

Status quo und Massnahmen zu rassistischer Hassrede im Internet: Übersicht und Empfehlungen

Dr. Lea Stahel
Soziologisches Institut, Universität Zürich

August 2020



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI

Generalsekretariat EDI

Fachstelle für Rassismusbekämpfung

INHALTSVERZEICHNIS

1	ZUSAMMENFASSUNG	1
2	EINLEITUNG	4
3	BEGRIFFSERKLÄRUNGEN	5
3.1	Hassrede	5
3.2	Rassismus	8
3.3	Social Media	9
3.4	Abgrenzung der rassistischen Online-Hassrede	10
3.4.1	Von struktureller digitaler Diskriminierung	10
3.4.2	Von sonstigen direkten Online-Angriffen	10
4	HÄUFIGKEIT VON RASSISTISCHER ONLINE-HASSREDE	11
4.1	Datenlage und internationale Situation	11
4.2	Situation in der Schweiz	13
4.3	Sondergruppe Kinder und Jugendliche	14
5	DIE DIGITALE UMWELT DER RASSISTISCHEN HASSREDE	16
5.1	Organisationsgrad von Täterschaften	16
5.2	Individuelle, soziale und gesellschaftliche Einflussfaktoren von Täterschaft	17
5.3	Relevante Plattformen	19
5.4	Rassismus-begünstigende Online-Infrastruktur	22
5.4.1	Kommunikationsbedingungen	22
5.4.2	Architektur von Online-Plattformen	24
6	NEGATIVE FOLGEN VON RASSISTISCHER ONLINE-HASSREDE	29
6.1	Vorbemerkung zu digitaler Verstärkung	29
6.2	Unmittelbar Betroffene	29
6.3	Publikum und Gesamtgesellschaft	30
7	BESTEHENDE GEGENMASSNAHMEN, UMSETZUNG UND WIRKSAMKEIT	32
7.1	Gesetzgebung und Rechtsprechung zu rassistischer Hassrede	32
7.2	Social-Media-Anbieter	33
7.3	Klassische Medienunternehmen in der Schweiz	35
7.4	Forschung in der Schweiz	36
7.5	Zivilgesellschaft im In- und Ausland	37
7.5.1	Prävention und Sensibilisierung	38
7.5.2	Meldung und Beratung	41
7.5.3	Monitoring	43
7.5.4	Gegenrede	46
8	SCHWEIZER STELLEN: HERAUSFORDERUNGEN UND ANLIEGEN	50
8.1	Verwaltungsstellen	50
8.2	Öffentliche (oder öffentlich finanzierte) Beratungsstellen	51
8.3	Private Organisationen	53
9	PRÄVENTIONSMASSNAHMEN FÜR DIE SCHWEIZ	54
9.1	Kompetenzen und Zuständigkeiten	54
9.2	National ausgerichtetes Meldetool	55
9.3	Zielgruppen und Erreichbarkeit	55
9.4	Kriterien für die Beurteilung von Projekteingaben	57
10	BERATUNG UND INTERVENTIONEN IN DER SCHWEIZ	58
10.1	Notwendige Kenntnisse und Kompetenzen von Beratungsstellen	58
10.1.1	Beratung für kleinere Fälle	58
10.1.2	Komplexe Beratung und Intervention	58
10.2	Massnahmen für Bekanntmachen und Nutzbarmachen von Beratungsangebot	59
11	HANDLUNGSEMPFEHLUNGEN	60
12	VERZEICHNIS DER ABKÜRZUNGEN	62
13	LITERATURVERZEICHNIS	64

1 ZUSAMMENFASSUNG

Rassistische Hassrede im Internet: Wie soll die Schweiz in Zukunft damit umgehen? Der vorliegende Bericht, welchen die Fachstelle für Rassismusbekämpfung (FRB) in Auftrag gegeben hat, liefert dazu einen Beitrag. Er beinhaltet eine kritische Zusammenfassung der Datenlage, einen Einblick in die digitale Umwelt und begünstigenden Kommunikationsbedingungen rassistischer Online-Hassrede sowie einen Überblick zu bestehenden Gegenmassnahmen in der Schweiz und im Ausland. Schliesslich werden die Herausforderungen und Anliegen von Schweizer Verwaltungsstellen, Beratungsstellen und privaten Organisationen identifiziert und Handlungsempfehlungen für die Schweiz abgegeben.

Das Phänomen der rassistischen Hassrede im Internet ist komplex und dynamisch. Hassrede beschreibt grundsätzlich Äusserungen, die Gruppen und Mitglieder von Gruppen herabsetzen und verunglimpfen. Bei rassistischer Diskriminierung werden Menschen aufgrund ihrer ethnischen, nationalen, kulturellen oder religiösen Merkmale in angeblich naturgegebene Gruppen eingeteilt und aufgrund dieser Zugehörigkeit ungerecht behandelt, herabgesetzt oder gedemütigt. Die «rassistische Online-Hassrede» bezieht sich somit auf die Hassrede im Internet als direkte bzw. unmittelbare Form der Diskriminierung, die Menschen aufgrund ihrer Zugehörigkeit abwertet. Der Fokus des Berichts grenzt sich von indirekter, struktureller Diskriminierung im digitalen Raum oder von direkten Online-Angriffen wie Cyber-Mobbing oder Cyber-Stalking ab.

Die rassistische Hassrede hat sich zunehmend im Online-Raum ausgebreitet und löst in der Öffentlichkeit Interesse, Empörung, aber auch Ratlosigkeit aus. Historische, kulturelle und rechtliche Faktoren beeinflussen, ob eine Äusserung als Hassrede eingeschätzt wird. Je nach Kontext kann dieselbe Äusserung in einer Situation unproblematisch sein, in einer anderen jedoch als Hassrede gelten. Hassrede ist somit kontextsensitiv, weshalb keine allgemeingültige Definition der rassistischen Online-Hassrede existiert. Diese heterogenen Verständnisse sind für eine effektive Herangehensweise an rassistische Online-Hassrede zu berücksichtigen.

Wie oft tritt rassistische Online-Hassrede auf? Eine abschliessende Antwort ist aufgrund der Kontextabhängigkeit schwierig. Oft werden unterschiedliche Definitionen und Datenquellen genutzt, was die Vergleichbarkeit der erhobenen Daten erschwert. Auch kann rassistische Online-Hassrede durch eine langfristige «Normalisierung» des Phänomens entweder unterschätzt oder aufgrund seiner hohen medialen Sichtbarkeit überschätzt werden. Eine wichtige Datenquelle bilden zur Zeit die Auswertungen von Online-Inhalten (z.B. Kommentare, Videos). Der Anteil von Hassrede an der Grundgesamtheit schwankt über Länder und Plattformen zwischen 1 und 20 Prozent. Andererseits zeigen länderübergreifende Umfragen bei der Bevölkerung, dass bis zu zwei Drittel der Menschen Hassrede im Internet beobachten, wobei rassistische Hassrede den grössten Teil davon ausmacht. Von zehn Personen wurden zwischen einer bis vier Personen schon persönlich angegriffen. Kinder und Jugendliche sind unter den Beobachtenden, Täter*innen und Betroffenen überdurchschnittlich hoch vertreten. Für die Schweiz gibt es keine vergleichbaren, systematisch erhobenen Daten. Ein Vergleich von Daten zahlreicher Quellen zeigt jedoch, dass auch in der Schweiz online viel passiert, aber wenig gemeldet wird. Die offiziell erhobenen Beratungs- und Rechtsfälle liegen im tiefen zweistelligen Bereich. Doch Umfragen zu allgemeinen (nicht spezifisch rassistischen) Online-Angriffen ergeben, dass viel mehr Personen betroffen sind.

Wer sind die Täter*innen? Im vielschichtigen System digitaler Plattformen verschwimmen die Übergänge zwischen organisierten Hassgruppen, «fluiden» Netzwerken wie zum Beispiel der US-amerikanischen Alt-Right-Bewegung oder Einzelpersonen. Trotz des breiten Spektrums von «Typen» von Täter*innen legen wissenschaftliche Studien nahe, dass ihnen oft gewisse Merkmale gemeinsam sind: Sie sind meistens männlich, tendieren zu bestimmten Persönlichkeitseigenschaften wie Gewissenlosigkeit oder Impulsivität, zu bestimmten Wahrnehmungen wie Vorurteile und Misstrauen, sie informieren sich in einer umfangreichen rechtspopulistischen Medienlandschaft, bewegen sich auf Hassrede-Plattformen oder wurden selbst Opfer von Hassrede, suchen Aufmerksamkeit und Unterhaltung oder wollen ihre Gruppe verteidigen. Über ihre soziostrukturelle Einbettung (z.B. Familie, Einkommen, Bildung) ist wenig bekannt. Klar ist jedoch, dass gesellschaftliche Entwicklungen (z.B. ökonomische, politische oder kulturelle Umwälzungen) Offline- und Online-Hassrede längerfristig beeinflussen und Einzelereignissen wie beispielsweise Attentate diese situativ verstärken.

Rassistische Hassrede kann prinzipiell in allen Interneträumen vorkommen, doch manche Räume sind besonders geeignet, um Hassrede und Rassismus auszudrücken und zu verbreiten, darunter

Webseiten und Blogs, Foren (z.B. 4chan), Social-Media-Plattformen (z.B. Facebook, Twitter), Computerspiele, E-Mails und Chats (z.B. WhatsApp). Die Kommunikationsbedingungen im virtuellen Raum schwächen Hindernisse, die im realen Umfeld Hassrede beschränken. Autor*innen sind enthemmter, da eine Distanz zum Gegenüber besteht und sie sich weniger sichtbar fühlen können. Zudem sind Konflikte wahrscheinlicher, da Menschen und Meinungen aus unterschiedlichen sozialen und geografischen Räumen aufeinandertreffen. Autor*innen können ihre Ressentiments unmittelbar und impulsiv veröffentlichen, und simplifizierend-dramatisierende Inhalte bescheren ihnen umso mehr Aufmerksamkeit.

Im Kontext dieser Kommunikationsbedingungen stellen Plattform-Architekturen ihren Nutzenden vielfältige technische Verbreitungselemente (z.B. Videos, «Hashtags») bereit, womit auch rassistisches Gedankengut rasch produziert und weit verbreitet werden kann. Täter*innen von Online-Hassrede nutzen die hochdynamischen, interaktiven und partizipativen Strukturen von Social-Media-Plattformen, Messenger-Diensten und Foren optimal, um rassistische Hassrede zu verlinken, sich zu vernetzen und gegenseitig zu inspirieren. Sie sind oft auf mehr als einer Plattform unterwegs und bauen dichte Netzwerke auf, in welchen sie Hassrede und Propaganda praktisch kostenlos, rasch und weit verbreiten können. Sie können zudem auf Fake-Profilen und Bots (Computerprogramme) zurückgreifen, um eine grosse Mehrheit vorzutäuschen (metrische Manipulation), oder auf emotional-simplifizierende Fake News und Verschwörungstheorien, die viele Klicks und somit grosse Reichweiten generieren (inhaltliche Manipulation). Oft tarnen sie rassistische Inhalte, damit sie von Nutzenden und Erkennungsalgorithmen nicht erkannt werden (z.B. humoristische «Memes»). Das rassistische Gedankengut wird schliesslich subtil in die digitale, breite Öffentlichkeit getragen, indem es sich in die als legitim angesehene digitale Landschaft integriert («Mainstreaming»).

Rassistische Online-Hassrede hat nicht nur negative Folgen für die Betroffenen, sondern auch für die Zuschauenden und die Gesamtgesellschaft. Für direkt Betroffene sind die Folgen mit denjenigen von Offline-Rassismus vergleichbar; gleichzeitig sind manche digitalen Aspekte potentiell belastender. So sind Betroffene vor einem grösseren Publikum exponiert, können sich den Angriffen räumlich und zeitlich kaum entziehen und sind mit dauerhaft abrufbaren Hassnachrichten konfrontiert. Ein signifikanter Teil gibt an, mit emotionalen, aber auch körperlichen Symptomen zu kämpfen. Zudem ziehen sie sich regelmässig aus dem digitalen Leben zurück. Online-Hassrede schwächt laut Experimenten aber auch das soziale Miteinander, indem sie die Zuschauenden dazu verleitet, feindseliger zu denken und zu handeln. Dies schadet der demokratischen Diskurskultur im Internet. Schliesslich gibt es Anzeichen dafür, dass Online-Hassrede einen Nährboden für Offline-Hassverbrechen bereitet.

Um Online-Hassrede entgegenzuwirken, ergreifen Akteure wie Medienunternehmen, Social-Media-Anbieter, die Wissenschaft, rechtliche Institutionen und die Zivilgesellschaft in der Schweiz und im Ausland unterschiedliche Gegenmassnahmen. Klassische Medien können Hassrede durch ihre Berichterstattung, welche institutionalisierten Richtlinien unterliegt, und dem Management ihrer kommentierenden «Community» beeinflussen. Social-Media-Anbieter sind weniger stark reguliert, jedoch können Nutzende Hassrede direkt auf den Plattformen melden. Die Transparenz der Prüfungskriterien und die Qualität der automatisierten Prüfung müssen heute aber noch kritisch beurteilt werden. Die Forschung trägt kontinuierlich zum Verständnis des Phänomens bei – wobei Daten über den Schweizer Kontext (noch) Mangelware ist. In der Schweizer Rechtsprechung entwickelt sich die Anwendung der relevanten Rassismus-Strafnorm (Art. 261bis StGB) auf die digitale Hassrede laufend. Doch da sich Online-Hassrede oft im Raum internationaler Netzwerke abspielt, kann dies eine grosse Herausforderung für die Rechtsdurchsetzung darstellen.

Die Zivilgesellschaft im In- und Ausland setzt Gegenmassnahmen in vier zentralen Bereichen um: Prävention und Sensibilisierung, Beratung und Meldung, Monitoring sowie Gegenrede. Im Ausland sind die meisten Ansätze etabliert; in der Schweiz stecken sie noch in den Kinderschuhen. Präventions- und Sensibilisierungsansätze sollen durch digitale Medienkompetenzbildung Online-Hassrede verhindern oder ihre Wirkung eindämmen. Das Ziel ist eine digital versierte Zivilgesellschaft. Obwohl ihre längerfristige Wirksamkeit schwierig einzuschätzen ist, legen Studien kurzfristige positive Effekte nahe. Melde- und Beratungsstellen bieten bei konkreten Angriffen psychologische, soziale und/oder juristische Beratung sowie Meldetools an. Diese sollten möglichst bekannt und niedrigschwellig sein. Monitoring verfolgt und beschreibt Online-Hassrede im lokalen Kontext. Es liefert einen quantitativen Überblick zu Hassrede-Typen oder Trends und tiefere, qualitative Einblicke in die Eigenheiten von Hassrede. Quantitatives Monitoring ist informativ, die effektive Umsetzung jedoch herausfordernd. Bei Gegenrede setzen Nutzende Texte, Bilder oder Videos ein, um Hassrede mit sachlichen

Gegenargumenten zu schwächen. Normen wie Respekt und Sachlichkeit sollen gestärkt, und gegenüber Betroffenen soll Solidarität signalisiert werden. Grundsätzlich scheint Gegenrede wirksam, wobei auch hier die Mühen einer effektiven und langfristigen Umsetzung nicht zu unterschätzen sind.

In der Schweiz verfügen die relevanten Stellen über unterschiedlichste Kompetenzen und Zuständigkeiten zur Bekämpfung von Rassismus. Verwaltungsstellen sind insbesondere in der Lage zu informieren, sensibilisieren, vernetzen und finanziell zu unterstützen. Beratungsstellen können bei kleineren Fällen Unterstützung leisten und Wissen an Akteure und Institutionen vermitteln. Private Organisationen können durch ihre finanziell unabhängige Position ein grösseres Spektrum von Aufgaben wahrnehmen, darunter Gegenrede oder Engagement in grösseren Fällen mit politischer Dimension. In Bezug auf rassistische Online-Hassrede verfügen die Stellen über punktuelle Sensibilität und Kompetenz. Gleichzeitig mangelt es an Grundlagenwissen um die digitale Dimension von Rassismus und an entsprechenden Beratungskompetenzen.

Um Beratungsstellen mit den notwendigen Kenntnissen und Kompetenzen auszustatten, muss zwischen einfacher Beratung und komplexer Beratung bzw. Intervention unterschieden werden. Bei Ersterem ist insbesondere der Aufbau von inhaltlichem, technischem und juristischem Grundwissen empfohlen. Bei Zweitem werden darüberhinausgehende Kompetenzen verlangt, unter anderem eine gute Vernetzung und Kenntnisse von Mediendynamiken.

Der Bericht empfiehlt verschiedene Massnahmen zur Verhinderung und Eindämmung rassistischer Online-Hassrede in der Schweiz. Die aktuellen Massnahmen in der Schweiz werden grundsätzlich als stark ausbaufähig eingeschätzt. Breit zu sensibilisieren und Kompetenzen auf inhaltlicher, technischer und rechtlicher Ebene zu stärken, ist ein Anfang. Zukünftige Massnahmen gegen Rassismus sollten den Online-Raum jeweils mitdenken: Konkrete Projekte gegen Rassismus sollten auch nach Kriterien beurteilt werden, die auf Hassrede im digitalen Raum zugeschnitten sind. Denn diese Projekte verlangen beispielweise unterschiedliche (z.B. technische) Expertise, Kanäle für die Erreichung von Zielgruppen, Öffentlichkeitsarbeit sowie Schutzmassnahmen. Zusätzlich sollten Massnahmen spezifisch für den digitalen Raum entworfen werden. Um Massnahmen an vulnerable Zielgruppen anzupassen, müssen diese sorgfältig nach ihrer Rolle (Betroffene, Täter*innen, Zuschauende), ihrem Alter (insbesondere Kinder und Jugendliche) und sonstigen Gruppenmerkmalen (z.B. Migrant*innen, Politiker*innen, NGOs) unterschieden und ihre oft dynamisch ändernden Mediengewohnheiten identifiziert werden. Dazu sind Investitionen in Forschung und Monitoring nötig. Staatliche und nicht-staatliche Organisationen sollten enger zusammenarbeiten: Die Umgangsweise mit der komplexen Online-Hassrede verlangt einen gesamtheitlichen Lösungsansatz, der die Bedürfnisse und Verantwortlichkeiten verschiedenster Institutionen berücksichtigt und von ihren Handlungsmöglichkeiten profitiert. Das Vorgehen sollte dementsprechend möglichst holistisch sein und wissenschaftlich basierte Massnahmen berücksichtigen. Schliesslich wird zum Ausbau und zur Bewerbung von Beratung und Interventionen auf lokaler und nationaler Ebene geraten, darunter ein nationales Meldetool. Alle diese Massnahmen könnten dazu beitragen, das Ungleichgewicht zwischen der digital versierten Hassrede und den oft noch analog ausgerichteten Gegenmassnahmen längerfristig auszugleichen.

2 EINLEITUNG

«Was mich am meisten beunruhigt ist, was auch immer wir da erschaffen haben, wir haben dem Rassismus freien Auslauf auf dem Planeten erlaubt, und die Folgen für die Zivilisation und die Demokratie sind sehr, sehr unerfreulich, falls sie nicht angegangen werden.»

Berners-Lee, Begründer des World Wide Web¹

Heutzutage ist es unumgänglich, das Internet zu verstehen, um den Rassismus der Gegenwart zu begreifen. Denn Rassismus tritt nicht mehr (nur) offen als «Glatzköpfe mit Springerstiefeln» zutage – sondern im Mantel vielfältigster und getarnter Inhalte und Symbole im Internet. Dabei ist das Internet nicht per se destruktiv: Seine Kommunikationsmöglichkeiten haben Informationen und Meinungen prinzipiell vielfältiger gemacht. Man denke dabei an Social-Media-gestützte demokratische und emanzipatorische Bewegungen wie den Arabischen Frühling, Occupy Wall Street oder #metoo. Auch die Schweizer*innen kommunizieren immer mehr über Internet: Im Jahr 2019 nutzten 92 Prozent das Internet und 66 Prozent soziale Netzwerkseiten.² Verbreitet sich in diesen digitalen Räumen zunehmend rassistisches Gedankengut, kann dieses «gesellschaftsfähig» werden. Dabei sind die negativen Auswirkungen rassistischer Online-Hassrede für Betroffene, Zuschauende und die Gesellschaft mittlerweile gut belegt. Online-Hassrede muss jedoch anders angegangen werden als Offline-Hassrede. Wie das konkret geschehen soll, wird von Regierungen, Social-Media-Anbietern, zivilgesellschaftlichen Organisationen und Forschenden kontrovers diskutiert.

Heutige Informations-, Präventions- und Beratungsaufträge gegen Rassismus sollten die Online-Dimension umfassen. Denn rechtlich-repressive Massnahmen alleine kommen gegen rassistische Online-Hassrede nicht an: Diese bewegt sich oft unter der Schwelle eines strafrechtlichen Tatbestandes – wobei sie das gesellschaftliche Miteinander nichtsdestotrotz bedroht. Basierend auf diesen Einsichten zielt die Fachstelle für Rassismusbekämpfung (FRB) darauf ab, die auf Rassismus und Diskriminierung spezialisierten Beratungsstellen mit den nötigen Kompetenzen zum Umgang mit rassistischer Online-Hassrede auszustatten. Zudem sollen im Rahmen einer Schwerpunktsetzung der Finanzhilfen der FRB verschiedenste Massnahmen verstärkt unterstützt werden, darunter Projekte zur Sensibilisierung der Öffentlichkeit, Präventionsmassnahmen und Initiativen von staatlichen und privaten Organisationen sowie Interventionsangebote (z.B. in Gemeinden). Um über Grundlagen für die Erreichung dieser Ziele zu verfügen, hat die FRB den vorliegenden Bericht mandatiert. Er soll die folgenden Fragen beantworten:

1. Wie tritt rassistische Hassrede im Internet auf, wie wird sie durch das Internet gefördert, und wie wirkt sie sich auf Betroffene, Zuschauende und die Gesellschaft aus?
2. Welche Gegenmassnahmen bestehen in der Schweiz und im Ausland, wie werden sie umgesetzt und wie wirksam sind sie?
3. Mit welchen Herausforderungen in Bezug auf rassistische Online-Hassrede sind staatliche und nicht-staatliche Stellen in der Schweiz konfrontiert? Welche Anliegen haben sie?
4. Welche Präventionsmassnahmen könnten in der Schweiz von wem umgesetzt werden? Anhand welcher Kriterien können Projekteingaben beurteilt werden?
5. Welche Kenntnisse und Kompetenzen benötigen Beratungsstellen, um rassistischer Online-Hassrede zu begegnen? Wie können sie ihr Angebot bekannt und nutzbar machen?

Der Bericht, begleitet von einer Expert*innengruppe³, berücksichtigt die aktuelle Forschung, Praxisberichte, Online-Informationen relevanter Organisationen und Interviews mit Expert*innen aus dem In- und Ausland.⁴

¹ Berners-Lee, The Guardian vom 12. März 2017: [I invented the web. Here are three things we need to change to save it.](#)

² Latzer, Büchi und Festic, 2019a.

³ Eva Wiesendanger (FRB), Michele Galizia (FRB), Alma Wiecken (EKR), Nora Refaeil (Anwältin und Mediatorin, Vizepräsidentin EKR), Nina Hobi (BSV), Magdalena Küng (FRB).

⁴ Michael Bischof (Stadt Zürich, Integrationsförderung), Eveline Lüönd (Geschäftsstelle Diskriminierungsschutz Zentralschweiz), Estelle Rechsteiner (Cardis – Centro Ascolto Razzismo e Discriminazione, Ticino), Jolanda Spiess-Hegglin (#Netzcourage), Dominic Pugatsch (GRA), Stéphane Koch (intelligentzia, Genf), Claire Pershan (Renaissance Numérique, Paris), Jonathan Birdwell (Institute of Strategic Dialogue, London), Catherine Blaya

3 BEGRIFFSERKLÄRUNGEN

3.1 Hassrede

Für den Begriff der «Hassrede» (engl. hate speech; fr. discours de haine; it. discorsi d'odio) gibt es keine einheitliche, universell anerkannte Definition.⁵ Dieses Unterkapitel versucht somit darzustellen, wie die Wissenschaft, rechtliche und internationale Institutionen sowie Social-Media-Unternehmen den Begriff behandeln. Aufgrund der heterogenen Verständnisse geht der vorliegende Bericht grundsätzlich von einer breiten Definition von Hassrede aus, nämlich als Sammelbegriff für Äusserungen, die Gruppen und Mitglieder von Gruppen herabsetzen und verunglimpfen. Der Begriff der Hassrede macht die Diskussion um die Herabsetzung und Verunglimpfung von Gruppen bzw. Mitgliedern von Gruppen universeller und griffiger, obwohl das Risiko der Verwässerung besteht (im Vergleich zu spezifischeren Begriffen wie Antisemitismus, Rassismus etc.).⁶

Die *Wissenschaft* hat sich insbesondere mit der Bedeutung des Begriffs und seinen definitorischen Schwierigkeiten auseinandergesetzt. Der Begriff setzt sich zusammen aus «Hass» und «Rede». «Hass» meint ein «nachhaltiges Gefühl starker Antipathie im Sinne einer 'feindseligen Grundstimmung' als intensivste Form der Abneigung».⁷ Er zielt auf die Vernichtung des «Gehassten» ab und geht über situativ hochkochenden und abklingenden Zorn hinaus. «Rede» dagegen beschreibt einen Sprechakt, also eine «sprachliche Äusserung» durch «Wort, Schrift, Gebärde, Bild, Gestik, Mimik, die in einem spezifischen sozialen Kontext eine Bedeutung aufweisen».⁸ Betroffene von Hassrede können unterschiedliche Gruppen auf Basis von Kategorien wie Rasse, Religion, sexuelle Orientierung etc. sein – insbesondere solche, die historisch und gegenwärtig Unterdrückung und Stigmatisierung erleben. Dies bedeutet jedoch nicht, dass diejenigen vor Hetze geschützt wären, die aktuell eine (z.B. politische) Machtposition einnehmen.⁹ Dabei können Personen auch aufgrund mehrerer Merkmale, z.B. Hautfarbe und Geschlecht, angegriffen werden («Intersektionalität»)¹⁰.

Hassrede ist unter anderem schwierig zu definieren, weil das Phänomen kontextsensitiv ist und sich unterschiedlich manifestieren kann. Abhängig vom Kontext können vielerlei Inhalte in die unterschiedlichsten Definitionen passen. Dieser Kontext findet jedoch selten Eingang in Definitionen.¹¹ Die prototypische Ausblendung des Kontexts zeigt sich im Versuch, Hassrede in Online-Inhalten¹² mithilfe von automatisierten Wörterlisten erkennen zu wollen. Solche Analysen werfen ein weiteres Problem auf: Wer definiert eine Aussage als Hassrede? Sind es die Betroffenen? Die Täter*innen? Oder unbeteiligte Aussenstehende (z.B. Forschende)? Speziell im digitalen Raum bewegen sich Menschen in unterschiedlichen Kontexten, was gemeinsame Interpretationen erschwert. Eine weitere Schwierigkeit ist, dass manche definitorischen Komponenten (z.B. eine «Absicht zur Schädigung») kaum eindeutig bestimmbar oder wegen ihrer Tarnung nicht erkennbar sind (siehe zu Tarnung Kap. 5.4.2). Letzteres kann nur von denjenigen als Hassrede erkannt werden, die in die Tarnung eingeweiht sind. Letztlich muss der Begriff auch nicht auf das verweisen, was landläufig angenommen wird: «Verbreiter*innen von Online-Hassrede», sei es durch kommentieren, teilen oder « liken », müssen die Emotion « Hass » prinzipiell nicht selbst verspüren – auch wenn ihre Äusserungen einen hassend-emotionalen Ton haben. Auch ist nicht alle Hassrede strafrechtlich relevant. Will man Hassrede also erkennen, greift die Regel « I know it when I see it » zu kurz.¹³ Jurist*innen, internationale Gremien und Social-Media-Anbieter¹⁴ haben unterschiedliche Definitionen vorgeschlagen:

(Universität Nice Côte; LASALE bei HEP, Vaud - Lausanne), Hansi Voigt (dasnetz.ch; bajour), Johannes Baldauf (Experte für Radikalisierung und Hassrede im Internet, ehemals Amadeu Antonio Stiftung, aktuell Facebook).

⁵ Naguib 2014: S. 13 und S. 17.

⁶ Rafael und Ritzmann 2018: S. 13.

⁷ Naguib 2014: S. 82.

⁸ Naguib 2014: S. 83.

⁹ Naguib 2014: S. 84.

¹⁰ Intersektionalität beschreibt die Überschneidung von Diskriminierungsformen, die dadurch eigenständige Diskriminierungserfahrungen hervorbringen. Beispielsweise kann die mit einer sexistischen Intention verknüpfte Handlung durch eine rassistische Begründung instrumentalisiert bzw. begründet werden (Naguib 2014: S. 26).

¹¹ Sellars 2016: S. 14 und S. 32.

¹² Online-Inhalte beschreiben Daten, die in digitaler Form her- und bereitgestellt werden inklusive Video- und Audioinhalte, digitale Texte, Kommentare, Bilder, Fotos.

¹³ Sellars 2016: S. 14.

¹⁴ Für eine Definition von Social Media, siehe Kap. 3.3.

Aus *rechtlicher Perspektive* ist Hassrede kein feststehender, juristisch definierter Begriff – auch in der Schweizer Behördenpraxis nicht.¹⁵ Die fehlende Definition sei «symptomatisch für den sich derzeit politisierten und in Entwicklung befindenden Diskurs».¹⁶ Obwohl alle Hassrede abwertend – und abhängig von der Perspektive moralisch verwerflich – ist, gilt nicht jede Hassrede als gefährlich genug, als dass sie rechtlich unzulässig wäre.¹⁷ Denn grundsätzlich gilt das Recht zur freien Meinungsäusserung (z.B. Art. 19 der Allgemeinen Menschenrechtserklärung), welches jedoch nicht uneingeschränkt ist. Im konkreten schweizerischen Recht können Hassrede-Handlungen unter bestimmten Straftatbestimmungen erfasst werden.¹⁸ Relevant sind die strafrechtlichen Verbote zur Rassendiskriminierung, wobei rechtlich gegen Hassrede in Bezug auf rassifizierte, ethnische oder religiöse Gruppen vorgegangen werden kann. Das Verbot der Diskriminierung aufgrund der sexuellen Orientierung ist seit dem 1. Juli 2020 in Kraft (Schweizerisches Strafgesetzbuch, Art. 261bis StGB). Die konkreten Straftatbestände kriminalisieren den Aufruf zu Hass oder zur Diskriminierung, die Verbreitung rassistischer Ideologie, Vorbereitungshandlungen im Bereich der rassistischen Propaganda sowie die diskriminierende Herabsetzung sowie das Leugnen, Rechtfertigen und Verharmlosen von Völkermord und Verbrechen gegen die Menschlichkeit.

Trotz der definitorischen Schwierigkeiten versuchen Länder zunehmend, Hassrede zu regulieren. Ihre Definitionen unterscheiden sich dahingehend, dass sie verschiedene Aspekte unterschiedlich stark gewichten. Dazu gehört beispielsweise, dass nur öffentliche Hassrede oder auch private Rede verboten sein soll; ob der Wahrheitsgehalt der Rede relevant ist; und ob der Schaden beabsichtigt sein muss oder auch fahrlässig sein kann.¹⁹ Dies äussert sich schliesslich in unterschiedlich «toleranten» Gesetzgebungen: während sie beispielsweise in den USA relativ permissiv ist, ist sie in Deutschland relativ restriktiv.²⁰

Ohne hier auf einzelne Länder einzugehen, gibt es Versuche verschiedenster *internationaler Gremien*, internationale Standards zu verabschieden.²¹ Diese Definitionen stammen teilweise aus dem Zeitalter vor Social Media; nichtsdestotrotz sind sie nicht nur auf die analogen, sondern auch auf die digitalen Formen von Hassrede gleichermaßen anwendbar. Eine breite Definition hat der Ministerrat des Europarats vorgeschlagen, worauf auch heute in wissenschaftlichen und nicht-wissenschaftlichen Publikationen oft Bezug genommen wird:

«Jegliche Ausdrucksformen, welche Rassenhass, Fremdenfeindlichkeit, Antisemitismus oder andere Formen von Hass, die auf Intoleranz gründen, propagieren, dazu anstiften, sie fördern oder rechtfertigen, einschliesslich der Intoleranz, die sich in Form eines aggressiven Nationalismus und Ethnozentrismus, einer Diskriminierung und Feindseligkeit gegenüber Minderheiten, Einwanderern und der Einwanderung entstammenden Personen ausdrückt.»²²

Neben nationalistischer, rassistischer oder religiöser Hetze sind hier auch andere mögliche Gruppen mitgedacht. Die Europäische Kommission gegen Rassismus und Intoleranz (ECRI) hingegen spezifiziert die Ausdrucksformen und listet noch mehr konkrete Gruppen auf:

«(...) das Befürworten und Fördern von oder Aufstacheln zu jeglicher Form von Verunglimpfung, Hass oder Herabwürdigung einer Person oder Personengruppe, ebenso wie jegliche Belästigung, Beleidigung, negative Stereotypisierung, Stigmatisierung oder Bedrohung einer Person oder Personengruppe und die Rechtfertigung der genannten Ausdrucksformen – aufgrund einer nicht vollständigen Liste von persönlichen Eigenschaften und Statusmerkmalen, darunter 'Rasse', Hautfarbe, Sprache, Religion oder Überzeugung, Staatsangehörigkeit oder

¹⁵ Naguib 2014: S. 90.

¹⁶ Naguib 2014: S. 91, wobei der Autor auf den Rabat Plan of action (2012) verweist. Siehe: United Nations Human Rights. [Rabat Plan of action on the prohibition of advocacy of national, racial or religious hatred that constitute incitement to discrimination, hostility or violence](#). Rabat (4. und 5. Oktober 2012).

¹⁷ George 2015: S. 1.

¹⁸ Naguib 2014.

¹⁹ Für einen detaillierteren Einblick in die Ähnlichkeiten und Unterschiede von Gesetzgebungen unterschiedlicher Länder, siehe Sellars 2016: S. 18.

²⁰ Hawdon, Oksanen und Räsänen 2017: S. 258.

²¹ Auch Hassrede im Internet wurde thematisiert, siehe Gagliardone et al. 2015.

²² [Council of Europe's Committee of Ministers' Recommendation 97\(20\)](#) vom 30. Oktober 1997.

nationale oder ethnische Herkunft sowie Abstammung, Alter, Behinderung, biologisches oder soziales Geschlecht, Geschlechtsidentität und sexuelle Orientierung.»²³

Im Internationalen Übereinkommen zur Beseitigung jeder Form von Rassendiskriminierung (CERD)²⁴ werden kontextuelle Aspekte aufgeworfen, die bei der Entscheidung über die Strafbarkeit einer spezifischen Äusserung von Relevanz sein können. Berücksichtigt wird der Inhalt und die Form der Äusserung (Direktheit, Art der Verbreitung, Stil), das bestehende wirtschaftliche, soziale und politische Klima (bestehende Diskriminierungsmuster), die Position des Redenden (z.B. Meinungsführer*in), die Reichweite der Rede (Art des Publikums und Kommunikationskanal) und das Ziel der Rede (z.B. sollen Reden zum Schutz von Menschenrechten nicht strafbar sein).

