



LANDESANSTALT FÜR MEDIEN NRW  
Der Meinungsfreiheit verpflichtet.

# UNSICHTBARE STIMMEN – AUSSCHLUSS UND RÜCKZUG AUS ONLINE-DISKURSEN

Forschungsmonitor Informationsintermediäre

Ausgabe 7: Dezember 2020

Realisiert von:

**iRights.Lab**  
Think Tank für die  
digitale Welt

# RÜCKBLICK UND ÜBERBLICK

## UNSIHTBARE STIMMEN – AUSSCHLUSS UND RÜCKZUG AUS ONLINE-DISKURSEN

Falsche und irreführende Informationen, Hass und Polarisierung kennzeichnen einen Teil der Kommunikation und Information online. Das verdeutlichte die Corona-Krise eindrücklich. Gerade zu Beginn, im Frühjahr 2020, wurden insbesondere Desinformation, Verschwörungstheorien und Polarisierung zu weltweit wiederkehrenden Kennzeichen der Online-Diskurse. Die anfangs wenigen gesicherten Erkenntnisse sowie die außergewöhnliche gesellschaftliche und politische Situation verstärkten die Verbreitung von Desinformation und Hass. Denn die Menschen suchten in dieser Situation nach verlässlichen Informationen und Erklärungen, was einige Akteure ausnutzten, um ihre Ansichten zu verbreiten, indem sie auf die Pandemie Bezug nahmen. Diesen Zusammenhängen widmete sich die letzte Ausgabe des Forschungsmonitors ausführlich und behandelte Untersuchungen hierzu aus der ganzen Welt. Ebenso wenig wie die Pandemie verschwunden ist, sind es diese Dynamiken, auch wenn die Informationslage gegenwärtig weniger diffus ist. Weiterhin kursieren zahlreiche Verschwörungserzählungen und irreführende Darstellungen, die etwa in sozialen Netzwerken gezielt gestreut werden. Die Informationslage zur Corona-Krise scheint in der Normalität gegenwärtiger Diskurse angekommen – mit ihren typischen Herausforderungen (Insight 1).

Wissenschaftliche Untersuchungen stellen fest: Desinformation, Verschwörungstheorien und Hass verbreiten sich zunehmend auch über Messengerdienste wie etwa WhatsApp oder Telegram und verlagern sich damit heraus aus öffentlich zugänglichen Informationsintermediären und hinein in private Gruppenchats (Insight 1, 2 und 3). Dieser Vorgang nennt sich Deplatforming: Personen, die in klassischen sozialen Medien gegen geltendes Recht oder die darüberhinausgehenden Plattformregeln verstoßen, werden teilweise von diesen Anwendungen ausgeschlossen. Sie ziehen dann häufig um zu alternativen Netzwerken oder zu Messengern. Medienberichte über den Ausschluss verhelfen diesen Personen und Gruppen zu mehr Aufmerksamkeit. Sie selbst inszenieren sich derweil häufig als Opfer angeblich mangelnder Meinungsfreiheit auf etablierten sozialen Plattformen, die sie ausschließen. Ihre Inhalte verbreiten sie über private Gruppen in Messengerdiensten ähnlich wie zuvor auf gängigen sozialen Netzwerken (Insight 1). Häufig handelt es sich hierbei um Anhängerinnen und Anhänger rechtsextremer Weltanschauungen (Insight 1, 2 und 10).

Informationsintermediäre sind im digitalen Raum vielfach zwischen Rezipientin beziehungsweise Rezipienten und Information geschaltet. Sie erfüllen demnach eine vermittelnde Funktion. Dabei schaffen sie ein jeweils personalisiertes Angebot für die Nutzerin bzw. den Nutzer. Dieses Angebot ist für die allgemeine Öffentlichkeit nicht einsehbar. Algorithmische Systeme<sup>1</sup> kuratieren zunehmend die jeweilige mediale Wirklichkeit jedes Individuums. Die Grundlage dafür bilden massenhaft gesammelte Daten von Nutzerinnen und Nutzern. Mit ihrer Hilfe bestimmen algorithmische Systeme etwa, was auf die individualisierten Ergebnisseiten der Suchmaschinen gelangt und gestalten die jeweiligen Newsfeeds von Social-Media-Profilen. Sie wählen nicht nach Relevanz für die Öffentlichkeit aus, sondern selektieren durch automatisierte algorithmische Prozesse nach anderen, weitgehend unbekanntem Kriterien. Die Entwicklerteams hinter den algorithmischen Systemen wählen die Datengrundlage aus, die sogenannte „ground truth“, oder bestimmen, welche Datenpunkte in diesen Systemen welche Relevanz erhalten.

Mit solchen Umzügen entzieht sich die betreffende Kommunikation weitgehend wissenschaftlichen Untersuchungen sowie aufsichts- und strafrechtlichen Bewertungen, da die Diskussionen nunmehr komplett im Privaten stattfinden. Aus Gründen des Datenschutzes und der Verschlüsselung von Nachrichten, die viele Messenger anbieten, sind auch Korrekturen von falschen Informationen durch die Anbieter nicht mehr möglich, ebenso wie andere Maßnahmen (Insight 1 und 2). Indes finden in privaten Gruppen bei Messengerdiensten wie WhatsApp oder dem Facebook Messenger Korrekturen durch Nutzerinnen und Nutzer tendenziell häufiger statt als in (halb-)öffentlichen sozialen Netzwerken. Das zeigt, dass falsche Informationen oft nicht unbemerkt bleiben und dass als geschützt empfundene Kommunikationsräume eher zu Korrekturen beitragen (Insight 3). Zugleich verweist das auf eine weitere wichtige Beobachtung: Viele Nutzerinnen und Nutzer trauen sich nicht, sich öffentlich frei zu äußern. Einige ziehen sich ganz aus öffentlichen Online-Debatten zurück, da sie Anfeindun-

<sup>1</sup> Der Begriff algorithmisches System umfasst sowohl einfachere als auch komplexere Softwareanwendungen. Dazu zählen auch lernende Systeme, die häufig als „Künstliche Intelligenz“ bezeichnet werden.

gen befürchten (Insight 5 und 9). Insbesondere die gewaltsamen Äußerungen und Drohungen gegenüber bestimmten Bevölkerungsgruppen wie Mädchen und Frauen oder People of Colour führen dazu, dass deren Stimmen öffentlich weniger sichtbar sind (Insight 2 und 5).

