

# Hass im Netz

Anregungen und Empfehlungen für Redaktionen.

Von *Leif Kramp* und *Stephan Weichert*

**Abstract** Bei Berichterstattung über gesellschaftliche Reizthemen entfesselt sich häufig eine explosive Stimmung in den Kommentarbereichen von Nachrichtenangeboten im Internet. Viele Redaktionen haben ihre Kommentarfunktion wegen Überforderung bereits abgeschaltet. Gerade in den Sozialen Medien geht es – nicht nur für Journalist\_innen – inzwischen um den gesellschaftlichen Zusammenhalt. Der Beitrag fasst die Ergebnisse einer Studie zu Hasskommentaren auf Online-Nachrichtenangeboten und daraus folgende Implikationen für redaktionsseitige Moderationsstrategien zusammen, die im Auftrag der Landesanstalt für Medien Nordrhein-Westfalen erstellt wurde. Die Autoren raten Redaktionen, systematisch Ansätze für den Nutzerdialog zu entwickeln – auch mit hassgetriebenen Kommentator\_innen.

Das Kommentarverhalten von Mediennutzer\_innen zu journalistischen Veröffentlichungen im Internet verändert die öffentliche Debattenkultur: Potentiell gibt es, zumal für junge Zielgruppen, verbesserte Teilhabemöglichkeiten an gesellschaftlichen Diskursen sowie an politischer Willens- und Meinungsbildung. Zugleich birgt diese Debattenkultur auch Risiken für die demokratische Diskursrationalität, wenn dysfunktionale Kommunikationspraktiken in den Vordergrund treten. Gerade für Nachrichtenanbieter und deren Redaktionen stellt der Dialog mit dem Publikum, insbesondere die Sichtung, Moderation, Prüfung und Freischaltung von Kommentaren, nicht nur eine riesige Gestaltungsaufgabe dar (vgl. Kramp/Loosen 2018), sondern sie ist vor allem frustrierend. Seit einigen Jahren ist im Zuge der verstärkten Nutzung von Social-

*Dr. Leif Kramp ist Wissenschaftlicher Mitarbeiter und Forschungskoordinator am Zentrum für Medien-, Kommunikations- und Informationsforschung der Universität Bremen.*

*Prof. Dr. Stephan Weichert ist Leiter des berufsbegleitenden Masterstudiengangs „Digital Journalism“ und Direktor des „Digital Journalism Fellowship“ an der Hamburg Media School.*

Media-Plattformen eine Dynamisierung und Unterminierung öffentlicher Debatten erkennbar, die mittels einer „besseren Sichtbarkeit von öffentlicher und halböffentlicher Kommunikation [...] in unterschiedlichen Zuspitzungen unter dem Stichwort ‚Hassrede‘ oder auf Englisch ‚hate speech‘“ beschrieben wird, wie das wissenschaftliche Gutachten zum Medien- und Kommunikationsbericht der Bundesregierung feststellt (Hasebrink u. a. 2017, S. 212).

Das Netz hat dabei schwer durchschau- und nachvollziehbare Dynamiken im demokratischen Staatsgefüge freigesetzt: Nicht nur Hassrede, Fake News oder Hetze allgemein, auch Tendenzen zu konkreter Volksverhetzung, Denunziation sowie Aufrufe zu Straftaten und Gewalt sind zur alltäglichen Routine

einer in Teilen verwahrlosten Debattenkultur geworden. Vor allem Social-Media-Redaktionen journalistischer Medienanbieter müssen sich in ihren Kommentarbereichen und vor allem ihren Social-Media-Ablegern u. a. bei Facebook, Twitter, YouTube oder Instagram

*Damit ein konstruktiver Nutzerdialog gelingt, muss meist ein hoher personeller und finanzieller Aufwand betrieben werden.*

mit der Frage auseinandersetzen, was sie tun, wenn sich einige Nutzer\_innen nicht an die vom jeweiligen Medium aufgestellten digitalen Benimmregeln (Netiquette) halten (vgl. Lilienthal u. a. 2014a und b; Springer/Kümpel 2018; Weichert 2014a und b, 2015).

Formen lärmender, zügelloser und teils strafbarer Individual- und Massenkommunikation in den Sozialen Netzwerken haben massive Implikationen für die digitale Transformation der Gesellschaft. Diese neuen Qualitäten öffentlicher Debatten erfordern eine möglichst dichte Beschreibung von Diskursverläufen. Damit ein konstruktiver Dialog zwischen Nutzer\_innen in Gang kommen kann, muss derzeit fast immer aufseiten von Nachrichtenorganisationen ein hoher personeller und finanzieller Aufwand betrieben werden. Aber wie können journalistische Redaktionen die wachsende Zahl und Intensität von Hassrede bewältigen?