Seit kurzem haben auch *Social-Media-Anbieter* Definitionen entwickelt, auf deren Grundlage nutzergenerierte Inhalte auf ihren Plattformen moderiert werden. Die Video-Plattform YouTube definiert in ihren Richtlinien (2020) «Hasserfüllte Inhalte» als

«(...) Inhalte, die Gewalt gegen Einzelpersonen oder Gruppen aufgrund von ethnischer Zugehörigkeit, Religion, Behinderung, Geschlecht, Alter, Nationalität, Veteranenstatus, Gesellschaftsklasse, sexueller Orientierung oder Geschlechtsidentität billigen oder aktiv befürworten, oder die basierend auf diesen Merkmalen zu Hass gegen Einzelpersonen oder Gruppen aufrufen.»²⁵

Ähnlich definiert Facebook in seinen Gemeinschaftsstandards (2020) «Hassrede» als einen

«direkten Angriff auf Personen aufgrund geschützter Eigenschaften: ethnische Zugehörigkeit, nationale Herkunft, religiöse Zugehörigkeit, sexuelle Orientierung, Kaste, Geschlecht, Geschlechtsidentität, ernsthafte Erkrankung oder schwere Behinderung. Auch der Einwanderungsstatus ist in gewissem Umfang eine geschützte Eigenschaft. Wir definieren Angriff als gewalttätige oder entmenslichende Sprache, Aussagen über Minderwertigkeit oder Aufrufe, Personen auszuschliessen oder zu isolieren.»²⁶

Abschliessend fordert auch Twitter in seiner «Richtlinie zu Hass schürendem Verhalten» (2020) seine Nutzenden auf:

«Du darfst keine Gewalt gegen andere Personen fördern, sie direkt angreifen oder ihnen drohen, wenn diese Äusserungen aufgrund von Abstammung, ethnischer Zugehörigkeit, nationaler Herkunft, sexueller Orientierung, Geschlecht, Geschlechtsidentität, religiöser Zugehörigkeit, Alter, Behinderung oder ernster Krankheit erfolgen.»²⁷

Um die gemeinsamen Nenner in den rechtlichen und unternehmensbezogenen Definitionen auszumachen, hat der US-amerikanische Forscher A. Sellars einen integrierenden Definitionsversuch gewagt.²⁸ Reden, die all diesen Charakteristiken entsprächen, würden mit aller Wahrscheinlichkeit von den meisten Ländern und Social-Media-Anbietern als Hassrede eingeordnet:

- zielt auf eine Gruppe oder auf Individuen als Mitglieder einer Gruppe ab
- drückt auf die eine oder andere Art Hass aus
- Sprecher*innen beabsichtigen zu schaden
- verursacht Schaden (physische Gewalt oder auch strukturell, z.B. soziale Beziehungen)
- hetzt auf (z.B. Aufstachelung zu Gewalt)
- ist entweder öffentlich oder an ein Mitglied der anvisierten Gruppe gerichtet
- der sozial-politische Kontext macht gewalttätige Antworten wahrscheinlich

²³ [ECRI General Policy Recommendation N°15 on Combating Hate Speech](#) vom 8. Dezember 2015.

²⁴ [CERD, Allgemeine Bemerkungen Nr. 35 betreffend «Combating racist hate speech»](#) vom 26. September 2013.

²⁵ Youtube. [Richtlinien und Sicherheit](#). Pfad: Youtube > About > Richtlinien und Sicherheit.

²⁶ Facebook. [Hassrede](#). Pfad: Facebook Gemeinschaftsstandards > Anstössige Inhalte > Hassrede.

²⁷ Twitter. [Richtlinie zu hassschürendem Verhalten](#). Pfad: Twitter > Hilfe-Center > Twitter Regeln und Richtlinien
Richtlinie zu hassschürendem Verhalten.

²⁸ Sellars 2016: S. 24-31.

- hat keinen «legitimen» Zweck über den Hass hinaus

Zusammenfassend ist zu bemerken, dass jede der bestehenden Definitionen ihre eigenen Priorisierungen und fehlenden Elemente aufweist, wobei signifikante Überlappungen bestehen.²⁹ Daher könnte man zwar von einer «definitorischen Unschärfe» sprechen, die das bestehende Wissen über Hassrede und deren Gegenmassnahmen trübt.³⁰ Jedoch bleibt festzuhalten, dass die Kontextabhängigkeit von Hassrede eine allgemeingültige Definition und Regulierung prinzipiell kaum zulässt. Die bestehenden heterogenen Verständnisse liegen also in der grundsätzlichen Natur von Hassrede. Um Hassrede effektiv entgegenzutreten zu können, sollte diese Tatsache mitberücksichtigt werden.

3.2 Rassismus

Der vorliegende Bericht legt seinen Fokus auf Hassrede im Bereich Rassismus. Der Begriff «Rassismus» ist aufgrund seiner politischen und rechtlichen Dimensionen komplex und wird daher an dieser Stelle nur kurz erläutert. Rassismus beschreibt laut der FRB

«eine Ideologie, die Menschen aufgrund ihrer ethnischen, nationalen oder religiösen Zugehörigkeit in angeblich naturgegebene Gruppen (sogenannte 'Rassen') einteilt und diese hierarchisiert. Damit werden Menschen nicht als Individuen, sondern als Mitglieder solcher pseudo-natürlicher Gruppen mit kollektiven, als unveränderbar betrachteten Eigenschaften beurteilt und behandelt.»³¹

«Rasse» ist somit ein soziales Konstrukt, welches anhand äusserlicher Merkmale und/oder angenommener kultureller, religiöser oder herkunftsmässiger Gruppendifferenzen erbaut wird. Die Zugehörigkeit zu diesen Gruppen wird in der Folge mit bestehenden sozioökonomischen oder bildungsbezogenen Ungleichheiten in der Gesellschaft in Verbindung gesetzt und diese dementsprechend biologisch begründet. Manifestieren kann sich Rassismus auf interpersonaler, struktureller, institutioneller und kultureller Ebene.³²

Die rassistische Hassrede reflektiert primär eine interpersonale, direkte bzw. unmittelbare Form der «rassistischen Diskriminierung», d.h. jede Praxis, «die Menschen aufgrund physiognomischer Merkmale, ethnischer Herkunft, kultureller Merkmale (Sprache, Name) und/oder (realer oder angenommener) religiöser Zugehörigkeit Rechte vorenthält, sie ungerecht oder intolerant behandelt, demütigt, bedroht oder an Leib und Leben gefährdet».³³ Verschiedene Personen in vergleichbaren Situationen werden dabei «aus nicht zulässigen Gründen weniger vorteilhaft behandelt».³⁴ Diese Behandlung muss jedoch nicht zwingend ideologisch begründet sein. Dies grenzt die Diskriminierung somit vom Konzept des Rassismus ab: Die möglicherweise dahinterliegenden rassistischen Einstellungen können stereotyp, müssen jedoch nicht per se ideologisch begründet sein.

Der vorliegende Bericht wird die indirekte oder mittelbare Diskriminierung, d.h. «Praktiken, Politiken, oder gesetzliche Grundlagen (...), welche trotz scheinbarer Neutralität im Ergebnis zu einer nicht zulässigen Ungleichbehandlung führen»³⁵, grösstenteils aussen vorlassen.³⁶ Mit «Online-Rassismus» ist im vorliegenden Bericht primär die direkte, rassistische Hassrede im Internet («rassistische Online-Hassrede») gemeint.

²⁹ Sellars 2016: S. 24 und 31.

³⁰ Siegel 2020: S. 5.

³¹ Fachstelle für Rassismusbekämpfung 2019: S. 10.

³² Bliuc et al. 2018: S. 75; Krieger 1990.

³³ Fachstelle für Rassismusbekämpfung 2019: S. 11.

³⁴ Fachstelle für Rassismusbekämpfung 2019: S. 11.

³⁵ Fachstelle für Rassismusbekämpfung 2019: S. 12.

³⁶ Wie sie sich im Internet manifestieren kann, zeigt eine Kurzübersicht in Kap. 3.4.1.

3.3 Social Media

Der vorliegende Bericht fokussiert auf rassistische Hassrede im Internet, insbesondere auf Social Media. Social Media beschreiben «mehr oder weniger offene, interaktive und partizipative Plattformen, welche es Nutzenden ermöglichen zu kommunizieren, Beziehungen aufzubauen und diese zu pflegen. Zudem können Nutzende sozialer Netzwerke mit geringem Aufwand Informationen und Inhalte von und über Dritte austauschen sowie selbst oder in Gemeinschaft mit anderen Inhalte generieren und diese anderen Nutzenden zugänglich machen».³⁷ Solche Inhalte können Texte, Bilder, Fotos, Symbole, Videos und Musik, jedoch auch Hyperlinks und Downloads sein. Social Media sind in der Schweiz verbreitet: Während Schweizer*innen im Jahr 2019 durchschnittlich 3 Stunden und 33 Minuten pro Tag im Internet verbrachten, waren über 70 Prozent täglich oder mehrmals täglich auf Social Media unterwegs. Bei den 14- bis 19-Jährigen steigt dieser Wert auf 99 Prozent.³⁸

Das Angebot von Social-Media-Plattformen ist mehr als umfangreich. Wie man sie nutzen kann, hängt von der Plattform-Architektur und den Vernetzungsmöglichkeiten ab.³⁹ Im Vergleich zur Kommunikation in Zeiten vor Social Media stellen Expert*innen bei diesen Diensten eine Verschiebung von Grenzen fest. Die Rollen der Produzent*innen und Konsument*innen von Inhalten verschwimmen: Heute sind Laien nicht mehr nur zum passiven Konsum von Inhalten offizieller Publikationsmedien «verdammt». Sie können eigenständig Inhalte produzieren und diese auf Social-Media-Plattformen verbreiten. Dies macht sie zu «Prosument*innen», also sowohl Produzent*innen als auch Konsument*innen. Die Kanäle für private und öffentliche Kommunikation sind zudem weniger getrennt; sie findet oft auf derselben Plattform statt. Daten sind nicht mehr nur lokal abgelegt, sondern werden ins Internet eingespeist und können über den Moment der Veröffentlichung hinaus jederzeit «digital ausgegraben» werden. In diesem Sinne verschmilzt sozusagen die Vergangenheit mit der Gegenwart und Zukunft. Auch soziale und räumliche Grenzen werden abgebaut, da die globale Kommunikation über Gruppenzugehörigkeit und nationale Grenzen hinweg möglich gemacht wird.⁴⁰

Da sich Social-Media-Plattformen dynamisch entwickeln, ist ihre klare Einordnung schwierig. Manche Plattformen sind darauf ausgelegt, Wissen und Informationen zu liefern (z.B. Wikis, Blogs, Special-Interest-Foren), sich zu unterhalten oder virtuelle Welten zu erfahren (z.B. Computerspiele, YouTube), Beziehungen zu pflegen (z.B. Facebook, Dating-Plattformen) oder die eigene Identität und Reputation zu managen (z.B. Blogs, LinkedIn). Im Kontext von Online-Rassismus sind insbesondere, wenn auch nicht ausschliesslich, die folgenden Plattform-Arten⁴¹ bedeutend:

- Social Media (z.B. Facebook, LinkedIn)
- Mikroblogs (z.B. Gab, Twitter, Weibo)
- Foren (z.B. 4chan, 8chan, Reddit)
- audiovisuelle Seiten (z.B. YouTube, Instagram, TikTok)
- Social-News-Plattformen, auf denen Nutzende Inhalte einstellen und/oder anbieten können (z.B. BuzzFeed)
- Messenger-Plattformen (z.B. Threema, WhatsApp, Telegram)
- Plattformen zur Partnersuche (z.B. Tinder, Parship)

Darüber hinaus können Social-Media-Elemente auch in altbekannten Webseiten integriert sein. Ein Beispiel sind Online-Zeitungen, die selbst keine Social-Media-Anbieter sind, jedoch interaktive Kommentarspalten oder Share/Teilen-Buttons anbieten. Dementsprechend sei hier anzumerken, dass Rassismus kein reines «Social-Media-Problem» ist, sondern sich vielmehr in einem System von verlinkten Social-Media-Plattformen und altbekannten Webseiten einbettet. Alle können Rassismus theoretisch fördern, sofern sie dafür nützliche Kommunikationsbedingungen und Plattform-Architekturen anbieten (siehe mehr in Kap. 5.4.1 und 5.4.2).

³⁷ Der Bundesrat 2013: S. 7.

³⁸ Latzer, Büchi und Festic 2019b: S. 8.

³⁹ Der Bundesrat 2013: S. 7.

⁴⁰ Boyd 2010: S. 10.

⁴¹ Für weitere allgemeine Arten von Plattformen, siehe Der Bundesrat 2017: S. 7.

3.4 Abgrenzung der rassistischen Online-Hassrede

3.4.1 Von struktureller digitaler Diskriminierung

Der Begriff «Online-Rassismus», in diesem Bericht äquivalent zu «rassistische Online-Hassrede» gebraucht, kann andernorts nichtsdestotrotz breiter verstanden werden. Rassismus kann nämlich auch in die grundlegenden Strukturen des Internets eingebaut sein. In der Vergangenheit wurde diese strukturelle Diskriminierung beispielsweise diskutiert in Bezug auf die Hightech-Industrie (z.B. ethnische Verzerrung in der Führungszusammensetzung). Auch gewisse aus ihr entstehende Plattformen und Algorithmen haben diskriminierendes Potential. Dazu gehören Plattform-Designs (z.B. rassistische Kategorien kodiert in Benutzeroberflächen und Dropdownmenüs), künstliche Intelligenz (z.B. Gesichtserkennung) und prädiktive Analysen (z.B. Predictive Policing⁴²). Strukturelle digitale Ungleichheit kann auch bei der Zusammensetzung von Internetnutzenden beobachtet werden. Dies kann insbesondere beobachtet werden bei Computerspielen (z.B. Überrepräsentation weisser Männer bei Designer*innen und Spielenden).⁴³ Diese strukturelle Diskriminierung wird hier nicht vertieft. Sie ist jedoch – obwohl weniger sichtbar als die explizite, bewusste Hassrede – mindestens genauso tiefgreifend, da sie unbewusste rassistische Handlungen und Gewohnheiten fördern kann.

3.4.2 Von sonstigen direkten Online-Angriffen

Zusätzlich gilt es, rassistische Online-Hassrede (sowie allgemeine Online-Hassrede) von ähnlichen direkten Online-Angriffen abzugrenzen. Auch diese Angriffe verletzen Umgangsformen, überschreiten Grenzen oder sind gar gewalttätig – aber nicht primär diskriminierend. Beispiele sind Cybermobbing (z.B. Belästigung über digitale Medien, insbesondere bekannt aus Schulen), Digilantismus (Selbstjustiz über das Internet, meist Denunzierung oder öffentliches Beschämen), Cyberstalking (das Verfolgen und Belästigen von Individuen, Gruppen, und Organisationen durch Diffamierung oder Identitätsdiebstahl), und Trolling (meist unterschwelliges Provozieren und Stören von Online-Debatten).⁴⁴ Von Online-Rassismus unterscheiden sich diese Phänomene in ihrer «Reinform» insbesondere darin, dass im Prinzip alle Menschen und Institutionen in gleichem Masse betroffen sein können. Mögliche Abwertungen sind prinzipiell individuell zu verstehen (hier stehen öfter persönliche Konflikte oder das konkrete Verhalten der Zielperson im Vordergrund). Dabei kann eine grössere Brandbreite von inhaltlichen Angriffen und Absichten genutzt werden. Online-Rassismus dagegen richtet sich primär gegen Personen, weil sie einer bestimmten Gruppe zugeordnet werden. Angriffe sind somit auf die Gruppenmitgliedschaft bezogen. In der Praxis sind die Grenzen zwischen Online-Rassismus und sonstigen direkten Online-Angriffen jedoch fließend: Ein Schüler kann im Klassenchat gemobbt werden, wobei seine Hautfarbe abgewertet wird. Oder «Trolle» verbreiten rassistische Inhalte, erklären dabei jedoch, apolitisch zu sein und damit «nur provozieren» zu wollen.⁴⁵

⁴² Beschreibt die «vorhersagende Polizeiarbeit», also «die Datenanalyse zur Wahrscheinlichkeitsschätzung zukünftiger Straftaten, um eine möglichst konkrete Unterstützung für die polizeiliche Einsatzplanung zu erhalten». Basis ist die Verarbeitung großer Datenmengen. Siehe: Öffentliche IT (2015). [Vorhersagende Polizeiarbeit](#) (März 2015). Pfad: Öffentliche IT > Trendschau > Vorhersagende Polizeiarbeit. / Siehe auch Shapiro 2019.

⁴³ Z.B. Banaszczuk 2019: S. 5; Daniels 2012: S. 700.

⁴⁴ Für eine detailliertere Beschreibung von Online-Angriffen, siehe Fortuna und Nunes 2018: S. 8.

⁴⁵ Marwick und Lewis 2017: S. 4.

4 HÄUFIGKEIT VON RASSISTISCHER ONLINE-HASSREDE

Social-Media-Plattformen werden in der Öffentlichkeit nicht selten als Bastionen des Hasses beschrieben.⁴⁶ Journalist*innen, Forschende und Politiker*innen begründen ihren Fokus auf das Thema oft damit, dass es über die Zeit zugenommen habe. Systematische empirische Belege werden bei solchen Aussagen jedoch selten präsentiert. Dieses Kapitel diskutiert daher, wieso es so schwierig ist, die Häufigkeit von Online-Rassismus zu erfassen, und präsentiert vorliegende Zahlen.

4.1 Datenlage und internationale Situation

Wie oft Online-Rassismus tatsächlich auftritt, ist aus den folgenden Gründen schwierig festzustellen:

- *Fehlende einheitliche Definition:* Das zu messende Phänomen wird in Studien unterschiedlich benannt und definiert. Es reicht von Straftatbeständen wie Volksverhetzung zu der undefinierten, aber häufig von der Meinungsfreiheit gedeckten Hassrede. Dies beeinflusst die Zahl der gemessenen Fälle, welche dadurch nur schwer vergleichbar sind. Dies ist nicht prinzipiell problematisch, da die Kontextabhängigkeit von Hassrede eine allgemeingültige Definition ausschliesst. Beim systematischen Vergleich von Studienergebnissen sollte jedoch der Kontext berücksichtigt werden.
- *Unterschiedliche Interpretationen von Online-Hassrede:* Selbst wenn mit dem gleichen Begriff («Online-Hassrede») gearbeitet wird, sind Vergleiche schwierig, denn der Begriff unterliegt zu einem gewissen Grad einer subjektiven Erfahrung. Dies manifestiert sich darin, dass befragte Personen trotz gleicher Erfahrungen unterschiedlich oft angeben, «Hassrede» erfahren zu haben, oder dass Kodierer*innen von Online-Inhalten sich uneinig sind, welche Inhalte als Hassrede gelten. Länder, Gruppen und Individuen unterscheiden sich dementsprechend stark darin, ob und wie intensiv sie Online-Inhalte als Hassrede wahrnehmen.⁴⁷ Die Übereinstimmung ist bei extremeren Inhalten höher. Sie sinkt dagegen bei mittelstarken Inhalten, was am häufigeren Vorliegen von getarnten Inhalten, Sarkasmus und Humor liegen könnte.⁴⁸ Darüber hinaus scheinen manche Gruppen gewisse Reden rascher als Hassrede einzuordnen als andere. Umfragen bei US-Amerikaner*innen finden solche höheren Sensibilitäten beispielsweise bei Frauen, «politisch Moderaten und Liberalen» und Mitgliedern von historisch marginalisierten oder benachteiligten Gruppen wie Afroamerikaner*innen.⁴⁹
- *Unterschiedliche Datenquellen:* Studienergebnisse basieren auf Umfragedaten, plattform-spezifischen oder -übergreifenden Online-Inhalten sowie gemeldeten Online-Beiträgen aus dem gesamten Internet. Dieser Flickenteppich erfordert, bei Generalisierungen die Limitationen jeder Datengrundlage zu berücksichtigen.
- *Überschätzung durch hohe Sichtbarkeit:* Aufgrund der hohen medialen Aufmerksamkeit für Online-Hassrede, insbesondere für kollektive Hasskampagnen, ist es vorstellbar, dass die breite Gesellschaft ihre Auftretenswahrscheinlichkeit überschätzt. Belastbare Ergebnisse gibt es für diese Annahme nicht, ähnliche Prozesse sind jedoch aus der Forschung zur Wahrnehmung von Terrorismus bekannt. Wurde Hassrede früher am privaten Stammtisch geäußert, wird sie heute in eine digitale Öffentlichkeit getragen und ist somit sichtbarer – ob sie jedoch öfter auftritt, ist eine andere Frage.
- *Langfristige Wahrnehmungsveränderung:* Sollte sich Rassismus im Internet stark verbreiten, kann eine «Normalisierung» eintreten. Rassismus wird nicht mehr als solcher wahrgenommen.

⁴⁶ Lizza, New Yorker vom 19. Oktober 2016: [Twitter's anti-semitism problem](#).

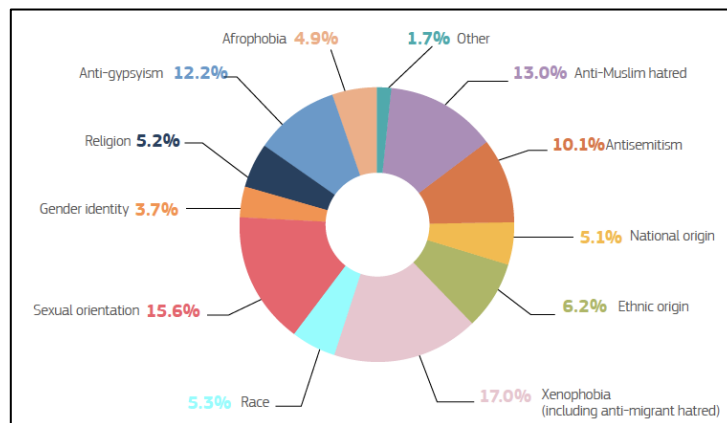
⁴⁷ Dies zeigen 18'125 Beurteilungen von Online-Hassrede-Kommentaren durch Internetnutzende in 50 Ländern. Siehe: Salminen et al. 2018.

⁴⁸ Salminen et al. 2019: S. 213 und S. 216.

⁴⁹ Kenski, Coe und Rains 2017: S. 809; Costello et al. 2019; Tynes und Markoe 2010.

Personen können in Umfragen aufgrund dieser «Abstumpfung» niedrigere Werte angeben – obwohl Rassismus unter Umständen kontinuierlich zunimmt.⁵⁰

Wie oft rassistische Online-Hassrede auftritt, wird insbesondere anhand von Inhalten auf Online-Plattformen und Umfragen in diversen Ländern festgestellt. Dabei wird meist Hassrede im Allgemeinen erhoben und nur selten rassistische Hassrede im Speziellen. Da jedoch Rassismus wiederholt als häufigste Form von Hassrede festgestellt wird (siehe Grafik 1), ist es naheliegend, sich in der folgenden Situationsbeschreibung auf Daten zu allgemeiner und rassistischer Hassrede zu beziehen.⁵¹



Grafik 1. Typen-Einteilung von 4392 Online-Beiträgen, die Social-Media-Anbietern im Rahmen des *Code of Conduct* von Nutzer*innen und «Trusted Flaggen» zwischen dem 5. November und 14. Dezember 2018 gemeldet wurden (mehr zum Code of Conduct in Kap. 7.2).

Inhalte auf Online-Plattformen. Online-Hassrede wird auf Plattformen automatisiert erkannt und ausgewertet. Beispielsweise haben Siegel et al.⁵² in über einer Milliarde Tweets amerikanischer Twitter-User in etwa ein Prozent Hassrede gefunden. Ähnliche Zahlen (0.7%) zeigen sich auf äthiopischen Facebook-Seiten. In der italienischen Twitter-Sphäre dagegen wurden ungewöhnlich hohe Werte (15%) festgestellt; davon waren zehn Prozent gegen Immigrant*innen, sechs Prozent gegen Muslim*innen, und sechs Prozent gegen Juden und Jüdinnen gerichtet.⁵³ Diese Verhaltensdaten legen nahe, dass sich der Anteil von Hassrede an der Gesamtheit der Online-Inhalte über Länder und Plattformen hinweg unterscheidet. Hassrede scheint sich zudem auf bestimmte Online-Gemeinschaften zu konzentrieren. Auf dem Thread /pol/ («politically incorrect») des populären Internetforums 4chan beinhalteten beispielsweise 12 Prozent der Beiträge Hassrede.⁵⁴ Darüber, wie oft diese Beiträge angeschaut werden, sagen diese Zahlen jedoch nichts aus. Zudem sind direkte Vergleiche zwischen Ländern oder Plattformen schwierig, da sich deren Moderations- und Löschraxis unterscheidet. Wie viele Inhalte also öffentlich sichtbar sind und somit ausgewertet werden können, sagt nur bedingt etwas darüber aus, wie viel rassistische Hassrede ursprünglich zur Verbreitung hochgeladen wurde. Neben Online-Inhalten werden schliesslich auch Social-Media-Profile ausgewertet: Guhl et al.⁵⁵ schätzen auf dieser Basis, dass 15'000 bis 50'000 deutschsprachige Personen mit rechtsextremen Meinungen auf alternativen Plattformen wie Gab oder Reddit aktiv sind, wobei sie unterschiedlich viele Online-Inhalte beitrugen.

Daten aus Umfragen. Zusätzlich werden repräsentative Bevölkerungsstichproben und nicht-repräsentative Stichproben von Internet-Nutzenden danach befragt, wie oft sie Hassrede beobachtet,

⁵⁰ Barnidge et al. 2019: S. 8.

⁵¹ Jourová 2019a. Hier sei einschränkend bemerkt, dass die Online-Beiträge in der Grafik eine Selbstselektion der gemeldeten Beiträge darstellen und nicht notwendigerweise ein repräsentatives Abbild der Häufigkeit von Hassrede-Typen in Social Media.

⁵² Siegel et al. 2019.

⁵³ Gagliardone et al. 2016; Lingardi et al. 2019.

⁵⁴ Hine et al. 2017.

⁵⁵ Guhl, Ebner und Rau 2020: S. 8.

persönlich erfahren oder selbst verbreitet haben. Die vorliegenden Ergebnisse legen nahe, dass die meisten Menschen schon einmal Hassrede im Internet beobachtet haben. Von 2592 jungen Erwachsenen aus sechs Ländern berichteten 70 Prozent, Hassrede im Internet gesehen zu haben.⁵⁶ Dieser Wert variiert länderabhängig (am höchsten in Finnland und Spanien und am tiefsten in Frankreich und Grossbritannien). Die rassistische Hassrede wurde dabei am häufigsten beobachtet (Grafik 2). Dass vier von fünf Befragten zufällig bzw. ohne danach zu suchen, auf diese Inhalte gestossen sind, veranschaulicht, wie schnell es im Internet zum passiven Konsum von Hassrede kommt.

	Total sample	Finland	France	Poland	Spain	U.K.	U.S.A.
Ethnicity or Race	964	90	143	147	102	201	281
Nationality or Immigrant Status	37.2%	36.4%	29.4%	37.8%	26.4%	39.0%	49.6%
Sexual Orientation	853	125	138	149	122	155	164
Religious Conviction and Belief	32.9%	50.6%	28.3%	38.3%	31.5%	30.1%	28.9%
Political Views	889	93	146	153	163	133	201
Sex/Gender/Gender Identity	34.3%	37.7%	30.0%	39.3%	42.1%	25.8%	35.4%
Disability Status	640	81	123	118	52	124	142
Appearance	24.7%	32.8%	25.3%	30.3%	13.4%	24.1%	25.0%
Sample Size	615	59	84	107	117	87	161
	23.7%	23.9%	17.3%	27.5%	30.2%	16.9%	28.4
	686	57	99	75	165	112	178
	26.5%	23.1%	20.4%	19.3%	42.6%	21.7%	31.4%
	223	16	32	34	46	33	62
	8.6%	6.5%	6.6%	8.7%	11.9%	6.4%	10.9%
	470	49	91	83	85	64	98
	18.1%	19.8%	18.7%	21.3%	22.0%	12.4%	17.3%
	2,592	247	487	389	387	515	567

Grafik 2. Frage an alle Personen, die Hassrede beobachtet hatten: Auf welche sozialen Merkmale richteten sich die Angriffe?

Im nahen Deutschland wurden schon mehrere Umfragen zu Hassrede lanciert. In der bisher grössten, national repräsentativen Untersuchung zu Hassrede im Internet von Geschke et al.⁵⁷ berichteten von 7349 deutschen Internetnutzenden vier von zehn, Hassrede bereits gesehen zu haben. Die allermeisten der gesehenen Kommentare bezogen sich auf Migrant*innen, Politiker*innen, Muslim*innen und geflüchtete Menschen. Etwas weniger Angriffe wurden gegen Juden und Jüdinnen und Sinti oder Roma beobachtet. Grundsätzlich stehen also historisch benachteiligte Gruppen auch im Internet im Zentrum von Angriffen. Eine etwas kleinere Umfrage⁵⁸ zeigt eine signifikante Zunahme in der beobachteten Hassrede über die Zeit (von 65% in 2016 auf 78% in 2018).

Ein kleinerer Prozentsatz von Personen berichtet, persönlich von Hassrede betroffen zu sein. In der deutschen Befragung von Geschke et al. waren dies acht Prozent, wobei Personen mit Migrationshintergrund von doppelt so vielen Angriffe berichteten wie solche ohne Migrationshintergrund. Die Werte variieren länderabhängig: Für die Türkei oder die USA finden Studien höhere Werte von 25 bis 40 Prozent.⁵⁹

4.2 Situation in der Schweiz

Für die Schweiz gibt es keine systematisch erhobenen Prävalenzraten zu Hassrede und Rassismus im Internet. Daher ist davon abzuraten, die aktuell limitierten, nichtsdestotrotz informativen Befunde zu generalisieren. Ein vorläufiges Bild legt nahe, dass online viel passiert, jedoch wenig gemeldet oder angezeigt wird. Laut dem Bericht «Rassismussvorfälle aus der Beratungspraxis: Januar bis Dezember 2019»⁶⁰ haben von den 352 Fällen von rassistischer Diskriminierung mit Beratungsleistung nur 23 Fälle (zwei Prozent mehr als im Vorjahr) im Internet stattgefunden. Auch die Zahl der Rechtsfälle zu rassistischen Äusserungen im 2019 ist mit bisher sechs Fällen auf Social Media-Plattformen und sechs Fällen in Blogs und Kommentarspalten klassischer Medien bescheiden.⁶¹

⁵⁶ Reichelmann et al. 2020: S. 4.

⁵⁷ Geschke et al. 2019: S. 19.

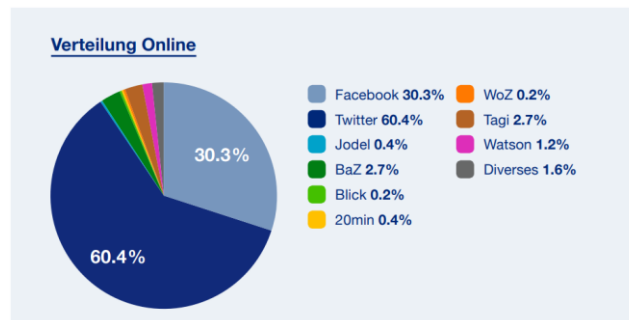
⁵⁸ Landesanstalt für Medien NRW 2018: S. 2.

⁵⁹ Duggan 2017; Celik 2019: S. 1456.

⁶⁰ Beratungsnetz für Rassismussopfer 2020: S. 11.

⁶¹ Eidgenössische Kommission gegen Rassismus. Vorläufige Zahlen, Stand: Mai 2020.

Diese tiefen Zahlen verwundern vor dem Hintergrund, dass Hassaufrufe im Internet von relevanten Stellen als «Dauerproblem» wahrgenommen werden.⁶² Auch legen einzelne Erfahrungsberichte eine signifikante Menge von Hasskommentaren in der Schweizer Internetsphäre nahe. Mitglieder der selbstorganisierten Facebook-Gruppe «Meldezentrale für Eidgenossen», die systematisch Hasskommentare auf Facebook melden, berichten beispielsweise, nur schon im Jahr 2017/2018 ganze 7000 Hasskommentare erfolgreich zur Löschung gebracht zu haben.⁶³ Zudem haben von den erhobenen 523 antisemitischen Fällen im «Antisemitismusbericht für die Deutschschweiz»⁶⁴ ganze 96 Prozent online stattgefunden – die meisten davon auf Social Media (siehe Grafik 3).



Grafik 3. Verteilung der Online-Fälle auf Social Media und klassische Medien (GRA & SIG 2019).

Hier sind auch Studien zu anderen direkten Online-Angriffen zu berücksichtigen – denn auch solche Angriffe können Rassismus beinhalten. Im Jahr 2019 gaben von über 2000 Personen einer Bevölkerungsumfrage⁶⁵ in der Schweiz ca. acht Prozent an, über das Internet verspottet, beleidigt, beschimpft, bedroht oder Opfer von Gerüchten oder veröffentlichten privaten Informationen geworden zu sein; sieben Prozent wurden sexuell belästigt. Ob Personen einen Migrationshintergrund hatten oder nicht, hatte keinen Einfluss. Es scheint zudem, dass insbesondere (teil-)öffentliche Personen solchen Angriffen ausgesetzt sind: Im Jahre 2017 gaben von 637 Journalist*innen in der Schweiz⁶⁶ über die Hälfte an, im vorangegangenen Jahr von ihrem Publikum beleidigt, bedroht oder sonstig angegriffen worden zu sein (9 von 10 davon über digitale Medien). Davon berichteten sieben Prozent, dass die Angriffe auf ihren Migrationshintergrund oder ihre Herkunft zielten.

4.3 Sondergruppe Kinder und Jugendliche

Die jüngere Alterskohorte verdient besonderes Augenmerk: Sie bewegt sich alltäglich und versiert im Internet. Social Media gehört zu ihrem Leben. Dies macht sie sichtbar, aber auch angreifbarer. Über Länder hinweg wird festgestellt, dass sie am häufigsten (rassistische) Hassrede beobachten und Täter*innen oder Betroffene sind. In einer Umfrage in Deutschland geben die 14- bis 24-Jährigen deutlich häufiger an, «persönlich Hasskommentare im Internet gesehen» zu haben als ältere Personen (96 Prozent gegenüber 85 Prozent der 25- bis 44-Jährigen, 75 Prozent der 45- bis 59-Jährigen und 60 Prozent der über 60-Jährigen).⁶⁷ Werden die vergangenen drei Monate abgefragt, geben bei den 15- bis 30-Jährigen in Deutschland ein Drittel an, Online-Hass mindestens einmal gesehen zu haben; in Grossbritannien und Finnland sind diese Werte etwas höher (39% respektive 48%), und am höchsten sind sie in den USA (53%).⁶⁸ Eine kleinere Gruppe ist persönlich von Hassrede betroffen, Täter*innen machen die kleinste Gruppe aus: In einer Umfrage in Frankreich haben von fast 2000 Jugendlichen sieben Prozent angegeben, schon einmal «erniedrigende oder hasserfüllte» Nachrichten, Kommentare

⁶² Humanrights.ch, 23. August 2017: [Hassaufrufe im Internet - Schweizer Fälle und Politik der Newsportale](#).

⁶³ Germann, WOZ vom 20. September 2018: [7000 gesperrte Kommentare](#).

⁶⁴ Stiftung gegen Rassismus und Antisemitismus (GRA) und Schweizerischer Israelitischer Gemeindeverband (SIG) 2019. Hier fliessen Fälle ein, die von beobachtenden oder unmittelbar betroffenen Personen gemeldet werden, in den Medien erwähnte Vorfälle und eigene Recherchen auf Social Media.

⁶⁵ Baier 2019: S. 39.

⁶⁶ Stahel und Schoen 2019: S. 11.

⁶⁷ Landesanstalt für Medien NRW 2018: S. 2.

⁶⁸ Hawdon et al. 2017: S. 260.

oder Bilder gegen eine Person oder eine Gruppe veröffentlicht oder geteilt zu haben. Elf Prozent sind Opfer geworden.⁶⁹ Für die USA wurden höhere Zahlen gefunden: 20 Prozent waren Täter*innen und fast ein Drittel Betroffene.⁷⁰

Für die Schweiz existieren nur Zahlen für das Beobachten von Online-Hassrede. Verglichen mit anderen Ländern sind sie relativ tief: In der Studie *EU Kids Online* gab im Jahr 2019 von den 12- bis 16-jährigen Internetnutzenden im Schnitt eine von acht Personen an, mindestens monatlich Online-Hassrede zu sehen. Eine von fünf sieht sie weniger oft und zwei Drittel nie.⁷¹ Die Älteren beobachten Hassrede häufiger, wie die gleiche Studie zeigt: Nach Altersgruppe aufgeteilt beobachten unter den 15- bis 16-Jährigen doppelt so viele Nutzende (21 Prozent) Hassrede als unter den 12- bis 14-Jährigen (11 Prozent).

⁶⁹ Blaya und Audrin 2019: S. 6.

⁷⁰ Costello und Hawdon 2018: S. 58.

⁷¹ Smahel et al. 2020: S. 66 und S. 67.

5 DIE DIGITALE UMWELT DER RASSISTISCHEN HASSREDE

Rassistisches Gedankengut im Internet entfaltet sich in einem zunehmend komplexen und vielschichtigen System von digitalen Plattformen und Anwendungen. Es wird innerhalb und zwischen Netzwerken von Nutzenden verbreitet.⁷² Dabei ist ein grosses Spektrum von Äusserungen zu beobachten, von expliziten Vernichtungsdrohungen bis zu weiter verbreiteten stereotypen Vorurteilen. In diesem Kapitel interessiert, wie rassistische Online-Hassrede organisiert ist und auf welchen Plattformen sie auftritt. Zudem wird beleuchtet, welche individuellen, sozialen und gesellschaftlichen Faktoren ihre Verbreitung begünstigen und wie Kommunikationsbedingungen und Architekturen von Plattformen dazu beitragen.