Organisierte Aktionen der Gegenrede (Counter Speech) wollen diesen Entwicklungen begegnen und Diskurse online depolarisieren. Bislang gelang es wissenschaftlichen Studien kaum, die Effekte solcher Aktionen genauer zu untersuchen. Nun veröffentlichte ein Forschungsteam Ergebnisse, die darauf hindeuten, dass organisierte Counter Speech dazu beitragen kann, Diskurse zu depolarisieren und ein breiteres Meinungsspektrum zu generieren (Insight 4). Diese Ergebnisse lassen sich jedoch nicht auf jede Form der Gegenrede anwenden, denn die untersuchten Counter-Speech-Aktionen waren gut organisiert und von enormem Umfang. Darüber hinaus braucht es weitere Begegnungsformen. Eine davon ist die automatisierte Inhaltmoderation, beispielsweise das Sortieren, Filtern, Sperren oder Löschen von Inhalten, durch die Anbieter sozialer Netzwerke. Das kann einerseits zum bereits beschriebenen Ausschluss mancher Personen führen, die gegen Recht und Richtlinien verstoßen oder dazu, Inhalte zu kennzeichnen, die irreführend oder falsch sind. Andererseits birgt automatisierte Inhaltmoderation weitere Möglichkeiten für die Anbieter selbst: Sie können ihre hierfür entwickelten Algorithmen potenziell auf die gesamte Kommunikation anwenden, die über die jeweilige Plattform läuft. Das erlaubt es ihnen, die Inhalte nicht nur im Hinblick auf Verstöße zu analysieren, sondern auch hinsichtlich kommerzieller Erwägungen. Letztere könnten wiederum ihre automatisierten Moderationspraktiken beeinflussen, ohne dass das von außen ersichtlich ist (Insight 6). Algorithmische Selektion spielt nicht nur bei automatisierten Moderationspraktiken eine Rolle. Auch die Auswahl von Inhalten, die Nutzerinnen und Nutzer in sozialen Netzwerken personalisiert ausgespielt bekommen, wird algorithmisch erstellt. Dieser Auswahlprozess wirkt sich mitunter auf die Arbeit von Redaktionen sowie Journalistinnen und Journalisten aus. Dabei geht es um die Relevanz, die Algorithmen Inhalten durch diesen Auswahl- und Sortierprozess zuweisen. Redaktionen beziehen ihre Annahmen über diese automatisierten Vorgänge neben anderer Erwägungen zunehmend mit ein, wenn sie Inhalte produzieren (Insight 7).

Zahlreiche Untersuchungen widmen sich nicht nur den Problematiken, die in Zusammenhang mit der Informationslage und der Meinungsbildung entstehen, sondern auch Begegnungsansätzen. Zentrale Themen sind dabei Transparenz und Nachvollziehbarkeit, alternative am Gemeinwohl orientierte Angebote und Datenschutz seitens der Plattformanbieter, was sich in entsprechenden Regulierungen widerspiegeln muss (Insight 2, 5, 6, 8 und 9). Denn die Anbieter alleine können diesen Entwicklungen nicht begegnen. Vielmehr ist das eine gesamtgesellschaftliche Aufgabe, die ebenfalls breite Ansätze und eine geöffnete Debatte notwendig macht (Insight 2, 4, 5, 8 und 9). Der Blick der Wissenschaft weitet sich demnach zunehmend. Denn die Diskurse online folgen zentralen gesellschaftlichen Entwicklungen. Dazu zählen Nationalismus, Rassismus und Sexismus (Insight 2 und 5). Online erhalten sie neue Spielregeln und Ausprägungsformen (Insight 4, 5 und 6). Das wiederum strahlt zurück in die Debatten – auch solche fernab der Onlinekommunikationsräume (Insight 2 und 7).

# INSIGHTS

<b>I. VERÖFFENTLICHUNGEN</b>	<b>05</b>
Insight 1: Extreme ziehen um	05
Insight 2: Desinformation als Alltagsproblem	06
Insight 3: Korrekturen im Privaten	07
Insight 4: Dem Hass entgegenreden	08
Insight 5: Gewalt geht online – Frauen gehen offline	09
Insight 6: Automatisierte Kommerzialisierung	10
Insight 7: Quantifizierte Relevanz	11
<b>II. AKTUELLE FORSCHUNGSPROJEKTE</b>	<b>12</b>
Insight 8: Gemeinwohl im algorithmischen System	12
Insight 9: Hate Speech verfolgen und bestrafen	13
<b>III. VERANSTALTUNGSANKÜNDIGUNGEN</b>	<b>14</b>
Insight 10: Zwischen den Online-Diensten	14

# I. VERÖFFENTLICHUNGEN

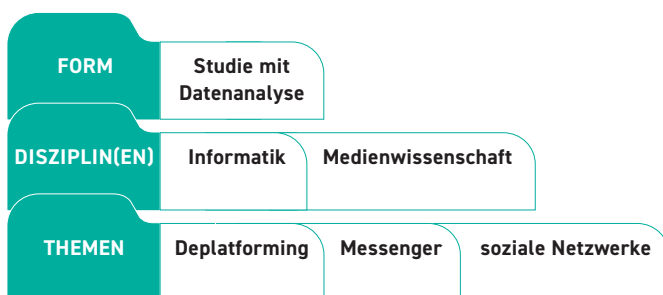
## INSIGHT 1: EXTREME ZIEHEN UM

Deplatforming beschreibt den Ausschluss Einzelner aus herkömmlichen öffentlichen sozialen Medien. Sie verbreiten ihre extremen Botschaften dann über alternative Netzwerke und Messengerdienste, beispielsweise in dafür eingerichteten Gruppenchats und Kanälen.

Mit diesen Zusammenhängen beschäftigt sich Richard Rogers in „Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media“. Die im Mai erschienene Studie untersucht vor allem die Aktivitäten von Personen des öffentlichen Lebens aus Großbritannien und den USA, die extreme Meinungen äußern und damit auf herkömmlichen Plattformen gegen entsprechende Richtlinien verstoßen, was zu ihrem Ausschluss führte. Das wirkt sich auch auf die Forschung aus, denn eine Vielzahl der Informationsverbreitungswege verlagert sich damit ins Verborgene. Deshalb findet laut dem Autoren auch eine gewisse Form des Deplatforming der Wissenschaft statt.