## Methodisches Vorgehen

Zur Beantwortung dieser Frage haben die Autoren dieses Beitrags das Diskussionsverhalten der Nutzer\_innen führender Nachrichtenanbieter im Netz und deren konkrete Moderationsstrategien untersucht. Analysiert wurden die Websites bzw. Social-Media-Auftritte des „Deutschlandfunks“, der „Rheinischen Post“, von „RTL Aktuell“ und der „Tagesschau“. Im Zentrum

des Erkenntnisinteresses stand die Frage, wie journalistische Medien den Dialog und die Interaktion mit ihrem Publikum vorantreiben können. Das Sample der Kommentaranalyse umfasst insgesamt 24 unterschiedliche Online-Diskursverläufe zu 16 einschlägigen journalistischen Beiträgen. Dabei wurden die relevanten Plattform- bzw. Moderationsstrategien der Kooperationspartner aufgegriffen: Bereits in Vorgesprächen mit den Redaktionsverantwortlichen stellte sich heraus, dass sich die strategischen Maßnahmen der Nachrichtenanbieter im Social Web vor allem auf Facebook konzentrieren. Analysiert wurden deshalb die Nutzerdiskurse bzw. das Kommentaraufkommen auf den Facebook-Ablegern der genannten Medienangebote sowie bei „RP Online“ und „tagesschau.de“ zusätzlich in den Kommentarbereichen der redaktionseigenen Websites.<sup>1</sup>

*Die im Rahmen der Studie analysierten Nutzerdiskurse lassen sich in fünf unterschiedliche Charakterisierungen unterteilen.*

Insgesamt wurden im Rahmen der Studie im Zeitraum April bis Dezember 2017 6876 Kommentare zu den 16 ausgewählten journalistischen Veröffentlichungen auf den Facebook-Seiten und den Websites der Nachrichtenanbieter analysiert. Die auf den unterschiedlichen Plattformen der betreffenden Medien geführten Nutzerdiskurse lassen sich in fünf unterschiedliche Gesamtcharakterisierungen unterteilen (vgl. Abb. 1):

- 1 Nutzerdiskurse auf der Social-Media-Plattform Facebook unterliegen grundlegend unterschiedlichen Bedingungen als solche auf den eigenen Websites der untersuchten Nachrichtenanbieter: Während die Nutzerkommentare zu den journalistischen Beiträgen auf Facebook automatisch freigeschaltet werden und erst im Anschluss von der Redaktion moderiert – und auch gelöscht – werden können, müssen die Beiträge in den Kommentarsektionen der eigenen redaktionellen Website im Untersuchungssample vor der Veröffentlichung erst von der Redaktion freigegeben werden. Ein weiterer Unterschied bei den technischen Vorbedingungen der Nutzerdiskurse liegt in der Antwortfunktion: Auf Facebook können Nutzer\_innen direkt auf spezifische Kommentare antworten, auf den eigenen redaktionellen Websites der Nachrichtenanbieter ist dies nur in Form eines Verweises durch das in Sozialen Medien geläufige @-Zeichen mit dem Namen der jeweils anzusprechenden Nutzer\_innen oder sonstiger namentlicher Bezugnahme aufeinander möglich. Dies hat jedoch keine Auswirkung auf die Sortierung der betreffenden Beiträge: Auf den eigenen Websites der Redaktionen erscheinen die Kommentare strikt chronologisch, während sie auf Facebook von einem Algorithmus in unterschiedlichen Sortierungen angezeigt werden können.

- ▶ *Hass- bzw. konfliktgeladene Diskurse zu Beiträgen über gesellschaftspolitische Reizthemen stark geprägt von vielen destruktiven Kommentaren:* Diese Nutzerdiskurse sind gekennzeichnet von einer überdurchschnittlichen Zahl an Nutzerkommentaren, die Elemente von Hassrede enthalten. Zudem sind in der Regel starke Umdeutungsmuster des redaktionell gesetzten Themas zu beobachten. Eine weitere Auffälligkeit ist ein hoher Grad an eskalierender Interaktion zwischen unterschiedlichen Nutzer\_innen.
- ▶ *Von starker Negativität und singulären Artikulationen geprägte Diskurse ohne Eskalationstendenz:* Diese Nutzerdiskurse fallen durch das vermehrte Auftreten von Hassrede auf, die jedoch selten andere Diskursteilnehmer\_innen adressiert. Eine Eskalation im Sinne einer sich zuspitzenden Streitkommunikation ist nicht zu beobachten.
- ▶ *Diskurse mit nutzerseitigem Selbstregulierungseffekt:* Diese Nutzerdiskurse zeugen von einem lebendigen Diskurscharakter im Sinne eines ausgeglichenen Verhältnisses zwischen Hassredner\_innen und ihren Gegenredner\_innen. Elementen von Hassrede setzen andere Kommentierende im Sinne einer diskursiven Selbstregulierung konstruktive, lösungsorientierte Äußerungen und Argumentationen entgegen.
- ▶ *Diskurse mit überwiegend konstruktiven, lösungsorientierten und/oder affirmativen Kommentaren:* Diese Nutzerdiskurse sind überwiegend von einem Empfehlungs- und Nutzwertcharakter gekennzeichnet. Der journalistische Beitrag stimuliert bei kommentierenden Nutzerinnen und Nutzern eine meist konstruktive Auseinandersetzung mit dem redaktionell gesetzten Thema.
- ▶ *Diskurse mit lebensweltlichem Alltagsbezug und hohem, weitgehend neutralem Kommentaraufkommen:* Diese Nutzerdiskurse entspinnen sich zu journalistischen Beiträgen, die ein nur geringes Konfliktpotential aufweisen und überwiegend themenbezogene und sachliche Nutzerkommentare evozieren.