5.1 Organisationsgrad von Täterschaften

Verbreiter*innen von rassistischer Online-Hassrede lassen sich nach ihrem Organisationsgrad unterscheiden – von hierarchischen Hassgruppen über «fluide» Netzwerke bis zu Einzelnutzenden. Personen können dynamisch zwischen diesen Einheiten wechseln, die Übergänge sind fließend. Der Organisationsgrad dieser Einheiten beeinflusst jedoch prinzipiell, auf welchen Plattformen, mit welchen Zielen und auf welche Art sie kommunizieren.⁷³

Am ältesten und bekanntesten sind die sogenannten *Hassgruppen* («Hate groups»)⁷⁴. Damit sind gemeinhin hierarchisch organisierte Gruppen gemeint, die bereits existierten, bevor es Internet gab. Seit den neunziger Jahren sind sie auch online und dabei meist explizit rassistisch. Durch die digitale Weiterentwicklung hin zu Social Media und unter dem zunehmenden Druck, der sozialen Stigmatisierung expliziter rechtsextremer Inhalte zu entgehen, haben sich in den letzten Jahren neuere, weniger organisierte und weniger explizite Einheiten entwickelt.

Diese *digitalen Netzwerke* sind informeller und beherbergen Nutzer*innen, die sich auf einer oder mehreren Plattformen bewegen und unterschiedlich stark miteinander verknüpft sind.⁷⁵ Sie interagieren bei Gelegenheit, z.B. bei koordinierten Angriffen. Diese Netzwerke scheinen fluide und ambivalent: Klare organisatorische Zentren fehlen.⁷⁶ Ein Beispiel ist die digital entstandene US-amerikanische Alt(ernative)-Right Bewegung.⁷⁷ Sie wurde einer breiteren Öffentlichkeit spätestens bekannt, als sie Donalds Trumps Präsidentschaft aktiv unterstützte. Sie ist prinzipiell führerlos, abgesehen von ihren dynamisch wechselnden Meinungsführern auf Social Media. Diese werden auch «the Intellectual Dark Web» genannt und reichen von «Influencer*innen» auf Twitter über YouTube-Vlogger zu Podcasters.⁷⁸ Sie beeinflussen emotional – anstatt durch Befehle. Sie fechten «politisch korrekte Medien» und die öffentliche Meinung mitsamt den Personen, die diese akzeptieren (genannt «Normies»⁷⁹), an. Sie sehen sich als Opfer und als gefährdete Gemeinschaft, die sich nur mit einer radikalen Alternative und mit radikalen Lösungen zur Wehr setzen kann. In ihrer Anhängerschaft verschwimmen die Grenzen zwischen organisierten Aktivist*innen und Unterstützer*innen. Hochaktiv sind jedoch ihre Trolle: Sie streben grösstmögliche Aufmerksamkeit durch das Stören von Debatten an und schrecken vor der Diffamierung unliebsamer Personen und Organisationen nicht zurück.⁸⁰

Bei manchen Organisationsformen schliessen sich Hierarchie und Netzwerkcharakter jedoch nicht aus. Eine solche hybride Form war die rechtsextreme deutsche «Reconquista Germanica»⁸¹ aus dem

⁷² Guhl et al. 2020: S. 5.

⁷³ Bliuc et al. 2018: S. 76.

⁷⁴ Winter 2019: S. 54.

⁷⁵ Fielitz und Marcks 2019: S. 16.

⁷⁶ Fielitz und Marcks 2019: S. 1.

⁷⁷ Beschrieben als “a set of far-right ideologies, groups and individuals whose core belief is that ‘white identity’ is under attack by multicultural forces using ‘political correctness’ and ‘social justice’ to undermine white people and ‘their’ civilization”. Southern Poverty Law Center 2020. Pfad: www.splcenter.org > Fighting Hate > Extremist Files > Ideology > [Alt-Right](#).

⁷⁸ Winter 2019.

⁷⁹ [Hier](#) ist eine umfängliche Selbstbeschreibung der Bewegung zu finden: Anglin, Daily Stormer vom 31. August 2016.

⁸⁰ Marwick und Lewis 2017: S. 4.

⁸¹ Die Gruppe hat im November 2019 über [Youtube](#) ihre Auflösung erklärt. Pfad: Youtube > Reconquista Germanica meldet sich ab.

Umfeld der «Identitären Bewegung»: In diesem nicht-öffentlichen Online-Forum hatten Nutzende nach strengen Befehlsketten Kommentarspalten und Social-Media-Profile mit ihren eigenen Inhalten «überflutet».⁸²

Der Übergang zu den kaum oder nicht organisierten *einzelnen Nutzer*innen* ist fließend. Sie verbreiten rassistische Inhalte relativ isoliert, sei dies unbewusst oder bewusst, aufgrund mangelnder Medienkompetenz oder strategischer Absicht, aus situationsbedingtem Impuls oder tiefster ideologischer Überzeugung – auf jeden Fall nicht im Namen einer Gruppe. Sie sind häufig auf den weitverbreiteten Social-Media-Plattformen wie Facebook oder Twitter zu finden.

5.2 Individuelle, soziale und gesellschaftliche Einflussfaktoren von Täterschaft

Wer verbreitet Rassismus im Internet und wieso? Untersuchungen zu Verbreiter*innen von Hassrede im Allgemeinen können hierbei Aufschlüsse geben. Zum einen sind *individuelle Merkmale* relevant. Beispielsweise besitzen Internetnutzende, die regelmässig in Online-Debatten stören und provozieren (bzw. «trollen»), stärkere (subpathologische) sadistische und psychopathische Eigenschaften, verglichen mit den Internetnutzenden, die nicht trollen. Sie sind impulsiver, gewissenloser, suchen stärkeren Nervenkitzel und können Mitgefühl schlechter nachempfinden.⁸³ Auch Erfahrungen im Internet sind einflussreich. Zwar ist umstritten, ob Täter*innen mehr Zeit im Internet verbringen als Nicht-Täter*innen. Jedoch bewegen sie sich zumindest häufiger auf Plattformen und in Gemeinschaften, in welchen Hassrede propagiert wird.⁸⁴ Gut belegt ist die signifikante Überlappung von Täter*in und Opfer in der gleichen Person: Aus Opfer wird Täter*in und/oder umgekehrt.⁸⁵ Wie Computerspiele Hassrede beeinflussen, ist in der Wissenschaft bisher ungeklärt. Die gefundenen Effekte sind stark gemischt, wobei Metaanalysen (d.h. Überblicksstudien zu bisherigen Studien) tendenziell minimale positive bis keine Effekte finden.⁸⁶ Insbesondere positive Effekte werden durch zwei zentrale Mechanismen erklärt.⁸⁷ Personen mit einer schon vorhandenen Aggressionstendenz sind stärker von bestimmten Spielen angezogen (gäbe es keine Spiele, hätten sie ihre Aggression woanders ausgelebt). Gleichzeitig könnten Shooter-Spielende gewaltsames Verhalten durch das Beobachten von gewalttätigen Spielfiguren lernen. Nutzende könnten diese Verhaltensweisen ins analoge Verhaltensrepertoire aufnehmen und im realen Umfeld ausleben. Schliesslich beeinflussen auch persönliche Ziele und Einstellungen, ob Nutzende Hassrede verbreiten. Die Motive von Täter*innen sind auf einem breiten Spektrum zu finden. Laut Interviews und Umfragen mit Täter*innen suchen sie durch Hassrede Aufmerksamkeit, Anerkennung, Rache, Einfluss, Macht und Unterhaltung; sie wollen Andere aufklären, wahrgenommene Ungerechtigkeiten beseitigen, ihre soziale Gruppe verteidigen und ihre Überzeugungen validieren.⁸⁸ Diese Heterogenität widerspiegelt das breite Spektrum unterschiedlicher «Typen»: von der «frustrierten Wutbürgerin», die sich in einer hitzigen Online-Debatte einen einmaligen «Hass-Ausrutscher» erlaubt, bis zum Hass-Influencer, der Hassrede strategisch und langfristig vor einem grösseren Online-Publikum als politisches Mittel einsetzt. Darüber hinaus gelten auch Vorurteile und Misstrauen als Nährboden für Hassrede. In einer Umfrage bei Jugendlichen in Frankreich wurde festgestellt, dass diese eher Hassrede verbreiteten, wenn sie gegenüber Gewalt und Rassismus positiv eingestellt waren und wenn sie den Institutionen Schule und Politik misstrauten.⁸⁹

Über die *sozialen Merkmale* von Täter*innen und ihre gesellschaftliche Einbettung ist bisher wenig bekannt. Auf jeden Fall sind Männer signifikant häufiger Täter.⁹⁰ Auch jüngere Menschen scheinen relativ häufiger Täter*innen zu sein (siehe Kap. 4.3). Systematische Vergleiche mit älteren Kohorten stehen jedoch aus. Es ist vorstellbar, dass unterschiedliche Generationen themen- und opferspezifischer «hassen»: Beispielsweise sind die meisten mithilfe des Schweizer Vereins #Netzcourage angezeigten Kommentierenden Senioren, wobei in diesen Fällen meist weibliche, öffentliche Personen wie Politikerinnen betroffen sind.⁹¹ Einen klaren Einfluss scheint das soziale

⁸² Fielitz und Marcks 2019: S. 13.

⁸³ Buckels et al. 2014.

⁸⁴ Costello und Hawdon 2018: S. 58.

⁸⁵ Blaya und Audrin 2019: S. 11.

⁸⁶ Elson und Ferguson 2014.

⁸⁷ Elson und Ferguson 2014: S. 3.

⁸⁸ Z.B. Craker und March 2016: S. 83; Erjavec und Kovačić 2012: S. 912; Guhl et al. 2020: S. 47.

⁸⁹ Blaya und Audrin 2019.

⁹⁰ Z.B. Blaya und Audrin 2019: S. 3; Costello und Hawdon 2018: S. 58.

⁹¹ Interview mit Jolanda Spiess-Hegglin im März 2020.

Umfeld zu haben: Junge Täter*innen gehören oft einer devianten (z.B. kriminellen) Jugendgruppe an, haben ansonsten relativ wenige offline Beziehungen und viele online Kontakte, z.B. in Online-Gemeinschaften.⁹² In «soziale Klassen» sind sie bislang nicht klar verortbar; die wenigen bestehenden Daten scheinen kaum Effekte von Bildung und damit zusammenhängender Arbeitstätigkeit und Einkommen zu zeigen.⁹³ Dies überrascht, da diese soziostrukturellen Merkmale allgemein Verbrechen miterklären. Die Autor*innen der Studien spekulieren jedoch, dass aufgrund zunehmender Medienbildung in Zukunft soziostrukturelle Effekte häufiger zu beobachten sein könnten.⁹⁴

Wird bedacht, dass das reale Umfeld und der virtuelle Online-Raum zunehmend verschmelzen, verwundert es nicht, dass sich auch *gesellschaftliche Entwicklungen* in rassistischer Online-Hassrede niederschlagen.⁹⁵ Werden soziale, ökonomische, politische oder kulturelle Machtstrukturen umgewälzt, wie es im Falle von Migration, postmaterieller Werte, Gleichberechtigung oder (Un-)Gleichheit geschieht, können gesellschaftliche Konflikte geschürt werden. Vormalig privilegierte Gruppen können Statusverluste befürchten oder sie tatsächlich erleben. Um dieser Bedrohung entgegenzutreten, können sie auf analoge oder digitale Hassrede zurückgreifen, um die klare Grenze zwischen denjenigen, die sie als Norm und übergeordnet, und denjenigen, die sie als anders und untergeordnet ansehen, wiederherzustellen – und Letztere daran zu erinnern, wo sie hingehören.⁹⁶ Solche Konflikte schlagen sich auch in der Berichterstattung nieder: Berichten Journalist*innen in der Schweiz beispielsweise zu Identitäts-bezogenen Themen wie Religion oder Geschlechtergleichheit, werden sie von ihrem Publikum besonders oft angegriffen.⁹⁷ Auch gelten polarisierte Gesellschaften als idealer Nährboden für Hassrede⁹⁸: Das Internet dient dabei als erweiterte Kampfzone um Ressourcen, Macht und Lebenschancen. Die mit der Alternative für Deutschland (AFD) assoziierten Trolle der ehemaligen Reconquista Germanica sind nur ein Beispiel für den strategischen Einsatz von Online-Hassrede zu solch politischen Zwecken.

Diese gesellschaftlichen Strukturen können sich in *Einzelereignissen* niederschlagen, die wiederum rassistische Online-Hassrede auslösen. Die öffentliche Bühne von Social Media erlaubt insbesondere bei situativen Ereignissen, rassistische Hassrede massenhaft zu produzieren und zu konsumieren. Grossflächige Analysen von Online-Kommentaren legen nahe, dass Attentate auf Menschen im digitalen Raum Hassrede gegen die als schuldig befundenen Gruppen verstärken. Dies wird unter anderem damit erklärt, dass Attentate Gefühle der Bedrohung und Unsicherheit auslösen, die wiederum Feindseligkeit gegen «Fremde» schüren. Die soziale Zugehörigkeit der Opfer scheint jedoch entscheidend zu sein: Auf Twitter stieg nach islamistischen Attentaten – wie jenem auf die Konzertgänger in der Manchester Arena im Jahr 2017 mit 23 Toten – die Zahl von Hassreden an (und interessanterweise auch die dazugehörige Gegenrede, die sich gegen pauschalisierende Verurteilungen der muslimischen Gemeinschaft wehrte). Bei Attentaten, die muslimfeindlich motiviert waren – wie der Erschiessung zweier fälschlicherweise für Iraner gehaltene Inder in Kansas, USA, im Jahr 2017 – waren dagegen kein Anstieg von Hassreden feststellbar.⁹⁹ Online-Hassrede kann auch während sonstigen politischen Ereignissen wie Wahlen zunehmen¹⁰⁰, jedoch flacht sie teilweise rasch wieder ab.¹⁰¹

Diese Studien – allesamt ausserhalb der Schweiz – geben einen ersten Eindruck in die komplexe Erklärungslage rassistischer Online-Hassrede. Insbesondere der Zusammenhang mit sozialen

⁹² Z.B. Blaya und Audrin 2019: S. 11; Ribeiro et al. 2018.

⁹³ Z.B. Costello und Hawdon 2018: S. 58; Lowry et al. 2016: S. 978.

⁹⁴ Dies weil moralisches Verhalten oft durch Bildung gelernt werde; digitale Medienbildung, die moralisch-empathisches Verhalten im Internet und das Abschätzen der Folgen des eigenen Verhaltens lehrt, ist heute oft Teil der Schulausbildung. Personen, die in Zukunft weniger Bildung geniessen, würden somit weniger medienkompetent sein, was sie anfälliger zur Verbreitung von Online-Hassrede macht (wobei auch eine erfolgte Medienbildung keine «Impfung gegen Hassrede» ist). Dass bei den heutigen Erwachsenen kaum Effekte gefunden werden, könnte daran liegen, dass digitale Medienbildung innerhalb der Bildung dieser gesamten Generation grösstenteils fehlte; der Bildungsstand alleine kann die Anfälligkeit zur Verbreitung von Hassrede kaum beeinflussen.

⁹⁵ Siehe dazu auch die Länderabhängigkeit in den konkreten Äusserungen von rassistischer Online-Hassrede in sCAN 2018a.

⁹⁶ Quent 2018: S. 49 ff.

⁹⁷ Stahel, NZZ vom 7. Mai 2019: [Gehässige Leserreaktionen können die Qualität der Berichterstattung auch erhöhen](#).

⁹⁸ Interview mit Jonathan Birdwell (Institute of Strategic Dialogue) im April 2020.

⁹⁹ Olteanu et al. 2018: S. 228.

¹⁰⁰ Lim 2017.

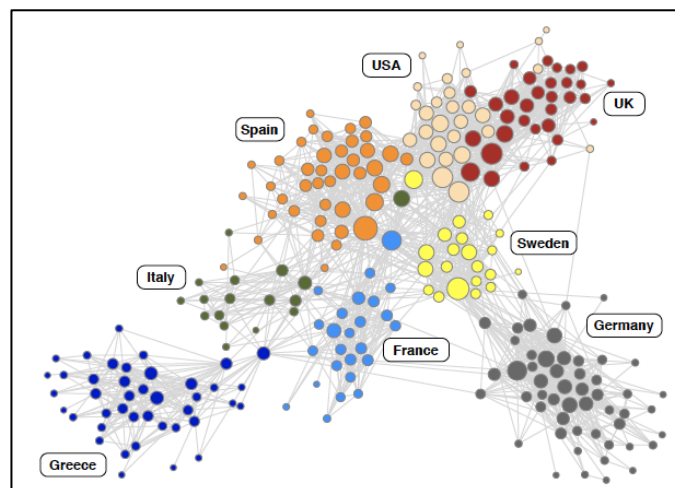
¹⁰¹ Siegel et al. 2019: S. 32.

Faktoren (Einkommen, Bildung, soziale Position etc.) und gesellschaftlichen Ereignissen (z.B. Wahlen) sollte in der Schweiz erforscht werden, um Interventionen zielgruppenfokussiert ausrichten zu können.

5.3 Relevante Plattformen

Es gibt viele Online-Räume, in denen rassistische Ideen verbreitet werden können. Jedoch sind manche Plattformen besonders zentral, weil sie entweder in der Gesellschaft breit genutzt werden (z.B. Facebook) oder alternative Plattformen darstellen. Letztere wurden entweder speziell für extremistische Zwecke gegründet, oder sie tolerieren aufgrund ihres breiten Verständnisses von Meinungsfreiheit Hassrede oder sie wurden als ursprünglich unpolitische Plattformen zweckentfremdet (z.B. Computerspiele).¹⁰² Die relevantesten Plattformen werden hier separat vorgestellt.

Grundsätzlich sind Verbreiter*innen von Hassrede in der Praxis jedoch auf mehreren Plattformen unterwegs und verlinken, vernetzen und inspirieren sich über Plattformen hinweg. Strategisch agierende Akteure wie die muslimfeindliche deutsche PEGIDA-Gruppe, die ethnonationalistische, in Europa angesiedelte Identitäre Bewegung (IB) oder die internationale, rechtsterroristische Atomwaffen Division (AWD) bedienen sich des Internets, um Propaganda und Desinformation transnational zu verbreiten, Attentäter zu verherrlichen, koordiniert Politiker*innen anzugreifen und «Meme»-Kampagnen (siehe 5.4.2) zu starten.¹⁰³ Solche Gruppen sind national und international vernetzt, wobei eine gemeinsame Sprache wie in den USA und den UK die Vernetzung fördert (siehe Grafik 4). Sie informieren sich in einer umfangreichen rechtspopulistischen Medienlandschaft, welche die als «Mainstream» bezeichneten Nachrichtenmedien diskreditiert und Hassrede legitimiert.¹⁰⁴



Grafik 4. Vernetzung von Profilen rechtsextremer Gruppen aus acht Ländern auf der Plattform Twitter (eine Verbindungslinie kommt zustande, wenn das Twitter-Profil einer Gruppe das Profil einer anderen in Tweets erwähnt oder deren Tweets teilt).

Webseiten und Blogs

Für rassistisches Gedankengut sind Webseiten und Blogs wichtige Informationszentralen. Dieses wird dort meist offen kommuniziert, beispielsweise auf [The Daily Stormer](#) und [Vanguard News Network](#), oder, in abgeschwächter Form, auf rechtspopulistischen Webseiten wie [Breitbart](#) oder dem deutschen [PI News](#). Vergleichbare, an Rassismus oder Populismus grenzende oder vertretende Webseiten in der Schweiz sind [Hammerschweizer.ch](#) (Deutschschweiz) oder [LesObservateurs.ch](#) (Romandie). Obwohl im Tessin verbreitete fremdenfeindliche Neigungen festgestellt werden (was mit seiner Grenze zu Italien zusammenhängen mag)¹⁰⁵, scheinen sich dortige Sympathisant*innen besonders auf italienischen

¹⁰² Guhl et al. 2020: S. 5.

¹⁰³ Guhl et al. 2020: S. 7.

¹⁰⁴ Guhl et al. 2020: S. 7.

¹⁰⁵ Haymoz et al. 2019: S. 11.

Webseiten zu informieren.¹⁰⁶ Beispiele sind die Webseiten der rechtsextremen Bewegung [Casapound Italia](#) (die auch einen Twitter- und Telegram-Kanal unterhält) und eine der grössten und am besten organisierten italienischen, nationalsozialistischen Gemeinschaften, die [Comunità Militante Dei Dodici Raggi](#) (Do.Ra.). Insbesondere durch die massenhafte Verbreitung des Internets haben solche Webseiten in den letzten zwanzig Jahren stark zugenommen.¹⁰⁷ Ihre konkrete Zahl ist jedoch schwer bestimmbar, da sie oft schliessen oder ihre Namen ändern.¹⁰⁸ Um sichtbarer zu werden, interagieren sie umfangreich, indem sie sich gegenseitig verlinken und zitieren. Sie sprechen unterschiedlich breite Zielgruppen an. Beispielsweise wird Frauen durch das Betonen ihrer Rolle in der Gründung einer Familie, sei sie arisch oder jihadistisch, Bedeutung zugewiesen.¹⁰⁹

Foren

Foren sind hingegen partizipativer, indem Nutzende eigene rassistische Äusserungen beitragen können. Das älteste Forum ist das rechtsextreme [Stormfront](#), welches im Jahr 1995 in den USA online ging. Dort kann man neben Radio hören, Chatten und Blogs anschauen vorrangig Inhalte posten und sich miteinander unterhalten. Die Inhalte drehen sich um die «Weisse Überlegenheit» und beinhalten sogenannte «Threads» bzw. separate Diskussionen, beispielsweise zu «History and revisionism». Aktuellere Foren sind [Gab](#), [Reddit](#) und [8chan](#), die weltweit Anhänger*innen haben.¹¹⁰ Rassismus konzentriert sich dort häufig in einzelnen Threads, z.B. der populäre Pro-Trump-Thread [the_donald](#) (auf Reddit) oder der «politisch unkorrekte /pol/» (auf [4chan](#)). Rassismus kann jedoch auch Debatten begleiten, wie beispielsweise jene über Gen-Tests, um die eigene «genetische Reinheit» zu beweisen.¹¹¹ Diese Foren sind mehrheitlich anonym, oft humoristisch und weniger kontrolliert als etablierte Social-Media-Plattformen. Rassismus und misogyne Bewegungen wie die der «Incels» gehen dort Hand in Hand.¹¹² Viele dieser Foren bezeichnen sich selbst als rechtsextreme Seiten. Dies zieht Menschen dieser Gesinnung an, da sie «sichere Häfen» erwarten, in denen ihre Äusserungen nicht sanktioniert werden. Bewegen sie sich einmal in diesen Foren, stossen sie rasch auf dort verbreitete Links zu extremistischen Webseiten, was einer weiteren Radikalisierung Vorschub leistet.¹¹³

Schliesslich können auch Foren, die nicht zur Verbreitung von Hassrede initiiert wurden, sondern beispielsweise zu Sport¹¹⁴, für die Verbreitung von Rassismus zweckentfremdet werden.

Etablierte Social-Media-Plattformen

Das soziale Netzwerk Facebook, die audiovisuelle Plattform YouTube oder der Microblog Twitter – auf solchen etablierten Plattformen ist eine Mehrheit der Schweizer Bevölkerung unterwegs. Hier trifft der oder die Durchschnittsbürger*in meist zufällig auch auf rassistische Hassrede.¹¹⁵ Denn diese Plattformen sind durch ihre Vernetzungsgeflechte prädestiniert, Inhalte global zu verbreiten – und damit besonders attraktiv für rassistische Akteur*innen, um Sichtbarkeit zu erlangen. Sie streuen beispielsweise Inhalte auf Twitter, die dann von Journalist*innen aufgenommen werden.¹¹⁶ So wurde medial über die Twitter-Kampagne «#120db» der rechtsradikalen Identitären Bewegung berichtet: Diese nutzte das Narrativ von Frauenrechten, um gegen sogenannte «Migrantengewalt» gegen Frauen

¹⁰⁶ Information des Dipartimento delle istituzioni (DI), Piattaforma di prevenzione della radicalizzazione e dell'estremismo violento, Bellinzona.

¹⁰⁷ Winter 2019: S. 40; Perry und Olsson 2009: S. 188.

¹⁰⁸ Perry und Olsson 2009: S. 188.

¹⁰⁹ Musial 2017.

¹¹⁰ Hine et al. 2016.

¹¹¹ Mittos et al. 2019.

¹¹² «Incel» ist ein Akronym für «Involuntary Celibate», also «unfreiwilliges Zölibat». Die Mitglieder dieser Online-Gemeinschaft, meist Männer, geben an, dass ihnen ihr Recht auf Geschlechtsverkehr verweigert werde. Sie hängen der Ideologie einer hegemonialen Männlichkeit an und drücken ihren Hass auf sexuell aktive Personen und insbesondere Frauen offen aus. Siehe: Pfeiffer, Belltower News vom 25. März 2020: [Hate Speech in der Incel-Szene](#).

¹¹³ Marwick und Lewis 2017.

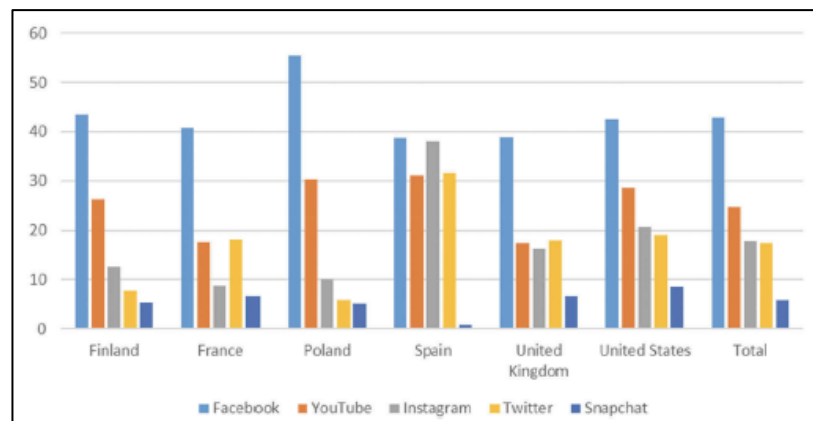
¹¹⁴ Cleland 2013.

¹¹⁵ Celik 2019; Reichelmann et al. 2020: S. 6.

¹¹⁶ Marwick und Lewis 2017: S. 26.

zu protestieren – und gegen männliche Zuwanderer zu hetzen.¹¹⁷ Oder Täter*innen laden auf YouTube rassistische Musikvideos hoch, durch die sich die Nutzenden klicken – und darunter rassistisch kommentieren. Auf diesen Plattformen sind auch Influencer*innen vertreten, die als Brücke zwischen «Mainstream» und Extremismus fungieren. Sie brillieren im «Aufmerksamkeits-Hacking», wobei sie die Sichtbarkeit ihrer Ideen durch die strategische Nutzung von Plattformen verstärken.¹¹⁸ Stefan Molyneux, ein YouTube-Kommentator, verbreitet beispielsweise erfolgreich rassistische Thesen im wissenschaftlichen Mantel. Die Videos auf seinem YouTube-Kanal, die Namen tragen wie «Human Biodiversity and Criminality», sind bisher fast 300 Millionen Mal aufgerufen worden.¹¹⁹

Eine länderübergreifende Umfrage¹²⁰ zeigt, dass Nutzende auf Facebook am meisten Hassrede sehen – was mit seiner Position als meistgenutzte soziale Netzwerk-Seite zusammenhängen mag (siehe Grafik 5). In teils öffentlichen, teils privaten Gruppen wird beispielsweise gegen Flüchtende und Muslim*innen gehetzt: in Deutschland in «Nein zum Heim»-Gruppen und in der Schweiz in nationalistisch ausgerichteten Gruppen wie der Facebook-Gruppe «New Swiss Journal». Hier bekommen auch diejenigen, die sich für Flüchtlinge einsetzen, den Stempel des «Volksverräters».



Grafik 5. Anteil der 18- bis 25-Jährigen in sechs Ländern, die berichteten, in den vorangegangenen drei Monaten Hassrede auf der jeweiligen Internetplattform gesehen zu haben (N= 2'592).

Computerspiele

Auch der Wettbewerbscharakter von Computerspielen (Online-Games), die starke Identifikation mancher Spielenden und eine spezifische Spielkultur kann Hassrede begünstigen.¹²¹ Computerspiele sind mittlerweile weitverbreitet und diversifiziert, was die Endgeräte, Genres, spielenden Menschen, und Spielarten (Single-/Multiplayer-Spiele) angeht. Obwohl die Bandbreite der Interaktionen zwischen Spielenden gross ist, stellen Computerspiele laut Expert*innen seit jeher auch einen lebendigen Raum für Rassismus dar.¹²² Hassgruppen haben schon früh Hate-Games produziert, d.h. Hass-Versionen von populären Online-Games zur Unterhaltung und Rekrutierung neuer Mitglieder. Hate-Games reproduzieren Stereotypen und Vorurteile, verwischen Realität und Fantasie und präsentieren meist gewalttätige «Lösungen». Daneben existiert die allgemeine, international vernetzte Gaming-Gemeinschaft.¹²³ Manche Forschende setzen diese sehr breite Gruppe von Spielenden mit kulturellen Gemeinschaften gleich, in denen ein Verständnis gemeinsamer Symbole, Bedeutungen und Praktiken herrscht. Eine verbreitete, exkludierende Praktik ist die des «Trash Talk», d.h. verbale – darunter rassistische – Beleidigungen zwischen den Spielenden.¹²⁴ Diese Praktik wurde besonders deutlich in

¹¹⁷ Jetzt, 20. Februar 2018: [Rechtsextreme „Feministinnen“ stören die Berlinale](#).

¹¹⁸ Marwick und Lewis 2017: S. 1.

¹¹⁹ Southern Poverty Law Center, 19. April 2018: [McInnes, Molyneux, and 4chan: Investigating pathways to the alt-right](#). Siehe auch den Kanal von S. Molyneux auf [Youtube](#).

¹²⁰ Reichelmann et al. 2020.

¹²¹ Breuer 2017: S. 107; Banaszczuk 2019.

¹²² Z.B. Ortiz 2019.

¹²³ Banaszczuk 2019.

¹²⁴ Ortiz 2019.

der #Gamergate-Kampagne. Im August 2014 wurden diverse Frauen, die Misogynie und Rassismus in der Gaming-Kultur kritisiert hatten, auf 4chan, Reddit und Twitter mit Online-Hass überzogen.

Emails und Chats

Auch auf privaten Kanälen können rassistische Angriffe stattfinden. Über E-Mail können spezifische Individuen direkt angegriffen und eingeschüchtert werden. Solche Angriffe sind nicht selten. Vier von fünf Journalist*innen in der Schweiz, die im Jahr 2017 berichteten, im vorangegangenen Jahr von ihrem Publikum (rassistisch und/oder nicht rassistisch) angegriffen worden zu sein, wurden über nicht-öffentliche Online-Kanäle angegriffen.¹²⁵ Darunter fallen auch Chats. Medial wurden Fälle bekannt, in denen «Klassen-Chats» rassistische «Memes»¹²⁶ transportierten¹²⁷ und Jugendliche «Witze über Juden, Konzentrationslager und Hitlergrüsse» in Whatsapp-Chats mit dem Namen «FC NSDAP» teilten.¹²⁸ Es wird davon ausgegangen, dass diese nicht- oder teil-öffentliche Kommunikation auf WhatsApp (Gruppenchat-Limit von 256 Mitgliedern) oder Telegram (bis zu 200.000 Mitgliedern) für Rassismus zunehmend an Relevanz gewinnt: Sie kann von aussen kaum kontrolliert werden und zieht somit diejenigen an, die von etablierten Social-Media-Plattformen verdrängt werden.¹²⁹

5.4 Rassismus-begünstigende Online-Infrastruktur

Das System der vorgestellten Plattformen bietet eine Opportunitätsstruktur, die Hassrede und Rassismus begünstigt. Deren Kommunikationsbedingungen und Architekturen schwächen Hindernisse, die im realen Umfeld Hassrede beschränken.¹³⁰ Dementsprechend berichten Menschen, auf Social Media mehr Hassrede zu beobachten als im analogen Leben.¹³¹ In diesem Kapitel wird diese digitale Umwelt näher diskutiert; sie soll jedoch nicht als ursächlich verstanden werden, sondern als ein Baustein einer komplexeren Erklärungslage (siehe auch Kap. 5.2).

5.4.1 Kommunikationsbedingungen

Umweltbedingungen sind entscheidend für die Art und Weise, wie Menschen kommunizieren. Der digitale Raum bietet zwar nicht grundsätzlich neue Kommunikationsbedingungen im Sinne einer unverwechselbaren Qualität (Anonymität gab es beispielsweise schon immer). Er verstärkt jedoch unterschiedlichste Bedingungen, die kombiniert Hass-fördernd wirken.¹³² Dazu gehören:

Konflikte durch aufeinander treffende Räume. Im Internet «prallen divergente Alltagswelten aufeinander: (...) Feminist*innen auf Maskulinisten und liberale Grossstädter*innen auf Dorfnazis». ¹³³ Zwar kommen seit jeher Menschen mit unterschiedlichem ethnischen, sozialen oder kulturellen Hintergrund in Kontakt. Mit der globalen digitalen Vernetzung hat sich dies jedoch potenziert – wobei die Werte und Normen, die in jeder Teilöffentlichkeit gelebt werden, sich nicht aufgelöst haben. Für diese Teilöffentlichkeiten sind die im Internet zeitlos archivierte Inhalte, mitsamt den merkwürdigsten, konträrsten und «unmoralischsten» Lebensweisen anderer Teilöffentlichkeiten, nur einen Klick entfernt. Wer sich empören möchte, hat genug Material. Das soziale und geografische Aufeinandertreffen von Räumen kann dementsprechend dynamische und unkontrollierbare Konflikte auslösen¹³⁴, die sich auch in rassistischer Hassrede entladen.

¹²⁵ Stahel 2020. Nicht veröffentlichte Daten, erhoben 2017.

¹²⁶ Memes beschreiben eine Informationseinheit – oft eine Kombination von Text und Bild – typischerweise mit humoristischer Note, welche durch Verbreitung Einfluss gewinnt. Ein Beispiel ist die Cartoonfigur «Pepe the Frog»; sie wurde von der Anti-Defamation League in ihre Datenbasis rassistischer Online-Symbole aufgenommen. Siehe Spiegel.de vom 28. September 2016: [Freundlicher Frosch wird Hasssymbol](#).

¹²⁷ 20min vom 30. Dezember 2019: [Primarschüler machen sich über Dunkelhäutige lustig](#).

¹²⁸ 20min vom 19. März 2019: [«Rechtsextreme Chats gibt es an jeder Schule»](#).

¹²⁹ Guhl et al. 2020 : S. 18.

¹³⁰ Brown 2018: S. 301.

¹³¹ Barnidge et al. 2019.

¹³² Brown 2018: S. 306.

¹³³ Quent 2018: S. 48.

¹³⁴ Marwick 2010.

