilenia et al. (2021): *Communicating COVID-19 against the backdrop of conspiracy ideologies: How Public Figures discuss the matter of Facebook and Telegram*. Hg. v. Disinformation Research Lab. Universität Passau. Passau (Working Paper, 01/2021). URL: https://www.researchgate.net/publication/351698784_Communicating_COVID-19_against_the_backdrop_of_conspiracy_ideologies_HOW_PUBLIC_FIGURES_DISCUSS_THE_MATTER_ON_FACEBOOK_AND_TELEGRAM (besucht am 27.01.2022).

Holzer, Boris (2021): Zwischen Protest und Parodie: Strukturen der "Querdenken"-Kommunikation auf Telegram (und anderswo). In: Reichardt, Sven (Hg.), *Die Misstrauensgemeinschaft der "Querdenker". Die Corona-Proteste aus kultur- und sozialwissenschaftlicher Perspektive*. Frankfurt/New York: Campus Verlag, S. 125-157.

Holznagel, Bernd (2018): Phänomen „Fake News"– Was ist zu tun? *MultiMedia und Recht (MMR)*, 21(1), S. 18-22.

Hornung, Gerrit und Wagner, Bernd (2020): Anonymisierung als datenschutzrelevante Verarbeitung? Rechtliche Anforderungen und Grenzen für die Anonymisierung personenbezogener Daten. *Zeitschrift für Datenschutz (ZD),* S. 223-228.

Jalilvand, Asal und Neshati, Mahmood (2020): Channel retrieval: finding relevant broadcasters on Telegram. In: *Social Network Analysis and Mining* 10 (1), S. 1–16. DOI: 10.1007/s13278-020-0629-z.

Jarass, Hans D./Pieroth, Bodo (Hrsg.) (2020): *Grundgesetz für die Bundesrepublik Deutschland - Kommentar*. München: C.H.Beck.

Jünger, Jakob; Gärtner, Chantal (2020): Datenanalyse von rechtsverstoßenden Inhalten in Gruppen und Kanälen von Messengerdiensten am Beispiel Telegram. Düsseldorf: Landesanstalt für Medien NRW.

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

Kamocki, Pawel und Witt, Andreas (2020): Privacy by Design and Language Resources. *In Proceedings of the 12th Language Resources and Evaluation Conference*, S. 3423–3427, Marseille, France. European Language Resources Association.

Kim, Antino; Moravec, Patricia L. und Dennis, Alan R. (2019): Combating Fake News on Social Media with Source Ratings: The Effects of User and Expert Reputation Ratings. *Journal of Management Information Systems*, 36(3), S. 931–968. doi:10.1080/07421222.2019.1628921.

Kleeberg, Bernhard; Suter, Robert (2014): „Doing truth" Bausteine einer Praxeologie der Wahrheit. *Zeitschrift für Kulturphilosophie*, 2(8), S. 211-226.

Kluck, Jan P.; Schaewitz, Leonie und Krämer, Nicole C. (2019): Doubters are more convincing than advocates. The impact of user comments and ratings on credibility perceptions of false news stories on social media. *Studies in Communication and Media*, 8(4), S. 446–470. doi:10.5771/2192-4007-2019-4-446.

Knuutila, Aleksi; Herasimenka, Aliaksandr; Bright, Jonathan; Nielsen, Rasmus und Howard, Philip N. (2020): *Junk News Distribution on Telegram: The Visibility of English-language News Sources on Public Telegram Channels*. Project on Computational Propaganda. Oxford, UK (COMPROP Data Memo, 2020.5). URL: https://demtech.oii.ox.ac.uk/research/posts/junk-news-distribution-on-telegram-the-visibility-of-english-language-news-sources-on-public-telegram-channels/ (besucht am 27.01.2022).

Kühling, Jürgen (2021): »Fake News« und »Hate Speech« – Die Verantwortung der Medienintermediäre zwischen neuen NetzDG, MStV und Digital Services Act. *Zeitschrift für Urheber- und Medienrecht (ZUM),* S. 461-472.