### TAKEAWAYS:

- (1) Der Messengerdienst *Telegram* ist eine beliebte Alternative für Prominente, die extreme Inhalte verbreiten. Dabei nutzen sie den Messengerdienst auf eine vergleichbare Weise wie herkömmliche Plattformen: Sie streuen hauptsächlich ihre Inhalte, nutzen ihre Accounts aber weniger dazu, um neue Anhängerinnen und Anhänger zu rekrutieren oder Inhalte zu archivieren, die auf herkömmlichen Angeboten nicht mehr verfügbar sind.
- (2) Prominente, die extreme Inhalte verbreiten, erhalten durch den Ausschluss aus einem herkömmlichen Netzwerk zunächst verstärkt Aufmerksamkeit. Nutzerinnen- und Nutzerzahlen ihrer Accounts in alternativen Netzwerken steigen mit Medienberichten stark an, die den Ausschluss aufgreifen. Auch entsprechende alternative Plattformen erhalten hierdurch Zulauf. Auf einen längeren Zeitraum bezogen, zeigt sich jedoch, dass sich ihr Publikum mit dem Umzug verkleinert. Auch die Äußerungen wurden mit der Zeit milder.
- (3) Die Ausgeschlossenen bezeichnen und inszenieren sich häufig als Opfer der „Cancel Culture“. Sie beurteilen den Ausschluss als Unrecht und Ungleichbehandlung. Dabei bezeichnen sie insbesondere *Facebook* und *Instagram* als Plattformen, die die freie Meinungsäußerung nicht garantieren.



### Quelle(n):

Roger, Richard: Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media. In: European Journal of Communication, Vol. 35, Issue 3, 2020.  
Online unter: <https://doi.org/10.1177%2F0267323120922066> (04.12.2020).

### Hinweis:

In der Studie „Datenanalyse von rechtsverstoßenden Inhalten in Gruppen und Kanälen von Messengerdiensten am Beispiel Telegram“ analysieren Jakob Jünger und Chantal Gärtner von der Universität Greifswald im Auftrag der Landesanstalt für Medien NRW Rechtsverstöße und die Verbreitung von Desinformation auf Telegram. Die Untersuchung steht seit 5. November hier zur Verfügung: [https://www.medienanstalt-nrw.de/fileadmin/user\\_upload/Forschungsmonitor\\_Informationsintermediaere\\_4.Ausgabe.pdf](https://www.medienanstalt-nrw.de/fileadmin/user_upload/Forschungsmonitor_Informationsintermediaere_4.Ausgabe.pdf).

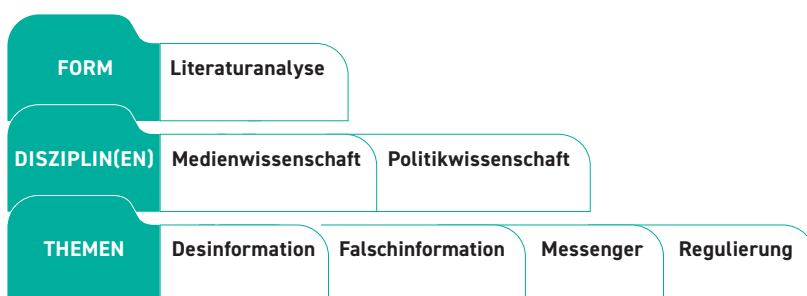
# INSIGHT 2: DESINFORMATION ALS ALLTAGSPROBLEM

Die Informationslage zum Coronavirus und der Pandemie in Deutschland macht vielen bewusst: Desinformation ist kein abstraktes oder entferntes Problem. Zunehmend verbreitet sich Desinformation auch über Messenger.

Das beschreibt Ann Cathrin Riedel in „Behind Closed Curtains“. Sie wertet zahlreiche Studien aus, um ein genaueres Bild dessen zu erzeugen, wie sich Desinformation unter anderem über Messenger wie *Telegram* oder *WhatsApp* in Deutschland, Brasilien und Indien verbreitet. Die im Juli erschienene Analyse kommt zu dem Schluss, dass entsprechende Messengerdienste es erleichtern, falsche und irreführende Informationen zu streuen. Jedoch reicht das alleine nicht aus, um die komplexe und diffuse Informationslage zu erklären. Zugleich spielen in allen Ländern mangelndes Vertrauen in die Regierung, zunehmender Nationalismus und begleitende Erscheinungen wie Rassismus und Sexismus eine übergeordnete Rolle. Die Analyse zeigt das insbesondere für Brasilien und Indien. Sie bezeichnet das Vorgehen gegen Desinformation als eine gesamtgesellschaftliche Aufgabe, die nicht alleine in der Hand entsprechender Anbieter liegen darf, und arbeitet sechs Empfehlungen hierfür aus. Im Ländervergleich ist Desinformation in Deutschland weniger stark verbreitet als in Brasilien und Indien. Hierzulande handelt es sich bei den untersuchten Inhalten häufig um irreführende Darstellungen, die als rechtsradikal einzustufen sind.

## TAKEAWAYS:

- (1) Begegnungsansätze sind Aufklärung und das Verbreiten entsprechender Kompetenzen in der Bevölkerung. Dazu zählt auch, korrekte Begriffe zu verwenden. „Fake News“ etwa ist kein klar definierter Begriff und wird missbräuchlich verwendet. Zwischen Falsch- und Desinformation<sup>2</sup> ist zu differenzieren. Entsprechende Institutionen müssen zudem über Intentionen und Verbreitungswege von Desinformation aufklären und Materialien für Schulen sowie die Weiterbildung in jedem Alter bereitstellen.**
- (2) Die an der Entstehung einer globalen Internet Governance Beteiligten müssen das Vorgehen gegen Desinformation mitdenken. Dabei sind die Plattformbetreiber zu adressieren. Zugleich können Reformen und das Durchsetzen des Kartell- und Datenschutzrechts zielführend sein. Die Grundlage für wirksame regulatorische Maßnahmen ist immer vorangegangene Forschung.**
- (3) Plattformen und Messengerdienste müssen ihre Angebote so umgestalten, dass sich Desinformation weniger stark verbreitet. Dafür sind Ansätze notwendig, die die Transparenz und Nachvollziehbarkeit fördern. Konkrete Möglichkeiten sind, das massenhafte Weiterleiten von Nachrichten einzuschränken und grundsätzliche vertrauenswürdige Informationen verfügbar zu machen. Zugleich ist es sinnvoll, den Journalismus finanziell zu stärken, um so die Qualität der Angebote und das Vertrauen in sie auszubauen.**



### Quelle(n):

Riedel, Ann Cathrin: Behind Closed Curtains. Desinformation auf Messengerdiensten. Friedrich Naumann Stiftung, 2020. Online unter: <https://shop.freiheit.org/#!/Publikation/908> (04.12.2020).

<sup>2</sup> Falschinformationen sind faktisch nicht korrekt, werden aber ohne Absicht verbreitet. Sie entstehen etwa aufgrund nicht ausreichender Recherchen. Auch Desinformation ist faktisch nicht korrekt, widerspricht also der Sachlage oder stellt sie ausschnittshaft dar. Akteure, die sie verbreiten, tun dies jedoch bewusst und verfolgen damit eine Absicht.