Leichte Zugänglichkeit. Menschen können jederzeit ortsunabhängig und praktisch kostenlos etwas zur digitalen Öffentlichkeit beitragen (vorausgesetzt Internetzugang, Kompetenzen und länderabhängige digitale Infrastruktur sind vorhanden). Dies kontrastiert zu früher, wo die Gatekeeper der klassischen Medien die Bereitstellung von Information nach ethischen Richtlinien kontrolliert haben. Die niedrigen Zugangsbarrieren zum Internet haben schon vor 20 Jahren Hoffnungen auf eine Demokratisierung des öffentlichen Diskurses geweckt. Diese ist teilweise tatsächlich eingetreten – mitsamt ihrer Kehrseite. Es äussern sich auch Nutzende, deren Meinungen von üblichen Kommunikationsnormen abweichen, weil sie inhaltlich falsch, strafbar oder «qualitativ schlecht» sind. Vom Aufbruch dieses «Kults der Amateure»¹³⁵ haben auch Hassgruppen profitiert: Ein US-amerikanischer Anführer des Ku Klux Klan erklärte im Jahr 2008: «We don't really need the media any more (...) the only thing we need is the internet».¹³⁶

Sofortige Veröffentlichung. Wollte man im vor-digitalen Zeitalter rassistisches Gedankengut vor ein grösseres Publikum tragen, musste man meist warten, bis aufwändige Flugblätter gedruckt waren oder eine Demonstration stattfand. Es gab also eine Zeitverzögerung zwischen gedanklichem Impuls und Veröffentlichung: Man schief gezwungenermassen nochmals über einen Auslöser und betrachtete diesen am nächsten Tag vielleicht gelassener. Heute dagegen ist diese Zeitverzögerung stark geschrumpft. Der innere Stammtisch wird über Social Media innert Sekunden einem Publikum ungefiltert zugänglich gemacht. Es ist daher vorstellbar, dass Social Media spontane, gefühlsmässige und ungefilterte Ausdrucksweisen von rassistischer Hassrede langfristig besonders fördert.¹³⁷

Distanz und wahrgenommene Unsichtbarkeit. Natürlich konnte man sich schon im vor-digitalen Zeitalter aus der Distanz und mit der Zuversicht, unsichtbar zu bleiben, rassistisch äussern (z.B. in Briefen, Flyern, grossen Demonstrationen). Der Internetkommunikation ist eine Distanz zum Gegenüber jedoch inhärent. Sprechende und Publikum sind durch die Technik getrennt und interagieren oft asynchron. Der Austausch ist entfremdeter und unpersönlicher.¹³⁸ Nonverbale Mimik und Gestik, die antisoziales Verhalten analog teilweise in Schach halten, werden herausgefiltert. Ob das Gegenüber emotional verletzt ist, kann kaum abgelesen und somit nicht gespiegelt werden. Dies macht Menschen weniger empathisch. Beim Online-Chatten reduzieren sich dementsprechend aggressive Äusserungen, wenn Augenkontakt (z.B. durch Webcam) hergestellt wird.¹³⁹ Denn je mehr (non-verbale) Information über das Gegenüber vorliegt, desto kleiner wird die Fläche, in die man eigene Stereotypen und Vorurteile projizieren kann. Im Internet wird also enthemmter kommuniziert, und damit werden auch rassistische Gedanken weniger zensiert.¹⁴⁰

Anonymität / Klarnamen. Ob Anonymität Hassrede verstärkt, ist kontextabhängig. Grundsätzlich können Nutzende technisch anonym sein, d.h. die faktische Identifizierung der Person ist durch Pseudonyme, falsche persönliche Angaben oder das Verbergen von IP-Adresse nicht möglich. Oder sie sind sozial anonym, d.h. andere Nutzende nehmen die Person ohne individuelle Identität wahr, trotz eventuellem Klarnamen.¹⁴¹ Die Anonymität kann Hassrede fördern. Denn es ist schwieriger, die Autor*innen zu sanktionieren. Im sozial anonymen Zustand orientieren sich Nutzende zudem stärker an Gruppen. Ist die vorherrschende Norm in einem Forum rassistisch, werden auch anonyme Neulinge rassistischer kommentieren.¹⁴² Gleichzeitig können Verbreiter*innen von Online-Hassrede auch Klarnamen präferieren, insbesondere wenn sie Anerkennung und Vertrauen suchen, von ihrem Handeln moralisch überzeugt sind oder von einer Community unterstützt werden. Studien finden daher unterschiedliche Zusammenhänge zwischen Anonymität und Aggression bzw. Hassrede.¹⁴³ Die Klarnamenspflicht im südkoreanischen Internet beispielsweise reduzierte Hassrede nur bei spezifischen Gruppen.¹⁴⁴

Schwache Sanktionsmechanismen. Die Wahrscheinlichkeit, im Internet für rassistisches Gedankengut bestraft zu werden, kann aktuell als gering eingeschätzt werden. Sind Täter*innen technisch anonym, können sie sich unter Umständen vor einer Identifikation durch die Strafbehörden schützen. Da die

¹³⁵ Keen 2007.

¹³⁶ Garland, Cult Education Institute vom 27. März 2008: [Klan's new message of cyber-hate](#).

¹³⁷ Brown 2018: S. 306.

¹³⁸ Brown 2018: S. 298 ff.

¹³⁹ Lapidot-Lefler und Barak 2012.

¹⁴⁰ Suler 2004.

¹⁴¹ Hayne und Rice 1997: S. 432.

¹⁴² Christopherson 2007: S. 3048.

¹⁴³ Mondal et al. 2018; Rost, Stahel und Frey 2016.

¹⁴⁴ Cho, Kim und Acquisti 2012.

angegriffenen Personen oder Gruppen physisch abwesend sind, können sich Täter*innen möglichen Konflikten oder Racheakten entziehen, indem sie sich «ausloggen». Zudem fehlt die Präsenz direkter Autoritäten; diese Rolle wird gemeinhin anderen Internetnutzenden zugeschrieben. Ihre möglicherweise missbilligenden Blicke sind nicht sichtbar für Täter*innen. Nutzende verurteilen rassistische Kommentare im öffentlichen Internet zudem häufig nicht durch Gegenkommentare. Werden Nutzende somit nicht zur Verantwortung gezogen, sinken psychologische und soziale Hemmschwellen. Auch kann der Eindruck eines rechtsfreien Raumes entstehen. Beides fördert das unkontrollierte Ausdrücken rassistischen Gedankenguts.¹⁴⁵

Aufmerksamkeit durch Simplifizierung und Dramatisierung. Hassrede gehört im Internet zu den Inhalten, die oft mit Aufmerksamkeit belohnt werden. Denn im informationsüberladenen Internet besteht grundsätzlich ein starker Wettbewerb um die Aufmerksamkeit des Publikums.¹⁴⁶ Wenn Tweets, Blogs und Posts auf Social Media gelesen werden sollen, ist der Unterhaltungswert entscheidend. Konflikte, Skandale oder Verstöße gegen rechtliche und moralische Normen sind besonders geeignet, Stimmungen anzusprechen und Gefühle auszulösen. Dazu gehören auch Beiträge, die emotional, extrem, simplifizierend und polarisierend sind. Ähnlich wie simpel gestrickte Gerüchte oder Verschwörungstheorien entspricht rassistische Hassrede diesem vereinfachten Kommunikationsmuster und verspricht mehr «Teilen»- und «Gefällt mir»-Angaben.¹⁴⁷

5.4.2 Architektur von Online-Plattformen

Hinter den diskutierten Kommunikationsbedingungen stehen Plattform-Architekturen, die Nutzenden die Möglichkeiten und Restriktionen für ihr Handeln vorgeben. Dies kann offensichtlich (z.B. durch Anonymitätsoptionen) oder weniger offensichtlich geschehen (z.B. welche Inhalte vom Algorithmus zuoberst auf der Webseite präsentiert werden und welche nicht). Nutzende können für ihr Verhalten auch belohnt werden, beispielsweise indem Algorithmen ihre Äusserungen aufgrund der gewählten Formulierung stark verbreiten. Sie werden auch bestraft, beispielsweise indem ihre Beiträge mithilfe von Meldeoptionen gelöscht werden können. Dieser «materialistischen» Sicht entsprechend entsteht Rassismus aus der komplexen soziotechnischen Interaktion zwischen Menschen, der Funktionsweise von Technologien und den Absichten der Plattform-Designer.¹⁴⁸ Die bestehende technische Umwelt, so argumentieren Forschende zunehmend, sei ein idealer Nährboden für Rassismus.¹⁴⁹ Einige zentrale Faktoren werden hier vorgestellt.

Vielfältige technische Verbreitungselemente. Social Media stellen ein reichhaltiges Spektrum von Elementen zur Verfügung, womit rassistisches Gedankengut produziert und zirkuliert werden kann. Dies sind klassischerweise Texte, Kommentare, Bilder, Fotos, Symbole, Videos und Musik. Dabei können Plattformvorgaben, wie eine maximale Zeichenanzahl von 280 auf Twitter, die Kommunikation subtil steuern, indem sie beispielsweise auf Simplifizierung drängen und längere Argumentationen erschweren. Grundsätzlich hilft die Vielfalt dieser Elemente, rassistisches Gedankengut und die unterschiedlichen Facetten von Rassismus optimal und überzeugend zu vermitteln: Texte helfen zur Übermittlung der rassistischen Ideologie und emotionale Videos zum Schüren von Ängsten. Laut der Medienreichhaltigkeitstheorie ist dies insbesondere gegeben bei personalisierten Inhalten, interaktiver Rückmeldung (z.B. in Chaträumen) und vielfältigen Hinweisreizen (z.B. YouTube-Videos).¹⁵⁰ Mithilfe des formatreichen Angebots können rassistische Inhalte schnell und unkompliziert auf Zielgruppen zugeschnitten werden, z.B. Computerspiele und Musik für Kinder und Jugendliche.¹⁵¹ Auch «Gefällt mir»-Angaben oder das Teilen von Hyperlinks zu rassistischen Webseiten können zur Verbreitung dienen.¹⁵² Selbst Hashtags (#) werden instrumentalisiert: Der Kampagnen-Hashtag #MAGA («Make America Great Again») von US-Präsident Donald Trump wurde von rassistischen Gruppen genutzt, um

¹⁴⁵ Kaspar 2017: S. 68.

¹⁴⁶ Marwick 2015: S. 138.

¹⁴⁷ Hine et al. 2016.

¹⁴⁸ Murthy und Sharma 2019: S. 196.

¹⁴⁹ Z.B. Matamoros-Fernández 2017.

¹⁵⁰ Daft und Lengel 1986; Bliuc et al. 2018: S. 81.

¹⁵¹ Siehe als Beispiel [Stormfront for Kids](#).

¹⁵² Ben-David und Matamoros-Fernández 2016.

sich orts- und zeitunabhängig zu organisieren und zu äussern.¹⁵³ Letztlich können auch Kontrollelemente (z.B. «Hassrede-Meldekнопf») über Hassrede / Nicht-Hassrede entscheiden.

Algorithmen. Algorithmen können Rassismus reproduzieren.¹⁵⁴ Beispielsweise liefern manche Suchmaschinen Suchresultate, die dem vorherigen (rassistischen) Nutzerverhalten ähneln, wobei die Nutzenden in ihren (rassistischen) Meinungen bestätigt werden. Ob diese algorithmische Logik im Kontext einer zunehmenden Personalisierung von Inhalten jedoch vielzitierte «Filterblasen»¹⁵⁵, d.h. algorithmisch gebildete Echokammern produzieren, ist umstritten. Lim¹⁵⁶ stellte in den Social-Media-Debatten während Wahlen in Indonesien fest, dass nicht Algorithmen alleine «algorithmische Enklaven» formieren. Vielmehr würden sich Nutzende und Algorithmen in der Sortierung, Klassifizierung und Hierarchisierung von Menschen, Informationen und politischen Präferenzen gegenseitig prägen. In den resultierenden Enklaven wird dann Rassismus ausgelebt. Ob Empfehlungsalgorithmen Nutzende grundsätzlich davon abhalten oder sie motivieren, extremistische Inhalte zu besuchen, ist umstritten.¹⁵⁷ Auch ein abschliessendes Urteil über den konkreten Einfluss von Algorithmen auf Rassismus steht aus.

Netzwerkstruktur. Netzwerke werden genutzt, um Hassrede besser zu verbreiten und «Hasskampagnen» zu koordinieren. Entsprechend der prinzipiellen Logik «Freiheit über Kontrolle» eliminiert das Internet räumliche Grenzen durch grossflächige Vernetzung.¹⁵⁸ Hassgruppen nutzen dies, um geografisch und sozial isolierte Individuen zu erreichen und die kritische Masse, die lokal fehlt, zu kompensieren. So tragen sie zu einer «globalen rassistischen Subkultur» bei.¹⁵⁹ Aufgrund ihrer dichten Vernetzung verbreiten sich ihre Inhalte besser. Eine Analyse¹⁶⁰ von 21 Millionen Posts von ca. 340'000 Nutzenden im Forum Gab kam zum Ergebnis, dass sich Inhalte derjenigen, die Hassrede verbreiteten (verglichen mit solchen, die keine Hassrede verbreiteten), rascher und weiter verbreiteten und ein grösseres Publikum erreichten. Solche dichten Hass-Netzwerke fördern wiederum Polarisierung, da skeptische Personen die Netzwerke zunehmend verlassen. In den entstehenden Echokammern sind die Gleichgesinnten alternativen Meinungen immer seltener ausgesetzt.¹⁶¹ Auch werden «Hasskampagnen» initiiert – denn nicht alles, was nach spontanen Empörungswellen einer breiten Öffentlichkeit aussieht, ist das auch. Bekannt dafür sind rechtsextreme «Trollfabriken».¹⁶² Reconquista Germanica beispielsweise hatte in der Vergangenheit mehrere tausend Netzwerkmitglieder und Trolle eingesetzt, um punktuell Online-Diskurse zu (zer-)stören. Solche «Troll-Armeen» agieren auch im Namen etablierter politischer Akteur*innen, wie dies der Fall der regierenden AKP (Partei für Gerechtigkeit und Aufschwung) in der Türkei zeigt.¹⁶³ Inwieweit organisierte Netzwerke auch in der Schweiz bestehen, ist bisher nicht bekannt. Erfahrungsberichte deuten jedoch auf Vernetzungen hin.¹⁶⁴

Metrische Manipulation. Mit der pluralistischen Öffnung der Medienlandschaft können nun auch rassistische Netzwerke ihre Inhalte zahlenmässig optimieren.¹⁶⁵ Sie machen es manchen Werbenden und politischen Akteur*innen gleich und kaufen sich «Likes» («Gefällt mir»-Angaben) oder nutzen Fake-Profiles (d.h. Profile mit gefälschten Angaben) oder automatisch generierte Inhalte mittels Computerprogrammen (Bots). Sie kämpfen im digitalen Wettbewerb um Follower und Sichtbarkeit, um politisch wirksam zu sein. Laut Fielitz und Marcks¹⁶⁶ spielt diese Möglichkeit dem radikalen Pragmatismus extremistischer und rassistischer Gruppen in die Hände. Diese fühlten sich ethischen Richtlinien weniger unterworfen und pflegten eine instrumentale Beziehung zur Wahrheit – die auch «verbogen» werden könne, solange damit Einfluss und Macht erreicht wird. Fake-Profiles und Bots werden insbesondere bei Hasskampagnen eingesetzt, um zahlenmässig Diskurse zu manipulieren und

¹⁵³ Eddington 2018.

¹⁵⁴ Daniel 2018.

¹⁵⁵ Pariser 2011.

¹⁵⁶ Lim 2017: S. 422.

¹⁵⁷ Für eine beispielhafte Studie, siehe Ledwich und Zaitsev 2019.

¹⁵⁸ Klein 2017.

¹⁵⁹ Perry und Olsson 2009: S. 185; Fielitz und Marcks 2019: S. 9.

¹⁶⁰ Mathew et al. 2019.

¹⁶¹ Marwick und Lewis 2017: S. 18; Sunstein 2000.

¹⁶² Kreissel et al. 2019; Hine et al. 2016.

¹⁶³ Bulut und Yörük 2017.

¹⁶⁴ Der Verein #Netzcourage beobachtet auf Social Media-Plattformen rund ein Dutzend relevante, Schweiz-bezogene Gruppen mit 800 bis 2000 Nutzenden, die sich teilweise stark überlappen. Interview mit Jolanda Spiess-Hegglin im Februar 2020.

¹⁶⁵ Fielitz und Marcks 2019: S. 12.

¹⁶⁶ Fielitz und Marcks 2019: S. 12.

andere Meinungen zu marginalisieren. Der Einsatz vieler Fake-Profile kann rassistisch unterlegte Hashtags zudem «trendig» machen – was die Wahrscheinlichkeit erhöht, dass sie von klassischen Medien aufgenommen und dadurch plattformübergreifend sichtbar werden. Eine lautstarke Minderheit kann dadurch eine grosse Mehrheit vortäuschen. Eine Auswertung von tausenden Kommentaren und über einer Million «Gefällt mir»-Angaben auf Facebook¹⁶⁷ kam zum Schluss, dass die aktivsten fünf Prozent der Nutzenden 50 Prozent aller «Gefällt mir»-Angaben von Hassbeiträgen generierte und das aktivste Prozent gar 25 Prozent. Und da Menschen sich gerne Mehrheiten anschliessen, können solche manipulierten Meinungen den Meinungsbildungsprozess verzerren.

Inhaltliche Manipulation. Neben Metriken werden auch Nachrichten inhaltlich manipuliert, man denke nur an Fake News und Verschwörungstheorien. Diese Triebfedern von Hassrede bieten simple Lösungen für komplexe Probleme an und greifen dabei auf abwertende Stereotypen zurück; insbesondere Fake News sollen verunsichern und desorientieren.¹⁶⁸ Sie zirkulieren insbesondere in einem alternativen Nachrichtensystem, aufgebaut durch rechtsextreme Kreise. Sie können auch in die breite Gesellschaft gelangen, wie das Gemenge an Fake News und rassistischer Hassrede gegen Chines*innen im aktuellen Bedrohungsklima von Corona zeigt (siehe Beispiel in Grafik 6).¹⁶⁹ Jede*r kann grundsätzlich manipulierte Inhalte eigenständig generieren, beispielsweise Verschwörungsvideos von Amateurfilmer*innen auf YouTube. Inhaltlich manipulierte Beiträge können erfunden oder real sein, jedoch sind sie stark in eine bestimmte Richtung verzerrt und oft mit fehlerhaften Statistiken unterlegt.¹⁷⁰ Sozial konstruierte Figuren (z.B. in Deutschland der «belästigende Ausländer»), Personalisierung (z.B. Parolen wie «Merkel muss weg») und Moralisierung kultivieren dabei eine «aggressive Affektivität».¹⁷¹ Strategisch ergibt dies Sinn: Laut Studien verbreiten sich Beiträge in Social Media weiter und rascher, wenn sie moralisch-emotionale Wörter und Falschinformationen beinhalten.¹⁷² Da sie mehr Klicks generieren, werden sie von Algorithmen potentiell priorisiert, weil sie nach der Logik der Monetarisierung sozialer Interaktionen den Internet-Unternehmen mehr Einnahmen versprechen.



Grafik 6. Online-Kommentar auf 20min.ch im März 2020 (gemeldet bei GRA).

Tarnung. Rassistisches Gedankengut kann zirkulieren, ohne dass es von Nutzenden oder automatisierten Erkennungsalgorithmen sofort als solches erkannt wird. Durch das neue Gewand sollen solche Inhalte nicht mit dem traditionellen Rechtsextremismus assoziiert werden und in einer medienkompetenten Gesellschaft überzeugender wirken. Bewegungen wie der Islamische Staat (IS) oder die Alt-Right-Bewegung sind von «schwarzer Propaganda» der direkten Hassrede zu «grauer Propaganda» übergegangen, d.h. Hassrede verpackt in einer verschlüsselten Sprache und in Anlehnung an eine satirisch-ironische Internetkultur.¹⁷³ Besonders populär sind Memes.¹⁷⁴ Sie holen das Publikum emotional ab und übermitteln Informationen schnell. An den «Nazi-Ponys» (Grafiken 7 und 8) zeigt sich, wie attraktive Inhalte für ahnungslose Rezipient*innen (hier Kinder) mit Nazisymbolik kombiniert werden – und von ihnen nichtsahnend weiterverbreitet werden können.¹⁷⁵ Mit humoristischen Elementen können zudem sozial nicht akzeptierte Inhalte im Deckmantel von Witzen verbreitet

¹⁶⁷ Kreissel et al. 2019: S. 12.

¹⁶⁸ Fielitz und Marcks 2019: S. 11.

¹⁶⁹ Priebe, FAZ vom 3. Februar 2020: [Wie Rassisten das Coronavirus für sich nutzen](#).

¹⁷⁰ Lanzke et al. 2013: S. 25 ff..

¹⁷¹ Fielitz und Marcks 2019: S. 10.

¹⁷² Z.B. Brady et al. 2017; Vosoughi, Roy und Aral 2018.

¹⁷³ Klein 2017: S. 27.

¹⁷⁴ Die «Memetische Kriegsführung» wird von Gruppen auch explizit als Strategie beworben, wie im folgenden «Strategiemanual» nachzulesen ist: Generation D vom 18. Mai 2017: [Das Shitposting 1×1](#).

¹⁷⁵ Winter 2019: S. 51.

The image consists of two side-by-side panels. The left panel features a brown pony with a black mane and tail, wearing a tan German military uniform. It has a black armband with a white swastika on its right arm and a black belt with a silver buckle. The pony is standing on a black background. The right panel features a blue winged pony with a yellow and red mane and tail, wearing a tan German military uniform. It has a black armband with a white swastika on its right arm and a black belt with a silver buckle. The pony is standing on a light blue background with a large white swastika on a red shield behind it.

[illegible]

Mainstreaming. Obwohl mit der Tarnung verwandt, ist «Mainstreaming» eher auf der strukturellen Ebene anzusiedeln. Rassistische Akteur*innen können von der Legitimation bestehender Plattform-Architekturen und Verknüpfungen des Internets profitieren, indem sie ihre Inhalte danach ausrichten und sich darin integrieren.¹⁸¹ Dies kann ganz ohne das Zutun dieser Akteure vonstatten gehen. Informationssuchende erreichen beispielsweise nicht-rassistische wie auch rassistische Webseiten oft durch den gleichen Kanal: Suchmaschinen präsentieren sie nebeneinander und somit gleichberechtigt. Rassistische Akteure können auch aktiv das Auftreten bzw. Design gemeinhin vertrauenswürdiger

181 Klein 2017.

Quellen übernehmen, sei dies von Wikipedia (z.B. [Metapedia](#)¹⁸²), wissenschaftlichen Institutionen (z.B. Webseiten von Holocaust-Leugnern, die auf den ersten Blick wie objektive und wissenschaftlich seriöse Forschungszentren wirken) oder zeitgenössischer Ästhetik (z.B. modernes Webseitendesign der [American Nazi Party](#)).¹⁸³ Auch Webseiten-Namen können in die Irre führen: Die Blogs [Council of Conservative Citizens](#) oder [world peace](#) werden entgegen möglicher Erwartungen dem Weissen Nationalismus zugeordnet.¹⁸⁴ Rassistisches Gedankengut wird dadurch aus der extremen Nische subtil in die digitale, breite Öffentlichkeit getragen.

¹⁸² Beschrieben als «rechtsextreme Online-Enzyklopädie» von [Wikipedia](#).

¹⁸³ Klein 2017: S. 28.

¹⁸⁴ Klein 2017: S. 53.

6 NEGATIVE FOLGEN VON RASSISTISCHER ONLINE-HASSREDE

6.1 Vorbemerkung zu digitaler Verstärkung

Es ist nicht einfach, die kurz- und langfristigen Folgen von Online-Rassismus auf die Opfer, die Zuschauenden und die breite Gesellschaft zu erfassen. Die bestehenden Studien legen jedoch nahe, dass diese vergleichbar sind mit denen des Offline-Rassismus. Der digitale Raum kann negative Folgen zudem verstärken, da er das Zusammenkommen unterschiedlicher Faktoren, die belastend sein könnten, wahrscheinlicher macht.¹⁸⁵ So können physische Angriffe durch ihre Unmittelbarkeit zwar mehr Stress auslösen, aber virtuelle Angriffe erlangen schneller eine grössere Öffentlichkeit, was die öffentliche Scham der Betroffenen verstärkt. So berichten Jugendliche in der Schweiz, öffentliches – verglichen mit nicht-öffentlichem – (Cyber-)Mobbing als besonders gravierend zu empfinden.¹⁸⁶ Online können Betroffene zudem innert geringer Zeit mit exponentiell steigenden Hassnachrichten konfrontiert werden. Letztlich können sich Betroffene räumlich und zeitlich den Angriffen kaum entziehen. Während Offline-Hassrede, sei es in Graffitis, Zeitungen oder physischer Interaktion, irgendwann durch Übermalen, Entsorgen oder die abnehmende Erinnerung verschwinden, sind Online-Beiträge rund um die Uhr und überall abrufbar. Sobald sie ins Internet gestellt sind, kann man sie kaum definitiv löschen, da sie jederzeit gespeichert und auf anderen Plattformen veröffentlicht werden können. All dies kann ein Gefühl der Unkontrollierbarkeit und Traumatisierung auslösen.

6.2 Unmittelbar Betroffene

Rassistische und allgemeine Online-Hassrede kann für die Betroffenen unmittelbar emotional belastend sein, wobei sich dies auch in körperlichen Symptomen niederschlägt. So gaben interviewte Muslim*innen in Grossbritannien an, dass Online-Angriffe ihre Verletzlichkeit und Unsicherheit verstärken und sie darauf mit Depressionen, Ängstlichkeit und Anpassungen an die Mehrheitsgesellschaft (z.B. Kopftuch weglassen, Bart rasieren) reagieren – in der Befürchtung, dass Online-Angriffe auch physisch ausgelebt werden.¹⁸⁷ Auch gaben in einer deutschen Umfrage¹⁸⁸ zwei Drittel der Personen, die persönlich mit Online-Hasskommentaren angegriffen wurden, an, darunter gelitten zu haben. Sie nannten psychische Probleme wie Abgeschlagenheit oder Lustlosigkeit, Angst und Unruhe, Depressionen und Probleme mit dem eigenen Selbstbild. Auch darauffolgende Probleme bei der Arbeit oder der besuchten Bildungseinrichtung wurden erwähnt. Gemäss Studien aus den USA treten solche Folgen unabhängig von möglichen gleichzeitigen Offline-Angriffen. Die Effekte sind somit tatsächlich auf Angriffe im Online-Raum zurückführbar.¹⁸⁹ Eine grössere Bandbreite emotionaler Reaktionen wurde in einer australischen Umfrage bei 103 direkt von Online-Rassismus Betroffenen festgestellt.¹⁹⁰ Sie empfanden Wut und Frustration (48%), Abscheu (38%), Amüsiertsein (25%), Machtlosigkeit oder Depression (18%), Beschämung (13%) oder Mitgefühl für die Täter*innen (12%). Ein kleiner Anteil (10%) berichtete von Magenproblemen, Kopfschmerzen und Herzrasen. Wie reagierten die Betroffenen auf diese Angriffe? Zwei Drittel gaben an, sich aktiv zu wehren, beispielsweise durch einen Gegenkommentar, das Melden des Beitrags oder das Blocken der Täter*innen. Jede siebte Person ignorierte dagegen den Beitrag.

Längerfristig kämpfen die Betroffenen beispielsweise mit Leistungsabfall und digitalem Rückzug. In einer Langzeitbefragung¹⁹¹ unter amerikanischen Jugendlichen berichteten diejenigen, die zunehmenden Online-Rassismus gegen sich beobachteten, gleichzeitig sinkende schulische Motivation. Dass Betroffene sich selbst «zum Schweigen zwingen», d.h. als Reaktion auf Angriffe weniger Inhalte im Internet veröffentlichen und somit ihre öffentliche Sichtbarkeit verringern, ist dagegen insbesondere von (halb-)öffentlichen Personen bekannt. Dies verwundert nicht, da diese Personengruppe überproportional von Online-Hassrede betroffen ist.¹⁹² Beispielsweise schreiben

¹⁸⁵ Brown 2018: S. 306 ff.

¹⁸⁶ Sticca und Perren 2013.

¹⁸⁷ Awan und Zempi 2015: S. 37.

¹⁸⁸ Geschke et al. 2019: S. 27.

¹⁸⁹ Tynes et al. 2008.

¹⁹⁰ Jakubowicz et al. 2017: S. 77-79.

¹⁹¹ Tynes, Torro und Lozada 2019.

¹⁹² Isbister et al. 2018.

Wissenschaftler*innen, die ihre Forschung aktiv auf Social Media verbreiten, dass sie dort rassistisch angegriffen werden und mit dem digitalen Rückzug ringen.¹⁹³ Auch Journalist*innen in der Schweiz haben angegeben, als Reaktion auf Online-Angriffe – unabhängig ob rassistisch oder nicht – ihr Publikum stärker zu vermeiden.¹⁹⁴ Diese Ergebnisse deuten darauf hin, dass Online-Hassrede Menschen in der Ausübung ihrer Menschenrechte wie der Meinungsäusserungsfreiheit, der Religionsfreiheit oder der Sicherheit der Person einschränken kann.

Schliesslich gibt es Indizien, dass Online-Hassrede manche Gruppen, insbesondere jüngere Personen, Frauen und Menschen mit mehreren Minderheitsmerkmalen stärker belastet.¹⁹⁵ Junge Menschen sind also nicht nur häufiger involviert, sondern leiden auch stärker darunter. Ob die stärkere Belastung für Frauen mit häufigeren und inhaltlich stärkeren Angriffen zusammenhängt oder mit anderen unterliegenden Faktoren¹⁹⁶, ist aktuell nicht abschliessend geklärt.

6.3 Publikum und Gesamtgesellschaft

Rassistische Online-Hassrede kann sich auch auf unbeteiligte Zuschauende, also das breite Publikum, und längerfristig auf die Gesamtgesellschaft auswirken. Hassrede kann soziale Spannungen verstärken – offline wie online. Sie gilt seit jeher als frühes Warnzeichen für politische Instabilität und Gewalt. Dies belegen Studien zum Effekt von Radio auf die Stärke der Gewalt im Genozid in Ruanda oder von Radiopropaganda in Nazi-Deutschland auf anti-semitische Gewalt.¹⁹⁷ Sie dient als «Nährboden für physische Übergriffe und Einschüchterungsaktionen», wobei «durch stereotypisierende Verunglimpfungen (...) die 'Entmenschlichung' vorbereitet [wird] und die Barriere zur Begehung von strafbaren Handlungen sinkt».¹⁹⁸ Diese «Vorbereitung» kann dahingehend geschehen, dass rassistische Hassrede, sofern sie sich im Internet verbreitet und selten gemeldet oder sanktioniert wird, gesellschaftsfähig werden könnte.¹⁹⁹ Beim Publikum könnte der Eindruck entstehen, dass «ja alle hassen» und «man nicht dafür bestraft» würde: Hassrede würde somit als Norm wahrgenommen, was wiederum motiviert, selbst auch öffentlich Hassrede auszuüben.

Somit kann Online-Hassrede «anstecken»: Sie kann das Publikum feindseliger denken und handeln lassen. Dies ist ersichtlich in Experimenten, in denen Diskussionen «fabriziert» werden, wobei manchen Teilnehmenden Hassrede-Kommentare präsentiert werden und anderen neutrale Kommentare.²⁰⁰ Teilnehmende in der ersten Gruppe hatten nach der Diskussion grössere Vorurteile und stärker polarisierte Meinungen zum diskutierten Thema. Zudem schätzten sie die breite Öffentlichkeit danach als polarisierter ein. Dies kann Personen wiederum motivieren, selbst stärker polarisierte Meinungen zu vertreten. Auch waren diese Teilnehmenden weniger bereit, Geld an Flüchtende zu spenden. Längerfristige negative Effekte zeigten sich auch in weiteren Experimenten: Haben Teilnehmende eines polnischen Experiments regelmässig Online-Hassrede gegen Minderheitengruppen (z.B. Muslim*innen) beobachtet, wurden sie gegenüber Hassrede gegen diese Gruppen desensibilisiert. Dabei fingen sie auch an, sich gegenüber diesen Gruppen innerlich zu distanzieren, sie weniger wertzuschätzen und vorurteiliger wahrzunehmen.²⁰¹ Hassrede, Online-Selbstjustiz und Desinformation sind letztlich eng verknüpft. Unbewiesene Vermutungen und Falschbeschuldigungen können im Internet zirkulieren und von Aufrufen zur Gewalt (z.B. Erhängen, Anzünden) begleitet werden. Aus der Praxis sind unzählige Fälle bekannt, in denen die Reputation der angeprangerten Person nachhaltig zerstört wurde – trotz Beweisen für ihre Unschuld. All dies trägt zu einem negativen Denk- und Diskussionsklima bei.

Zudem kann Hassrede die Diversität in der digitalen Öffentlichkeit gefährden. Es ziehen sich nicht nur Betroffene digital zurück, sondern auch Mitglieder des breiten Publikums. Fast die Hälfte aller

¹⁹³ Barlow und Awan 2016.

¹⁹⁴ Stahel und Schoen 2019.

¹⁹⁵ Z.B. Bucchianeri et al. 2014; Geschke et al. 2019.

¹⁹⁶ Diese Studien finden z.B. ähnlich viele Online-Angriffe gegen Frauen wie Männer: Celik 2019; Duggan 2017; Stahel und Schoen 2019. Die folgende Studie findet bei Frauen mehr Belastung durch sexistische Angriffe: Fox und Tang 2017.

¹⁹⁷ Gagliardone et al. 2016; Yanagizawa-Drott 2014; Adena et al. 2015.

¹⁹⁸ Humanrights.ch vom 6. Februar 2017: [Hassreden: Die Grenzen der Meinungsäusserungsfreiheit](#).

¹⁹⁹ Dies trifft auf Zuschauende zu, siehe Jakubowicz et al. 2017: S. 79; Landesanstalt für Medien NRW 2018: S. 4.

²⁰⁰ Siehe beispielhafte Experimente hier: Anderson et al. 2014; Hsueh et al. 2015; Ziegele, Koehler und Weber 2018.

²⁰¹ Soral, Michal und Winiewski 2018.

Internetnutzenden in einer deutschen Umfrage²⁰² berichtete, sich aufgrund (drohender) Hasskommentare seltener in politischen Online-Diskussionen zu beteiligen. 15 Prozent gaben an, wegen Hassrede ihr Online-Profil schon einmal deaktiviert oder gelöscht zu haben. Interessanterweise ziehen sich jedoch nicht alle Personengruppen mit gleicher Wahrscheinlichkeit zurück. Wähler*innen der «Alternative für Deutschland» (AFD) gaben in einer weiteren Umfrage²⁰³ signifikant seltener an, sich aufgrund von Hasskommentaren aus Diskussionen zurückzuziehen, als Wähler*innen anderer Parteien. Dies kann einen Hinweis darauf geben, welche Personengruppen sich im Falle von längerfristig zunehmenden Hasskommentaren aus der Öffentlichkeit zurückziehen und welche nicht.

Schliesslich hängt Online-Hassrede mit Offline-Hassverbrechen zusammen, wobei Kausalitäten nicht klar feststellbar sind. Von mehreren Attentätern ist bekannt, die sich im Internet radikalisiert haben: beispielsweise der 21-jährige Attentäter, der 2015 in Charleston (USA) neun Afroamerikaner in einer Kirche erschoss. Er zitierte die als rassistisch eingeordnete Gruppe [Council of Conservative Citizens](#) als seine erste rassistische Inspiration.²⁰⁴ Oder Anders Breivik, ein rechtsextremer und islamfeindlicher [Stormfront.org](#)-Nutzer, der in Norwegen 77 Menschen umbrachte. Auch die Täter der rechtsextremen Anschläge in Pittsburgh und San Diego und im neuseeländischen Christchurch in den Jahren 2018 und 2019 waren Mitglieder rechtsextremer Online-Subkulturen.²⁰⁵ Datenreiche Erkenntnisse aus der Wissenschaft unterstützen diese Beobachtungen. So zeigt eine deutsche Langzeitstudie, dass anschwellende Anti-Flüchtlingsstimmungen auf der Facebook-Seite «Alternative für Deutschland» mit Brandstiftungen und physischen Angriffe gegen Flüchtende einhergehen.²⁰⁶ Um den Effekt des Internets auf Offline-Hassverbrechen in den USA zu klären, wurde zwischen 1999 und 2008 die Anzahl der Breitbandanbietenden mit der Anzahl rassistischer Hassverbrechen verglichen.²⁰⁷ Es zeigte sich, dass eine Zunahme im Breitbandzugang in einem bestimmten Gebiet zu einer Zunahme von Hassverbrechen führte, wogegen andere Verbrechen wie Einbrüche oder Mord stabil blieben. Dieser Effekt war stärker für Gebiete, in denen mehr Segregation herrschte und mehr rassistische Wörter im Internet gesucht wurden. Laut Autor*innen würde das Internet bestehende Neigungen zu rassistischen Offline-Handlungen verstärken, weil sich Hassrede besser verbreitet und weil sich Gleichgesinnte besser vernetzen. Schliesslich kann auch das «Dog Whistling» (siehe Kap. 5.4.2) öffentlicher Personen Hassverbrechen motivieren. Eine Studie ergab, dass bei einer Zunahme der Twitter-Aktivität (bzw. des «Dog Whistling») derjenigen Mitglieder des US-Repräsentantenhauses, welche die grösste nationalistisch denkende Online-Anhängerschaften haben, auch die Anzahl von Offline-Hassverbrechen anstieg.²⁰⁸ Gemäss der Studie können Politiker*innen ihre Anhängerschaft subtil zu Gewalt aufstacheln, indem sie durch ihre Rhetorik extreme Ideologie bestätigen und diese Gemeinschaften ermächtigen.

Die kurz- und langfristigen Folgen von Online-Hassrede für Betroffene und die Gesamtgesellschaft sind signifikant. Sie sollten in Zukunft auch im Kontext der Schweiz studiert werden.

²⁰² Geschke et al. 2019: S. 28.

²⁰³ Eckes et al. 2018: S. 6.

²⁰⁴ Southern Poverty Law Center, 21. Juni 2015: [The Council of Conservative Citizens: Dylann Roof's gateway into the world of white nationalism](#).

²⁰⁵ Guhl et al. 2020: S. 7.

²⁰⁶ Müller und Schwarz 2019.

²⁰⁷ Chan, Ghose und Seamans 2013.

²⁰⁸ Chyzh, Nieman und Webb 2019.