Kümpel, Anna Sophie; Karnowski, Veronika und Keyling, Till (2015): News Sharing in Social Media: A Review of Current Research on News Sharing Users, Content, and Networks. *Social Media + Society*, 1(2). doi:10.1177/2056305115610141.

Kunda, Ziva (1990): The case for motivated reasoning. *Psychological Bulletin*, 108(3), S. 480–498. doi:10.1037/ 0033-2909.108.3.480.

Lamberty, Pia und Holnburger, Josef (2021): *Die Bundestagswahl 2021. Welche Rolle Verschwörungsideologien in der Demokratie spielen*. Hg. v. CeMAS - Center für Monitoring, Analyse und Strategie gGmbH. Berlin. URL: https://cemas.io/publikationen/die-bundestagswahl-2021-welche-rolle-verschwoerungsideologien-in-der-demokratie-spielen/ (besucht am 27.01.2022).

Lazer, David M.J.; Baum, Matthew A.; Benkler, Yochai; Berinsky, Adam J.; Greenhill, Kelly M.; Menczer, Filippo; Metzger, Miriam J.; Nyhan, Brendan; Pennycook, Gordon; Rothschild, David; Schudson, Michael; Sloman, Steven A.; Sunstein, Cass R.; Thorson, Emily A.; Watts, Duncan J. und Zittrain, Jonathan L. (2018): The science of fake news. *Science*, *359*(6380), S. 1094–1096. doi:10.1126/science.aao2998.

Lee, Chei Sian und Ma, Long (2012): News sharing in social media: The effect of gratifications and prior experience. *IComputers in Human Behavior*, 28(2), S. 331–339. doi:10.1016/j.chb.2011.10.002.

Lee, Kenton; He, Luheng und Zettlemoyer, Luke (2018): "Higher-Order Coreference Resolution with Coarse-to-Fine Inference." *In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)* (pp. 687–692). Association for Computational Linguistics.

Lipowicz, Irena und Szpor, Grażyna (2021): Neue Aspekte der Desinformation. *Datenschutz und Datensicherheit (DuD)*, 45(6), S. 381-384.

Lou, Chen; Tandoc Jr., Edson C.; Hong, Li Xuan; Pong, Xiang Yuan; Lye, Wan Xin und Sng, Ngiag Gya (2021): When Motivations Meet Affordances: News Consumption on Telegram. In: *Journalism Studies*, 22(7), S. 934-952. DOI: https://doi.org/10.1080/1461670X.2021.1906299 (besucht am 27.01.2022).

Mafi-Gudarzi, Nima (2019): Desinformation: Herausforderung für die wehrhafte Demokratie. *Zeitschrift für Rechtspolitik (ZRP)*, 52(3), S. 65-68.

Mena, Paul (2020): Cleaning Up Social Media: The Effect of Warning Labels on Likelihood of Sharing False News on Facebook. *Policy and Internet*, 12(2), S. 165–183. doi:10.1002/poi3.214.

Metzger, Miriam J. und Flanagin, Andrew J. (2013): Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics*, 59, S. 210–220. doi:10.1016/j.pragma.2013.07.012.

Michailidou, Asimina und Trenz, Hans-Jörg (2021): Rethinking journalism standards in the era of post-truth politics: from truth keepers to truth mediators. In: *Media, Culture & Society*, 43(7), S. 1340-1349. DOI: 10.1177/01634437211040669.

Nachtwey, Oliver; Schäfer, Robert und Frei, Nadine (2020): *Politische Soziologie der Corona-Proteste. Grundauswertung*. Hg. von der Universität Basel. Basel. URL: https://osf.io/preprints/socarxiv/zyp3f/ (besucht am 27.01.2022).

Nickerson, Raymond S. (1998): Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), S. 175–220. doi:10.1037/1089-2680.2.2.175.