Wie soziale Partizipation an öffentlichen Auseinandersetzungen zu medialer Berichterstattung sowohl auf der zurzeit populärsten sozialen Netzwerkplattform Facebook als auch auf ausgewählten Nachrichtenwebsites praktiziert wird und wie Nutzerkommentare einer redaktionellen Dialogisierung zuträglich sein könnten, wird in den analysierten Nutzerdiskursen



deutlich.<sup>2</sup> Sie zeigen allerdings auch die Omnipräsenz von Inzivilität (vgl. Beckert/Ziegele 2018, S. 67) in Nutzerkommentaren als Ausdruck von Aversion und Herabsetzung, die sich insbesondere bei gesellschaftlichen Reizthemen in Form von Hassrede äußern. Dies stellt Inhalteanbieter wie vor allem Nachrichtenredaktionen, die sich mit ihren Veröffentlichungen um einen Nutzerdialog im Netz bemühen, vor große Herausforderungen.

Die in Deutschland mit einem Fokus auf Jugendliche und junge Erwachsene im Rahmen der Europaratsinitiative gegen Rassismus und Diskriminierung im Internet gestartete No-Hate Speech-Kampagne versteht unter Hassrede eine intendierte gesellschaftlich diskriminierende Kommunikation als Ausdruck „gruppenbezogener Menschenfeindlichkeit“ (Groß/Zick/Krause 2012; vgl. auch No Hate Speech Movement 2017, S. 3f.). Diese

Abbildung 1:  
Gesamtcharakterisierung der analysierten Online-Nutzerdiskurse (unter Verwendung von Illustrationen von iStock.com/leremy)

2 Aus forschungspragmatischen Gründen beschränkt sich das im Rahmen der Online-Diskursanalyse codierte Material jeweils auf die ersten 150 Kommentare (zuzüglich der dazugehörigen Replies/Antwortkommentare im Falle von Facebook). Bei Kommentaren auf den eigenen Websites der Nachrichtenanbieter handelt es sich dabei um die 150 chronologisch zuerst veröffentlichten Kommentare, bei Kommentaren auf Facebook handelt es sich um die 150 zuerst gelisteten „Top-Kommentare“, die von Facebook aufgrund ihres Interaktionsgrades (z. B. Nutzerreaktionen) als besonders relevant gekennzeichnet werden.

Beschreibung wurde auch von staatlichen und zivilgesellschaftlichen Organisationen wie u. a. der Bundeszentrale für politische Bildung oder der Amadeu-Antonio-Stiftung aufgegriffen. Für die Online-Diskursanalyse wurde darauf rekurrend ein differenziertes Muster von Hassrede zugrunde gelegt, das von der Arbeitsgemeinschaft Kinder- und Jugendschutz (AJS) für die Landesanstalt für Medien NRW in Kooperation mit der klicksafe-Kampagne im Rahmen des CEF-Telecom-Programms der Europäischen Union entwickelt wurde (Tab. 1).

Tabelle 1:  
Muster von Hate  
Speech (AJS/LfM  
NRW 2016)

Muster von Hate Speech	
1)	Bewusste Verbreitung (uninformierter oder) falscher Aussagen
2)	Tarnung als Humor oder Ironie
3)	Herabwürdigende und verunglimpfende Begriffe; sexistische und rassistische Beleidigungen
4)	Bedienen von Stereotypen und Vorurteilen durch bestimmte Begriffe und Sprachmuster
5)	Verallgemeinerungen
6)	Wir/Die-Rhetorik
7)	Verschwörungstheorien
8)	Plakative Bildsprache
9)	Gleichsetzung
10)	Befürwortung oder Androhung sexualisierter Gewalt – oft in konzentrierter Form
11)	Befürwortung von oder Aufruf zu Gewalttaten

Standardmäßig wird den Facebook-Nutzer\_innen unter den redaktionellen Postings ein Ranking nach sogenannten „Top-Kommentaren“ angezeigt: Eine Sortierung nach der vermeintlichen Relevanz der Kommentare, gemessen an ihrem Interaktionsgrad und dem verifizierten Status der Nutzer\_innen, einschließlich der Redaktion als Inhaberin der Fanpage. Dieser Vorsortierung der Kommentare ist durch die voreingestellte

Rangfolge ein maßgeblicher Einfluss auf den Verlauf von Diskursen zu attestieren, auch wenn die Reihenfolge durch die Nutzer\_innen manuell geändert werden kann. Manuell kann eine Sortierung nach „neuesten Kommentaren“ (chronologisch) oder nach „Top-Komentaren (ungefiltert)“ (hier einschließlich Spam und Kommentaren in anderen Sprachen) ausgewählt werden.<sup>3</sup> Inhaber von Facebook-Seiten, den sogenannten Fanpages, haben die Möglichkeit, als problematisch eingestufte Kommentare nicht zwingend zu löschen, sondern nur auszublenken, so dass die betreffenden Kommentare für die jeweiligen Verfasser\_innen sowie deren Freunde weiterhin sichtbar bleiben, für alle übrigen Facebook-Nutzer\_innen jedoch nicht. Die Kommentare enthalten häufig Rechtschreib- und/oder Zeichensetzungsfehler, die wir eins zu eins übernommen haben.