7 BESTEHENDE GEGENMASSNAHMEN, UMSETZUNG UND WIRKSAMKEIT

Die zusehends erkennbaren Auswirkungen von Online-Rassismus bewegen Entscheidungsträger*innen, Social-Media-Anbieter, Forschende, NGOs und die Zivilgesellschaft in vielen Ländern zur Entwicklung von Gegenmassnahmen. Diese aktive Herangehensweise entspricht dem grossen Handlungsbedarf, der von einer breiten Bevölkerung in Bezug auf Online-Hassrede gesehen wird; zumindest wurde dieser in Umfragen im nahen Deutschland festgestellt.²⁰⁹ Dabei kommen präventive als auch intervenierende Ansätze zum Tragen. Nicht immer ist deren Abgrenzung klar. Beispielsweise kann «Gegenrede» im weitesten Sinne (siehe Kap. 7.5.4) längerfristig zur Schaffung eines gewaltreduzierten Diskurses eingesetzt werden. Gleichzeitig kann sie auch spezifische Verbreiter*innen von Online-Hassrede situativ und öffentlich konfrontieren.

Das folgende Kapitel gibt einen Überblick über die bestehenden Gegenmassnahmen relevanter Akteur*innen und Institutionen, sowohl in der Schweiz wie im (insbesondere nahen) Ausland. Zuerst wird diskutiert, wie Online-Hassrede vom Rechtssystem, von Social-Media-Anbietern, klassischen Medien und der Forschung adressiert wird. Sie stellen ein strukturelles Umfeld dar; manche haben wichtige Kontrollfunktionen. Ihren Handlungsspielraum im Hinblick auf Online-Hassrede zu verstehen, ist grundlegend, um in einem nächsten Schritt Massnahmen durch die Zivilgesellschaft und ihre Wirksamkeit kritisch zu beleuchten.²¹⁰ Der vorliegende Überblick erhebt jedoch keinen Anspruch auf Vollständigkeit. Zudem werden die hinter den Diskussionen stehenden grösseren Fragen der Haftbarkeit und der Verantwortlichkeit nur angeschnitten, jedoch nicht vollumfassend behandelt. Schliesslich sei auf das grundsätzliche Problem der Aktualität und Flexibilität hingewiesen: Lösungsansätze hinken oft hinter der Gefahr der Online-Hassrede her, sowohl was ihre Menge als auch ihre Geschwindigkeit angeht, da sich die Verfasser*innen dem Internet agil anpassen.²¹¹

7.1 Gesetzgebung und Rechtsprechung zu rassistischer Hassrede

Der nachfolgende Abschnitt bietet eine kurze Übersicht darüber, in welchem rechtlichen Rahmen sich Gesetzgebung und Rechtsprechung zur rassistischen Hassrede bewegen und welche zusätzlichen Fragestellungen sich bei deren digitalen Verbreitung stellen. Eine präzise Darstellung der Rechtslage in der Schweiz wird demnächst 2020 publiziert und auf der Webseite der FRB aufgeschaltet werden. Zudem gibt der [Online-Rechtsratgeber](#)²¹² der FRB Grundinformationen zu Recht und Rechtsmitteln im Falle von rassistischer Diskriminierung.

Die Gesetzgebung zur rassistischen Hassrede muss immer zwei Menschenrechten gerecht werden, die in einem Spannungsfeld zueinanderstehen: die Meinungsäusserungsfreiheit und der Schutz vor Diskriminierung. Alle internationalen menschenrechtlichen Vereinbarungen²¹³ lassen für die Abwägung zwischen diesen beiden Rechten einen gewissen Freiraum, damit jedes Land seinen spezifischen Gegebenheiten Rechnung tragen kann. So wird etwa im anglo-sächsischen Sprachraum die Meinungsäusserungsfreiheit stärker gewichtet als im europäischen Raum. Letzterer wiederum ist tendenziell toleranter gegenüber sexuellen oder gar sexistischen Inhalten als dies zumindest in den USA der Fall ist.

Die einzelnen Staaten entscheiden also bis zu einem gewissen Grad selbst, welche Äusserungen (in Wort, Gestik, Bild oder Ton) sie als so gravierend betrachten, dass sie strafrechtlich verfolgt werden müssen, bei welchen Äusserungen zumindest der privatrechtliche Rechtsweg offenstehen sollte und welche Verantwortung die Medien bzw. Social-Media-Anbieter für die Inhalte übernehmen müssen, die sie veröffentlichen, moderieren oder verbreiten.

Aber nicht nur die Gesetzgebung, sondern auch deren Auslegung im konkreten Fall ist vom zeitlichen und kulturellen Kontext abhängig. Inhalte, die vor zwanzig Jahren noch akzeptabel schienen, sind im

²⁰⁹ Geschke et al. 2019: S. 32.

²¹⁰ Für eine informative Kurzevaluation bisheriger Massnahmen in Deutschland, siehe Dittrich et al. 2020: S. 74.

²¹¹ Baldauf, Ebner und Guhl 2018.

²¹² Pfad: www.frb.admin.ch > Recht und Beratung > [Rechtsratgeber](#).

²¹³ Für die Schweiz relevant sind insbesondere die Menschenrechtskonventionen der UNO (Allgemeine Erklärung der Menschenrechte, Internationaler Pakt über wirtschaftliche, soziale und kulturelle Rechte (Pakt UNO I), Internationaler Pakt über bürgerliche und politische Rechte (UNO Pakt II) sowie die Europäische Menschenrechtskonvention.

heutigen gesellschaftlichen Empfinden klar verpönt. Umgekehrt ist der rassistische Diskurs oder die Hetze gegen einzelne Bevölkerungsgruppen enorm anpassungsfähig und findet immer wieder Mittel und Wege, sich «unter dem Radar» gesetzlicher Schranken weiter zu verbreiten.²¹⁴ Somit muss die Einschätzung eines problematischen Inhaltes immer kontextbezogen erfolgen. Gerade bei codierten, impliziten oder subtilen Formen von rassistischer Hassrede gibt es keine allgemeingültigen Rezepte.

In der Schweiz ist der Schutz vor rassistischer Diskriminierung in der Verfassung verankert (Art. 8 BV). Am stärksten konkretisiert wird dieses Grundrecht in der Rassismus-Strafnorm, die öffentlich geäußerte rassistische Inhalte von einer gewissen Intensität verbietet (Art. 261bis StGB). Bei nicht öffentlichen Äusserungen oder solchen, deren rassistischer und gravierender Charakter nur schwer zu belegen ist, kann strafrechtlich auf die Ehrverletzungsdelikte zurückgegriffen werden (Art. 173, 174 oder 177 StGB). Privatrechtlich können sich Betroffene auf den Persönlichkeitsschutz berufen (Art. 28 ff. ZGB) und beispielsweise mittels Klage die Löschung eines Inhalts fordern. Für Vorfälle in den klassischen Medien bestehen mit dem Schweizer Presserat und der Unabhängigen Beschwerdeinstanz für Radio und Fernsehen (UBI) institutionalisierte Selbstregulierungsorgane, die angerufen werden können (Art. 4 Abs. 1 RTVG).²¹⁵

Diese rechtlichen Mittel gegen rassistische Hassrede sind bekannt, deren Anwendung auf die digitale Hassrede entwickelt sich laufend. Die Rechtsprechung muss beispielsweise beurteilen, was im Internet als «öffentliche Äusserung» gilt, etwa im Fall einer geschlossenen Facebook- oder WhatsApp-Gruppe.²¹⁶ Mit der Vermischung zwischen Autor*in, Produzent*in, Verbreitenden und Nutzenden muss geklärt werden, wer welche Verantwortung für rassistische Inhalte und deren Verbreitung zu übernehmen hat. Werden die Inhalte über klassische Medien digital verbreitet, müssen auch medienrechtliche Bestimmungen berücksichtigt und auf die digitalen Gegebenheiten übertragen werden. Zusätzlich spielt das internationale Recht eine grössere Rolle, denn häufig kommentieren die Nutzenden von einem Land aus, während der Sitz des digitalen Mediums oder der Social-Media-Anbieter in einem anderen Land liegt. Dies kann eine grosse Herausforderung für die Rechtsdurchsetzung sein.

Wie auch bei der rassistischen Offline-Hassrede kann häufig nicht mit rechtlichen Mitteln gegen rassistische Online-Hassrede vorgegangen werden. Vielfach sind, wie erwähnt, die Inhalte so formuliert, dass sie nicht genügend eindeutig als rassistische Diskriminierung im strafrechtlichen Sinn qualifiziert werden können. Und die Hürden für ein zivilrechtliches Verfahren sind hoch.

Immer möglich ist es jedoch, rassistische Inhalte im Internet zu melden. Nebst den erwähnten Beschwerdeinstanzen der klassischen Medien können Inhalte bei Social-Media-Anbietern gemeldet werden. Die Transparenz der Prüfungskriterien (die sich häufig auf Kontext und Rechtsprechung des Landes abstützen, in dem der Anbieter seinen Sitz hat) und die Qualität der teils automatisierten Prüfung müssen heute aber noch kritisch beurteilt werden (siehe mehr dazu in Kap. 7.2). Zudem reagieren die Anbieter auch nicht systematisch auf alle Meldungen – oder können dies angesichts der grossen Menge an Inhalten und Meldungen gar nicht tun. Einige Plattformen haben deshalb ein System von «Trusted Flaggers» eingerichtet. Mit diesem Status werden die Meldungen von Stellen oder Organisationen, die den Status eines Trusted Flaggers haben, prioritär behandelt. Die Fachstelle für Rassismusbekämpfung bietet auf ihrer Webseite eine Übersicht zu den Meldestellen und Trusted Flaggers in der Schweiz ([Link](#)).

7.2 Social-Media-Anbieter

Facebook, Twitter und andere Social-Media-Anbieter tragen zur Regulierung von rassistischer Online-Hassrede primär mit der Moderation ihrer Inhalte bei. Moderation soll die Wahrscheinlichkeit senken, dass alltägliche Nutzende versehentlich Hassrede ausgesetzt sind.²¹⁷ Die Plattformen können Inhalte, die ihre Richtlinien (siehe Kap. 3.1) verletzen, löschen oder in Abstimmung mit nationalen Gesetzgebungen länderspezifisch sperren: Facebook beispielsweise verbietet Naziinhalte in

²¹⁴ Eindrückliches Beispiel dafür ist der Ausdruck «Get back to work!», der sich im Bürgerkrieg von Ruanda von 1994 als eindeutiger Aufruf zum Genozid an den Tutsi etablierte: S. dazu: Schabas, 2001, S. 161.

²¹⁵ Grundsätzlich sind die Medien aufgrund ihrer in der Verfassung verbrieften Freiheit und Unabhängigkeit (Art. 17 BV) selbst dafür zuständig, für die Einhaltung der Erklärung der «Pflichten der Journalistinnen und Journalisten» zu sorgen

²¹⁶ Refaeil und Wiecken 2018.

²¹⁷ Siegel 2020: S. 27 ff.

Deutschland (und sperrt sie dort), aber erlaubt sie in den USA (und schaltet sie dort frei). Plattformen können auch Einzelprofile und ganze Gemeinschaften sperren. Beispielsweise löschte Facebook im Juni 2020 ca. 190 Social Media-Profilen, die mit «White supremacy»-Gruppen verlinkt waren und konkrete Aktivitäten planten.²¹⁸ In diesem Zug löschte Twitter einige rechtsextreme Profile mit vielen Followers. Für die Erkennung von Hassrede setzen Plattformen Algorithmen mit Worterkennungssoftware ein und/oder sie vertrauen auf Nutzende und Beschwerdestellen, die Inhalte melden bzw. «flaggen». Die Mehrheit der Social-Media-Anbieter wie Facebook, YouTube, Twitter, Instagram, Snapchat oder TikTok haben eine Meldefunktion integriert.²¹⁹ Meist interne Moderator*innen entscheiden dann auf Basis von Richtlinien über eine Löschung oder Sperrung. Um Löschungen voranzutreiben, hat die Europäische Kommission in Zusammenarbeit mit Facebook, Twitter, YouTube und Microsoft im Jahr 2016 einen freiwilligen «Code of Conduct on Countering Illegal Hate Speech Online»²²⁰ herausgegeben. Damit haben sich diese Plattformen (und mittlerweile auch andere) öffentlich verpflichtet, Hate Speech (nach der EU-Definition) auf ihren Plattformen zu löschen.

Was spricht für die Regulierung durch Social-Media-Anbieter? Brown²²¹ argumentiert, dass Social-Media-Anbieter ironischerweise sowohl die besten Übermittler als auch die besten Regulierer von Hassrede sind. Ihr zentraler Status erlaubt ihnen die nötige Flexibilität, um heutiger Hassrede bestmöglich zu begegnen. Moderator*innen können durch die ständige Einbindung in die Inhalte Innovationen beim Ausdruck von Online-Rassismus erkennen. Änderungen lassen sich rasch in die Richtlinien übersetzen. Inhalte können zudem theoretisch rasch nach ihrer Veröffentlichung gelöscht werden – auch weil sich die Unternehmen öffentlich nicht für die Löschungen rechtfertigen müssen, im Gegensatz zu Gerichten. Brown fügt ein moralisches Argument hinzu: Diejenigen, die eine Technologie, die schädliche und rechtswidrige Handlungen (wenn auch unabsichtlich) fördern, erfinden und unterhalten, sollen auch verantwortlich sein für die Einführung von Massnahmen zur Verhinderung oder Begrenzung dieser Handlungen, speziell wenn sie von dieser Technologie profitieren.

Wie gut funktioniert die Moderation durch die Plattformen? Grundsätzlich werden diese Prozesse trotz Verbesserungspotential als zunehmend positiv beurteilt. Dittrich et al.²²² stellen fest, dass sich die Umgangsweise der Plattformen mit Online-Hassrede verbessert hat, was die Richtlinien und Meldemöglichkeiten (z.B. Trusted Flagger), die Umsetzung der Richtlinien, die Verstärkung von Gegenrede (z.B. durch Werbung) und die bessere Fortbildung und Arbeitsbedingungen für Moderator*innen angeht. Auch die EU beurteilt die Umsetzung des «Code of Conduct» grundsätzlich als positiv²²³: Anbieter bearbeiteten im Schnitt 89 Prozent der in diesem Rahmen gemeldeten Beiträge innerhalb von 24 Stunden, wobei sie 72 Prozent der Beiträge löschten (was als «zufriedenstellend» eingeschätzt wird). Gleichzeitig könnten Rückmeldungen an Nutzende und Trusted Flaggers verstärkt werden.

Wie gut Moderation Online-Hassrede tatsächlich reduzieren kann, ist nicht abschliessend geklärt. Forschende, Journalist*innen und Politiker*innen machen regelmässig auf Kritikpunkte aufmerksam. Auch wenn beispielsweise der Anteil der gelöschten an allen gemeldeten Beiträgen bekannt ist, bleibt oft unklar, wie gross der Anteil der gemeldeten Hassrede an der gesamten Hassrede ist.²²⁴ Auch wurde die automatisierte, algorithmische Erkennung für ihr Potential kritisiert, neutrale Inhalte oder Gegenrede als Hassrede zu klassifizieren. Sie sei ungenügend an lokale Kontexte und Sprachen angepasst²²⁵ – das Schweizerdeutsch ist ein Paradebeispiel. Unter anderem aus diesen Gründen wurde international wiederholt mehr Transparenz gefordert²²⁶, speziell in Bezug auf die Moderator*innen (Expertise, Anzahl Personen etc.), wobei deren schlechte Arbeitsbedingungen und psychische Belastung in der

²¹⁸ Klepper, ABC News vom 6. Juni 2020. [Facebook removes nearly 200 accounts tied to hate groups.](#)

²¹⁹ Da die Meldesysteme der Anbieter sehr unterschiedlich aufgebaut sind, verlinkt die Stiftung gegen Rassismus und Antisemitismus GRA in ihrem Meldetool direkt zu den Meldestellen der wichtigsten Anbieter (z.B. Facebook, Twitter), Pfad: [www.gra.ch > Bildung > Hate speech > Vorfall melden.](#)

²²⁰ Jourová 2019a.

²²¹ Brown 2018: S. 310.

²²² Dittrich et al. 2020.

²²³ Jourová 2019a; Jourová 2019b. Für eine ähnliche, wenn auch etwas kritischere Bewertung, siehe sCAN. 2019.

²²⁴ Siegel 2020.

²²⁵ Siegel 2020: S. 27.

²²⁶ Siehe dazu mehr in INACH, 19. November 2019: [The state of policy on cyber hate in the EU.](#)

Vergangenheit wiederholt kritisiert wurden.²²⁷ Auch wurde eine transparentere Umsetzung von Moderationsrichtlinien verlangt (z.B. unter Berücksichtigung nationaler Gesetze).

Erste wissenschaftliche Studienergebnisse legen grundsätzlich einen längerfristigen Nutzen von Löschungen und Sperrungen nahe – jedoch mit möglichen Nebenwirkungen. Sperrungen einzelner Hass-Threads auf Reddit brachte viele ehemalige Thread-Nutzende dazu, Reddit im Allgemeinen nicht mehr zu besuchen. Die gebliebenen Nutzenden veröffentlichten signifikant weniger Hassrede und interagierten seltener miteinander.²²⁸ Hassrede verschob sich also nicht einfach in andere Threads. Auch die oft beschworene Abwanderung auf weniger moderierte Plattformen (De-Platforming) ist nur in begrenztem Masse feststellbar.²²⁹ Dies könnte daran liegen, dass gesperrte Nutzende bei jedem neuen Profil Einbrüche in ihren Follower-Zahlen erfahren. Grundsätzlich könnten Sperrungen also sinnvoll sein, sofern sie gut begründet sind und somit «Zensur-Vorwürfe» vorwegnehmen. Jedoch sollten schädliche Nebenwirkungen nicht ignoriert werden. Beispielsweise kann das «De-Platforming» Nutzende dazu drängen, sich auf extremere Plattformen zurückzuziehen, wo sie moderaten Meinungen seltener ausgesetzt sind («Echokammern»). Dies fördert ihre Radikalisierung.²³⁰

7.3 Klassische Medienunternehmen in der Schweiz

Klassische Medien bilden gesellschaftliche Debatten ab und tragen dazu bei.²³¹ Sie spielen daher auch eine zentrale Rolle in der Verhinderung oder Verstärkung von rassistischer Hassrede durch die Veröffentlichung ihrer eigenen Inhalte und durch das Management ihrer Community, d.h. ihrer Nutzenden (für rechtliche Aspekte in Bezug auf Medienunternehmen, siehe Kap. 7.1). Die Community kommentiert Medieninhalte direkt in den Kommentarspalten der Medien-Webseiten und auf den Social-Media-Profilen, die von den Medien unterhalten werden. Diese Orte unterscheiden sich in den Beitragsbedingungen für Nutzende und den Moderationsmöglichkeiten der Medien. Auf Medienseiten müssen Nutzende oft identifizierende Mailadressen oder Namen angeben. Auf Social-Media-Seiten dagegen können sie mit einem Fake-Profil anonym kommentieren. Dabei sind ihre Beiträge mit ihren Social-Media-Profilen und Netzwerken verknüpft, was ihr Kommentarverhalten beeinflusst. Auf ihren eigenen Webseiten können Medien die Inhalte von Nutzenden vor der Veröffentlichung zudem kontrollieren oder je nach Situation ihre Kommentarspalten schliessen, um Hassrede zu beschränken. Beides ist auf Social Media weniger gut möglich.

Welche Art von Inhalten publiziert wird und wie die Beiträge der Community gemanagt werden, ist nicht bei allen Medienunternehmen gleich. Medien können beispielsweise durch die Darstellung und den Ton ihrer eigenen Inhalte die Ausgrenzung von sozialen Gruppen fördern. Ähnlich wie öffentliche Personen subtil zu Hassrede aufstacheln können²³², können auch Medien durch ihre Berichterstattung ausgrenzende Denkmuster bestätigen und für Nutzende indirekt eine Legitimation liefern, diese Denkmuster öffentlich zu vertreten. Nicht zu ignorieren sind auch unterschiedliche Verständnisse von Communities.²³³ Denn Medien unterscheiden sich darin, wie stark sie einseitig oder zweiseitig kommunizieren bzw. wie intensiv sie Community-Management betreiben. Der Schweizer Journalist und Unternehmer Hansi Voigt²³⁴ empfiehlt zur Verhinderung von Hassrede, dass Moderator*innen in Kommentarspalten aktive Präsenz zeigen. Dies verhindere, dass sich die Diskussionen verselbstständigen, und reduziere Projektionsflächen. Er nennt zudem die Möglichkeit, dass auch Kommentierende die Inhalte anderer Kommentierenden bewerten und/oder moderieren können und so aktiv in die Gestaltung des Diskussionsklima einbezogen werden. Theoretisch sind diese Ansätze vielversprechend, da sich Nutzende gehört fühlen und sich bewusst werden, dass sie sich in einem Umfeld bewegen, in dem Kommunikationsnormen durchgesetzt werden: Es ist ein reales Gegenüber vorhanden, was Online-Enthemmung reduziert.²³⁵ Alle diese Massnahmen können sich auch auf

²²⁷ Newton, The Verge vom 25. Februar 2019: [The Trauma Floor](#).

²²⁸ Z.B. Chandrasekharan et al. 2017.

²²⁹ Guhl et al. 2020: S. 11.

²³⁰ Marwick und Lewis 2017: S. 18; Sunstein 2000.

²³¹ Dieses Unterkapitel basiert insbesondere auf dem Bericht der Fachstelle für Rassismusbekämpfung. 2019: S. 55ff, für konkrete Äusserung, siehe S. 55.

²³² Chyzh et al. 2019.

²³³ Die folgenden Aussagen basieren unter anderem auf einem Interview mit Hansi Voigt (dasnetz.ch; bajour) im April 2020.

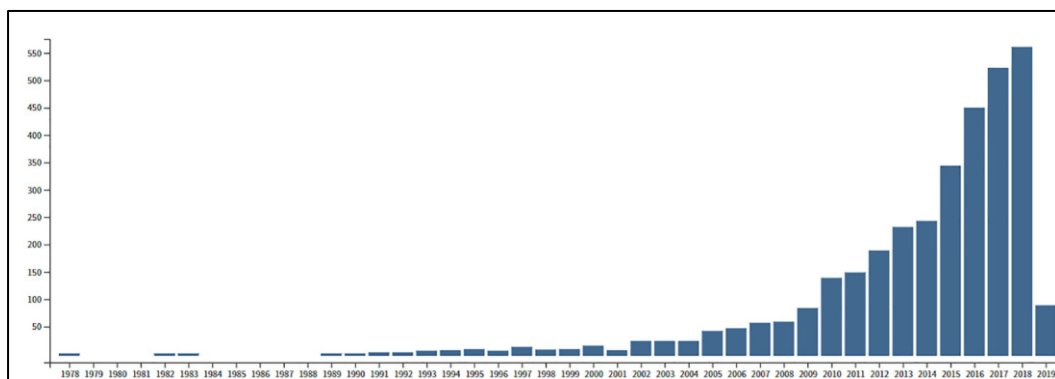
²³⁴ Interview mit Hansi Voigt (dasnetz.ch; bajour) im April 2020.

²³⁵ Suler 2004.

Journalist*innen und deren Arbeit auswirken. Denn auch in der Schweiz kämpfen sie zuweilen mit Angriffen und Hasskommentaren aus dem Publikum.²³⁶ Aus Studien²³⁷ ist bekannt, dass sie sich einerseits unter Druck fühlen, online präsent zu sein. Die damit einhergehende öffentliche Exposition verstärkt aber das Risiko von Angriffen. Andererseits erhalten Journalist*innen nicht immer die diesbezüglich nötige Unterstützung ihres Mediums. Hassrede zu verhindern oder zu reduzieren, kann also nicht nur ein zivilisiertes Klima in der Medienberichterstattung und bei Publikumsdebatten fördern, sondern auch dem Wohl von Journalist*innen dienen.

7.4 Forschung in der Schweiz

Die Forschung, sei sie an Universitäten, Fachhochschulen, Think Tanks oder anderen Institutionen durchgeführt, spielt eine zentrale Rolle in der Generierung von Wissen über Online-Rassismus. Die wissenschaftliche Erforschung von Online-Hassrede im Allgemeinen hat – parallel zur steigenden Medienaufmerksamkeit für das Thema – exponentiell zugenommen. Wurden im Jahr 2005 international noch ca. 50 Artikel dazu publiziert, waren es im Jahr 2018 ganze 550 (Grafik 10).²³⁸



Grafik 10. Entwicklung der Anzahl Publikationen zu «online hate» von 1978 bis 2019 (März).

Die bestehende Forschung konzentriert sich darauf, allgemeine sowie rassistische Online-Hassrede zu beschreiben und zu erklären. Dies beginnt mit Abhandlungen über die relevanten Begriffe, die durch ihre Komplexität (siehe dazu Kap. 3.1 und 3.2) eine multidisziplinäre Herangehensweise verlangen. Damit zusammenhängend werden insbesondere folgende Aspekte untersucht: Wie oft und in welcher Form (z.B. offen oder getarnt) tritt sie auf? Gegen wen richtet sie sich? Wie wirkt sie sich auf die Gesellschaft aus? Dabei greifen Forschende methodisch auf qualitative Methoden wie Interviews mit Betroffenen und Täter*innen zurück, um erste Erkenntnisse zu generieren. Seit wenigen Jahren führen sie vermehrt Umfragen durch um zu erfahren, wer als Betroffene, Beobachtende und Täter*innen involviert ist, wie oft und wieso. Im Gegensatz zu diesen Selbstberichten, die teilweise aufgrund sozial erwünschter Antworten verzerrt sind, eignen sich neuere Ansätze mit Nutzung grosser Datenmengen, um das Online-Verhalten von Millionen Internetnutzenden direkt zu beobachten. Sie erlauben es, relativ einfach grosse Datenmengen über eine längere Zeitperiode zu erheben (wobei hier auf die Relevanz von Datenschutzrichtlinien verwiesen ist). Weiter wird heute untersucht, wie technische Infrastrukturen und Plattform-Architekturen Hassrede hervorbringen. Schliesslich setzen Forschende Experimente ein, um mehr über die Folgen von Online-Hassrede zu erfahren. Dort werden Situationen mit und ohne Hassrede variiert, um deren Auswirkungen auf das Denken und Handeln von Teilnehmenden zu analysieren.

Obwohl viele dieser Erkenntnisse auf die Schweiz übertragbar sind, finden sich kaum Studien, die explizit rassistische Online-Hassrede im Kontext der Schweiz untersuchen. Zwar gibt es diverse Institutionen, an denen sich Forschende mit Hassrede und aggressivem Verhalten im Internet befassen haben, darunter die Universität Zürich, die Zürcher Hochschule für Angewandte Wissenschaften

²³⁶ Stahel und Schoen 2019.

²³⁷ Z.B. Wolfe 2019.

²³⁸ Waqas et al. 2019: S. 5.

(ZHAW), die haute école pédagogique (HEP) Vaud oder die Università della Svizzera italiana.²³⁹ Grösser angelegte Studien über allgemeine und insbesondere rassistische Online-Hassrede in der Schweiz stehen jedoch aus. Priorität sollte eine Bevölkerungsumfrage zur Erhebung von Betroffenen, Beobachtenden und Täter*innen haben. Dabei wird empfohlen, auch soziostrukturelle Merkmale wie Bildung, Arbeitstätigkeit oder soziale Netzwerke zu erfassen. So werden Risikogruppen identifiziert, um Zielgruppen-spezifische Gegenmassnahmen zu entwerfen. Insbesondere der Zugang zu Täter*innen wäre entscheidend, um ihre Motivlage und Netzwerke festzustellen. Langzeitstudien zur Verfolgung von Entwicklungen, Trends und Auswirkungen sind zu bevorzugen.

Wird das bestehende wissenschaftliche Wissen auf die Praxis angewendet, sollte darüber hinaus eine zu starke Vermengung von rassistischer Online-Hassrede und sonstigen Hassrede-Typen vermieden werden. Selbstverständlich teilen unterschiedliche Hassreden gewisse Grundlagen, wie beispielsweise die Logik der Simplifizierung und der Verbreitung. Vermischungen sind jedoch problematisch, wenn Erkenntnisse über Hassrede im Allgemeinen direkt auf rassistische Hassrede übertragen werden, obwohl sie sich in ihren Ursachen, Auswirkungen und Gegenmassnahmen unterscheiden können. Beispielsweise sind von rassistischer Hassrede andere Gruppen betroffen als von Hassrede aufgrund sexueller Orientierung. Jede dieser Gruppen hat ihre eigene Diskriminierungsgeschichte, die in den Gegenmassnahmen berücksichtigt werden sollte. Darüber hinaus kann Hassrede direkt oder indirekt sein, offen oder getarnt, einmalig oder wiederholt, und unterstützt durch mächtige Institutionen und Personen oder nicht.²⁴⁰ All dies beeinflusst, wie gross der Schaden von Hassrede für die Opfer ist.

Grundsätzlich kann die Forschung einen wichtigen Beitrag zur informierten und – bisher leider wenig – evidenzbasierten Auswahl und Umsetzung von Gegenmassnahmen leisten. Dialog und Wissensaustausch zwischen Forschenden und Praktiker*innen schafft hier Abhilfe.

7.5 Zivilgesellschaft im In- und Ausland

Die Zivilgesellschaft kann vielfältig gegen rassistische Online-Hassrede vorgehen: insbesondere dort, wo rechtliche Massnahmen nicht greifen, Social-Media-Anbieter nicht in lokale Kontexte vordringen oder die Mobilisierung von vielen Menschen nötig ist. Im vorliegenden Unterkapitel werden zivilgesellschaftliche Projektansätze (im Folgenden «Ansätze») vorgestellt, die sich mit Online-Hassrede und teilweise konkret mit Online-Rassismus auseinandersetzen.²⁴¹ Sie sind vielfältig und beinhalten institutionalisierte Projekte, wenig koordinierte Spontanaktionen, spezialisierte Organisationen oder individuelle Experten und Tools. Aus der Vielzahl der Ansätze finden hier nur diejenigen Erwähnung, die als besonders relevant eingestuft werden. Beratungsstellen, Jugendorganisationen und NGOs bilden die Projektträgerschaften. Die Ansätze werden in vier Bereiche eingeteilt:

- Prävention und Sensibilisierung
- Beratung und Meldung in konkreten Fällen
- Monitoring
- Gegenrede

Oft decken einzelne Ansätze mehr als einen Bereich ab. In der folgenden Strukturierung der Projektlandschaft werden sie daher in den Bereichen aufgelistet, in denen sie von der Autorin als beispielhaft eingestuft werden. Die Wirksamkeit der Projekte ist mit Skepsis zu betrachten, da

²³⁹ Beispielsweise Catherina Blaya (Lasale, HEP), Dirk Baier (Institut für Delinquenz und Kriminalprävention, ZHAW), Lea Stahel (Soziologisches Institut, UZH), Elisabeth Stark (Romanisches Seminar, UZH), Katharina Lobinger und Eleonora Benecchi (Istituto di media e giornalismo und Istituto di tecnologie digitali per la comunicazione, USI), Stand: Mai 2020.

²⁴⁰ Delgado und Stefancic 2009: S. 361.

²⁴¹ Trotz Überlappungen bleiben hier Ansätze unbeachtet, die sich mit allgemeiner digitaler (De-)Radikalisierung oder explizit anderen Formen von Online-Hassrede auseinandersetzen (z.B. sexistische Hassrede). Auch Ansätze, die sich mit Rassismus im Allgemeinen oder nur mit Offline-Rassismus beschäftigen, jedoch keinen Direktbezug zur Online-Dimension haben, werden nicht diskutiert.

systematische Evaluationen einzelner Projekte kaum existieren²⁴² und «in-house»-Evaluationen selbstreferentiellen Charakter haben.

Neben den wenigen Evaluationsstudien fliessen Erfahrungsberichte aus den geführten Interviews ein. Der dargestellte Wissensstand kann nur einen groben Überblick geben. Auch soll die hier eingeschätzte Wirksamkeit von Ansätzen nicht als allgemeingültig verstanden werden, da die Wirksamkeit konkreter Projekte meist von deren Kontext und konkreter Umsetzung abhängt. Insbesondere da Online-Hassrede diverse Gesellschaftsbereiche tangiert (z.B. Politik, Recht, Technik-Industrie), betonen viele Praktiker*innen die Notwendigkeit einer Verlinkung, Vernetzung und Kooperation zwischen Projekttragenden, Social-Media-Anbietern, Politik, Technik-Expert*innen, Strafverfolgungsbehörden und Anderen. Zur Umsetzung solcher Vernetzungsversuche sei hier exemplarisch die deutsche Vernetzungsplattform [Das NETTZ](#) genannt. Auch das [Institute for Strategic Dialogue](#) und die deutsche [Amadeu Antonio Stiftung](#) sind bekannt für ihre Aktivitäten in Forschung, Praxis und Vernetzung. In der Schweiz gibt es laut Wissen der Autorin kaum explizite Ansätze gegen Online-Rassismus; die hier erwähnten stammen meist aus dem Ausland.

7.5.1 Prävention und Sensibilisierung

Grundsätzlich haben Präventions- und Sensibilisierungsansätze (Tabelle 1 für die Schweiz und Tabelle 2 für das Ausland) eine Langzeitperspektive. Sie sollen durch Aufklärung Resilienz schaffen und Online-Hassrede von vornherein verhindern oder alternativ ihre Wirkung präventiv eindämmen. Sie stellen nur einen Teil einer weit grösseren Bemühung dar, eine digital versierte Zivilgesellschaft aufzubauen. Die Schwerpunkte sind unter anderem:

- Förderung von Medienkompetenz (z.B. zu Kommunikationsbedingungen, Plattform-Architektur)
- Stärkung einer vielfältigen Debattenkultur und einer konstruktiven Konfliktaustragung im Internet
- Digitale Sicherheit und Vorsorge (z.B. Tipps, um sich zu schützen, Resilienzaufbau)
- Wissen um Definitionen, Hassrede-Formen und deren Erkennung, Ursachen, Hintergründe, Auswirkungen, Schädlichkeit und Vorgehensweise von Verbreiter*innen von Online-Hassrede
- Kontextualisierung bzw. Einbettung ins Thema der Menschenrechte und Diskriminierung
- Erkennbarmachung von Verantwortung («Täter*innen sind verantwortlich, nicht Betroffene»)
- Rechte und Rechtslage
- Umgangsstrategien für Betroffene und Zuschauende (z.B. Moderationsstrategien)
- Lernen von digitaler Vernetzung für potentiell oder tatsächliche Betroffene, um Ansprechpersonen, Austausch und Unterstützung zu erhalten
- Unterstützung beim Planen, Produzieren und Bewerben eigener Projekt

Die angebotenen Hilfsmittel und technischen Formate, welche diese Inhalte übermitteln, sind vielfältig. Zum Einsatz kommt Informationsmaterial²⁴³, welches entweder direkt auf Webseiten zu finden oder herunterladbar ist (z.B. Flyers, Broschüren, Berichte und Studien). Angeboten werden auch individuelle Beratung, Workshops/ Trainings / Referate / Webinars (online und offline), Petitionen (zur Positionierung bzw. Aufbau öffentlichen Drucks), öffentliche Kampagnen, Podiumsdiskussionen und Konferenzen. Diverse technische Formate werden eingesetzt: Texte, Videos (z.B. Video-Produktion von Jugendlichen für Online-Blogs), Apps, Memes, YouTube-Kanäle, TED-Talks oder Podcasts. Die Projektansätze sind auf unterschiedliche Zielgruppen ausgerichtet: die breite Öffentlichkeit, Kinder und Jugendliche, Multiplikator*innen (z.B. Pädagog*innen, NGOs etc., aber auch Nutzende mit grossem Online-Netzwerk), Eltern, Medienschaffende, und teilweise auch Unternehmen, Politik und Online-Plattformen. Dabei schlägt sich die Wahl der Zielgruppe auf das genutzte Design, die benutzte Sprache und angebotene Involvierung nieder (z.B. mehr Interaktion bei Jugendlichen).

²⁴² Blaya 2019: S. 163.

²⁴³ Für allgemeines Material zu Medienkompetenz für Kinder und Jugendliche in der Schweiz, siehe [Zischtig.ch](#), und für Fokus auf Hassrede, siehe [Jugend und Medien](#). Pfad: Jugend und Medien > Themen > Diskriminierung und Hate Speech.

Tabelle 1: Ansätze zu Prävention und Sensibilisierung in der Schweiz²⁴⁴

Projekt	(Haupt-)Projekttragende / Organisation	Sprache	Hervorzuhebendes Merkmal
Stop Hate Speech	Frauendachverband; Alliance F	DE (FR)	Informationsmaterial für Öffentlichkeit (in Planung)
#NetzCourage #WirGegenHass	#NetzCourage	DE	Öffentliche Aufklärungsarbeit, u.a. Informationsflyer für Politiker*innen
Hate Speech / Hassrede im Internet ²⁴⁵	Stiftung gegen Rassismus und Antisemitismus (GRA)	DE FR	Flyer zu rassistischer Hassrede und Gegenmassnahmen, v.a. für Schulen
JASS gegen HASS	JASS	DE	Situationsanalyse (2017) und Positionierungsvideos

²⁴⁴ Nach dem Wissen der Autorin existieren im Tessin keine speziellen diesbezüglichen Ansätze (abgesehen von der Bildung in Schulen).