Oeldorf-Hirsch, Anne; Schmierbach, Mike; Appelman, Alyssa und Boyle, Michael P. (2020): The Ineffectiveness of Fact-Checking Labels on News Memes and Articles. In: *Mass Communication and Society*, 23(5), S. 682–704. doi:10.1080/15205436.2020.1733613.

Paal, Boris P. und Pauly, Daniel A. (2021): *Kommentar Datenschutz-Grundverordnung/Bundesdatenschutzgesetz.* 3. Aufl. München: C.H.Beck.

Pantenburg, Johannes; Reichardt, Sven und Sepp, Benedikt (2021): Wissensparallelwelten der "Querdenker". In: Reichardt, Sven (Hg.), *Die Misstrauensgemeinschaft der "Querdenker". Die Corona-Proteste aus kultur- und sozialwissenschaftlicher Perspektive*. Frankfurt/New York: Campus Verlag, S. 29-65.

Pennycook, Gordon und Rand, David G. (2021): The Psychology of Fake News. *Trends in Cognitive Sciences*, 25(5), S. 388–402. doi:10.1016/j.tics.2021.02.007.

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.

Pennycook, Gordon; McPhetres, Jonathon; Zhang, Yunhao; Lu, Jackson G. und Rand, David G. (2020): Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention. *Psychological Science*, 31(7), S. 770-780.

Pörksen, Bernhard (Hrsg.): (2015) *Schlüsselwerke des Konstruktivismus*. Springer: Wiesbaden.

Potthast, Martin; Kiesel, Johannes; Reinartz, Kevin; Bevendorff, Janek und Stein, Benno. (2018): A Stylometric Inquiry into Hyperpartisan and Fake News. *In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, S. 231–240, Melbourne, Australia. Association for Computational Linguistics.

Rathje, Jan (2021): Das souveränistische Milieu als Auffangbecken für Enttäuschte. In: *Die Bundestagswahl 2021. Welche Rolle Verschwörungsideologien in der Demokratie spielen* (S. 54-59). Hg. v. CeMAS - Center für Monitoring, Analyse und Strategie gGmbH. Berlin. URL: https://cemas.io/publikationen/die-bundestagswahl-2021-welche-rolle-verschwoerungsideologien-in-der-demokratie-spielen/ (besucht am 27.01.2022).

Reinhard, Marc-André und Sporer, Siegfried L. (2010): Content Versus Source Cue Information as a Basis for Credibility Judgments *Social Psychology* 41(2), S. 93–104. doi:10.1027/1864-9335/a000014.

Resende, Gustavo; Melo, Philipe; Sousa, Hugo; Messias, Johnnatan; Vasconcelos, Marisa; Almeida, Jussara und Benevenuto, Fabrício (2019): (Mis)Information Dissemination in WhatsApp: Gathering, Analyzing and Countermeasures. *The World Wide Web Conference*. New York, NY, USA: ACM, 818–828. doi:10.1145/3308558.3313688.

RND (05.01.2022). *Medienrecherchen dokumentieren Hunderte Tötungsaufrufe in Telegram-Chats*. Hg. vom Redaktionsnetzwerk Deutschland. URL: https://www.rnd.de/politik/telegram-hunderte-toetungsaufrufe-in-chats-dokumentiert-GJWHRWIB6UTO656LRZJTCR3VGA.html (besucht am 27.01.2022).

Rogers, Richard (2020): Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media. In: *European Journal of Communication*, 35(3), S. 213-229. DOI: https://doi.org/10.1177/0267323120922066 (besucht am 27.01.2022).

Roßnagel, Alexander (2018): Pseudonymisierung personenbezogener Daten. Ein zentrales Instrument im Datenschutz nach der DS-GVO. *Zeitschrift für Datenschutz (ZD)*, S. 243-247.