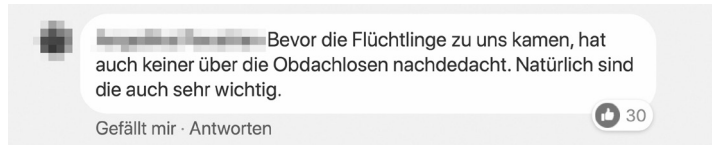
## Zusammenfassung der Ergebnisse

Die Analyse der Online-Diskurse auf den Fanpages der ausgewählten Nachrichtenangebote bei Facebook und auf den eigenen Websites der Medienhäuser ergab, dass der Aktivitätsgrad der Redaktionen in Bezug auf die Moderation von Kommentaren eher moderat bis nichtexistent ist: In den Kommentaren zu den untersuchten journalistischen Beiträgen gibt es wenig bis keine schriftlichen Moderationsbeiträge, so findet sich zum Beispiel in einem der Beiträge von „RP Online“ unter den ersten 150 Kommentaren nur ein einziger Moderationseintrag. Das Moderationsteam von „tagesschau.de“ belässt es dabei, elfmal auf Kritik mit Postings im selben Wortlaut zu reagieren. Davon abgesehen werden allenfalls Kommentare gelöscht bzw. ausgeblendet (Facebook) oder nicht freigeschaltet (Websites). Durch diese Zurückhaltung seitens der Redaktionen als sichtbarer Diskursakteur haben redaktionelle Eingriffe nur einen stark eingeschränkten Einfluss auf den Verlauf der öffentlichen Diskurse zu ihren Nachrichtenangeboten.

Aus der Untersuchung wird zudem deutlich, dass bestimmte Themen bei einigen Kommentierenden offenbar stärker Hassrede provozieren als andere: Ganz besonders auffällig ist

3 Facebook erklärt die voreingestellte Relevanzbemessung wie folgt: „Wenn die Rangfolge für eine Seite aktiviert ist, werden Kommentare mit den meisten ‚Gefällt mir‘-Angaben oder Antworten sowie Kommentare von Freunden oder verifizierten Profilen und Seiten standardmäßig ganz oben angezeigt“ (<https://www.facebook.com/help/539680519386145>).

Abbildung 2:  
Beispiel für „Wir/  
Die“-Rhetorik in ei-  
nem Nutzerkommen-  
tar auf der Facebook-  
Fanpage von „RTL  
Aktuell“



das fast schon „obligatorische“ Gegeneinander-Ausspielen von Minderheiten bzw. Personengruppen (Obdachlose vs. Flüchtlinge, Muslime vs. Buddhisten, Frauen vs. Männer etc.). Es handelt sich bei dieser verbalisierten „Wir/Die“-Rhetorik offenbar um einen dominanten Trend, der sich wie ein roter Faden durch die meisten Diskussionen zieht (vgl. Abb. 2).

Überraschend ist, dass sich im Schnitt nur etwa höchstens ein Drittel der analysierten Kommentare explizit auf Themen bezieht, die im Redaktionsbeitrag vorkommen: Meist gibt es ein zentrales Thema, das nicht im Artikel erwähnt wird, um das aber viele Kommentare kreisen. Diese Kommentierungen lassen vermuten, dass die meisten Kommentatoren vor allem ihre Meinung kundtun möchten, wobei die Kommentare tendenziell negativer Natur sind, also durch themen-fremde Verunglimpfungen und eine negative Haltung bis zur Hetze geprägt sind: Kaum ein Kommentar spricht sich für etwas aus (z. B. für Flüchtlinge oder für Frauen, die belästigt wurden). Eine Moderation seitens der Redaktionen bleibt häufig aus.

Die Analyse erlaubt zwar keine Rückschlüsse auf die persönlichen Motive der kommentierenden Nutzer\_innen, allerdings ist ein deutlicher Bezug zu Deutschland und „den Deutschen“ festzustellen, der häufig auch in Form von Pronomen („unsere Senioren“) geäußert wird. Konstruktiv sind solche Bezugnahmen zu Deutschland in der Regel nicht, d. h. es geht den Kommentierenden offenbar nicht darum, eine konstruktive Diskussionsebene wiederherzustellen, sondern Applaus anderer Diskussionsteilnehmer\_innen zu evozieren.

Es gibt in jedem Diskurs einige wenige dominante Nutzer\_innen, die sich bei den unterschiedlichen Medien in unterschiedliche Diskussionen in Kommentarverläufen zu unterschiedlichen redaktionellen Beiträgen (erneut) einmischen: Ihre Kommentare ähneln sich sehr und haben mitunter tendenziell appellativen Charakter. Diese notorisch negativ kommentierenden Nutzer\_innen „propagieren“ dabei eine klar erkennbare Botschaft bzw. Haltung, so dass der Eindruck entsteht, es handle sich um eine automatisierte Kommunikationsform. Dass es sich dabei um Fake-Profile handelt, ist wahrscheinlich, jedenfalls scheint diese Art der Kommentierung auf ein typisches



Abbildung 3:  
Beispiel für nutzerseitige Selbstregulierung auf der Facebook-Fanpage von „RP Online“:  
Hassredner\_innen und Gegenredner\_innen im Dialog

„Troll“-Verhalten hinzudeuten: Diese Wiederholungstäter\_innen sind spezielle Charaktere, deren Motive zwischen Geltungsdrang und missionarischem Eifer oszillieren.