²⁴⁵ Pfad: www.gra.ch > Bildung > [Hate Speech](#).

Tabelle 2: Ansätze zu Prävention und Sensibilisierung im Ausland.

Projekt	(Haupt-)Projekttragende / Organisation	Land	Hervorzuhebendes Merkmal
Hate Speech: Hass Im Netz	Landesanstalt für Medien NRW; AJS NRW; Klicksafe	DE	Handbuch mit Informationen für Fachkräfte und Eltern
Hate Speech ²⁴⁶	Deutsche Gesetzliche Unfallversicherung (DGUV)	DE	Detailliertes Lehrmaterial zum Herunterladen
Streitkultur 3.0	Berghof Foundation / Friedenspädagogik Tübingen	DE	Streitkultur stärken durch digitale «Dialog-Labore»
#NoHateNoFake	WERK 2 – Kulturfabrik Leipzig e.V.	DE	Videoblog-Projekt für Schulen
Werte leben - online	Juuuport e.V.	DE	Vermittlung von digitalen Diskussionswerten durch jugendliche «Scouts» per Webinar
Hate Speech im Netz stoppen!	Campact	DE	Petition mit Forderungen an Justiz
Tarik Tesfu	Tarik Tesfu	DE	Influencer mit YouTube-Kanal (Positioniert sich gegen Hassrede)
No Hate Speech Movement	Europarat, Neue deutsche Medienmacher	DE / EU	Umfangreiches, detailliertes Informationsangebot mit ansprechender digitaler Aufbereitung
ZARA	ZARA	AT	Trainings für Erwachsene und Jugendliche zu Zivilcourage und Argumentieren bei Online-Rassismus im Speziellen
Rete nazionale per il contrasto ai discorsi e ai fenomeni d'odio	Diverse inkl. UNAR (National Office for Racial Discrimination)	IT	Netzwerk zur Förderung und Unterstützung von insbes. Forschung, Sensibilisierungskampagnen, Informationsaustausch
CO.N.T.R.O	UNAR (National Office for Racial Discrimination), in partnership with IRS (Institute for Social Research)	IT	Studie und Forschung, Sensibilisierungs- und Informationsaktivitäten
BRICKS – Against Hate Speech	Cospe, Zaffiria	IT	Aufklärungsmaterial zu Hassrede
#NOHATESPEECH Giornalisti e lettori contro i discorsi d'odio	Associazione Carta di Roma	IT	Petition zu Online-Hassrede gegen Journalist*innen
#SilenceHate – Giovani digitali contro il razzismo	Cospe, Zaffiria, Priscilla associazione	IT	Bekämpfung der Verbreitung von Hassreden durch Medienerziehung junger Menschen
No Hate Speech Movement, auch auf Facebook	Dipartimento della Gioventù e del Servizio Civile Nazionale, Europarat	IT / EU	Fernseh- und Radiospots und Aufklärungskampagnen insbes. für Jugendliche, zusätzliches Aufklärungsmanual
Facing facts	Jewish contribution to an inclusive Europe (CEJI-A)	8 EU-Länder inkl. IT	Aufklärende Online-Kurse für diverse Zielgruppen (u.a. Polizei, zivilgesellschaftliche Organisationen, Entscheidungsträger*innen) in deutsch, französisch, italienisch, und englisch
LICRA	Ligue Internationale Contre le Racisme et l'Antisémitisme (LICRA)	FRA	Fokus auf Rassismus im Allgemeinen mit «Combattre la haine sur Internet» als ein Schwerpunkt
Mouvement contre le discours de haine	Nicht angegeben; Europarat	FRA / EU	Aufklärungskampagnen insbes. für Jugendliche
SELMA	European Schoolnet, EU	EU	Umfangreiches pädagogisches Material für junge Menschen
Campaign Toolkit	Institute for Strategic Dialogue (ISD)	UK	Hilfs-Tool zum Entwerfen eigener Projekte

Über die Wirksamkeit von Präventions- und Aufklärungsansätzen ist wenig bekannt. Dies liegt auch daran, dass Langzeiteffekte von Aufklärung, insbesondere auf das Verhalten der Aufgeklärten, schwer zu messen sind. Die nichtsdestotrotz bestehenden Evaluationen beziehen sich fast ausschliesslich auf Medienkompetenz bei Jugendlichen und finden positive Effekte. Projekte stärken beispielsweise das eigene Vertrauen darin, fremde Online-Inhalte (z.B. Desinformation) korrekt einzuschätzen und eigene Inhalte auf verantwortungsvolle Art zu erstellen und zu verbreiten. Auch wussten die Teilnehmenden nach Interventionen mehr über den sicheren und widerstandsfähigen Umgang mit dem Internet.²⁴⁷ Eine Meta-Analyse von 51 Projekten²⁴⁸ bestätigt dies: Projekte wirkten sich positiv auf das Wissen über

²⁴⁶ Pfad: www.dguv.de > Sekundarstufe I > Sucht- und Gewaltprävention > Hate Speech.

²⁴⁷ Z.B. Gatewood und Boyer 2019.

²⁴⁸ Jeong, Cho und Hwang 2012.

Medien und (auch, aber weniger) auf Einstellung, Selbstwirksamkeit oder Verhalten aus. Jedoch wird geraten, stärker das Verhalten zu trainieren, anstatt nur Wissen zu vermitteln. Schliesslich legen Studien nahe²⁴⁹, dass Online-Hassrede reduziert werden kann, wenn Nutzende die strafrechtlichen Konsequenzen und negativen Auswirkungen kennen.

Damit Projekte wirksam sind, raten diverse Expert*innen²⁵⁰ zu folgenden Punkten:

- Anpassung an Zielgruppe in Bezug auf Mediengewohnheiten und digitale Lebensrealitäten, digitale Bedürfnisse, Schwierigkeitsgrad und Involvierung. Insbesondere für Jugendliche: Gleichgewicht zwischen Bildungsinhalt und engagierenden, spielerischen Übungen
- Fokus auf vulnerable Gruppen, d.h. insbesondere wahrscheinliche und/oder tatsächliche Betroffene und Täter*innen, anstatt bereits Sensibilisierte zu belehren
- Didaktisches Drei-Schritt Vorgehen: 1. Sensibilisieren und informieren, 2. analysieren und reflektieren, 3. anwenden. Diese Vorgehensweise erlaubt das Erlernen neuen Wissens, den Erfahrungsaustausch und das mutige Umsetzen von Strategien
- Stärkung von Teilnehmenden, sodass selbstständige, moralisch-ethische Entscheide resultieren; vom moralisierenden Vorschreiben «korrekter» Meinungen und Haltungen wird abgeraten
- Bei Umgangsstrategien spezifische Problem-Fokussierung, inklusive Verhaltenstraining
- Nutzung populärer sozio-kultureller Referenzen aus der Internet-Welt (um an die Lebenswelt von Jugendlichen anzuknüpfen)
- Begleitendes, altersentsprechendes schriftliches Material
- Weiterbildung von Multiplikator*innen (z.B. Lehrer*innen) für eine kompetente Weitergabe von Inhalten

7.5.2 Meldung und Beratung

Werden Internetnutzende in konkreten Situationen rassistisch angegriffen oder beobachten sie solche Angriffe, können sie diese Vorfälle melden und/oder sich beraten lassen (für Ansätze, siehe Tabelle 3 für die Schweiz und Tabelle 4 für das Ausland). Grundsätzlich können Beiträge direkt auf den jeweiligen Social-Media-Plattformen gemeldet werden (sofern man vermutet, dass sie gegen deren Richtlinien verstossen). Alternativ können sie sich an Stellen wenden, die weitere Meldetools zur Verfügung stellen. Nutzen sie letztere Meldesysteme, prüfen die Stellen die gemeldeten Beiträge üblicherweise auf strafrechtliche Relevanz und leiten sie dann an die Plattform-Anbieter weiter, um eine Löschung zu fordern, und/oder an Strafbehörden, um eine Strafverfolgung einzuleiten (letztere Möglichkeit bezieht sich insbesondere auf Deutschland). In der Schweiz melden GRA und #Netzcourage regelmässig Online-Inhalte an Social-Media-Anbieter. #Netzcourage nutzt dafür eine Online-Gemeinschaft, die durch koordinierte Massmeldungen Löschungen wahrscheinlicher macht. Die bestehenden Meldetools bei den relevanten Stellen unterscheiden sich in ihrem Aufbau, z.B.:

- ob Meldungen nur über Internetseiten oder auch über Apps getätigt werden können;
- wie viel Hintergrundinformation zum Melden angeboten wird (z.B. was kann/soll man melden und was nicht, Vorteile einer Meldung);
- wie viele Informationen Nutzende angeben sollen/können (z.B. Screenshot, Links);
- ob anonym gemeldet werden kann oder nicht;
- ob meldende Personen Rückmeldung zum Lösch-/Verfahrensverlauf erhalten oder nicht;
- ob Berichte Statistiken, welche auf Meldungen basieren (Anzahl, Art, Beiträge), veröffentlicht werden.

Die bestehenden Beratungen sind psychologisch, sozial und/oder juristisch ausgerichtet. Nicht alle Beratungen beziehen alle diese Aspekte ein; um eine umfangreiche Unterstützung zu garantieren, wäre dies jedoch empfohlen. Beratungen werden meist kostenlos über diverse anonyme und nicht-anonyme Kommunikationskanäle angeboten (z.B. Telefon, Chat, persönlich). Auch besteht teilweise

²⁴⁹ Aizenkot und Kashy-Rosenbaum 2018..

²⁵⁰ Z.B. Blaya 2018; Blaya 2015; Dittrich et al. 2020; Gatewood und Boyer 2019; Wachs et al. 2019.

Unterstützung bei der Einleitung einer Anzeige bei der Polizei oder dem Gang zur Staatsanwaltschaft.²⁵¹ Betroffene werden juristisch und psychosozial im Rahmen von Vergleichs- oder Gerichtsverhandlungen begleitet. Daneben bieten Stellen Soforthilfe an, indem sie das Social-Media-Profil der Betroffenen managen (in der Schweiz primär #NetzCourage). Prozesskostenfinanzierung wird angeboten, aber laut Wissen der Autorin nur von HateAid in Deutschland. Schliesslich werden externe Therapeut*innen für die Verarbeitung von Angriffen vermittelt, wodurch unter anderem verhindert werden soll, dass Betroffene «digital verstummen».

Tabelle 3: Meldung und Beratung in der Schweiz²⁵²

Projekt	(Haupt-)Projekttragende / Organisation	Sprache	Hervorzuhebendes Merkmal
Meldetool ²⁵³	Stiftung gegen Rassismus und Antisemitismus (GRA)	DE FR	Meldetool für Online-Rassismus, teilweise Beratung
#NetzCourage	#NetzCourage	DE	Umfangreiche Begleitung (Beratung zu Möglichkeiten, Meldung, Unterstützung bei Anzeige etc.). Teilweise Kontaktherstellung mit Täter*innen.

Tabelle 4: Meldung und Beratung im Ausland

Projekt	(Haupt-)Projekttragende / Organisation	Land	Hervorzuhebendes Merkmal
Hass melden	Reconquista Internet	DE	Meldemöglichkeit mit App; strafrechtliche Prüfung und Weiterleitung an Strafbehörden
Eco Beschwerdestelle	Eco	DE	Meldemöglichkeit mit Weiterleitung an Strafbehörden und Plattformen; starker Einbezug juristischer Expertise; öffentliche Jahresberichte
HateAid	Campact e.V., Fearless Democracy e.V. (Anna-Lena von Hoderberg)	DE	Begleitung im gesamten Prozess (insbes. Beratung und Prozesskostenübernahme, ermöglicht durch Spendensystem)
Ban hate	Antidiskriminierungs-stelle Steiermark	AT	Melde-App mit Abruf des aktuellen Meldeverlaufs
Hass im Netz melden ²⁵⁴	ZARA	AT	Beratung und Melden speziell für Online-Rassismus
Segnala il razzismo online	Regione Abruzzo (nur Italien)	IT, EU	Bericht mit Informationen und Vorschläge für praktisches Vorgehen
INACH cyber hate complaints base	International Network Against Cyber Hate (INACH)	22 Länder inkl. FRA & IT	Verarbeitung von Meldungen in länderübergreifenden Berichten
Signaler un fait de racisme / d'antisémitisme	Ligue Internationale Contre le Racisme et l'Antisémitisme (LICRA)	FRA	Melde-Webseite für Rassismus (Offline- und Online-Formen)
Réagis au racisme	equal.brussels	BEL	Informationen zu Rassismus und Meldemöglichkeiten (Offline- und Online-Formen; u.a. in französisch und englisch)
No hate speech youth campaign ²⁵⁵	Council of Europe Youth Department	EU	Detaillierte Informationen zu Meldeverfahren bei einzelnen Social-Media-Unternehmen

Grundsätzlich ist davon auszugehen, dass Bekanntheit und Niedrigschwelligkeit der Angebote entscheiden, ob und wie oft Internetnutzende von Melden und Beratungen Gebrauch machen. Prinzipiell

²⁵¹ In Deutschland wurde ein unzureichendes Problembewusstsein bei diesbezüglichen Stellen beobachtet, was zu Forderungen für Polizeidienststellen mit Schwerpunkt für Hassrede, die Fortbildung von Polizeikräften, die Sensibilisierung der Polizei, Gerichte und Staatsanwaltschaft und die Möglichkeit von Online-Anzeigen führte, z.B. von: Neue deutsche Medienmacher & No Hate Speech Movement Deutschland 2019.

²⁵² Nach dem Wissen der Autorin existieren im Tessin keine diesbezüglichen Ansätze.

²⁵³ Pfad: www.gra.ch > Bildung > Hate Speech > [Vorfall melden](#).

²⁵⁴ Pfad: www.zara.or.at > Beratung > [Hass im Netz anonym melden](#).

²⁵⁵ Pfad: www.coe.int > Democracy > No Hate Speech Youth Campaign > Taking Action > [Reporting Hate Speech](#).

scheinen Nutzende Beiträge selten zu melden.²⁵⁶ Dieses Underreporting wird damit erklärt, dass Opfer kein Vertrauen in die Polizei oder Angst vor Vergeltung haben, Gewalt als gegeben akzeptieren und Verletzungen ihrer Privatsphäre oder Sprachbarrieren befürchten.²⁵⁷ Menschen melden Hassbeiträge jedoch eher, je weniger Medienbrüche sie erwarten müssen, d.h. je weniger sie zwischen Medien hin und her wechseln müssen. Optimal sind hier Meldungen direkt auf den Plattformen, auf denen Hasskommentare beobachtet wurden. Dabei kann jedoch nicht vorausgesetzt werden, dass Menschen die strafrechtliche Relevanz von Beiträgen korrekt einschätzen können und dementsprechend eine Meldung tätigen oder nicht. Die deutsche Eco-Beschwerdestelle²⁵⁸ stellt in ihrer Meldestatistik fest, dass im Bereich Rassismus und Hate Speech nur acht Prozent der Meldungen von Nutzenden tatsächlich strafrechtlich relevant waren. Dies unterstreicht gemäss den Autor*innen die Wichtigkeit juristischer Expertise bei der Prüfung der gemeldeten Inhalte, sodass «erlaubte Äusserungen nicht einfach als ungewünscht entfernt werden».²⁵⁹ Kommt es tatsächlich zu juristischen Prozessen, dauern diese lange, sind aufwändig und teuer. Es muss daher abgewogen werden, ob entsprechende Ressourcen vorhanden sind. [HateAid](#) begründet die Übernahme von Prozesskosten damit, dass die Bekämpfung von Gewalt im Internet nicht vom Einkommen der Betroffenen abhängig sein sollte. Schaffen es relevante Stellen, die hier angedeuteten Hindernisse zu reduzieren, kann dies Meldungen theoretisch fördern. In der Praxis sieht man denn auch grosse Unterschiede in Meldezahlen: Wo die schweizerische GRA pro Woche durchschnittlich drei Meldungen (nur aus der Schweiz) bekommt²⁶⁰, sind es beim deutschen Meldetool [Hass melden](#) 500 Meldungen (gesamthaft wurden bisher 16'000 Beiträge angezeigt).

Das Veröffentlichen von Meldestatistiken und Verurteilungen könnte zur Sichtbarkeit, Sensibilisierung, und Abschreckung beitragen – sodass das Internet weniger als rechtsfreier Raum wahrgenommen wird. Auch Vergleichsverhandlungen, wo sich Betroffene und Täter*innen gegenüber treten (in der Schweiz organisiert durch #Netzcourage), könnten Mechanismen wie Distanz, wahrgenommene Unsichtbarkeit und Sanktionslosigkeit (siehe Kap. 5.4.1) entgegenwirken und dadurch Online-Enthemmung reduzieren.

7.5.3 Monitoring

Das Ziel von Monitoring-Ansätzen (Tabelle 5 für die Schweiz und Tabelle 6 für das Ausland) ist grundsätzlich, Online-Hassrede in seinen lokalen Gegebenheiten zu verfolgen und zu beschreiben. Die Wissenschaft, Zivilgesellschaft und Think Tanks haben bisher unterschiedliche Ansätze lanciert, um insbesondere längerfristige Entwicklungen zu erfassen. Die Ergebnisse können zivilgesellschaftliche Gruppen über die lokal vorherrschenden Hassrede-Gruppen informieren, um Gegenrede-Kampagnen strategisch anzupassen – oder auch Behörden, um gefährdende oder gefährdete Gruppen zu erkennen und eventuell Hassverbrechen zu verhindern. Dabei lassen sich quantitative Ansätze (siehe Punkte 1 bis 3) und qualitative Ansätze (Punkt 4) unterscheiden.²⁶¹ Beide Ansätze ergänzen sich: Die quantitativen Ansätze können einen Überblick («Mapping») zu Häufigkeit und Typen von Online-Hassrede in einem grösseren Kontext geben – sofern das Sammeln und Analysieren der Daten systematisch geschieht und Daten somit vergleichbar sind (was bei unsystematisch gesammelten Meldungen nicht zutrifft). Qualitative Ansätze dagegen liefern tiefere Einblicke in die Natur und Funktionsweise von Hassrede.

1. *Kontinuierliches Echtzeit-Monitoring:* Hier wird eine grosse Menge an Online-Beiträgen mittels automatisierter Methoden analysiert, z.B. Natural Language Processing, Machine Learning, Big Data-Analyse.²⁶² Dabei werden bestimmte Internet-Räume (z.B. Social Media, Online-Zeitungen) und Plattformen (moderierte oder nicht-moderierte) konstant nach Hassrede-

²⁵⁶ Eine Umfrage der Landesanstalt für Medien NRW (2018) zeigt, dass von den Nutzenden, die berichten, Hassbeiträge im Netz schon mal beobachtet zu haben, nur eine Minderheit den Online-Beitrag beim Social Media-Anbieter (25%) oder der Polizei (1%) gemeldet haben.

²⁵⁷ Cerase, D'Angelo und Santoro 2015: S. 4.

²⁵⁸ Eco – Verband der Internetwirtschaft 2019.

²⁵⁹ Eco – Verband der Internetwirtschaft 2019: S. 25.

²⁶⁰ Interview mit Dominik Pugatsch (Stiftung gegen Rassismus und Antisemitismus, GRA) im März 2020.

²⁶¹ Für einen Überblick, siehe Lucas 2014.

²⁶² Siehe dazu Fortuna und Nunes 2018.

Inhalten durchsucht. Der Ansatz bietet ein Frühwarnsystem, kann Trends entdecken und erlaubt Reaktionen auf Ereignisse während ihrer Entstehung. Diesbezügliche Projekte sind aufgrund des Aufwands (z.B. Involvierung vieler Freiwilliger) jedoch selten.

2. *Retrospektives Monitoring:* Für dieses Monitoring kommen meist automatisierte Methoden für die Analyse grosser Mengen an Online-Beiträgen zum Einsatz. Beiträge werden dabei über eine relativ kurze Zeit ausgewählt und analysiert.
3. *Sammeln von gemeldeten Beiträgen:* In Meldesystemen eingegangene Online-Beiträge werden kontinuierlich (live) oder in zeitlichen Abständen gesammelt. Meldungen von der Öffentlichkeit werden teilweise durch interne, aktiv zusammengesuchte Vorfälle erweitert. Hier ist der Datenumfang überschaubar.
4. *Qualitative Analysen:* Hier werden eine überschaubare Menge an Beiträgen oder auch ganze Plattformen qualitativ analysiert, um tiefere Erkenntnisse über Themen, Motive oder Ideologien von Verbreiter*innen zu gewinnen.

Tabelle 5: Ansätze zu Monitoring in der Schweiz.

Projekt	(Haupt-)Projekttragende / Organisation	Sprache	Hervorzuhebendes Merkmal
Stop Hate Speech	Frauendachverband Alliance F	DE (FR)	Algorithmisches Monitoring von Hassrede auf Zeitungsplattformen und in Social Media für die Gesamtschweiz mit Freiwilligen (in Planung)
Chronologie «Rassismus in der Schweiz»	Stiftung gegen Rassismus und Antisemitismus (GRA)	DE FR	Erfasst seit 1992 alle öffentlich gewordenen Vorfälle, die auf rassistische oder fremdenfeindliche Motivationen zurückgehen (Offline- und Online-Fälle)
Antisemitismusbericht	GRA und Schweizerischer Israelitischer Gemeindebund (SIG)	DE FR	Jährlicher Bericht zu von der Öffentlichkeit gemeldeten und aktiv gesammelten antisemitischen Vorfällen; separate Auflistung von Online-Vorfällen
Rapports Antisémitism	Coordination Intercommunautaire Contre l'Antisémitisme et la Diffamation (CICAD)	FR	Jährlicher Bericht zu antisemitischen Vorfällen in der Romandie (Offline- und Online-Fälle).
Rassismuvorfälle aus der Beratungspraxis	Verein Humanrights.ch, Eidgenössische Kommission gegen Rassismus (EKR)	DE FR IT	Jährlicher, schweizweiter Bericht zu Beratungsfällen bei Rassismus; separate Auflistung von Online-Vorfällen

Tabelle 6: Ansätze zu Monitoring im Ausland.

Projekt	(Haupt-)Projekttragende / Organisation	Land	Hervorzuhebendes Merkmal
Debate//de:hate ²⁶³	Amadeu Antonio Stiftung	DE	Primär qualitative Monitoring-Berichte mit Fallbeispielen
Mapping Hate (siehe Anwendung hier)	Institute for Strategic Dialogue (ISD)	UK, USA	Methodologisch robustes, wissenschaftlich informiertes Monitoring mit Geolokalisierung. Feststellung von Treibern, Themen, Schlüsselfiguren, Gruppierung von Hass-Gruppen. Anwendbar in vielen Sprachen. ²⁶⁴
Task force contro i discorsi d'odio	Amnesty International	IT	Überwachung von und Reaktion auf Online-Hassreden. Zusätzlicher Bericht Il barometro dell'odio zu Online-Hassrede von Politiker*innen
LIGHT ON	Regione Abruzzo, ISIG, Progetti Sociali (nur Italien)	IT, EU	Visuelle Dokumentation von Fällen von Offline- und Online-Hassrede
Cartographie de la Haine en Ligne Tour d'horizon du discours haineux en France	l'Institut pour le Dialogue Stratégique (ISD) und Facebook	FRA	Bericht über Formen von Online-Hassrede in Frankreich
hatemeter	eCrime, Universität Trento, Rechtswissenschaftliche Fakultät	IT, FRA, UK	EU-weites, automatisiertes Live-Monitoring von anti-muslimischer Hassrede. Dieses soll Hassrede-Trends identifizieren. Auch soll automatisierte Gegenrede eingesetzt werden.
Research – Report – Remove: Countering Cyber Hate Phenomena	International Network Against Cyber Hate (INACH)	6 EU-Länder inkl. DE & FRA	Halbjährliche Berichte über länderabhängige Hassrede-Typen, Trigger-Ereignisse, Trends und relevante Plattformen auf Basis gemeldeter und wenig systematisch gesammelter Beiträge
Scan	LICRA (International League against Racism and Antisemitism)	10 EU-Länder inkl. DE, FRA, IT	Berichte und Trainings zu Tools und Wissen für Monitoring

Die Ziel von Monitoring-Ansätzen liegt insbesondere darin, das Auftreten von Online-Hassrede und ihrer langfristigen Muster innerhalb und zwischen Gesellschaften effektiv zu messen (sozusagen als Barometer gesellschaftlicher Spannungen). Viele bisher veröffentlichte Monitoring-Berichte sind reich an Informationen zu spezifischen Vorfällen und versuchen, Entwicklungen einzufangen. Der Einsatz algorithmischer Erkennung hilft zudem, massive Informationsmengen, die manuell nicht bearbeitbar sind, innert kurzer Zeit zu kategorisieren. Trotz der Vorteile solcher Verhaltensdaten (gegenüber beispielsweise Umfragedaten), ermöglichen auch sie es nicht, die Menge von Online-Hassrede in einem bestimmten Land zuverlässig einzuschätzen oder längerfristige Entwicklungen verlässlich vorherzusagen zu können. Denn das Monitoring deckt oft nur einzelne Online-Plattformen ab, nicht das breite Internet. Da Algorithmen und Moderationsstrategien von Plattformen zudem meist intransparent sind, bleibt unklar, auf was mögliche Veränderungen in beobachteter rassistischer Online-Hassrede zurückgeführt werden können: auf tatsächlich verändertes Nutzerverhalten oder auf geänderte Löschregeln der Plattformen? Daher müssen die Forschenden und involvierten Stellen die zugrundeliegenden Faktoren zu einem gewissen Grad auch interpretieren und kontextualisieren.²⁶⁵ Ein ähnliches Dilemma stellt sich bei der Interpretation gemeldeter Beiträge, weil sie eben nur repräsentativ für das Gemeldete sind. Es ist wahrscheinlich, dass sie nur die Spitze des Eisbergs darstellen.

Die Umsetzung von Monitoring ist für zivilgesellschaftliche Organisationen herausfordernd.²⁶⁶ Oft ist der Zugang zu Online-Daten nicht garantiert. Software zur automatisierten Erkennung ist teuer.²⁶⁷ Für die Programmierung von Algorithmen braucht es jedoch genügend linguistische Kompetenz, IT-Qualifikation und Arbeitskraft, damit die Algorithmen Hassrede in grossen Datenmengen verlässlich erkennen. Dieser Prozess ist aufgrund der komplexen Natur von Hassrede und subjektiven

²⁶³ Pfad: www.amadeu-antonio-stiftung.de > Projekt > de:hate > [Monitoring und Analyse](#).

²⁶⁴ Interview mit Jonathan Birdwell (Institute of Strategic Dialogue) im April 2020.

²⁶⁵ Schmidt und Wiegand 2017.

²⁶⁶ Siehe dazu mehr in sCAN 2018b.

²⁶⁷ sCAN 2018b: S. 18. Es gibt in der Schweiz Versicherungen gegen Online-Angriffe für Personen und Unternehmen. Diese beobachten den Social Media-Raum und übernehmen Löschungen oder Abmahnungen (z.B. [Silencio](#)). Aufgrund der Ausrichtung und Kosten bietet sich dieser Ansatz für die breite Rassismusbekämpfung jedoch weniger an.

Auffassungen darüber, was überhaupt als Hassrede gilt und was nicht, nicht zu unterschätzen.²⁶⁸ Auch der lokale Kontext und die Sprache sollten berücksichtigt werden. Werden diese Umweltfaktoren ignoriert, können Algorithmen leicht zu Fehleinschätzungen führen. Komplexere Algorithmen suchen nicht nur nach einzelnen Hassrede-Wörtern, sondern beziehen auch Metadaten (z.B. Geolokalisierung) ein. Zusätzliches qualitatives, manuelles Validieren der Ergebnisse ist nötig, um Fehlklassifikationen entgegenzuwirken. Um den geänderten Ausdrucksformen von Hassrede über die Zeit (siehe Kap. 5.4.2) Rechnung zu tragen, sind längerfristige und flexible Anpassungen nötig. Schliesslich sollten bei der Datensammlung Datenschutz- und Privatsphäre-Richtlinien berücksichtigt werden.²⁶⁹ Die Vorgehensweise bezüglich Definition des zu untersuchenden Phänomens, den Daten, des Algorithmus und der Ergebnisdeutung sollte transparent kommuniziert werden. Wird solchen Anforderungen entsprochen, erlauben Monitoring-Berichte hochinformativ Einblicke in länderspezifische Trigger-Ereignisse, Trends, Hassrede-Typen und Gruppierungen.

7.5.4 Gegenrede

«Don't feed the trolls»²⁷⁰ – dies stellt bei Gegenrede kein Leitprinzip mehr dar. Gegenrede (Counter Speech) sind Texte, Bilder oder Videos, die der Hassrede entgegentreten und einen «positiven» Grundton in die Debatten bringen sollen. Die Idee ist, Normen wie Respekt und Sachlichkeit zu stärken und gegenüber den Betroffenen Solidarität zu signalisieren. Verbreiter*innen von Online-Hassrede sollen konfrontiert werden: Dies soll sie (idealerweise) zum Umdenken bewegen, aber auch – und noch wichtiger – ihre Wirkung auf die Zuschauenden mindern. Gegenrede soll dem Publikum signalisieren, dass Hassrede nicht toleriert wird und nicht die dominierende Meinung darstellt. Das Bild der «schweigenden Mehrheit», das Verbreiter*innen von Online-Hassrede gemeinhin von sich zu übermitteln versuchen, soll also widerlegt werden. Diese Ziele können mit dem reinen Löschen von Inhalten nicht erreicht werden.

Die relevanten Ansätze (Tabelle 7 für die Schweiz und Tabelle 8 für das Ausland) unterscheiden sich insbesondere im Hinblick auf Zielgruppe, Plattformen und konkrete Umsetzung.²⁷¹ Die Zielgruppen werden ausgewählt nach Primärprävention (Minderjährige und Erwachsene in Bildungskontexten), Sekundärprävention (vulnerable Gruppen) oder Tertiärprävention (Verbreiter*innen von Online-Hassrede, denen direkt entgegentreten wird). Diese Zielgruppen schliessen sich innerhalb einzelner Initiativen jedoch nicht gegenseitig aus. Die Gegenrede wird insbesondere von Community-Manager*innen und ehrenamtlichen Internet-Nutzenden durchgeführt. Letztere versammeln sich in Aktionsgruppen und leisten Gegenrede. Auch bekunden sie ihre Solidarität mit Betroffenen – um diese zu ermutigen und vom «digitalen Rückzug» abzuhalten. Die Aktionsgruppen sind permanente Gemeinschaften oder bilden sich vorübergehend. Sie richten sich gegen eine oder mehrere Typen der (rassistischen) Online-Hassrede. Es gibt zudem unterschiedliche Stile des Eingreifens: konfrontativ oder gemeinschaftsorientiert, mit sachlicher, logischer Argumentation²⁷², Satire und Humor, mithilfe von Kommentaren, «Likes» und «Shares».²⁷³ Insbesondere «Memes» (Bilder, Cartoons, GIFs oder Sprüche, die neben einer künstlerischen Verwirklichung eine Gegenbotschaft enthalten) sind beliebtes Material. Ansätze für Gegenrede haben somit vielfältige Fokusse, darunter:

- Beratungs- und Fortbildungsangebote in Bezug auf Anpassung von Strategien auf Hassrede-Typen, Entwicklung von Gegenredematerial und Umsetzung von Gegenrede
- Bereitstellung von innovativem Gegenredematerial. Kann sich orientieren an Typen von Hassrede oder an rhetorischen Strategien (z.B. «Whataboutism»)
- Vernetzungsunterstützung für Gegenrededegemeinschaften, z.B. Facebook-Gruppe.
- eigene Umsetzung von Gegenrede

²⁶⁸ Salminen et al. 2018.

²⁶⁹ sCAN 2018b: S. 21 ff.

²⁷⁰ Lorentzen, TEDx Talk 2017. [Don't feed the trolls - Fight them](#). Publiziert auf Youtube am 7. Januar, 2020.

²⁷¹ Tuck und Silverman 2016.

²⁷² Für die Anwendung von Gegenrede existiert ein mehrsprachiger, Expert*innen-basierter Datensatz von 4078 Paaren von Hassrede und Gegenargumenten. Siehe: Chung et al. 2019.

²⁷³ Z.B. Ziegele et al. 2019.

- «Naming and shaming» bzw. Exponieren von insbesondere öffentlichen Personen, die als Verbreiter*innen von Online-Hassrede oder Trolls identifiziert werden (z.B. «[Troll of the month](#)»²⁷⁴)
- indirekte, künstlerische Ansätze wie innovative Spendenaktionen ([Hass hilft](#)), interaktive Kunstwerke ([Tools of Subversion](#)²⁷⁵) oder Satire-Veranstaltungen ([Hate Poetry](#))²⁷⁶
- Nicht-organisierte Ansätze, wobei sich Internet-Nutzende situativ und spontan gegen Hassrede stellen. Beispielsweise haben sich nach der Einführung von drei Klammern um jüdische Namen auf Social Media (siehe «Tarnung» in Kap. 5.4.2) nichtjüdische Nutzer*innen solidarisch gezeigt. Sie haben ihre Benutzernamen ebenfalls so gekennzeichnet, um die Suche nach antisemitischen Inhalten zu erschweren.²⁷⁷ Daneben sei hier die selbstmotivierte Gegenrede von Online-Nutzenden erwähnt: Viele Nutzende stellen sich nicht als Teil von Gruppen gegen Hassrede, sondern sozusagen im Alleingang.²⁷⁸

Mögliche Umsetzungen werden hier anhand von zwei bekannten Gegenredeprojekten veranschaulicht:

Das Tool [Seriously](#)²⁷⁹ des französischen Think Tanks *Renaissance Numérique* kombiniert einen wissenschaftlichen und einen pädagogischen Ansatz. Einerseits entwerfen und sammeln die Verantwortlichen potentiell Gegenredematerial (z.B. Statistiken, Fakten, Zitate). Dieses aktualisieren sie kontinuierlich, um aktuelle gesellschaftliche Debatten abzubilden. Bevor das Material interessierten Gegenredner*innen auf der Plattform zur Verfügung gestellt wird, wird es von einer unabhängigen Kommission von Wissenschaftler*innen auf seine Eignung und Qualität geprüft. Andererseits hat das Tool einen pädagogischen Charakter: Interessierte können sich bei Bedarf im kritischen Argumentieren und in Strategien der Deeskalation coachen lassen, um effektiv in Hassrede-Debatten einzugreifen.

Die selbsternannte «überparteiliche Aktionsgruppe» [#ichbinhier](#)²⁸⁰ konzentriert sich dagegen auf die Vernetzung von mittlerweile 45'000 freiwilligen Gegenredner*innen (Stand: Mai 2020). Die Organisator*innen stellen täglich Links zu aufkeimenden Debatten auf reichweitestarken Medienkanälen in ihre Facebook-Gruppe. Damit organisieren sie spontane Gegenrede. Dem Publikum soll ein ausgewogenes und sachliches Meinungsbild vermittelt werden. Dafür unterstützen Gruppenmitglieder ihre Gegenrede gegenseitig mit dem Geben von «Likes». So rückt der Algorithmus diese in der Kommentarspalte nach oben, da Algorithmen Interaktionen mit hohem Verkehr im «Relevanz-Ranking» bevorzugen. Dies macht Gegenrede sichtbarer – und Hassrede idealerweise weniger sichtbar.

Tabelle 7: Ansätze zu Gegenrede in der Schweiz²⁸¹

Projekt	(Haupt-)Projekttragende / Organisation	Sprache	Hervorzuhebendes Merkmal
Stop Hate Speech	Frauentachverband Alliance F	DE FR	Koordinierte Gegenrede durch Bürger*innen (in Planung)
#NetzCourage	#NetzCourage	DE	Koordinierte Gegenrede und Meldeaktionen auf Facebook
Meldezentrale für Eidgenossen	10 anonyme Gründer*innen	DE	Koordinierte Gegenrede auf Facebook seit 2017, mit Veröffentlichung grösserer Fälle auf begleitendem Blog

²⁷⁴ Pfad: www.getthetrollsout.org > What we do > [Troll of the month](#).

²⁷⁵ Pfad: www.goldextra.com > Projekte > [Tools of subversion](#).

²⁷⁶ Facebook: Hate Poetry.

²⁷⁷ Gunaratna, CBS News vom 10. Juni 2016: <https://www.cbsnews.com/news/neo-nazis-tag-jews-on-twitter-harassment-hate-speech-politics/>.

²⁷⁸ Jones und Benesch 2019.