Sängerlaub, Alexander, Meier, Miriam und Rühl, Wolf-Dieter (2018): *Fakten statt Fakes: Das Phänomen "Fake News". Verursacher, Verbreitungswege und Wirkungen von Fake News im Bundestagswahlkampf 2017* (Abschlussbericht Projekt "Measuring Fake News"). Berlin. URL: https://www.stiftung-nv.de/sites/default/files/snv_fakten_statt_fakes.pdf (besucht am 27.01.2022).

Schaewitz, Leonie und Krämer, Nicole C. (2020): Combating Disinformation: Effects of Timing and Correction Format on Factual Knowledge and Personal Beliefs. In: van Duijn, M., Preuss, M., Spaiser, V., Takes, F., Verberne, S. (eds) *Disinformation in Open Online Media. MISDOOM 2020*. Springer, Cham. DOI https://doi.org/10.1007/978-3-030-61841-4_16

Schaewitz, Leonie, Kluck, Jan Philipp, Klösters, Lukas und Krämer, Nicole (2020): When is Disinformation (In)Credible? Experimental Findings on Message Characteristics and Individual Differences. *Mass Communication and Society*, 23(4), 484-509.

Scheffler, Tatjana; Solopova, Veronika und Popa-Wyatt, Mihaela (2021): The Telegram Chronicles of Online Harm. In: *Journal of Open Humanities Data*, 7(8), S. 1-15. DOI: https://doi.org/10.5334/johd.31 (besucht am 27.01.2022).

Schleipfer, Stefan (2020): Pseudonymität in verschiedenen Ausprägungen. Wie gut ist die Unterstützung der DS-GVO?. *Zeitschrift für Datenschutz (ZD)*, S. 284-291.

Schmalenbach, Kirsten (2005): Wahrheit und Lüge unter der Herrschaft der Grundrechte. *Juristische Arbeitsblätter (JA)* S. 749-752.

Schwarzenegger, Christian (2022): Understanding the Users of Alternative News Media – Media Epistemologies, News Consumption, and Media Practices. In: *Digital Journalism*, S. 1-19. DOI: https://doi.org/10.1080/21670811.2021.2000454, zuletzt geprüft am 27.01.2022.

Shu, Kai; Sliva, Amy; Wang, Suhang; Tang, Jiliang und Liu, Huan (2017): Fake News Detection on Social Media: A Data Mining Perspective. *SIGKDD Explor. Newsl. 19, 1 (June 2017)*, S. 22–36. DOI:https://doi.org/10.1145/3137597.3137600 (besucht am 17.03.2022).

Simitis, Spiros; Hornung, Gerrit und Spieker genannt Döhmann, Indra (Hrsg.) (2019): *Kommentar Datenschutzrecht (DSGVO mit BDSG)*. Baden-Baden: Nomos.

Steinebach, M., Lutz, S., & Liu, H. (2020). Privacy and Robust Hashes. Privacy-Preserving Forensics for Image Re-Identification. *Journal of Cyber Security and Mobility*, 111–140. https://doi.org/10.13052/jcsm2245-1439.914 (besucht am 17.03.2022).

Steinebach, Martin; Bader, Katarina, Rinsdorf, Lars; Krämer, Nicole und Roßnagel, Alexander (2020): *Desinformation aufdecken und bekämpfen. Interdisziplinäre Ansätze gegen Desinformationskampagnen und für Meinungspluralität*. Baden-Baden: Nomos.

Steinebach, Martin; Klöckner, Peter; Reimers, Nils; Wienand, Dominik und Wolf, Patrick. (2013): Robust Hash Algorithms for Text. In *14th International Conference on Communications and Multimedia Security (CMS)*, S. 135-144. Springer.

Steinebach, Martin; Lutz, Sebastian Liu, Huajin (2019): Privacy and robust hashes. In *Proceedings of the 14th International Conference on Availability, Reliability and Security*. https://doi.org/10.1145/3339252.3340105 (besucht am 17.03.2022).

Sundar, S. Shyam (2008): The MAIN Model: A Heuristic Approach to Understanding Technology Effects on Credibility. In: Metzger, Miriam J./Flanagin, Andrew J. (Hrsg.): *Digital media, youth, and credibility*. Cambridge, MA, USA: The MIT Press, S. 73–100. doi:10.1162/dmal.9780262562324.073.