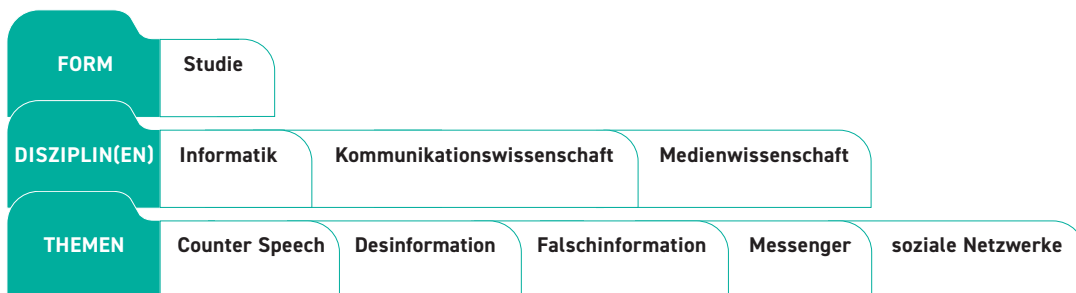
# INSIGHT 3: KORREKTUREN IM PRIVATEN

Desinformation spielt nicht nur in Demokratien des globalen Nordens eine Rolle. Der Großteil der Forschung hierzu konzentriert sich derzeit jedoch auf die USA sowie West- und Mitteleuropa.

Die Autorinnen der Studie „Dysfunctional information sharing on WhatsApp and Facebook“ nehmen die Informationslage auf Online-Plattformen in Brasilien in den Blick. Die Forscherinnen vergleichen, wie sich falsche Informationen über das soziale Netzwerk *Facebook* und über den Messengerdienst *WhatsApp* verbreiten. Dafür nutzen sie eine repräsentative Stichprobe mit Verhaltensdaten von 1.615 Nutzerinnen und Nutzern. Die im Juni erschienene Analyse arbeitet unter anderem Unterschiede der Informationsverbreitung zwischen (halb-)öffentlichen und privaten Plattformen heraus.

## TAKEAWAYS:

- (1) Es gibt einen Zusammenhang zwischen Teilhabe am Diskurs und dem Verbreiten falscher Informationen: Diejenigen, die sich stark in politische Debatten einbringen, tauschen mit signifikant höherer Wahrscheinlichkeit falsche Informationen aus als solche, die sich weniger stark einbringen. Zugleich erleben Personen mit hoher Teilhabe häufiger Korrekturen, was darauf hindeutet, dass die Verbreitung falscher Inhalte in den Netzwerken nicht unbemerkt bleibt.
- (2) Solche sozialen Korrekturen finden wesentlich häufiger bei *WhatsApp* als auf *Facebook* statt. Darüber hinaus sind die Bindungen zwischen den Nutzerinnen und Nutzern bei *WhatsApp* stärker. Darin sehen die Forscherinnen ein Anzeichen dafür, dass diese engeren Bindungen ein Sicherheitsgefühl vermitteln, welches solche Verhaltensweisen unterstützt. Unbekannt ist, wie sich diese Korrekturen auswirken.
- (3) Anbietern von Messengerdiensten ist es nicht möglich, die über sie verbreiteten Informationen auf Korrektheit zu prüfen. Dennoch erscheint es sinnvoll, dass die Dienste ihre Schnittstellen so ausrichten, dass sie die Nutzerinnen und Nutzer dabei unterstützen, die Glaubwürdigkeit einer Information zu beurteilen. Dafür ist der Kontext entscheidend, in den eine Information eingebettet ist, wozu etwa Verlinkungen und Quellenangaben zählen.



### Quelle(n):

Rossini, Patrícia et al.: Deplatforming: Dysfunctional information sharing on WhatsApp and Facebook: The role of political talk, cross-cutting exposure and social corrections. In: *New Media & Society*, 2020. Online unter: <https://doi.org/10.1177%2F1461444820928059> (04.12.2020).

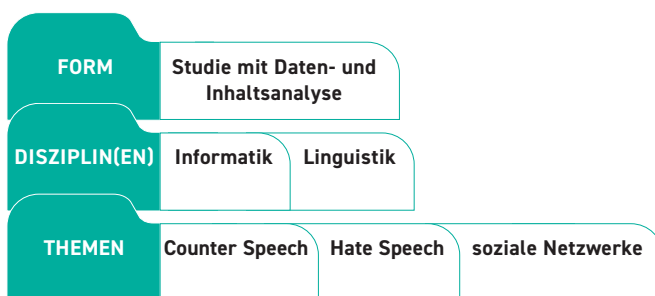
# INSIGHT 4: DEM HASS ENTGEGENREDEN

Organisierte Gegenrede (Counter Speech) kann dazu beitragen, dass sich hasserfüllte Diskussionen weniger stark polarisieren und weitere Nutzerinnen und Nutzer dazu anhalten, ebenfalls Gegenrede auszuüben.

Anhaltspunkte dafür veröffentlicht ein Forschungsteam verschiedener US-amerikanischer Universitäten und des *Max-Planck-Instituts for Mathematics in the Sciences* in Leipzig im Juni. Dem Team gelang es mithilfe eines Klassifikationsalgorithmus 135.000 Diskussionen, die zwischen 2013 und 2018 auf *Twitter* stattfanden, hinsichtlich Äußerungen des Hasses und der organisierten Gegenrede zu untersuchen. Mit dem Konzept der Gegenrede setzten sich bislang wenige empirische Untersuchungen auseinander, unter anderem weil ein geeigneter Datenkorpus fehlte. Den fand das Forschungsteam in Deutschland, wo „Reconquista Germanica“ (RG) bei organisierten Aktionen mit Hate Speech politische Debatten stört und der 2018 entstandenen Gegenbewegung „Reconquista Internet“ (RI). Letztere versuchte, mithilfe von aktiver Gegenrede den Diskurs zurück ins Gleichgewicht zu bringen. RG verzeichnete zu Spitzenzeiten zwischen 1.500 und 3.000 Aktive, während sich RI bereits in der ersten Woche 45.000 Mitglieder anschlossen.