²⁷⁹ Informationen über das Tool stammen insbesondere aus einem Interview mit Claire Pershan (Renaissance Numérique) im April 2020.

²⁸⁰ Facebook: [#ichbinhier](#).

²⁸¹ Nach dem Wissen der Autorin existieren im Tessin keine solchen Ansätze.

Tabelle 8: Ansätze zu Gegenrede im Ausland

Projekt	(Haupt-)Projekttragende / Organisation	Land	Hervorzuhebendes Merkmal
#ichbinhier und Hatecontrol Tool	ichbinhier e.V.	DE / international	Organisiert Gegenrede (siehe oben); zusätzliches Tool zur temporären Schliessung von Facebook-Kommentarspalten bei Hassrede
No Hate Speech Movement	Neue deutsche Medienmacher, Europarat	DE	Bereitstellung von (insbes. humorvollem) Gegenredematerial gegen unterschiedliche Typen rassistischer Hassrede
Democratic Meme Factory	La Red	DE	Kreativer Ansatz der Produktion von «Memes»
Love Storm	Bund für Soziale Verteidigung e.V.	DE	Training und Aktionen von Gemeinschaft; Fokus auf Solidaritätsbekundung
Online Civil Courage Initiative (OCCI) Rapport d'Observation de l'OCCI	Facebook (Gründer)	DE, UK, FRA	Unterstützung von NGOs und Aktivitäten; Entwicklung von Gegenredematerial
#jesuislà	Xavier Brandao (Gründer)	FRA	Organisiert Gegenrede; Frankophones Equivalent zu #ichbinhier
Seriously	Renaissance Numérique	FRA	Wissenschaftlich geprüftes Gegenredematerial und pädagogischer Ansatz (siehe oben)
ZARA	ZARA	AT	Wissen über und Training von Gegenrede für Jugendliche und Erwachsene
Youth Civil Activism Network ²⁸²	Institute for Strategic Dialogue (ISD)	Global	Globales Jugendnetzwerk für Gegenrede

Wie wirksam ist Gegenrede? Ihr Grundprinzip ist auf jeden Fall einleuchtend. Besonders bei Inhalten, die weder illegal sind noch gegen die Gemeinschaftsstandards der Plattformen verstossen, aber nichtsdestotrotz von Vielen als problematisch empfunden werden, kann Gegenrede die normierende Funktion der Gesellschaft übernehmen. Insbesondere bei den unspezifischen, Gruppen-bezogenen Angriffen, wo sich keine direkt betroffene Einzelperson zur Reaktion gezwungen fühlt, könnte Gegenrede einer Verantwortungsdiffusion entgegenhalten. Wird Verbreiter*innen von Online-Hassrede nicht «die Plattform überlassen», führt dies zu Hassrede-hemmenden Erkenntnissen («hassende Nutzende sind ja doch eine Minderheit»; «Hassrede wird ja von anderen sanktioniert»). Dies hat Signalwirkung, wobei sich Zuschauende tendenziell anpassen. Diese informelle soziale Kontrolle konnte in Experimenten²⁸³ bestätigt werden. Die Effekte verstärken sich, wenn die «strafende» Person aus der eigenen Gruppe stammt und hohen Status oder Vertrauen genießt.²⁸⁴ Dies spricht für die Involvierung öffentlicher Personen. Die Wirksamkeit von Gegenrede wurde auch am Praxisbeispiel von #ichbinhier festgestellt.²⁸⁵ Deren Gegenredekommentare machten die darauffolgende Diskussion weniger aggressiv und motivierten Nicht-Mitglieder von #ichbinhier, aus ihrer passiven Rolle herauszusteigen und ebenfalls gegenzureden.

Ob Ansätze für Gegenrede wirksam sind, hängt stark von ihrem Kontext und ihrer konkreten Umsetzung ab. Laut Expert*innen²⁸⁶ sind sie tendenziell wirksamer, wenn:

- das Material und Tool zusammen mit der Zielgruppe aufgebaut und an ihr getestet wird. Die Zielgruppe variiert je nach Ausrichtung des Ansatzes und beinhaltet meist das (noch) passive Publikum und die potentiellen Gegenredner*innen;
- Gegenrede-Aktionen auf die (Online-)Gemeinschaft, ihre Kultur, Werte und Plattform zugeschnitten sind. Idealerweise werden geläuterte Verbreiter*innen von Online-Hassrede aufgrund ihres Verständnisses involviert und als weitere Zielgruppe behandelt;
- das Online-Publikum eher mittelgross, inaktiv und unentschieden ist;
- die Gegenrede konstruktiv und zivilisiert ist;
- Gegenredende mit viel digitaler Reichweite involviert werden;

²⁸² Pfad: www.isdglobal.org > Programmes > Grassroots Networks > [Youth Civil Activism Network](#) (YouthCAN).

²⁸³ Álvarez-Benjumea und Winter 2018.

²⁸⁴ Munger 2017.

²⁸⁵ Ziegele et al. 2019.

²⁸⁶ Z.B. Laubenstein und Urban 2018; Reynolds und Tuck 2016; Schieb und Preuss 2016.

- die individuelle Bereitschaft von Nutzenden zur aktiven und regelmässigen Gegenrede berücksichtigt wird (diese hängt vom sozialen Hintergrund und von individuellen Erfahrungen ab);
- mindestens ein Kern von Gegenredner*innen über eine längere Zeit (natürlich freiwillig) verpflichtet werden können;
- Gegenredematerial aktuellen Debatten konstant angepasst wird.

Gleichzeitig sollten die folgenden Risiken nicht ignoriert werden:

- Selbstgefährdung durch eigene Exposition; Schutzmassnahmen sind hier einzuplanen.
- Humoristische Gegenrede ist vielversprechend, jedoch unter Umständen kontraproduktiv, da es als herablassend oder arrogant empfunden werden kann.
- Unterwanderung durch Fake-Profile (physisches Kennenlernen der Mitglieder als Abhilfe).
- Die (kontinuierliche) Produktion von Gegenredematerial kann zeitaufwändig sein.
- Die grundsätzliche Gefahr, durch die Bezugnahme auf Hassrede diese zu legitimieren.
- Die Schwierigkeit, ideologisch überzeugte Verbreiter*innen von Hassrede zu bekehren.

Um den Erfolg von Gegenredekampagnen zu quantifizieren, sind folgende Kriterien relevant²⁸⁷: Aufmerksamkeit (d.h. digitale Reichweite, Anzahl der aufgetauchten Tweets, Videoaufrufe etc.), Interaktion (d.h. Austausch der Kampagne mit Nutzenden über Nachrichten) und Wirksamkeit (d.h. Einstellungs- und Verhaltensänderungen von Nutzenden). Grundsätzlich wird jedoch beobachtet, dass die wenigsten Ansätze eine signifikante Social-Media-Präsenz oder eine durchdachte Strategie vorweisen. Nichtsdestotrotz kann man mit bescheidenen Kosten eine hohe Reichweite erlangen²⁸⁸: Eine Facebook-Gegenredeinitiative aus dem Vereinigten Königreich erreichte mit einem Budget von \$3750 über 670'000 Nutzende.²⁸⁹ Ressourcen können zudem durch das Zurückgreifen auf bestehende Online-Gemeinschaften tief gehalten werden.

²⁸⁷ Silverman et al. 2016: S. 12.

²⁸⁸ Laubenstein und Urban 2018.

²⁸⁹ Silverman et al. 2016: S. 7.

8 SCHWEIZER STELLEN: HERAUSFORDERUNGEN UND ANLIEGEN

In diesem Kapitel wird auf staatliche und nicht-staatliche Stellen eingegangen, die sich in der Schweiz mit Rassismus beschäftigen. Welche Herausforderungen und Fragen stellen sich diesen Verwaltungsstellen, öffentlichen bzw. öffentlich unterstützten Beratungsstellen und privaten Organisationen in Bezug auf rassistische Online-Hassrede? Was brauchen sie zukünftig, um (besser) auf rassistische Online-Hassrede reagieren zu können? Die folgende Situationsanalyse basiert auf Informationen aus einem Workshop-Bericht der FRB²⁹⁰ und aus Interviews mit zuständigen Personen. Die hier dargestellten Aspekte sollten auf die meisten relevanten Stellen zutreffen; dies schliesst nicht aus, dass einzelne Aspekte auf spezifische Stellen nicht oder nur teilweise zutreffen.

Welche Stellen wie mit rassistischer Online-Hassrede umgehen, ist auch von der Dynamik der vorliegenden Fälle abhängig. Einerseits gibt es kleinere Fälle: Sie beschreiben lokal begrenzte Situationen, in der beispielsweise eine unbekannte Einzelperson auf einem nicht-öffentlichen Kanal wie WhatsApp ein einziges Mal rassistisch angegriffen wird. Andererseits gibt es grössere Fälle: Sie beschreiben «Shitstorm» oder Hasskampagnen mit vielen rassistischen Online-Beiträgen, die sich innert kürzester Zeit verbreiten. Die grossen Fälle sind potentiell schwer kontrollierbar. Oft berichten auch klassische Medien darüber, was die Sichtbarkeit und Dynamik dieser Fälle noch verstärkt. Beispiele sind die hitzigen Diskussion um einen Transitplatz für ausländische Fahrende bei Wileroltigen im Jahr 2019/2020²⁹¹, der Aufruf in einer WhatsApp-Gruppe von Sekundarschülern zum Tragen von Edelweisshemden im Jahr 2015²⁹² oder Hassrede gegen Asylsuchende in einer spontan errichteten Facebook-Gruppe im Jahr 2014²⁹³, die deren Unterbringung in einem Schulhaus verhindern wollte.

8.1 Verwaltungsstellen

Die Verwaltungsstellen in der Schweiz sind föderalistisch aufgebaut. Auf Bundesebene gibt es eine fachspezifische Stelle (Fachstelle für Rassismusbekämpfung FRB), in den Kantonen und Gemeinden sind aus Rassismusbekämpfung spezialisierte Stellen meist an die Integrationsstellen angegliedert (die Integrationsdelegierten sind hier die relevanten Ansprechpartner). Die Integrationsstellen haben im Rahmen des Schwerpunkts Diskriminierungsschutz einen Auftrag zur Bekämpfung von Rassismus. Sie sind im Speziellen zuständig für Information und Sensibilisierung. Die rein staatliche Finanzierung dieser Stellen schränkt ihre Autonomie in Bezug auf ihr Tätigkeitsfeld ein.

Die FRB als zentrale Anlaufstelle für das Thema Online-Rassismus stellt vier grundsätzliche Ziele auf²⁹⁴:

1. *Die auf Rassismus spezialisierten Beratungsstellen sollen befähigt werden, mit Online-Rassismus umzugehen.* Die Stellen sollen insbesondere kleine Fälle eigenständig bearbeiten können.
2. *Die breite Öffentlichkeit soll über rassistische Online-Hassrede informiert und sensibilisiert werden.* Die öffentliche Auseinandersetzung soll ein Bewusstsein für die Einflussfaktoren und Auswirkungen von Online-Hassrede generieren und dadurch die Bevölkerung befähigen, mit rassistischer und diskriminierender Hassrede im Internet kritisch und verantwortungsvoll umzugehen.
3. *Die Präventionsmassnahmen gegen Online-Rassismus sollen verstärkt werden.* Prävention von Rassismus soll heute analog als auch digital erfolgen, wobei Massnahmen und Stellen im Internet beworben werden sollen. Projekte sollten nach ihrer erwarteten Wirksamkeit gefördert, eine zentrale Meldestelle bedacht und die Vernetzung relevanter Akteure verstärkt werden.

²⁹⁰ [Fachstelle für Rassismusbekämpfung 2019](#). Pfad: www.frb.admin.ch > Tätigkeitsfelder > Medien und Internet > Prävention im Netz > Workshops.

²⁹¹ Schweiz aktuell, SRF vom 9. Februar 2020: [Ja zum Transitplatz für Fahrende in Wileroltigen](#). Ein beispielhafter Kommentar, der im Rahmen dieses Falles bei der GRA gemeldet wurde, lautete: «Für was 1000 vo franke usgä. Wesi glich überau ihre müll wärde verteile, s het platz für 36 wohnwäge, u der rest? U ner chames au jahr saniere weu ds pack eh aues kaputt macht, bi üs i de läde chömesi cho chlaue bis zum geht nicht mehr, u ner schiebtmene no weiss ni was i arsch, nei sry».

²⁹² Serafini, Aargauer Zeitung vom 21. Dezember 2015: [Edelweiss-Streit: Auf Whatsapp blüht der jugendliche Patriotismus](#).

²⁹³ Humanrights.ch vom 23. August 2017: [Hassaufrufe im Internet - Schweizer Fälle und Politik der Newsportale](#).

²⁹⁴ Pfad: www.frb.admin.ch > Tätigkeitsfelder > Medien und Internet > [Internet](#).

4. *Interventionsangebote (z.B. in Gemeinden) sollen verstärkt werden.* Lokale Stellen sollen unterstützt werden, um in grösseren Fällen intervenieren zu können.

Punktuelle Einblicke in die Verwaltungsstellen ausserhalb der FRB, wie beispielsweise in die Integrationsförderung der Stadt Zürich²⁹⁵, ergeben ein Bild, worin der Umgang mit Online-Rassismus bisher noch keine allzu grosse Rolle spielt. Obwohl einzelne Personen sehr wohl sensibilisiert sind für das Thema, scheinen die Stellen in ihrer Gesamtheit in der Praxis wenig mit Fällen von rassistischer Online-Hassrede konfrontiert zu sein bzw. sich kaum explizit damit zu befassen. Dies könnte damit zusammenhängen, dass das Bewusstsein für das Thema innerhalb der Verwaltungsstellen und im relevanten Umfeld noch wenig verbreitet ist.

Tabelle 9 gibt einen Einblick in die Herausforderungen und Anliegen von Verwaltungsstellen im Allgemeinen.

Tabelle 9: Bestehende Situation für Verwaltungsstellen

Herausforderungen und Fragestellungen in Bezug auf rassistische Online-Hassrede	Was brauchen sie, um effizienter zu handeln?	Wird im vorliegenden Bericht wie und wo adressiert
Lücken in Wissensstand	Verbesserung des Wissensstandes über Social Media und rassistische Online-Hassrede	Aufarbeitung von Wissensstand (Kap. 3-6)
Mangelndes Bewusstsein für rassistische Online-Hassrede	Mehr Bewusstsein, dass Rassismus heute auch online stattfindet	Aufarbeitung von Wissensstand als Basis für Sensibilisierung (Kap. 3-6)
Optimale Vermittlung von Wissen	Wissen über relevante Zielpublika und ihre Erreichbarkeit	Überblick zu Zielgruppen und Erreichbarkeit (Kap. 9.3)
Mangelndes Wissen über Erfolgsfaktoren für Projekte zu Online-Rassismus	Kriterien zur Beurteilung von Förderprojekten	Bestehende Gegenmassnahmen (Kap. 7.5); Empfehlung von Beurteilungskriterien für Projekteingaben (Kap. 9.4)
Förderung der Zusammenarbeit zwischen relevanten Stellen	Überblick zu spezialisierten Personen und Organisationen, um Vernetzung voranzutreiben	Überblick existierender Projektansätzen im In- und Ausland (Kap. 7)

8.2 Öffentliche (oder öffentlich finanzierte) Beratungsstellen

Beratungsstellen, die sich mit Rassismus beschäftigen und in jedem Kanton vorzufinden sind (siehe dazu das «Beratungsnetz für Rassismusopfer»²⁹⁶), sind entweder ganz oder teilweise staatlich finanziert, was sie in ihrem Tätigkeitsfeld einschränkt, abhängig vom Anteil der staatlichen Finanzierung am Gesamtbudget.

Beratungsstellen haben das Thema Online-Rassismus zwar «auf dem Radar», aber sehen verschiedenste Verbesserungsmöglichkeiten (siehe Tabelle 10). Grundsätzlich ist die bisher tiefe Meldezahl eine zu dünne Erfahrungsbasis, als dass man darauf systematisch Kompetenzen hätte aufbauen können. Laut dem Beratungsnetz für Rassismusopfer²⁹⁷ wurden im Jahr 2019 nur 23 Fälle von Online-Rassismus gemeldet, die als Diskriminierung bestätigt wurden – die Dunkelziffer dürfte um einiges höher liegen.²⁹⁸ Dies widerspiegelt die Grundproblematik des Underreporting bei Rassismus: Betroffene trauen sich nicht, sich zu melden oder wissen nicht, wo sie Rat bekommen. Sie ordnen ihre Erfahrungen teilweise zwar als verstörend, aber nicht sofort als diskriminierend ein, oder akzeptieren sie als gegeben und suchen somit keine Beratungsstelle auf.²⁹⁹ Die relativ neuere, digitale Dimension von Rassismus könnte diese Probleme noch verstärken.

²⁹⁵ Z.B. Interview mit Michael Bischof (Stadt Zürich, Integrationsförderung) im Februar 2020.

²⁹⁶ www.network-racism.ch

²⁹⁷ Beratungsnetz für Rassismusopfer 2020.

²⁹⁸ Bundesamt für Statistik 2019: Die Erhebung «Zusammenleben in der Schweiz» stellt fest, dass sich 28 Prozent der Bevölkerung als Opfer von Diskriminierung im Allgemeinen bezeichnen.

²⁹⁹ In Bezug auf die Fallbearbeitung wird zudem auf die als schwach wahrgenommene Gesetzesgrundlage (z.B. kein Verbandsklagerecht) hingewiesen. Dadurch wären die Hände gebunden, um beispielsweise eine private Person rechtlich zu repräsentieren und sie somit vor weiterer Exposition digitaler Hassrede zu schützen. Für

Auf jeden Fall können die Stellen mit den bestehenden Beratungskompetenzen und -instrumenten Vorfälle teilweise bearbeiten, wenn auch noch unsystematisch. Ihnen stellt sich die Frage, wie auf rassistische Online-Angriffe reagiert werden soll. Nach welchen Kriterien soll die Relevanz oder auch Schwere von Angriffen eingeordnet werden? Wie ist auf unterschiedliche Täterschaften, beispielsweise Einzeltäter*innen oder Netzwerke, zu reagieren? Wie können relevante Inhalte rasch identifiziert werden, und wie kann strategisch dagegen vorgegangen werden? Wo ansetzen, wann reagieren, wann nicht? Welche Rolle können die staatlichen oder vom Staat finanzierten Beratungsstellen übernehmen, und wo sind die Grenzen des staatlichen Handelns? Diese Fragen reflektieren auch bestehende Wissenslücken: Wie manifestiert sich rassistische Hassrede im Internet, wie funktioniert sie, welche technischen Gegebenheiten gehören dazu und wie ist sie rechtlich erfasst? Eine diesbezügliche Sensibilisierung sei laut den Interviewten auch auf höheren Hierarchiestufen und bei sonstigen Auftraggebern vonnöten, damit die Berater*innen Fälle effizient bearbeiten könnten.

Ein weiteres Hindernis – dem unvollständigen Wissen über rassistische Online-Hassrede zuzuschreiben – ist die praktisch fehlende digitale Bewerbung des Beratungsangebots. Damit ist gemeint, dass sich auf den Webseiten der meisten Beratungsstellen für Rassismus kaum Informationen zum Thema Online-Rassismus oder zu Beratungsleistungen in Bezug auf Online-Rassismus finden. Aktuell begrenzen sich die Informationen und Beispiele auf den relevanten Webseiten grösstenteils auf Rassismus auf dem Wohnungsmarkt und am Arbeitsplatz. Die Online-Dimension wird kaum angesprochen. Verbesserungsfähig wäre dementsprechend die Online-Präsenz von Beratungsstellen, d.h. dass Beratungsstellen beispielsweise auf ihren Webseiten und auf SocialMedia Online-Rassismus thematisieren und – sobald genügend Beratungskompetenzen in diesem Bereich vorhanden sind – auch diese bewerben.

Interventionen in grösseren Fällen übersteigen schnell die verfügbaren Kapazitäten von Beratungsstellen. Es fehlt den betroffenen Gemeinden oder Organisationen an inhaltlicher Expertise, um solche Situationen fachgerecht begleiten können, insbesondere in den Bereichen Recht, Kommunikation und psychosoziale Schutzmassnahmen.

Tabelle 10: Bestehende Situation für öffentliche (oder öffentlich finanzierte) Beratungsstellen

Herausforderungen und Fragestellungen in Bezug auf rassistische Online-Hassrede	Was brauchen sie, um effizienter zu handeln?	Wird im vorliegenden Bericht wie und wo adressiert
Teilweise mangelndes Bewusstsein für die digitale Dimension von Rassismus, auch im erweiterten Umfeld	Mehr inhaltliches Wissen über das Funktionieren von Social Media und rassistischer Online-Hassrede, in Beratungsstellen selbst, aber auch bei übergeordneten Stellen, Geld- und Auftraggebern	Aufarbeitung von Wissensstand als Basis für Sensibilisierung (Kap. 3-6)
Gestaltung von Beratung/Begleitung bei individuellen bzw. interpersonellen Vorfällen	Mehr technische und rechtliche Kenntnisse für Interventionen in kleineren Fällen	Definition notwendiger Kenntnisse und Kompetenzen für kleinere Fälle (Kap. 10.1)
Mangelndes Wissen darüber, welche Massnahmen von staatlichen und nicht-staatlichen Stellen übernommen werden können	Mehr Wissen zur Verteilung von Kompetenzen (und Grenzen des staatlichen Handelns)	Überblick zu den Zuständigkeiten relevanter Stellen (Kap. 9.1)
Keine Erfahrungswerte bei grösseren Fällen	Orientierung für Umgang mit grösseren Fällen; Backup durch Expert*innen	Definition notwendiger Kenntnisse und Kompetenzen für grössere Fälle (10.1.2)
Bekanntmachen von Beratungsangebot	Wissen darüber, wie das eigene Angebot die Zielgruppen erreichen kann	Überblick zu Zielgruppen und Erreichbarkeit (Kap. 9.3); Massnahmen für Bekanntmachen und Nutzarmachen von Beratungsangebot (Kap. 10.2)

8.3 Private Organisationen

Organisationen, die rein privat finanziert sind, können aufgrund der Unabhängigkeit von staatlicher Finanzierung am autonomsten agieren. Sie stehen nichtsdestotrotz in einem Auftragsverhältnis zu ihren sonstigen Auftraggebern, die ihre Rolle mitdefinieren. Die wenigen bestehenden privaten Organisationen, die sich in der Schweiz mit rassistischer Online-Hassrede beschäftigen (z.B. die GRA oder Netzcourage), schätzen laut Eindruck der Autorin deren Gefahr im Schnitt als relevanter und akuter ein als rein staatlich finanzierte Beratungsstellen – was damit zu tun haben mag, dass sie sich durch ihre unabhängige Rolle in einem Tätigkeitsfeld bewegen, in dem sie mit mehr Fällen konfrontiert sind. Aktuell engagieren sich diese Organisationen im Bereich Beratung, Monitoring und Gegenrede (siehe mehr dazu in Kap 7.5).

Ein zentraler Aspekt, der von dieser Seite aufgeworfen wird (siehe Tabelle 11 für alle Aspekte), ist die als mangelhaft wahrgenommene Sensibilität bei relevanten Stellen wie der Polizei.³⁰⁰ Prototypisch beobachtbar ist dabei der verbreitete Vorschlag an Betroffene, bei Angriffen das eigene Profil zu löschen. Dieser Vorschlag ignoriert die Tatsache, dass Menschen heute ihre persönlichen und geschäftlichen Kontakte (auch) in Social Media aufbauen, insbesondere junge Menschen, Politiker*innen oder Journalist*innen. Eine Profillöschung entspräche einer Löschung des sozialen Netzwerks und einem Rückzug, indem man sich selbst zum Schweigen bringt. Diese Reaktion packe das eigentliche Problem jedoch nicht an der Wurzel. Mangelnde Sensibilität führt zudem zu negativen Erfahrungen mit relevanten Stellen.³⁰¹ Betroffene würden sich beispielsweise von der Polizei nicht immer ernst genommen fühlen. Auch das Anzeigeverfahren sei aufwändig, da zwischen mehreren digitalen Medien und analogen Situationen hin und her gewechselt werden muss (z.B. Screenshot machen und ausdrucken, auf Polizeiposten gehen, Polizei tippt Linkadresse ab etc.).

Tabelle 11: Bestehende Situation für private Organisationen

Herausforderungen und Fragestellungen in Bezug auf rassistische Online-Hassrede	Was brauchen sie, um effizienter zu handeln?	Wird im vorliegenden Bericht wie und wo adressiert
Teilweise ungenügendes Wissen	Hilfe bei der Erarbeitung von Grundlagen und Instrumenten	Aufarbeitung von Wissensstand (Kap. 3-6)
Wahrgenommene mangelnde Sensibilität in erweiterten Umfeld	Mehr Sensibilisierung relevanter Stellen, z.B. Polizei	Aufarbeitung von Wissensstand als Basis für Sensibilisierung (Kap. 3-6); Handlungsempfehlungen (Kap. 11)
Ungenügende personelle und finanzielle Ressourcen (insbesondere für grössere Fälle)	Finanzielle und ideelle Unterstützung seitens der staatlichen Stellen (z.B. öffentliche Anerkennung, Logo, Stellenprozente)	Beratung und Interventionen (Kap. 10); Empfehlungen (Kap. 11)

³⁰⁰ Z.B. Interview mit Jolanda Spiess-Hegglin (#Netzcourage) im Februar 2020.

³⁰¹ Z.B. Interview mit Jolanda Spiess-Hegglin (#Netzcourage) im Februar 2020 und mit Dominic Pugatsch (Stiftung gegen Rassismus und Antisemitismus, GRA) im März 2020.

9.1 Kompetenzen und Zuständigkeiten

Der folgende Überblick zu den aktuellen und möglichen Kompetenzen und Zuständigkeiten von Verwaltungsstellen, öffentlichen oder öffentlich finanzierten Beratungsstellen und privaten Organisationen in Bezug auf rassistische Online-Hassrede greift die vorgestellten Massnahmen aus dem Kapitel 7.5 («Zivilgesellschaft») auf und ergänzt sie mit den Kompetenzen und Zuständigkeiten, welche die Stellen bereits für den Offline-Rassismus besitzen.³⁰² Es werden primär präventive, aber auch intervenierende Massnahmen berücksichtigt, da diese in der Praxis oft zusammenfallen.

Grundsätzlich fällt bei der Aufstellung (Tabelle 12) auf, dass private Organisationen mehr Massnahmen ergreifen, die sich spezifisch auf das gegebene Phänomen beziehen (z.B. Gegenrede, Flaggen). Staatliche oder staatlich unterstützte Stellen können diese Aufgaben theoretisch auch wahrnehmen, solange ihr Auftrag dies vorsieht und es keine rechtlichen Grundlagen gibt, die dagegen sprechen. Theoretisch könnten Verwaltungsstellen beispielsweise Monitoring betreiben. Aufgrund der Tatsache, dass der Staat jedoch nicht in die Meinungsfreiheit eingreifen darf, können sie keine Gegenrede machen (im Gegensatz zu politischen Akteur*innen, die den Verwaltungsstellen vorstehen und dies dürfen). Eine Ausnahme besteht da, wo staatlich und staatlich unterstützte Stellen eigene Social-Media-Präsenzen haben. Hier ist es relevant, dass die Stellen ihre Social-Media-Profile moderieren und Hassrede nicht tolerieren. Private Organisationen sind dagegen autonomer und können ein grösseres Spektrum von Aufgaben wahrnehmen, speziell wenn sie nicht mit den öffentlichen Erwartungen konfrontiert sind, die mit staatlicher (Mit-)Finanzierung einhergehen. Insbesondere bei Gegenrede kann der Charakter einer «Bewegung von unten» bzw. aus der breiten Bevölkerung zu einer grösseren Akzeptanz der Gegenrede führen.³⁰³ Dies stimmt wiederum mit der Situation überein, dass der Staat nicht in die Meinungsfreiheit eingreifen darf. Auch bei grösseren Vorfällen mit meist politischer Dimension gilt abzuklären, welche Stellen wie darauf reagieren können und dürfen. Beispielsweise ist es denkbar, dass eine staatliche Institution öffentlich Stellung dazu beziehen könnte, jedoch wäre es einer privaten Organisation überlassen, mit Gegenrede einzugreifen. Auch bei ihnen gilt es im Einzelfall zu prüfen, was in ihren Zuständigkeitsbereich fällt, abhängig vom Anteil der staatlichen Finanzierung, dem Auftragsinhalt, den Interessen der Auftraggebenden und den eigenen Kapazitäten.

Tabelle 12: Zuständigkeiten relevanter Stellen in der Schweiz

Verwaltungsstellen	• Prävention und Sensibilisierung • Koordinations- und Vernetzungsfunktion übernehmen • Wissenserarbeitung/-vermittlung (im Rahmen bestehender oder neuer Gefässe) • Position beziehen (im Rahmen ihres Auftrags) • Projekte privater Organisationen unterstützen; Entwicklung von Interventionskompetenzen unterstützen
Öffentliche (oder öffentlich finanzierte) Beratungsstellen	• Prävention und Sensibilisierung • Unterstützung von privaten Organisationen • Vermittlung von Know-how (z.B. über Social Media und Online-Rassismus) • Positionsbezug im Rahmen des Auftrags • Beratung und Begleitung bei individuellen bzw. interpersonellen Vorfällen, Rückberatung mit Expert*innen mit langjähriger Erfahrung, Übergabe von Fällen an Expert*innen
Private Organisationen	• Prävention und Sensibilisierung • Beratung und Weitervermittlung • Backup bei Interventionen in grösseren Fällen • Meldetool und/oder aktives Monitoring-Tool von Social Media und klassischen Medien • Gegenredematerial bereitstellen, Gegenrede koordinieren und leisten • Rassistische Inhalte an Social-Media-Anbieter flaggen, an Behörden weiterleiten

³⁰² Grundlage dieses Unterkapitels sind insbesondere Interviews mit den Zuständigen der Stellen sowie Diskussionsmaterial aus einem Workshop, siehe Fachstelle für Rassismusbekämpfung 2018.

³⁰³ Interview mit Claire Pershan (Renaissance Numérique) im April 2020. Sie ist Projektleiterin des Gegenrede-Tools [Seriously](#), das nicht staatlich finanziert ist.

9.2 National ausgerichtetes Meldetool

Auf nationaler Ebene bietet sich die Errichtung eines zentralen Meldetools an. Dieses könnte eine oder mehrere Aufgaben abdecken, u.a.:

- Informations- und Beratungsmaterial zu rassistischer Online-Hassrede
- Möglichkeit des Meldens für Betroffene und für Beobachtende
- Online-Beratung
- Einschätzung der Strafbarkeit von Beiträgen
- Flagging an Social Media-Unternehmen (mit Status von Trusted Flagger) oder Signalisieren an klassische Medien und Politik
- Unterstützung beim Formulieren von Strafanzeigen
- Nachverfolgung und Rückmeldung des Meldestatus an meldende Nutzende
- Aktives Monitoring auf Social Media und klassischen Medienplattformen
- Publizieren von Monitoring-Berichten mit Statistiken (zu Meldung, Löschung, Themen etc.)

Welche Aufgaben das Meldetool übernehmen könnte und dürfte, wäre davon abhängig, an welche Stelle es angesiedelt wäre und wieviel Kapazitäten diese hätte. Soll das Tool begrenzt bleiben, könnte sie beispielsweise an der Eidgenössischen Kommission gegen Rassismus (EKR) andocken. Wäre eine Meldestelle mit unterschiedlichsten Aufgaben vorgesehen, würde sich ein Zusammenschluss privater Organisationen anbieten. Die Stelle könnte von einem Netzwerk mehrerer Personen getragen werden. Reicht dies nicht aus, könnte auf einen externen Expertenpool (z.B. Psycholog*innen, Soziolog*innen, Jurist*innen) zurückgegriffen werden – ideal wären dabei Expert*innen mit theoretischer als auch praktischer Erfahrung. Auch auf bestehende spezialisierte Organisationen (z.B. GRA, #Netzcourage, Stop Hate Speech) und Beratungsstellen könnte zurückgegriffen werden. Das Tool wäre idealerweise in allen Landessprachen (und unter Umständen in weiteren Sprachen) zugänglich, benutzer*innenfreundlich sowie niedrigschwellig gestaltet (z.B. durch anonymes Melden auf Webseite und App). Neben rassistischer Online-Hassrede könnten theoretisch auch andere Dimensionen von Hassrede sowie analoger Rassismus gemeldet werden. Für den Aufbau und die Lancierung eines relativ weiten Tools wäre eine breite Allianz aus staatlichen und staatlich unterstützten Stellen, zivilgesellschaftlichen Organisationen, (Fach-)Hochschulen, Vertretenden des Mediensystems und Social-Media-Anbietern ideal, um die nötige Wissensgrundlage sicherzustellen und die Sichtbarkeit zu verstärken. Die Benennung des Tools sollte letztlich sein rasches Finden über Suchmaschinen sicherstellen.

Ein nationales Tool ist aus verschiedenen Gründen sinnvoll. Heute kann jede Person überkantonale – und internationale – Hassrede verbreiten. Kantonale Spezifitäten treffen nur bedingt zu. Ein zentrales Tool könnte einen Wildwuchs an Ansätzen, Zuständigkeiten und Verantwortlichkeiten, der im Ausland teilweise beobachtet wird, verhindern, indem relevante Akteur*innen – wo möglich und sinnvoll – schon früh kollaborieren. Der von den Kantonen unabhängige Status garantiert zudem, dass die Stelle auf Meldungen neutral reagieren könnte; beispielsweise bei Fällen von Hassrede lokaler Politiker*innen, zu denen kantonale Stellen in einem Abhängigkeitsverhältnis stehen können. Der Online-Charakter des Tools würde zudem zu Meldungen motivieren, da Betroffene weniger Medienwechsel auf sich nehmen müssen (verglichen mit physischen oder telefonischen Meldemöglichkeiten). Das Tool würde die bestehenden Beratungsstellen jedoch keinesfalls überflüssig machen (siehe ihre Rolle in Kap. 10).

9.3 Zielgruppen und Erreichbarkeit

Grundsätzlich sollte bei allen Massnahme-Ansätzen die *Zielgruppe* sorgfältig bedacht und ausgewählt werden. Bei rassistischer Online-Hassrede kommen diverse Zielgruppen in Frage, die auf vielfältigen Wegen erreicht werden können. Grundsätzlich sind zur Reduktion des Medienbruchs primär digitale Kanäle empfohlen. Das digitale Angebot sollte sich nach den Mediengewohnheiten der Zielgruppe richten (d.h. welche Plattformen werden wie, wann und wie lange genutzt). Da sich die Nutzung von Plattformen dynamisch ändert (beispielsweise hat sich die App «TikTok» innert kürzester Zeit zu einer der meistgenutzten Kurzvideoplattformen unter Jugendlichen entwickelt), sollten die Mediengewohnheiten der relevanten Zielgruppen zeitnah an die jeweilige Massnahme in Erfahrung gebracht werden. Dabei kann einerseits auf vorhandene Publikationen zur Mediennutzung

zurückgegriffen werden.³⁰⁴ Andererseits könnten unabhängige Studien zur Mediennutzung in der zu interessierenden Zielgruppe in Auftrag gegeben werden. Sichtbarkeit kann durch eine breite Kanalwahl, d.h. Webseiten, unterschiedliche Social-Media-Kanäle, Offline-Kanäle (Workshops, Vorträge) und Multiplikator*innen erhöht werden. Gleichzeitig können im Sinne eines Schneeballeffekts bestehende digitale Netzwerke unterschiedlicher sozialer Gruppen genutzt werden. Tabelle 13 präsentiert relevante Zielgruppen und Ideen, wie diese Gruppen erreicht werden können. Je nach gewählter Massnahme müssen Vorgehensweisen jedoch angepasst werden und sind somit als vorläufiges Brainstorming zu verstehen.