Sundar, S. Shyam; Molina, Maria D. und Cho, Eugene (2021): Seeing Is Believing: Is Video Modality More Powerful in Spreading Fake News via Online Messaging Apps? *Journal of Computer-Mediated Communication*, 26(6), S. 1–19. doi:10.1093/jcmc/zmab010.

Sunyaev, Ali; Schmidt-Kraepelin, Manuel; Thiebes, Scott (2021), in: Hornung, Gerrit/ Müller-Terpitz, Ralf (Hrsg.): *Rechtshandbuch Social Media*. Berlin: Springer.

Theile, Hans (2012): Wahrheit , Konsens und § 257c StPO. *Neue Zeitschrift für Strafrecht (NStZ)*, S. 666-671.

Thüsing, Gregor und Rombey, Sebastian (2021): Anonymisierung an sich ist keine rechtfertigungsbedürftige Datenverarbeitung. Eine Auslegung von Art. 4 Nr. 2 DS-GVO nach den Methoden des EuGH. *Zeitschrift für Datenschutz (ZD)*, S. 548-553.

Tuchtfeld, Erik (21.12.2021): Don't shoot the Messenger: Von Telegrammen und öffentlicher Kommunikation. URL: https://verfassungsblog.de/dont-shoot-the-messenger/ (besucht am 21.01.2021).

Vogel, Inna und Jiang, Peter (2019): "Fake News Detection with the New German Dataset 'GermanFakeNC'". In: *International Conference on Theory and Practice of Digital Libraries,* S. 288-295. Springer, Cham.

Vogel, Inna und Meghana, Meghana (2018): "Analyzing Linguistic Features of German Fake News: Characterization, Detection, and Discussion." *Sicherheitslagen und Sicher-heitstechnologien: Beiträge der ersten Sommerakademie der zivilen Sicherheits-forschung, 2018*, S. 273-296.

Vogel, Inna und Meghana, Meghana (2020): "Detecting Fake News Spreaders on Twitter from a Multilingual Perspective". *The 7th IEEE International Conference on Data Science and Advanced Analytics. Special Session-Fake News, Bots and Trolls (DSAA 2020)*, 6-9 October 2020, Virtual Event, Sydney, Australia, S. 599-606.

Vosoughi, Soroush, Roy, Deb und Aral, Sinan (2018): The Spread of True and False News Online. *Science*, 359(6380): 1146-1151. https://doi.org/10.1126/science.aap9559.

Weeks, B. E. (2015): Emotions, Partisanship, and Misperceptions: How Anger and Anxiety Moderate the Effect of Partisan Bias on Susceptibility to Political Misinformation. *Journal of Communication*, 65(4), 699-719.

Weismueller, Jason; Harrigan, Paul; Coussement, Kristof und Tessitore, Tina (2022): What makes people share political content on social media? The role of emotion, authority and ideology. *Computers in Human Behavior*, 129. doi:10.1016/j.chb.2021.107150.

Winter, Christian; Battis, Verena und Halvani, Oren (2019): Herausforderungen für die Anonymisierung von Daten. Technische Defizite, konzeptuelle Lücken und rechtliche Fragen bei der Anonymisierung von Daten. *Zeitschrift für Datenschutz (ZD),* S. 489-493.

Winter, Christian; Steinebach, Martin; Heereman, Wendy; Steiner, Simone; Battis, Verena; Halvani, Oren; Yannikos, York und Schüßler, Christoph: (2020): Privacy und Big Data. Darmstadt: Fraunhofer-Institut für Sichere Informationstechnologie SIT.

Winter, Stephan und Krämer, Nicole C. (2014): A question of credibility - Effects of source cues and recommendations on information selection on news sites and blogs. *Communications*, 39(4), S. 435–456. doi:10.1515/commun-2014-0020.