## TAKEAWAYS:

- (1) Der Anteil der Hassrede war während des gesamten Untersuchungszeitraums weitgehend stabil und nahm gegen Ende etwas zu. Die Äußerungen wurden über den Zeitraum hinweg immer extremer. Der Anteil der Gegenrede nahm ebenfalls zu und auch hier wurden die Äußerungen extremer.
- (2) Bei einer niedrigen Anzahl von Hasskommentaren, kam es geringfügig zu weiteren Äußerungen dieser Art, während Kommentare der Gegenrede seltener sind. Mit zunehmender Zahl an Hasskommentaren nahm auch die Gegenrede weiter zu. Das führte schlussendlich, ab einem gewissen Punkt dazu, dass die Zahl der Hassäußerungen wieder abnahm. Eine geringe Anzahl von Äußerungen der Counter Speech rief wenige Hassreaktionen hervor. Existierten in einer Diskussion jedoch viele Inhalte der Gegenrede, zogen sie auch mehr hasserfüllte Äußerungen nach sich.
- (3) Im Zeitraum vor der Gründung von RI und der damit beginnenden organisierten Gegenrede zogen hasserfüllte Kommentare weitere Hassäußerungen an und unterdrückten die bis dahin geringe Zahl der Counter Speech. Mit der Entstehung von RI zog Gegenrede, immer noch Hasskommentare auf sich, aber in einem deutlich geringeren Ausmaß. Zugleich rief sie weitere Äußerungen der Gegenrede hervor. Für alle Zeiträume gilt: Gegenrede stimuliert insgesamt eher neutrale oder nicht klassifizierbare Sprache als Hassrede, was für eine gewisse Depolarisierung der Diskussion spricht.



### Quelle(n):

Garland, Joshua et al.: Countering hate on social media: Large scale classification of hate and counter speech. 2020. Online unter: <https://arxiv.org/abs/2006.01974> (04.12.2020).



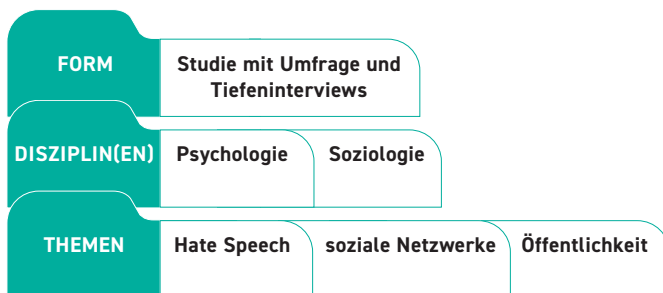
# INSIGHT 5: GEWALT GEHT ONLINE – FRAUEN GEHEN OFFLINE

58 Prozent der Mädchen und jungen Frauen weltweit erleben Belästigung und Gewalt in sozialen Medien. In Deutschland sind es 70 Prozent. Viele ziehen sich deshalb aus den digitalen Diskussionsräumen zurück.

Diese Ergebnisse stellt der im Oktober erschienene Bericht „The State of The World's Girls“ von *Plan International* vor. Er bereitet die Ergebnisse quantitativer Befragungen in 22 Ländern auf, darunter Deutschland. Hier nahmen 1.003 Mädchen und Frauen teil. In 16 Ländern wurden vertiefende Interviews durchgeführt. Die Untersuchung zeigt, dass Mädchen und junge Frauen regelmäßig Online-Angriffe erleben – unabhängig davon, wo sie leben. Dazu zählen Beschimpfungen und Beleidigungen, sexuelle Belästigung, persönliche Demütigung, Bodyshaming, rassistische Kommentare, Stalking, Kommentare gegen die sexuelle Orientierung oder Identität sowie das Androhen physischer Gewalt. 42 Prozent geben an, ein geringes Selbstbewusstsein oder Selbstwertgefühl zu haben. Bei 24 Prozent führen die Angriffe zu physisch empfundener Angst. Doch Konsequenzen ergeben sich nicht nur auf individueller Ebene, sondern wirken sich auch auf die Sichtbarkeit weiblicher Stimmen insgesamt aus.

## TAKEAWAYS:

- (1) 13 Prozent der Betroffenen weltweit nutzen soziale Medien aufgrund von Hass- und Gewalterfahrungen online weniger häufig. 13 Prozent veröffentlichen keine Inhalte mehr. Acht Prozent melden sich von den Plattformen ab. Am häufigsten erleben Mädchen und junge Frauen digitale Gewalt auf *Facebook* (39 Prozent) gefolgt von *Instagram* (23 Prozent).
- (2) In Deutschland nutzen 11 Prozent der Betroffenen die sozialen Medien weniger häufig. 9 Prozent veröffentlichen keine Inhalte mehr und 5 Prozent verlassen die Plattformen ganz. Hierzulande liegt *Instagram* mit 45 Prozent der Angriffe vor *Facebook* mit 35 Prozent. 52 Prozent der Betroffenen melden unangemessene Inhalte bei den Plattformen. Der Wert für Deutschland liegt damit 17 Prozentpunkte über dem weltweiten Durchschnitt.
- (3) Die Befragten sehen folgende Handlungsfelder, um die Situation zu verbessern: Bildung und Aufklärung, Befähigung von Mädchen und jungen Frauen, Auseinandersetzung mit den öffentlichen Darstellungen von Mädchen und Frauen, Hinterfragen der Geschlechternormen, die die Gesellschaft durchdringen, Einführen und Durchsetzen entsprechender Regulierungen, Nutzung von Technologie, wozu auch effektivere Meldemechanismen zählen. Der Bericht empfiehlt, diese für geschlechterspezifische Gewalt zu verbessern. Plattformanbieter, nationale Regierungen, die internationale Staatengemeinschaft und die Zivilgesellschaft stehen in der Pflicht entsprechend aufzuklären, zu schützen und insgesamt dafür zu sorgen, dass Menschenrechte in den sozialen Medien besser geachtet werden.



### Quelle(n):

Plan International: Free to be online? Girls' and young women's experiences of online harassment. The State of The World's Girls. 2020. Online unter: <https://www.plan.de/presse/pressemitteilungen/detail/welt-maedchenbericht-2020-digitale-gewalt-vertreibt-maedchen-und-junge-frauen-aus-den-sozialen-medien.html> (04.12.2020).

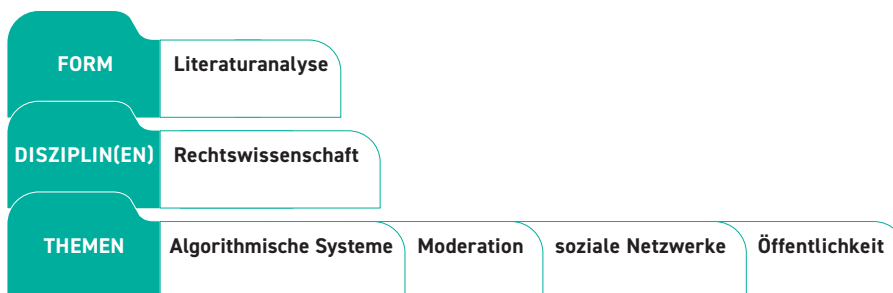
# INSIGHT 6: AUTOMATISIERTE KOMMERZIALISIERUNG

Aufgrund von Verstößen gegen geltendes Recht und Plattformregeln müssen die Anbieter sozialer Medien die dort zugänglichen Inhalte moderieren. Diese Moderation, beispielsweise das Sortieren, Filtern, Sperren oder Löschen von Inhalten, erfolgt zunehmend automatisiert.