Bei der für die Studie gewählten Kommentarsortierung bei Facebook handelt es sich um eine Reihung der Top-Kommentare, die durch hohe Interaktion und Reply-Anzahl gekennzeichnet sind. Diskussionen bzw. Dialoge unter Nutzer\_innen (in Form von Replies) sowie konzise Themenstränge entstehen bei den untersuchten Kommentaren der Facebook-Präsenzen fast ausschließlich im Verlauf der ersten zehn Kommentare („Top-Kommentare“), danach erhalten Kommentare allenfalls noch drei, höchstens fünf Antwort-Kommentare, selten mehr. In Summe

umfassen thematisch einschlägige Diskussionen also nur eine knappe Hand voll Kommentare, wobei die tatsächliche Chronologie der Diskussionsverläufe durch den Facebook-Algorithmus beeinflusst wird und die zeitliche Posting-Reihenfolge (durch die Nutzer\_innen) manuell rekonstruiert werden muss. So stehen die ersten Top-Kommentare unter einem Facebook-Beitrag in den meisten Fällen sinnbildlich für den weiteren Verlauf des Diskurses. Die thematischen Schwerpunkte und Argumentationsmuster werden schon zu Beginn des Diskurses gesetzt, das Gros der weiteren Kommentare orientiert sich daran.

Für die reguläre Chronologie gilt allerdings, dass fast alle Kommentare am Tag der Veröffentlichung eines Beitrags erstellt werden, danach kommentieren Nutzer\_innen nur noch vereinzelt und sporadisch. Dieser Umstand ist übergreifend festzustellen und an den einzelnen Veröffentlichungsdaten der Kommentare abzulesen. Ferner ist festzustellen, dass, je später Kommentare im Diskursverlauf erscheinen, diese umso kürzer sind und umso weniger Bezug zum journalistischen Beitrag aufweisen: Da Facebook nur

*Eine aktive Diskussionsbeteiligung der Redaktion wirkt sich direkt auf die Positionierung des Hauptkommentars aus.*

nach Interaktion, Länge des Kommentars etc. und nicht nach Inhalt sortiert, ist davon auszugehen, dass dieser Trend nichts mit der Relevanzbemessung und Sortierung durch Facebook zu tun hat.

Eine aktive Diskussionsbeteiligung der jeweiligen Redaktion wirkt sich unmittelbar auf die Positionierung des betreffenden Hauptkommentars und dessen Diskussionsstrang in der Sortierung des Facebook-Algorithmus aus (vgl. Abb. 4). Das heißt: Die entsprechenden Kommentare werden automatisch von Facebook höher gerankt, also selbst dann, wenn ein Nutzerkommentar nur verhältnismäßig wenige Replies oder Reaktionen von anderen Nutzer\_innen sammelt, kann er zu den Top-Kommentaren aufrücken, wenn die Redaktion als Moderator der Facebook-Fanpage einen Antwortkommentar verfasst. Auf diese Weise kann eine Redaktion u. E. durch gezielte Kommentierung bestimmten Diskussionssträngen zu mehr Aufmerksamkeit verhelfen.

Nutzer\_innen der analysierten Nachrichtenwebsites setzen sich scheinbar intensiver mit den kommentierten journalistischen Beiträgen auseinander als Nutzer\_innen auf Facebook, was nicht zuletzt an der Länge und am Faktengehalt der Kommentare auf Website und Facebook ersichtlich wird. Dass sich auf den analysierten Nachrichtenwebsites deutlich mehr Kom-



Abbildung 4:  
Beispiel für einen  
Moderationsbeitrag  
von „Deutschland-  
funk Kultur“, der  
vom Facebook-Algo-  
rithmus als relevan-  
tester Kommentar  
eingestuft wird.

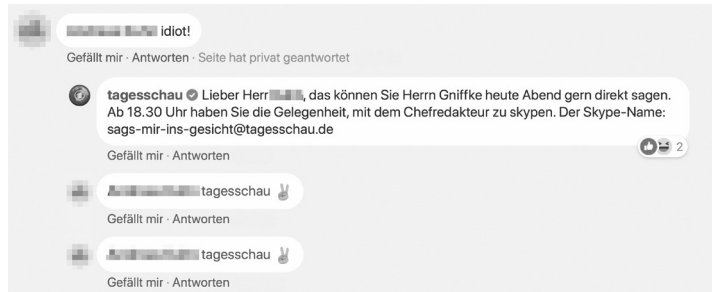
mentare auf ein Thema aus den jeweiligen journalistischen Beiträgen beziehen, untermauert diese Beobachtung.

Auch wenn die Profile der Kommentierenden im Sample bei der Codierung keine Rolle gespielt haben, ist festzustellen, dass zum einen fast alle Nutzer\_innen Klarnamen zu verwenden scheinen und zum anderen das Geschlecht weder auf die Häufigkeit noch auf die Härte der Kommentare Einfluss hat. Im Gegenteil entsteht oft der Eindruck, dass Frauen mehr abwertende und tendenziell mehr Hassrede enthaltende Kommentare schreiben als Männer. Besonders bei den Kommentaren zum Thema Sexismus wird dies deutlich: Hier unterstellen (vermeintliche) Nutzerinnen jeweils anderen Nutzerinnen, dass sie die Vorwürfe sexueller Belästigung frei erfunden oder sich bei Übergriffen nicht ausreichend gewehrt hätten.

Die scheinbar häufigste Form von Hassrede ist gleichzeitig auch die primitivste: das Verwenden von herabwürdigenden Begriffen und Beleidigungen (vgl. Abb. 5). Darüber hinaus sind vor allem die bewusste Verbreitung uninformativer oder falscher Aussagen, das Bedienen von Stereotypen und Vorurteilen durch bestimmte Begriffe und Sprachmuster sowie Verschwörungstheorien verbreitet. Der Vorwurf der Propaganda und Lügenpresse zieht sich durch nahezu alle analysierten Diskurse. Unabhängig vom Thema des journalistischen Beitrags werfen Nutzer\_innen den Medien bewusste Manipulation und interessengeleitete Berichterstattung vor.