Das im Oktober veröffentlichte Papier „Algorithmic Censorship by Social Platforms “ betrachtet, wie sich die automatisierte Moderation strukturell auswirkt. Die Autorin nennt als Gründe, warum diese Vorgänge automatisiert erfolgen, einerseits die Massen an zu prüfenden Veröffentlichungen sowie andererseits die enthaltenen oftmals gewaltsamen Darstellungen. Letztere können psychologische Auswirkungen auf menschliche Moderatorinnen und Moderatoren haben. Deshalb untersuchen zunehmend algorithmische Systeme die Inhalte nach vorher definierten Kriterien. Sie sortieren solche aus, die potentiell gegen Gesetze oder die Regeln der Plattformen verstoßen. Ihre genaue Funktionsweise ist weitgehend unbekannt. Bei einer zunehmenden Zahl sozialer Plattformen findet die automatisierte Inhaltsprüfung vor der Veröffentlichung statt.

## TAKEAWAYS:

- (1) Das Automatisieren der Moderation erlaubt es sozialen Plattformen, ihre entwickelten Systeme auf die gesamte Kommunikation anzuwenden. Die Moderations- und Selektionsalgorithmen ermöglichen potenziell eine größere Kontrolle über die Kommunikation von Millionen Menschen, als sie je zuvor ein privatwirtschaftlicher Akteur besaß.**
- (2) Die Plattformen erhalten mit der automatisierten Moderation mehr Möglichkeiten, bestimmte Ideen und Standpunkte eher zuzulassen und andere zu unterdrücken. Dabei können kommerzielle Erwägungen eine Rolle spielen und Unternehmenswerten kann Vorrang eingeräumt werden. Damit beeinflussen sie die strukturellen Bedingungen des öffentlichen Diskurses.**
- (3) Eine kleine Anzahl von Anbietern sozialer Medien dominiert den gesamten Markt. Deshalb ist es für Nutzerinnen und Nutzer schwierig, deren algorithmischen Zensur zu entgehen, auch wenn dies potentiell möglich ist. Dem können koordinierte Bemühungen begegnen, alternative soziale Plattformen und Nachrichtenkanäle einzurichten sowie Rechte auf Privatsphäre, Datenschutz und Meinungsfreiheit verstärkt geltend zu machen.**



## Quelle(n):

Cobbe, Jennifer: Algorithmic Censorship by Social Platforms: Power and Resistance. Philosophy & Technology, 2020. Online unter: <https://doi.org/10.1007/s13347-020-00429-0> (04.12.2020).

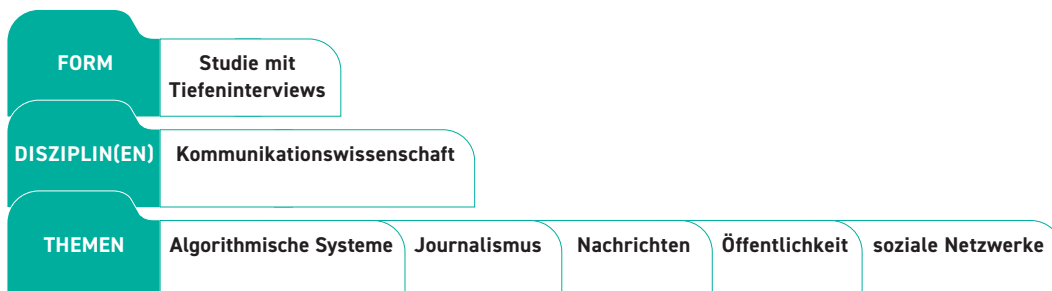
# INSIGHT 7: QUANTIFIZIERTE RELEVANZ

Die algorithmischen Systeme sozialer Plattformen wirken sich auf journalistisches und redaktionelles Arbeiten aus. Sie beeinflussen, welche Informationen und Inhalte Journalistinnen sowie Journalisten genauer betrachten und wie sie Nachrichten gestalten.

Anhaltspunkte dafür findet ein Forschungsteam der *Northwestern University*, das Tiefeninterviews mit 18 in den USA ansässigen Nachrichtenjournalistinnen und -journalisten sowie -redakteurinnen und -redakteuren führt. Dabei konzentriert sich die im Juli veröffentlichte Studie auf die Annahmen zu den Funktionsweisen von Social-Media-Algorithmen und die Frage, wie diese das journalistische Arbeiten beeinflussen. Die Untersuchung zeigt, dass sich die Art und Weise, wie die Algorithmen gestaltet sind und damit auch, wie sie Relevanz quantifizieren, auf journalistische Gate-keepingpraktiken auswirken. So beeinflussen sie mitunter auch, was das Publikum über klassische Medien rezipiert.

## TAKEAWAYS:

- (1) Die angenommene algorithmische Relevanz wirkt sich auf die redaktionelle Arbeit insbesondere dahingehend aus, wie Journalistinnen und Journalisten Inhalte gestalten und welchen Themen sie Ressourcen zuweisen. Das betrifft jedoch nicht den gesamten redaktionellen Entscheidungsprozess. Die algorithmische Relevanz wird mit dem traditionellen Verständnis eines Nachrichtenwerts abgeglichen und muss damit zu vereinbaren sein.
- (2) Zugleich ergeben sich Spannungen im Prozess dieses Abgleichs: Mit den klassischen Massenmedien bestand die Möglichkeit, Rezipientinnen und Rezipienten quasi zwangsweise mit den Inhalten zu bespielen, die die Redaktionen für gesellschaftlich bedeutsam hielten. Durch Informationsintermediäre besteht diese Möglichkeit nicht mehr. Damit treten zwei Aspekte in Konkurrenz miteinander: einerseits die Möglichkeit, dem Publikum das zu geben, was online potentiell die größte Aufmerksamkeit erhält sowie andererseits das Bestreben von Redaktionen, ihnen die für sie im journalistischen Sinne relevantesten Inhalte anzubieten.
- (3) Der Einfluss algorithmischer Systeme auf redaktionelle Entscheidungsprozesse nimmt insgesamt derzeit etwas ab. Die Redaktionen wenden sich vermehrt anderen Informationsintermediären und Anwendungen zu, deren Mechanismen durchschaubarer erscheinen, wie das beispielsweise bei Content-Aggregatoren<sup>3</sup> oder Suchmaschinenoptimierung<sup>4</sup> der Fall ist. Das ermöglicht es, ihre Inhalte dementsprechend anzupassen oder gezielter auszuspielen.



### Quelle(n):

Peterson-Salahuddin, Chelsea / Diakopoulos, Nicholas: Negotiated Autonomy: The Role of Social Media Algorithms in Editorial Decision Making. In: *Media and Communication*, Vol 8, No 3, 2020.

Online unter: <http://dx.doi.org/10.17645/mac.v8i3.3001> (04.12.2020).

<sup>3</sup> Content-Aggregatoren stellen automatisiert Inhalte aus verschiedenen Online-Quellen zusammen und spielen sie aus. Häufig handelt es sich dabei um Nachrichten. Beispiele dafür sind etwas *Google News* oder *Yahoo Nachrichten*.

<sup>4</sup> Suchmaschinenoptimierung dient dazu, die Sichtbarkeit einer Website bzw. ihrer Inhalte bei Suchmaschinenanfragen zu erhöhen, indem Benutzerinnen und Benutzer sie weiter vorne auf den Suchergebnisseiten angezeigt bekommen.

## II. AKTUELLE FORSCHUNGSPROJEKTE

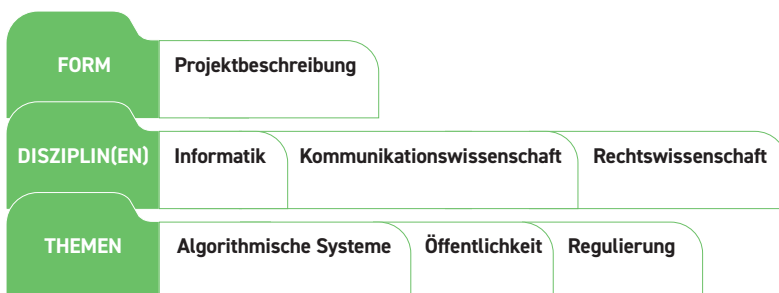
# INSIGHT 8: GEMEINWOHL IM ALGORITHMISCHEN SYSTEM

Wie kann Software dabei helfen, Medienpolitik zu gestalten? Dieser Frage widmet sich das interdisziplinäre Projekt „Coding Public Value“. Das *Hans-Bredow-Institut (HBI)* entwickelt mit verschiedenen Partnern eine öffentlich-rechtliche Medienplattform, die sich am Gemeinwohl orientiert.

Im Rahmen des von 2020 bis 2022 laufenden Projekts entwerfen und erproben die Beteiligten eine Medienplattform, die sich in ihrer technologischen Gestaltung an bestehenden institutionellen, politischen und organisationalen Bedingungen orientiert und die darin enthaltenen Ansätzen berücksichtigt, das Gemeinwohl zu stärken. Neben einem Prototyp für eine solche Software sieht das Projekt insbesondere eine Analyse von Stakeholdern im Mediensektor, der medienrechtlichen und regulativen Rahmenbedingungen sowie von Nutzungs- und Marktpotenzialen vor. Dabei zeichnet sich das HBI insbesondere für ein medienrechtliches Arbeitspaket verantwortlich.

### TAKEAWAYS (zugrundeliegende Annahmen):

- (1) Es gibt Handlungs- und Gestaltungsmöglichkeiten, damit sich Technologien, stärker am Gemeinwohl orientieren. Diese Möglichkeiten liegen nicht alleine in rein regulatorischen Maßnahmen, sondern auch darin, sie in die Gestaltung und Entwicklung von Technologien zu integrieren.
- (2) Im Fall von Medienplattformen ist ein wesentlicher Aspekt des gesetzlichen Rahmens für die öffentlich-rechtlichen Rundfunkanstalten der öffentlich-rechtliche Programmauftrag (§ 11 Rundfunkstaatsvertrag (RStV)). Für telemediale Angebote gelten zudem weitere spezifische Anforderungen und Verfahren (§§ 11d, 11f RStV). Die darauf basierenden Angebotskonzepte enthalten eine Vielzahl bestimmter und unbestimmter Konzeptvorgaben und Selbstverpflichtungen, die in der Regel die inhaltliche Gestaltung des Angebots betreffen.
- (3) Diese rechtlichen Vorgaben lassen sich ebenso wie gesellschaftlich-normative Verpflichtungen auf algorithmische Systeme anwenden, die bei Medienplattformen zum Einsatz kommen. Das ist bei bisherigen Plattformen jedoch nicht der Fall. Damit die rechtlichen Vorgaben in den Technologien wirken, gilt es, sie dafür zu konkretisieren.



#### Quellen:

<https://leibniz-hbi.de/de/projekte/coding-public-value>

<https://www.bidt.digital/bidt-foerdert-neun-forschungsprojekte-zur-digitalisierung/>

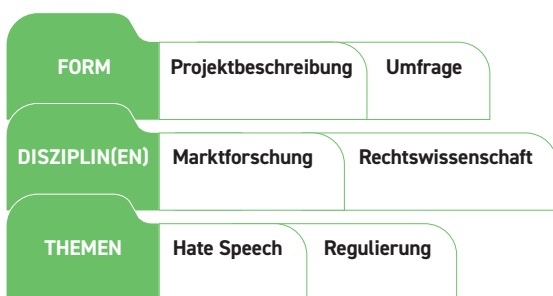
# INSIGHT 9: HATE SPEECH VERFOLGEN UND BESTRAFEN

Konkrete Vorschläge für den strafrechtlichen Umgang mit Online-Hassrede erarbeitet ein Forschungsprojekt der Universität Leipzig und nimmt dafür kriminologische, strafrechtliche und strafprozessuale Perspektiven genauer in der Blick.

Die erste Phase des Projekts „Der strafrechtliche Umgang mit digitalem Hass“ ist dem besseren Verständnis von Online-Hassrede gewidmet. Eine Analyse entsprechender Kommentare erfasst, gegen wen sich hasserfüllte Äußerungen im Internet richten und welchen Inhalten besonders häufig mit Hate Speech begegnet wird. Daneben plant das Forschungsteam Interviews mit den Adressatinnen und Adressaten von Hass-Kommentaren einerseits sowie den Verfasserinnen und Verfassern andererseits. Um einen ersten Überblick darüber zu erhalten, wie Nutzerinnen und Nutzer Hate Speech wahrnehmen, führte das Marktforschungsinstitut *Forschungsgruppe g/d/p* im Juni eine repräsentative Umfrage mit 1.069 Personen durch. Mit ihr startet das auf drei Jahre angelegte Forschungsprojekt, welches durch das Bundesministerium für Justiz und Verbraucherschutz gefördert wird.