Es gibt scheinbar ein einschlägiges Vokabular für Hassredner\_innen im Internet, das der Diskreditierung von Gegen-

Abbildung 5:  
Beispiel für eine  
Beleidigung und  
Entgegnung der  
redaktionellen  
Moderation auf der  
Facebook-Fanpage  
von „tagesschau.de“.



rede (Counter Speech) gilt. Zu den Sprachmustern, mit denen Nutzer\_innen andere Diskussionsteilnehmende „beschimpfen“, gehören Begrifflichkeiten wie SJW („Social Justice Warrior“), „Gutmensch“, „Solidaritätsbesoffene“, „Migrantenkuschler“ und „Teddybärenwerfer“.

Die Ergebnisse der Online-Diskursanalyse zeigen zudem, dass Redaktionen in den untersuchten Beispielen in ihrer Löschpraxis von öffentlicher Hassrede keine Unterscheidung zwischen der damit zum Ausdruck kommenden politischen Gesinnung der betreffenden Nutzer\_innen vornehmen. Zudem fällt auf, dass in den meisten Fällen Nutzer\_innen auf einer emotionalen Ebene diskutieren. Nur wenige Kommentare sind ausschließlich sachlich formuliert und/oder verzichten auf die Verwendung von Emojis. Hier wird die plattformabhängige Kommunikation auf Facebook anschaulich, die von Emotionen und persönlichen Äußerungen geprägt ist.

## Ansätze einer modularen Moderationsstrategie

In der Gesamtbetrachtung können in Bezug auf die Kommentarmoderation mindestens drei idealtypische Ansätze zur Eingrenzung von Hassrede und zur Stimulierung von konstruktiven Diskursen benannt werden, die sich übergreifend in eine *protektive* (Nutzer\_innen schützen), eine *disqualifizierende* (Nutzer\_innen ausschließen) und *unterstützende* Herangehensweise unterscheiden lassen. Dabei verstärkt die Bandbreite provokanter Herangehensweisen – zum Beispiel ein übermäßig raues durch die Moderation provoziertes Kommunikationsklima – das Unwesen von Hatern und Trollen nachweislich. Insgesamt lassen sich, je nach Intensität und Aktionsradius, zehn verschiedene Moderationselemente auf einem Kreuzdiagramm eintragen, das dem Vierfeldschema starkes/schwaches „Dis-Empowerment“ (bzw. Regulierung) vs. starkes/schwaches „Empowerment“ (Ermächtigung) und (eher) starker/(eher) schwacher „redaktioneller Aktivitätsgrad“ folgt (Abb. 6).

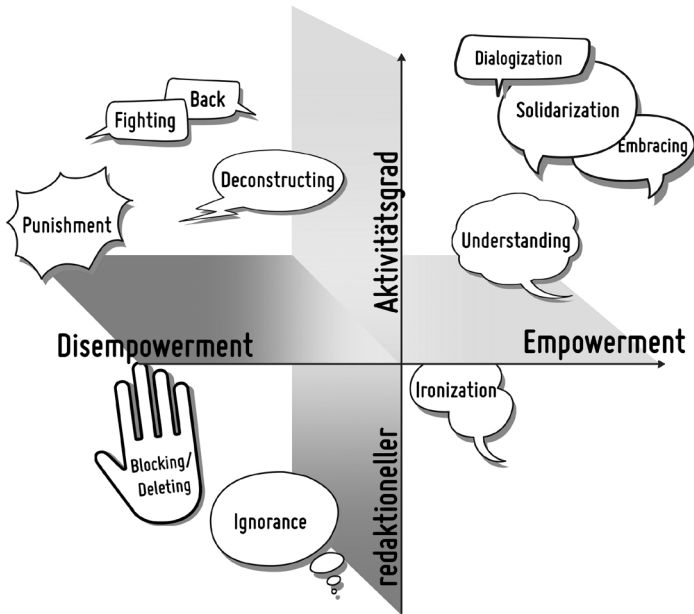


Abbildung 6: Elemente beschränkender und ermächtigender redaktioneller Moderationsstrategien (unter Verwendung von Illustrationen von iStock.com/gambarvektor).

## Beschränkende Strategie: Dis-Empowerment

- ▶ *Punishment (Bestrafung):*  
Straf-/ zivilrechtliche Maßnahmen gegen Hassrede

Diese Herangehensweise ist Teil einer Ausschlussstrategie, wenn Moderatorinnen und Moderatoren gefordert sind, Nutzerinnen und Nutzer abzumahnern oder bei entsprechenden Meinungsäußerungen strafrechtlich verfolgen zu lassen. Punishment (Bestrafung) zeichnet auf der y-Achse ein eher mittlerer redaktioneller Aktivitätsgrad, auf der x-Achse aber die höchste Form des Dis-Empowerments aus: Einer oder einem möglicherweise über die IP-Adresse aufgespürten Täterin oder Täter drohen strafrechtliche Sanktionen wie eine Freiheits- oder eine Geldstrafe nach §§185, 186, 187 StGB.