## TAKEAWAYS (Zwischenergebnisse):

- (1) 79 Prozent der Befragten nehmen Online-Kommentare im Verlauf der letzten fünf Jahre als etwas aggressiver (30 Prozent) oder deutlich aggressiver (49 Prozent) wahr. 18 Prozent gaben an, bereits von Hassrede betroffen gewesen zu sein. Der Anteil der Betroffenen steigt mit sinkendem Alter: Bei den 16- bis 30-Jährigen liegt der Anteil bei 32 Prozent.**
- (2) 42 Prozent geben an, dass sie eigene Online-Beiträge aufgrund von Hassrede vorsichtiger formulieren oder darauf verzichten, etwas zu veröffentlichen. Dieser Anteil der Befragten ist in der Altersgruppe der 31- bis 40-Jährigen mit 56 Prozent besonders hoch und sinkt mit steigendem Alter. Jedoch liegt er in allen Altersgruppen bei über 30 Prozent.**
- (3) Für härtere Strafen bei Online-Beleidigungen als bei persönlichen Beleidigungen sprechen sich 43 Prozent der Befragten aus, während 50 Prozent gleich hohe Strafen für angemessen halten. Lediglich vier Prozent halten eine mildere Strafe für richtig.**



### Quellen:

<https://medienstrafrecht.jura.uni-leipzig.de/forschungstaetigkeit/forschungsstelle-evaluation-von-rechtsanwendung-in-der-praxis/projekte-digitaler-hass/bmju-projekt-der-strafrechtliche-umgang-mit-digitalem-hass/>  
[https://medienstrafrecht.jura.uni-leipzig.de/download/0/0/1918901509/dc651fb36de80589e5f5f6ca-6eb6df991799323b/fileadmin/medienstrafrecht.jura.uni-leipzig.de/uploads/Veroeffentlichungen/gdp\\_Ergebnisse\\_HateSpeech\\_Kurzbericht.pdf](https://medienstrafrecht.jura.uni-leipzig.de/download/0/0/1918901509/dc651fb36de80589e5f5f6ca-6eb6df991799323b/fileadmin/medienstrafrecht.jura.uni-leipzig.de/uploads/Veroeffentlichungen/gdp_Ergebnisse_HateSpeech_Kurzbericht.pdf)  
[https://www.gdp-group.com/fileadmin/ms/presstext\\_hate\\_speech\\_final.pdf](https://www.gdp-group.com/fileadmin/ms/presstext_hate_speech_final.pdf)  
<https://www.jura.uni-leipzig.de/newsdetail/artikel/hate-speech-birgt-die-gefahr-wichtige-stimmen-aus-dem-oefentlichen-diskurs-zu-verdraengen-2020-07/>

# III. VERANSTALTUNGSANKÜNDIGUNGEN

## INSIGHT 10: ZWISCHEN DEN ONLINE-DIENSTEN

Sind herkömmliche soziale Plattformen für rechtsextreme Hassgruppen nicht mehr zugänglich, löst das Dynamiken wie etwa Kontensperrungen und Umzüge auf andere Kanäle aus. Das beeinflusst das Kommunikations-, Protest- und Gewalthandeln der extremen Rechten.

Diesen Zusammenhängen widmet sich die Tagung „Hate not found?! Das Deplatforming der extremen Rechten und seine Folgen“. Sie findet am 11. Dezember 2020 als hybrides Veranstaltungsformat statt. Ein Forschungsteam des *Institut für Demokratie und Zivilgesellschaft* untersuchte die Strategien rechtsextremer Akteure, Inhalte mit einem großen Publikum zu teilen und stellt auf der Veranstaltung seine Erkenntnisse dazu vor. Auf dieser Grundlage diskutieren internationale Expertinnen und Experten Mechanismen und Effekte des Deplatforming. Zu Gast sind unter anderem die Extremismusforscherin Cynthia Miller-Idriss und Justizministerin Christine Lambrecht. Die Tagung möchte Politik, Zivilgesellschaft und Plattformbetreibern neue Impulse für den Umgang mit Ausprägungen des Rechtsextremismus im Digitalen geben. Neben eines Livestreams gibt es ein internes Forum, in dessen Rahmen sich Forscherinnen und Forscher mit den Konferenzteilnehmerinnen und -teilnehmern austauschen können.

<b>FORM</b>	Veranstaltungsbeschreibung		
<b>DISZIPLIN(EN)</b>	Kommunikationswissenschaft	Medienwissenschaft	Psychologie
<b>THEMEN</b>	Deplatforming	Messenger	soziale Netzwerke

**Quelle(n):**

<https://www.idz-jena.de/vadet/digitale-tagung-hate-not-found-das-deplatforming-der-extremen-rechten-und-seine-folgen/>

# Impressum

## **Herausgeberin:**

Landesanstalt für Medien NRW  
Zollhof 2  
D-40221 Düsseldorf

info@medienanstalt-nrw.de  
www.medienanstalt-nrw.de

## **Projektleitung:**

Desiree Steppat

## **Realisierung:**

iRights.Lab GmbH  
Schützenstraße 8  
D-10117 Berlin  
www.irights-lab.de

## **Projektleitung:**

Philipp Otto (iRights.Lab)

## **Autorin:**

Jaana Müller-Brehm (iRights.Lab)

## **Redaktion:**

Dr. Wiebke Glässer (iRights.Lab)  
Annika Albert (iRights.Lab)

Der „Forschungsmonitor Informationsintermediäre“ wird im Auftrag der *Landesanstalt für Medien NRW* erstellt und durch den unabhängigen Think Tank *iRights.Lab* erarbeitet.

Ziel dieses Forschungsmonitors ist es, aktuelle Entwicklungen im Themenfeld digitale Informationsintermediäre und öffentliche Meinungsbildung so aufzubereiten, dass das Monitoring einen Überblick über neue wissenschaftliche Publikationen, den Stand aktueller Forschungsprojekte und kommende relevante Veranstaltungen im Themenfeld verschafft.

Das Team des *iRights.Lab* entwickelt Strategien und praktische Lösungen, um die Veränderungen in der digitalen Welt vorteilhaft zu gestalten. Wir unterstützen öffentliche Einrichtungen, Stiftungen, Unternehmen, Wissenschaft und Politik dabei, die Herausforderungen der Digitalisierung zu meistern und die vielschichtigen Potenziale effektiv und positiv zu nutzen. Dazu verknüpfen wir rechtliche, technische, ökonomische, sozialwissenschaftliche und gesellschaftspolitische Expertise.