- ▶ *Counter Speech (Gegenrede):*  
Argumentatives Gegensprechen gegen Hassrede

Im Moderationsprozess ist diese Strategie eine der aufwändigsten, weil sie eine aktive Kommunikation und vor allem eine hohe Aufmerksamkeit sowie dynamisches Handeln seitens der Moderation erfordert. Dementsprechend ist die Gegenrede durch

einen vergleichsweise hohen redaktionellen Aktivitätsgrad und starkes Dis-Empowerment definiert.

- ▶ *Deconstructing (Zerlegen):*  
Hassredner und Hassrede entkräften

Von Redaktionen fordert es einen hohen Aktivitätsgrad ab, Meinungsäußerungen von Hatern und Trollen bis ins Detail zu widerlegen, damit andere Nutzer\_innen nicht von den populistischen Falschaussagen beeinflusst werden. Das Dis-Empowerment ist hier etwas schwächer ausgeprägt, da Moderator\_innen sich mit den Argumentationsketten von Störern auseinandersetzen müssen.

- ▶ *Blocking/Deleting (Ausblenden):*  
Hassredner\_innen stummschalten

Diese Strategie ist in deutschen Medienhäusern vermutlich derzeit eine der gängigsten, weil zwar am preisgünstigsten und unaufwändigsten – aber auch eine der am wenigsten nachhaltigen. Ob sie intern von Redaktionen im Schichtbetrieb erledigt oder an einen Dienstleister ausgelagert wird: Hierbei bleibt die redaktionelle Aktivität vergleichsweise gering und das Dis-Empowerment ist relativ hoch.

- ▶ *Ignorance (Aufmerksamkeitsentzug):*  
Keine Reaktion auf Hassrede

Auch Negativkommentare komplett zu ignorieren, kann ein Moderationselement sein, wenn auch ein eher passives, das ein Minimum an redaktionellem Aufwand bei gleichzeitig begrenztem Dis-Empowerment bedeutet. Statt Hassrede ganz zu ignorieren bietet es sich für die Moderation an, einzelne Störenfriede bewusst zu übersehen und nicht auf ihre Hassbeiträge zu reagieren, um ihren Argumenten die Aufmerksamkeit zu entziehen.

## Bestärkende Strategie: Empowerment

- ▶ *Ironization (Ironie/Humor):*  
Hassrede ironisieren und mit Humor begegnen

In vielen Redaktionen begegnen Moderationsteams den Negativkommentaren von Nutzer\_innen mit Ironie, Sarkasmus und

Zynismus, teils aus Frustration, teils aus Hilflosigkeit. Das kann häufig jedoch kontraproduktiv sein: Im schriftlichen Austausch stößt Ironie oft auf Unverständnis, zumal sich hierbei das Empowerment in Grenzen hält.

- ▶ *Understanding (Verständnis zeigen):*  
Hintergründe von Hassrede ermitteln

Ein weiteres Strategieelement, das einen erhöhten redaktionellen Aktivitätsgrad erfordert, ist Verständnis gegenüber Hassredner\_innen zu zeigen. Im Spektrum des Empowerments ist dieser vor allem als Argumentetransfer gemeinte Ansatz nicht so zeitaufwändig wie ein intensiver Dialog, aber schon das gezielte Nachfragen nach den Hintergründen von hasserfüllten Kommentaren und den zugrundeliegenden Beweggründen mit einer gründlich moderierten Wertediskussion erfordern Beharrlichkeit, stabile Nerven und geschultes Personal.

- ▶ *Dialogization (Vermittelnder Dialog):*  
zwischen Gegenpositionen vermitteln

Bei der Dialogisierung geht es z. B. um die Vermittlung von Gegenpositionen, die mitunter auch Hassrede enthalten können. Anders als beim „Understanding“ versuchen Moderator\_innen nicht nur korrigierend, sondern auch vermittelnd einzugreifen, um einen konstruktiven Dialog anzuregen.

- ▶ *Solidarization (Gemein machen):*  
sich mit Gegenredner\_innen solidarisieren

Sich mit einer Sache gemein machen, auch mit einer guten, gilt im Journalismus spätestens seit Hanns-Joachim Friedrichs als unschicklich, kann aber in Nutzerdiskursen bestärkende Effekte zeitigen: Mit diesem Moderationselement können sich Kommentarbereiche binnen weniger Posts zu hassredefreien Zonen entwickeln, da Redaktionen Gegenredner\_innen demonstrativ den Rücken stärken.

- ▶ *Embracing („Umarmen“, motivieren):*  
konstruktive Kommentierende zielgerichtet stärken

Das umarmende Moderationselement bezieht sich auf konstruktiv kommentierende Nutzer\_innen, indem diese in ihrer

Argumentation gezielt gestützt und damit in ihrem Kommentarverhalten gestärkt werden. Der Diskurs wird dadurch – mit hohem redaktionellen Aktivitätsgrad – versachlicht und destruktive Kommentare in den Hintergrund gerückt, da sich diese durch ihre fragwürdigen Argumentationen im besten Fall selbst disqualifizieren.

## Fazit

Redaktionen werden den gestiegenen Anforderungen an ihre Angebots- und Kommunikationskonzepte im Social Web langfristig nur gerecht werden können, wenn sie der Moderation von Nutzerdiskursen zu ihren Nachrichtenangeboten einen (noch deutlich) höheren Stellenwert einräumen als bisher. Dass dies strukturelle Folgen für die Nachrichtenbranche haben wird, angefangen bei einem organisationalen Umbau in vielen Medien-

*Ein starkes Zeichen wäre: Wenn auch traditionelle Nachrichtenorganisationen den Mehrwert von Nutzerkommentaren anerkennen.*

häusern, für den Personaltableaus neu bestellt und redaktionelle Schwerpunktbereiche überdacht werden, ist wahrscheinlich. Entsprechend erscheint eine Allokation redaktioneller Ressourcen in das Plattformgeschäft zum Ausbau der dialogischen Journalismus-Publikums-Beziehung als sinnvoll. Insbesondere junge Medienunternehmen und Start-Ups entwickeln ihr Geschäft mit dem Journalismus über den kommunikativen Austausch mit ihren Nutzerinnen und Nutzern hauptsächlich unter Zuhilfenahme sozialer Mediendienste und -plattformen (vgl. Kramp/Weichert 2017). Wenn auch traditionelle Nachrichtenorganisationen verstärkt einen Mehrwert in der Moderation von Nutzerdiskursen für die redaktionelle Leistungsfähigkeit und Qualität der eigenen Berichterstattung erkennen und entsprechende Maßnahmen ergreifen, um diesen Mehrwert systematisch zu steigern und nicht nur aus Nutzerdiskursen resultierende Probleme zu bewältigen, wäre das ein starkes Zeichen gegen die Risiken einer „digitalen Verrohung“ (Schad 2018) der gesellschaftlichen Kommunikation.

## Literatur

- AJS Landesstelle NRW/LfM NRW (Hg.): *Hate Speech. Hass im Netz. Informationen für Fachkräfte und Eltern.* [http://publikationen.medienanstalt-nrw.de/modules/pdf\\_download.php?products\\_id=442](http://publikationen.medienanstalt-nrw.de/modules/pdf_download.php?products_id=442).
- Groß, Eva/Zick, Andreas/Krause, Daniela (2012): *Von der Ungleichwertigkeit zur Ungleichheit: Gruppenbezogene Menschenfeindlichkeit.* In: *Aus Politik und Zeitgeschichte*, 62. Jg., Nr. 16-17, S. 11-18.

- Hasebrink, Uwe et al. (2017): *Zur Entwicklung der Medien in Deutschland zwischen 2013 und 2016. Wissenschaftliches Gutachten zum Medien- und Kommunikationsbericht der Bundesregierung*. Hamburg.
- Kramp, Leif/Loosen, Wiebke (2018): *The Transformation of Journalism: From Changing Newsroom Cultures to a New Communicative Orientation?* In: Hepp, Andreas/Breiter, Andreas/Hasebrink, Uwe (Hg.): *Communicative Figurations. Transforming Communications in Times of Deep Mediatization*. London, S. 205-239.
- Kramp, Leif/Weichert, Stephan (2017): *Der Millennial Code. Junge Mediennutzer verstehen – und handeln*. Leipzig.
- Kramp, Leif/Weichert, Stephan (2018): *Hass im Netz. Steuerungsmechanismen für Redaktionen*. Leipzig.
- Lilienthal, Volker et al. (2014a): *Digitaler Journalismus. Dynamik – Teilhabe – Technik*. Leipzig.
- Lilienthal, Volker et al. (2014b): *Digitaler Journalismus. Dynamik – Teilhabe – Technik. Zusammenfassung der LfM-Schriftenreihe Medienforschung (Band 74)*. Düsseldorf: LfM. [http://www.lfm-nrw.de/fileadmin/lfm-nrw/Publikationen-Download/Band-74\\_Digitaler-Journalismus\\_Kurzfassung-final.pdf](http://www.lfm-nrw.de/fileadmin/lfm-nrw/Publikationen-Download/Band-74_Digitaler-Journalismus_Kurzfassung-final.pdf).
- No Hate Speech Movement (2017): *Leitfaden für Journalist\*innen im Umgang mit Hate Speech im Netz*. [https://no-hate-speech.de/fileadmin/user\\_upload/No\\_Hate\\_Speech\\_05-10-17\\_B.pdf](https://no-hate-speech.de/fileadmin/user_upload/No_Hate_Speech_05-10-17_B.pdf).
- Schad, Gina (2018): *Digitale Verrohung? Was die Kommunikation im Netz mit unserem Mitgefühl macht*. München.
- Springer, Nina/Kümpel, Anna Sophie (2018): *User-Generated (Dis)Content. Eine Literatursynopse zur Nutzung der Kommentarfunktion auf Nachrichtenwebsites im Internet*. In: Neuberger, Christoph/Nuernbergk, Christian (Hg.): *Journalismus im Internet. Profession – Partizipation – Technisierung*. Wiesbaden, S. 241-271.
- Weichert, Stephan (2014a): *Nutzerteilhabe als Mythos*. In: *Vocer vom 16.* 10., <http://www.vocer.org/nutzerteilhabe-als-mythos/>
- Weichert, Stephan (2014b): *Demokratie als Shitstorm? Implikationen zur politischen Debattenkultur durch Social Media*. In: *Communicatio Socialis*, 47. Jg., H. 2, S. 203-213.
- Weichert, Stephan (2015): *Keine polemischen Exzesse mehr! Appell zur Entstörung der Debattenkultur in Web-Foren*. In: *epd medien*, Nr. 3 vom 11. 9., S. 3-7.

Alle Internetquellen zuletzt aufgerufen am 30.12.2018.