Tabelle 13. Zielgruppen und Erreichbarkeit

Zielgruppe	Erreichbar über ...
Rolle in rassistischer Hassrede	
Betroffene	• Insbesondere Beratung (siehe dazu Kap. 10.2) und nationales Meldetool • Identifizierung und Anvisierung über Merkmale, die Betroffene laut Studien oft aufzeigen (siehe Kap. 6.2). Darauf basierende, Gruppen-spezifische Bewerbung auf Social Media.
Zuschauende bzw. Publikum	• Abhängig vom Ziel der vorliegenden Massnahme • Bei Gegenrede über die Beachtung der Wirksamkeitsfaktoren, Kap. 7.5.4
Verbreiter*innen von Online-Hassrede	• One-to-one Kontaktaufbau im Internet (bekannt aus Deradikalisierungsprojekten; jedoch zeitaufwändig) • Kontakt über Strafanzeigen und evtl. -verfahren (siehe #Netzcourage) • Identifizierung und Anvisierung über Merkmale, die Verbreiter*innen von Online-Hassrede laut Studien aufzeigen (siehe Kap. 5.2) • Bei Gegenrede über die Beachtung der Wirksamkeitsfaktoren, Kap. 7.5.4
Vulnerable Gruppen, d.h. potentielle Betroffene oder Verbreiter*innen von Online-Hassrede	• Identifizierung und Anvisierung über Merkmale, die laut Studien Risiko erhöhen (siehe Kap. 5.2 und 6.2) • Identifizierung von Subkulturen (Gamification, Satire, Kunst) und Mobilisierung «bevor Extremisten dies tun»³⁰⁵
Alterskategorie	
Kinder & Jugendliche	• Schulen; in Medienkompetenz-Modulen (nach Lehrplan 21/PER) • Abklärung von Mediennutzungsverhalten (z.B. aktuell TikTok) und Anpassung
Junge Menschen	• Abklärung von Mediennutzungsverhalten und Anpassung
Erwachsene, ältere Personen	• Abklärung von Mediennutzungsverhalten und Anpassung (grundsätzlich jedoch schwerer zu kontaktieren)
Sonstige sozialen Gruppen	
Nationale, kantonale und kommunale Politiker*innen (als Betroffene oder Täter*innen ³⁰⁶)	• Parteien und deren Distribution von Beratungsmaterial (z.B. Informationsflyer für Politiker*innen) • Hassrede-verbreitende öffentliche Figuren theoretisch kontaktierbar über nationales Meldetool
Migrant*innen, Schweizer*innen mit Migrationshintergrund, Flüchtlinge	• Insbesondere Beratung, siehe dazu Kap. 10.2
NGOs, Aktivist*innen (als Betroffene oder Multiplikator*innen)	• Insbesondere Beratung, siehe dazu Kap. 10.2

³⁰⁴ Z.B. auf den [IGEM-digiMONITOR](#) der Interessensgemeinschaft Elektronische Medien Schweiz (IGEM).

³⁰⁵ Ebner, zitiert in Heiderich, Fearless Democracy vom 26. April 2018: [Extremismusforscherin Julia Ebner: „Hasskampagnen folgen einem klaren Muster“](#).

³⁰⁶ Für eine Diskussion zur Verantwortung von Politiker*innen für Online-Hassrede, siehe sCAN 2020.

9.4 Kriterien für die Beurteilung von Projekteingaben

Für eine Einschätzung, ob Projekte grundsätzlich Anspruch auf Finanzierung haben, sei hier auf die Kriterien verwiesen, die auf der Website der FRB³⁰⁷ genannt sind. Darüber hinaus werden in Tabelle 14 zusätzliche Kriterien aufgelistet, die im Kontext der rassistischen Online-Hassrede als besonders relevant angesehen werden. Einige Kriterien sind relevanter für manche Projektansätze als für andere; ob ein Kriterium angewendet wird, muss im Einzelfall entschieden werden. Die Kriterien wurden aus Evaluationen vergangener Projekte im In- und Ausland und deren Empfehlungen abgeleitet und zusammengefasst (Kap. 7.5).

Tabelle 14. Kriterien für die Beurteilung von Projekteingaben

Expertise und Vorbereitung	• Liegt technische Expertise vor? Liegt Expertise für das Management von Marketing Tools auf Social Media vor? Ist die Expertise jederzeit verfügbar (insbes. bei Notfällen)? • Liegt ein theoretisches Verständnis von digitalen Medien vor? • Knüpft das Projekt an vorhandene Projekte im In- und insbesondere Ausland an?
Zielgruppe	• Wieviel Kenntnis über die Zielgruppe, ihre digitalen Gewohnheiten und Erfahrungen liegt vor? Wie wird diese Information bei der Umsetzung der Massnahme und Zielgruppenenerreichung eingesetzt? • Wird die Zielgruppe beim Aufbau und Testen der Massnahme involviert? • Wird auf eine eindeutige und sinnvolle Zielgruppe eingeschränkt? Wie werden noch nicht sensibilisierte oder vulnerable Gruppen berücksichtigt? • Wie sensibel wird mit Begrifflichkeiten und Perspektiven umgegangen? Wird ihr möglicher Effekt auf die Zielgruppe reflektiert? (z.B. Widerstand durch Nutzung von Begriff «Hassrede»)
Projekthalte	• Welche Typen von rassistischer Online-Hassrede (z.B. anti-muslimisch, antisemitisch) werden berücksichtigt? Wie wird mit der Herausforderung des Einbezugs von mehr als einer Hassrede-Form umgegangen? (Insbes. relevant für Monitoring- und Gegenrede-Projekte) • Nimmt das Projekt den lokalen und nationalen Kontext und gegenwärtige Ereignisse auf? • Ist die Herangehensweise simpel gestaltet? (Insbes. relevant für Medienkompetenz- und Gegenrede-Projekte) • Ist die Herangehensweise ungewöhnlich und kreativ? (erhöht u.a. Sichtbarkeit auf Social Media) • Welche Art von Engagement und wie viel davon wird von Teilnehmenden erwartet? Wie stellt sich das Projekt auf unerwartet viel Engagement organisatorisch ein? • Wie bedacht sind die eingeplanten Plattformen (Social Media, Webseiten, Apps) ausgewählt? Werden ihre Vorteile, Schwächen und Dynamiken im Hinblick auf die Zielerreichung reflektiert? Sind Offline-Ableger geplant?
Schutz	• Inwieweit ist der Schutz von involvierten Teilnehmenden vorgesehen? (Insbes. relevant für Monitoring- und Gegenrede-Projekte) • Werden ethische und Sicherheitsbedenken reflektiert? (z.B. Online-Privatsphäre, Verarbeitung von Online-Daten) • Inwieweit wird das mögliche Auftreten von Fake-Profilen und Bots thematisiert? • Werden sonstige potentielle Risiken thematisiert?
Trägerschaft	• Wie kann die ideologische Ausrichtung von Projekttragenden oder deren Bild in der Öffentlichkeit die Zielerreichung des Projekts beeinflussen? Inwiefern wird dies reflektiert und wie wird dem bei Notwendigkeit entgegengewirkt?
Öffentlichkeitsarbeit	• Wird in die Generierung medialer Aufmerksamkeit investiert, und wenn ja, wie? • Wird eine Vernetzung mit Vertreter*innen aus Politik, Wirtschaft, Medien, zivilgesellschaftlichen Organisationen und sonstigen Multiplikator*innen angestrebt? • Wird auf öffentliche Kritik oder digitale Angriffe vorbereitet? (z.B. Vorgehensweise planen, Verantwortlichkeit vergeben für Moderieren von Social-Media-Profil, FAQ mit Antworten auf erwartete Fragen vorformulieren) • Wird die Vorgehensweise transparent kommuniziert (z.B. Definition von Daten, Algorithmen etc.)? (Insbes. Relevant für Monitoring-Projekte)
Evaluation	• Wie wird der Erfolg bzw. die Wirksamkeit des Projekts gemessen? • Werden Metriken identifiziert, mit denen die Zielerfüllung bewertet werden kann? Ist eine klare Methode vorhanden, wie die gewählten Metriken beobachtet und erfasst werden? • Ist eine gewisse Unabhängigkeit der Evaluation gegeben?
Nachhaltigkeit	• Wird das Projekt unabhängig von einer zeitlich begrenzten Finanzierung durch die FRB über einen längeren Zeitraum ausgerichtet? Falls ja: Werden Leitfäden für die Weiterbetreuung ausgearbeitet?
Finanzierung	• Werden die Kosten für die Produktion von Material (insbesondere bei Gegenrede) und für die Verbreitung (z.B. Online-Werbung auf Social Media) eingeschätzt und begründet? • Werden die Ressourcen für Monitoring, Messung und Evaluation (z.B. mithilfe von Metriken) im Budget eingeplant? • Falls eine breite Mobilisierung über Social Media geplant ist: Werden zusätzliche Finanzierungskanäle wie Crowdfunding (d.h. eine Vielzahl von Internetnutzenden als Kapitalgebende) berücksichtigt?

³⁰⁷ www.frb.admin.ch > [Finanzhilfen](#)

10.1 Notwendige Kenntnisse und Kompetenzen von Beratungsstellen

Die Beratung von Fällen rassistischer Online-Hassrede verlangt Kenntnisse und Kompetenzen, die schon von der Beratung von Offline-Rassismus bekannt sind – diese sollten jedoch zusätzlich auf Internet- und Social-Media-basierte Aspekte ausgeweitet werden.³⁰⁸ Mit der zunehmenden Schwere der gemeldeten Fälle steigen die Anforderungen.

10.1.1 Beratung für kleinere Fälle

Bei kleineren Fällen sind grundsätzlich rasche Reaktionen gefragt, da nicht gelöschte Inhalte unter Umständen rasch verbreitet werden. Die folgenden Kenntnisse und Kompetenzen sind empfohlen:

- Kenntnis des prototypischen Ablaufs (1. Beratung, 2. Beweissicherung, 3. evtl. Unterstützung bei Meldung an Social-Media-Anbieter, 4. evtl. Unterstützung bei rechtlichem Verfahren)
- Technologisches Kontextwissen: Kenntnis von Social Media, deren Funktionsweise, Reichweite und Dynamiken
- Inhaltliches Wissen: Kenntnis der Auftretenshäufigkeit, Manifestation, Ursachen, Einflussfaktoren und Auswirkungen von rassistischer Online-Hassrede
- Psychosoziale Sensibilität im Hinblick auf «digitale» Form von Viktimisierung
- Kenntnis des präventiven oder reaktiven Selbstschutzes (z.B. Privatsphäre-Einstellungen)
- Juristisches Grundwissen zur Online-Dimension rassistischer Hassrede
- Kenntnis der Handlungsmöglichkeiten, deren Vor- und Nachteile sowie Umsetzung (z.B. Blocking, Meldung, Strafverfolgung, psychologische Netzwerkunterstützung)
- Anwendung verschiedener Handlungsmöglichkeiten auf unterschiedliche Fälle
- Kenntnis von spezialisierten Organisationen bzw. Personen zur Weiterleitung, z.B. für die Unterstützung beim «Massenmelden» (d.h. viele koordinierte Nutzende melden den Inhalt, was zur raschen Löschung führt), für Gegenrede bzw. Richtigstellung oder um eine Anzeige zu formulieren

10.1.2 Komplexe Beratung und Intervention

In den letzten Jahren sind grössere Fälle regelmässig aufgetreten, auch wenn sie nicht unbedingt bei Rassismusfachstellen gemeldet wurden. Ob in Schulklassen, der Politik oder anderen Orten: Diese Fälle können beträchtlichen Schaden anrichten und verlangen daher eine rasche und vorsichtige Herangehensweise. Zusätzlich zu den Kenntnissen und Kompetenzen der «einfachen Beratung» sind folgende Faktoren erforderlich, was entweder durch Beratungsstellen oder externe Expert*innen übernommen werden könnte:

- Kenntnis der Dynamik grösserer Fälle und der Umgangsweise (z.B. Identifikation von Nutzenden-Gruppen, Art der Kommunikation).
- Kenntnis des klassischen Mediensystems, seiner Mechanik und Einflussmöglichkeiten (z.B. Medienlandschaft, Reaktionen durch Gegenstellungs- und Einordnungsartikel).
- Kenntnis der präventiven und reaktiven Schutzmassnahmen gegen massenhafte Angriffe (z.B. Identifikation von Krisenthemen, Krisenplan für eine schnelle Orientierung, Monitoring von Person oder Institution im Internet, temporäre Übernahme von Social-Media-Profil unter Wahrung von Datenschutz, Gegenrede / Richtigstellung / Massenmeldung).
- Kenntnis von Rollen und Verantwortlichkeiten (z.B. Social-Media-Betreiber: Löschung von Inhalten; Staat: staatsrechtliche Verfolgung oder Positionierung; klassische Medien: Community-Management oder Gegenstellungs- und Einordnungsartikel).
- Nach Angriff: evtl. psychosoziale Unterstützung zur Verhinderung eines «digitalen Rückzugs».
- Kenntnis von Expert*innen mit guter Vernetzung (z.B. ehemalige Politiker*innen, Journalist*innen, Lehrer*innen), Wissen über die rassistische Online-Szene, einnehmende

³⁰⁸ Siehe dazu www.frb.admin.ch > Tätigkeitsfelder > Medien und Internet > Beratung.

Persönlichkeit («People person»): vertrauenswürdig, kontaktfreudig, verständnisvoll).³⁰⁹
Expert*innen-Netzwerk könnten spezifisch nach Sprachregion aufgebaut werden, wobei externe Spezialist*innen nicht in allen Sprachregionen gleich stark vertreten sind.

10.2 Massnahmen für Bekanntmachen und Nutzbarmachen von Beratungsangebot

Es gibt unterschiedlichste Massnahmen, wie die Angebote der Beratungsstellen zu rassistischer Online-Hassrede für Betroffene und Interessierte bekannt und sichtbar gemacht werden können:

- «Dorthin gehen, wo die Zielgruppe ist», d.h. die Bewerbung an digitale Gewohnheiten der Zielgruppe anpassen (insbes. Bewerbung auf digitalen Kanälen, da die Angebote nahe an Internet und Social Media sein sollten; hier kommt auch zielgruppenspezifisches Targeting durch die aktive Involvierung von Social Media-Betreibern infrage)
- breite digitale Applikationen nutzen (Beratungswebseite, Apps, Social-Media-Profile etc.), jedoch auch Offline-Kanäle, wo Vertrauen durch physischen Kontakt aufgebaut wird (Workshops, Vorträge, direkte Kontakte etc.)
- sich untereinander und mit Multiplikator*innen (z.B. Politiker*innen, klassische Medien) vernetzen, um die Reichweite der Angebote zu verstärken
- für Viralität: kreative «Outputs» in Social Media (z.B. Animationen, Kurzfilme) veröffentlichen
- Wissen über Online-Dimension einfließen lassen bei Informationen und Sensibilisierung zu Rassismus in Integrationszentren/-fachstellen und zivilgesellschaftlichen Organisationen

Sofern das Angebot von Betroffenen und Interessierten gefunden wurde: Wie kann es für sie so gut wie möglich nutzbar gemacht werden?

- Medienbruch reduzieren: Möglichst viele Angebote online zugänglich machen (z.B. Online-Beratung per Chat, Telefon oder Video)
- Benutzer*innenfreundlich und niedrigschwellig gestalten (z.B. Anonymität, rasches Zurechtfinden auf Webseite und/oder Handy-freundliche Version)
- Starke und kontinuierliche Online-Präsenz für Vertrauensaufbau und Signalsetzung («Wir sind hier und helfen jederzeit»), u.a. damit Zielgruppen nicht «wegbrechen» bzw. auf Social Media der Stelle nicht mehr folgen; hier ist allenfalls Expertise für die Social Media-Betreuung nötig
- Vertrauen durch gute Dienstleistung mit digital geschulten Berater*innen aufbauen (mindestens eine Person pro Stelle).
- Vertrauen durch Transparenz aufbauen: beispielsweise durch die Kommunikation gelöster Fälle und Erfolge auf Social Media-Kanälen
- Auf eventuelles zukünftiges nationales Meldetool verlinken

³⁰⁹ Interview mit Jonathan Birdwell (Institute of Strategic Dialogue) im April 2020.

11 HANDLUNGSEMPFEHLUNGEN

Der Bericht empfiehlt verschiedene Massnahmen zur Verhinderung und Eindämmung rassistischer Online-Hassrede in der Schweiz.³¹⁰ Grundsätzlich kann festgehalten werden, dass das aktuelle Gesamtangebot – insbesondere im Vergleich zum Ausland – stark ausbaufähig ist.

Breite Sensibilisierung und Kompetenzstärkung

Social Media sollte als reales Feld des gesellschaftlichen Zusammenlebens und somit als zentraler Kanal rassistischer Online-Hassrede ernst genommen werden – gleichzeitig sind an unterschiedlichsten Stellen ungenügendes Bewusstsein sowie fehlende Kenntnisse und Kompetenzen feststellbar. Es wird daher geraten, Sensibilisierung und Medienkompetenzen breit zu stärken: in Bezug auf Kenntnis des Phänomens, Technologie und Rechtliches – beispielsweise bei Akteur*innen direkt im Feld (z.B. Beratungsstellen) und bei Multiplikator*innen (z.B. Lehrer*innen) – und auf unterschiedlichen Hierarchiestufen in Gemeinden, Kantonen und Bund. Eine Gesamtsensibilität für die Online-Dimension von Rassismus und seinen Auswirkungen kann dazu beitragen, dass rassistische Online-Hassrede als anzugehendes Problem anerkannt wird. Informations- und Präventionsaufträge sollten daher prinzipiell den Online-Raum implizit und explizit miteinbeziehen. Für Praktiker*innen könnte dies ein unterstützendes Arbeitsumfeld als Grundlage für eine wirksame Vorgehensweise gegen rassistische Online-Hassrede schaffen. In diesem Rahmen sind bei öffentlichen Hassrede-Fällen auch sichtbare Positionierungen von offizieller Stelle (z.B. EKR) zu empfehlen.

Zusammenarbeit zwischen staatlichen, zivilgesellschaftlichen und privaten Organisationen

Rassistische Online-Hassrede ist dynamisch, vielschichtig und taucht oft ausserhalb fester organisatorischer Strukturen auf. Um diesem komplexen Phänomen entgegenzutreten, bedarf es eines gesamtheitlichen Lösungsansatzes, der die Bedürfnisse und Verantwortlichkeiten verschiedenster Institutionen berücksichtigt und von ihren Handlungsmöglichkeiten profitiert. Dazu gehören staatliche Institutionen, öffentliche Akteur*innen wie Politiker*innen, Bildungssystem und Hochschulen, Zivilgesellschaft, Social-Media-Anbieter, klassische Medienunternehmen (durch verantwortungsvolle Berichterstattung und Community-Management) sowie Justiz und Polizei (z.B. Verbesserung des Erstkontakts, Abklärung von Online-Anzeigen). Institutionenübergreifende Kooperationen können bestehendes Wissen bündeln und getroffene Massnahmen legitimieren. Denkbar wären hier Workshops oder ein «Runder Tisch»³¹¹, wo Wissens- und Erfahrungsaustausch sowie Vernetzung stattfindet; regelmässige Vernetzungs- und Koordinationsanlässe zur Stärkung der Zusammenarbeit; staatliche finanzielle Unterstützung von Projekten nach Bedarf; oder eine zentrale Online-Plattform (z.B. ein nationales Meldetool), die einen Überblick über relevante Akteure und ihre Rollen gibt und sie verlinkt.

Holistisches Vorgehen mit mehreren, möglichst evidenzbasierten Massnahmen

Wie im Bericht dargelegt, haben alle Ansätze ihre Stärken und Schwächen. Jeder deckt jedoch einen integralen Teil in einem präventiven und intervenierenden Lösungspuzzle ab. Daher wird zu einer Kombination ergänzender Massnahmen geraten: vom einfachen Meldeverfahren mithilfe eines nationalen Meldetools über die kompetente Beratung und, falls notwendig, effektive Strafverfahren sowie Gegenrede. Auf Seiten der Social-Media-Anbietenden ist das effiziente Löschen und die verbesserte Transparenz in Bezug auf Moderationsweisen denkbar, wobei der Staat im Rahmen der gegebenen Möglichkeiten Druck ausüben könnte, um die Prüfmechanismen unter Berücksichtigung der lokalen Schweizer Gesetze zu verbessern. Längerfristig im Blick behalten werden soll auch die Vermeidung struktureller Bedingungen, die Frustration, Hass und Gewalt fördern. Dazu gehören unter anderem Ausgrenzung und Demütigung, Ungleichverteilung von Chancen (z.B. Armut im Angesicht von Wohlstand), ungenügende Bildungsangebote und Wahrnehmungen von Unfairness, Ungerechtigkeit und Misstrauen.³¹² Denn was im Online-Raum geschieht, spiegelt vorherrschende gesellschaftliche Verhältnisse wider – wird dies ignoriert, reduzieren sich die Massnahmen auf kurzfristige

³¹⁰ Diese Empfehlungen sind aufgrund der teilweise ähnlichen Situation im Einklang mit Empfehlungen aus umliegenden Ländern, z.B. Österreich oder Deutschland. Siehe: Republik Österreich: Parlamentsdirektion 2016; Baldauf et al. 2018.

³¹¹ Hier sei der gemeinsam von FRB und «Jugend und Medien» organisierte Kick-off des Schwerpunkts «Hassrede» am 25. August 2020 zu nennen.

³¹² Bauer 2011.

Symptombekämpfung. Die Ausarbeitung von Massnahmen sollte sich in Bezug auf Wirksamkeit, Angemessenheit und Proportionalität an bestehenden wissenschaftlichen Erkenntnissen und langjährigen Erfahrungsberichten orientieren. So kann auch ein Gleichgewicht zwischen dem Schutz vulnerabler Gruppen vor Diskriminierung im Internet und dem zu schützenden Recht der Meinungsäusserungsfreiheit und Privatsphäre hergestellt werden. Denn Hassrede kann bekämpft werden – aber verschwinden wird sie in einer offenen Gesellschaft wohl nie. Daher legt eine evidenzbasierte Herangehensweise Wege zur längerfristigen wirksamen Behandlung des Phänomens offen.

Investition in Forschung und Monitoring

Aufgrund der aktuell dünnen Erkenntnisbasis zu rassistischer Online-Hassrede in der Schweiz ist es wesentlich, mehr systematisches Wissen zum Phänomen zu schaffen und finanzielle Mittel zur Verfügung zu stellen. Insbesondere Bevölkerungsumfragen können Täter*innen, Betroffene, Zuschauende und deren soziostrukturelle Hintergründe identifizieren, um Gegenmassnahmen zielgruppenspezifisch zu entwerfen. Dabei sollen grössere Ursachenzusammenhänge und Auswirkungen auf die Gesellschaft mitberücksichtigt werden. Die qualitative Analyse von Online-Hassrede und ihre Verbreitung in Netzwerken kann ein grundlegendes Verständnis der lokalen rassistischen «Online-Subkultur» fördern. Systematisches, durch die Wissenschaft informiertes Monitoring über dynamisch ändernde Auslöser, Themen, Schlüsselfiguren und Gruppierungen tragen schliesslich zur strategisch ausgerichteten Prävention und Intervention bei.

Adressierung ausgewählter Zielgruppen, insbesondere vulnerable Gruppen

Die Massnahmen sollten stark auf die «Endnutzer*innen» ausgerichtet werden. Kampagnen und Materialien zum Thema, die sich an die Allgemeinheit richten, gibt es im deutschsprachigen Raum mittlerweile zuhauf. Es besteht jedoch die Gefahr, dass insbesondere bei Online-Kampagnen vor einem bereits «konvertierten Publikum gepredigt wird». Es ist daher lohnenswert, tatsächliche oder potentielle Verbreiter*innen von Online-Hassrede und Betroffene zu identifizieren und die Inhalte spezifisch an ihrer Motivlage und ihrer digitalen, aber auch analogen Lebenswelt auszurichten. Zur Orientierung kann zurückgegriffen werden auf Literatur zu Merkmalen dieser Zielgruppen (siehe z.B. Kap. 5.2), auf (noch durchzuführende) Umfragen in den relevanten Kontexten (z.B. Schulen) und Interviews mit Betroffenen und idealerweise Täter*innen. Verbreiter*innen von Online-Hassrede (Kap. 5.2) stellen gemeinhin eine Minderheit dar; sie können jedoch die Herausbildung einer digital versierten Zivilgesellschaft gefährden, indem sie online sehr aktiv und sichtbar sind. Kinder und Jugendliche im Allgemeinen und mit spezifischen Merkmalen (Kap. 4.3 und 5.2) sollen über ihre starke Involvierung abgeholt werden, um wichtige Sozialisationsprozesse früh in Gang zu bringen. Auch Lehrer*innen sollen hier eingebunden werden. Eine sensible Vorgehensweise ist gefordert bei Personen, welche zu den Betroffenen und den Täter*innen gehören. Grundsätzlich sei geraten, Massnahmen spezifisch für den digitalen Raum zu entwerfen, anstatt Massnahmen, die sich gegen Offline-Rassismus bewährt haben, direkt in den digitalen Raum zu übertragen. Beispielsweise sieht der Aufbau von Unterstützungs- und Solidaritätsnetzwerken für potentiell Betroffene anders aus im realen Umfeld (insbesondere persönlicher Kontakt) als im digitalen Raum (zusätzlich grosse, lose Online-Netzwerke mit Personen mit grosser Reichweite, die nach Angriffen unmittelbar öffentliche Gegenrede machen).

Ausbau und Bewerbung von Beratung und Interventionen auf lokaler und nationaler Ebene

Da Rassismus heute nicht nur offline, sondern auch – und vielleicht schon jetzt primär – online stattfindet, wird Beratungsstellen geraten, die Online-Dimension stärker abzudecken. Ähnlich wie extreme Gruppen in den letzten 20 Jahren die Möglichkeiten digitaler Kommunikation für sich ausgeschöpft haben, gilt es nun hier aufzuholen, um das Ungleichgewicht von Hassrede gegenüber Gegenmassnahmen auszugleichen. Der Schlüssel liegt hier im Aufbau digitaler Kompetenz und digitaler Präsenz. Denn die Möglichkeit besteht, dass das grösstenteils mangelnde explizite Beratungsangebot für Online-Rassismus Betroffene davon abhält, sich bei relevanten Stellen zu melden – worauf wiederum ein Kompetenzaufbau und eine digitale Präsenz auf Seiten der involvierten Akteur*innen als nicht erforderlich gesehen wird. Dieser möglicherweise bestehende Kreislauf könnte mit einem bekannten und nutzbaren digitalen Angebot durchbrochen werden. Betroffene und Zuschauende werden so besser geschützt. Lokale Stellen spielen eine wichtige Rolle, weil sie durch ihre Vertrautheit mit den lokalen Strukturen relevante Zielgruppen erreichen können und bei konkreten Fällen mit dem politischen und kulturellen Hintergrund stärker vertraut sind. Eine zusätzliche nationale Koordination könnte jedoch Abhilfe schaffen, solange die Kompetenz in der kantonalen Breite noch nicht gegeben

ist. Zudem ist eine zentrale, möglicherweise auch längerfristige Online-Meldestelle denkbar, da rassistische Online-Hassrede mit einem Klick nationale Dimensionen annehmen kann.

12 VERZEICHNIS DER ABKÜRZUNGEN

AFD	Alternative für Deutschland
AJS NRW	Aktion Jugendschutz Nordrhein-Westfalen
AWD	Atomwaffen Division
BV	Bundesverfassung
CEJI-A	Jewish contribution to an inclusive Europe
CERD	Committee on the Elimination of Racial Discrimination
CICAD	Coordination Intercommunautaire Contre l'Antisémitisme et la Diffamation
DGUV	Deutsche Gesetzliche Unfallversicherung
ECRI	Europäische Kommission gegen Rassismus und Intoleranz
EKR	Eidgenössische Kommission gegen Rassismus
EU	Europäische Union
FRB	Fachstelle für Rassismusbekämpfung
GIF	Graphics Interchange Format
GRA	Stiftung gegen Rassismus und Antisemitismus
HEP	Haute École Pédagogique Vaud
IB	Identitäre Bewegung
INACH	International Network Against Cyber Hate
IS	Der Islamische Staat
ISD	Institute of Strategic Dialogue
LICRA	Ligue Internationale Contre le Racisme et l'Antisémitisme
NGO	Non-governmental organizations / Nichtregierungsorganisation
OCCI	Online Civil Courage Initiative (OCCI)
PEGIDA	Patriotische Europäer gegen die Islamisierung des Abendlandes
RTVG	Bundesgesetz über Radio und Fernsehen vom 24. März 2006 (SR 784.40)
sCAN	Platforms, Experts, Tools: Specialised Cyber-Activists Network
SIG	Schweizerischer Israelitischer Gemeindebund

StGB	Schweizerisches Strafgesetzbuch
UBI	Unabhängige Beschwerdeinstanz für Radio und Fernsehen
UNICRI	United Nations Interregional Crime and Justice Research Institute
USI	Università della Svizzera Italiana
UZH	Universität Zürich
ZARA	Zivilcourage und Anti-Rassismus-Arbeit
ZGB	Schweizerisches Zivilgesetzbuch
ZHAW	Zürcher Hochschule für Angewandte Wissenschaften

13 LITERATURVERZEICHNIS

Forschungsartikel, Berichte und Bücher

- Adena M. et al. 2015. Radio and the rise of the Nazis in prewar Germany. *The Quarterly Journal of Economics* 130(4): 1885-1939.
- Aizenkot D. und G. Kashy-Rosenbaum. 2018. Cyberbullying in whatsapp classmates' groups: Evaluation of an intervention program implemented in Israeli elementary and middle schools. *New Media & Society* 20(12): 4709-4727.
- Álvarez-Benjumea A. und F. Winter. 2018. Normative change and culture of hate: An experiment in online environments. *European Sociological Review* 34(3): 223-237.
- Anderson A. A. et al. 2014. The «Nasty effect»: Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication* 19(3): 373-387.
- Awan I. und I. Zempi. 2015. We fear for our lives: Offline and online experiences of anti-Muslim hostility (Oktober).
- Baier D. 2019. Kriminalitätsoffererfahrungen und Kriminalitätswahrnehmungen in der Schweiz: Ergebnisse einer Befragung. ZHAW (August).
- Baldauf J., Ebner J., und J. Guhl. 2018. *Hassrede und Radikalisierung im Netz. Der OCCI-Forschungsbericht Hassrede*. Institute for Strategic Dialogue, London.
- Barlow C. und I. Awan. 2016. «You need to be sorted out with a knife»: The attempted online silencing of women and people of Muslim faith within academia. *Social Media + Society* 2(4): 1-11.
- Barnidge M. et al. 2019. Perceived exposure to and avoidance of hate speech in various communication settings. *Telematics and Informatics* 44: 101263.
- Bauer J. 2011. *Schmerzgrenze: Vom Ursprung alltäglicher und globaler Gewalt*. Karl Blessing Verlag, München.
- Ben-David A. und A. Matamoros-Fernández. 2016. Hate speech and covert discrimination on social media: Monitoring the Facebook pages of extreme-right political parties in Spain. *International Journal of Communication* 10: 1167-1193.
- Beratungsnetz für Rassismuspfer. 2020. *Rassismuvorfälle aus der Beratungspraxis: Januar bis Dezember 2019*. Verein Humanrights.ch, Eidgenössische Kommission gegen Rassismus, Bern (April).
- Blaya C. 2019. Cyberhate: A review and content analysis of intervention strategies. *Aggression and Violent Behavior* 45: 163-172.
- Blaya C. 2015. *Cyberviolence et école – Les Dossiers des Sciences de l'Education*, 33. Presses universitaires du Midi, Toulouse.
- Blaya C. und C. Audrin. 2019. Toward an understanding of the characteristics of secondary school cyberhate perpetrators. *Frontiers in Education* 4(46).
- Bliuc A.-M. et al. 2018. Online networks of racial hate: A systematic review of 10 years of research on cyber-racism. *Computers in Human Behavior* 87: 75-86.
- Boyd D. M. 2010. Social network sites as networked publics: affordances, dynamics, and implications. In: *A Networked Self: Identity, Community, and Culture on Social Network Sites*, Papacharissi Z. (Hrsg.): 39-58. Routledge, New York.
- Brady W. J. et al. 2017. Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences* 114(28): 7313-7318.
- Breuer J. 2017. Hate Speech in Online Games. In: *Online Hate Speech: Perspektiven auf eine neue Form des Hasses*. Grimme-Institut – Gesellschaft für Medien, Bildung und Kultur mbH, Marl.
- Brown A. 2018. What is so special about online (as compared to offline) hate speech? *Ethnicities* 18(3): 297-326.
- Bucchianeri M. M. et al. 2014. Multiple types of harassment: Associations with emotional well-being and unhealthy behaviors in adolescents. *Journal of Adolescent Health* 54(6): 724-729.
- Buckels E. E. et al. 2014. Trolls just want to have fun. *Personality and Individual Differences* 67: 97-102.
- Bulut E. und E. Yörük. 2017. Digital populism: Trolls and political polarization of Twitter in Turkey. *International Journal of Communication* 11(25): 4093-4117.
- Bundesamt für Statistik. 2019. *Erhebung zum Zusammenleben in der Schweiz (ZidS): Ergebnisse 2018*. BfS, Neuchâtel.
- Celik S. 2019. Experiences of internet users regarding cyberhate. *Information Technology & People* 32(6): 1446-1471.
- Cerase A., E. D'Angelo und C. Santoro. 2015. Monitoring racist and xenophobic extremism to counter hate speech online: Ethical dilemmas and methods of a preventive approach. *VoxPol Workshop*, Brussels (19. Januar).
- Chan J., A. Ghose und R. Seamans. 2013. The internet and hate crime: Offline spillovers from online access, Working Papers 13-02, NET Institute.
- Chandrasekharan E. et al. 2017. You can't stay here: The efficacy of Reddit's 2015 ban examined through hate speech. *Proceedings of the ACM on Human-Computer Interaction* 1(CSCW): 1-22.
- Cho D., S. Kim und A. Acquisti. 2012. Empirical analysis of online anonymity and user behaviors: The impact of real name policy. *45th Hawaii International Conference on System Sciences* 3041-3050. IEEE.
- Christopherson K.M. 2007. The positive and negative implications of anonymity in Internet social interactions: On the internet, nobody knows you're a dog. *Computers in Human Behavior* 23(6): 3038-3056.

- Chung et al. 2019. CONAN – COunter NARratives through Nichesourcing: a multilingual dataset of responses to fight online hate speech. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 28. Juli – 2. August 2019: 2819-2829.
- Chyzh O., M. D. Nieman und C. Webb. 2019. The effects of dog-whistle politics on political violence. *Iowa State University, Political Science Publications* 59: 1-10.
- Cleland J. 2013. Racism, football fans, and online message boards: How social media has added a new dimension to racist discourse in English football. *Journal of Sport & Social Issues* 38: 415-431.
- Costello M. und J. Hawdon. 2018. Who are the online extremists among Us? Sociodemographic characteristics, social networking, and online experiences of those who produce online hate materials. *Violence and Gender* 5(1): 55-60.
- Costello M. et al. 2019. Social group identity and perceptions of online hate. *Sociological inquiry* 89(3): 427-452.
- Craker N. und E. March. 2016. The dark side of Facebook: The dark tetrad, negative social potency, and trolling behaviours. *Personality and Individual Differences* 102: 79-84.
- Daft R. L. und R. H. Lengel. 1986. Organizational information requirements, media richness and structural design. *Management Science* 32(5): 554-571.
- Daniels J. 2012. Race and racism in Internet studies: A review and critique. *New Media & Society* 15(5): 695-719.
- Daniels J. 2018. The algorithmic rise of the Alt-Right. *Contexts* 17(1): 60-65.
- Der Bundesrat. 2013. *Rechtliche Basis für Social Media. Bericht des Bundesrates in Erfüllung des Postulats Amherd 11.3912 (29. September 2011)*. Bern: Bundesrat.
- Der Bundesrat. 2017. *Rechtliche Basis für Social Media: Erneute Standortbestimmung. Nachfolgebericht des Bundesrates zum Postulatsbericht Amherd 11.3912 «Rechtliche Basis für Social Media» (10. Mai 2017)*. Bern: Bundesrat.
- Delgado R. und J. Stefancic. 2009. Four observations about hate speech. *Wake Forest Law Review* 44: 353-370.
- Dittrich M. et al. 2020. Alternative Wirklichkeiten. Monitoring rechts-alternativer Medienstrategien. *Amadeu Antonio Stiftung*, Berlin.
- Duggan M. 2017. Online Harassment 2017. *Pew Research Center* (Juli).
- Eckes C. et al. 2018. #Hass im Netz: Der schleichende Angriff auf unsere Demokratie. *Institut für Demokratie und Zivilgesellschaft*, Jena.
- Eco – Verband der Internetwirtschaft. 2019. Eco Beschwerdestelle: Jahresbericht 2018.
- Eddington S. M. 2018. The communicative constitution of hate organizations online: A semantic network analysis of «Make America Great Again». *Social Media + Society* 4: 1-12.
- Elson M. und C.J. Ferguson. 2014. Twenty-five years of research on violence in digital games and aggression. *European Psychologist* 19(1): 33-46.
- Erjavec K. und M. P. Kovačič. 2012. You don't understand, this is a new war! Analysis of hate speech in news web sites' comments. *Mass Communication and Society* 15(6): 899-920.
- Fachstelle für Rassismusbekämpfung. 2018. Workshop «Rassismus im Netz»: Blitzlichter auf die Inhalte der Referate und Diskussionen. Bern (11. Dezember).
- Fachstelle für Rassismusbekämpfung. 2019. Rassistische Diskriminierung in der Schweiz. Bern (September).
- Fielitz M. und H. Marcks. 2019. Digital fascism: Challenges for the open society in times of social media. *Berkeley Center for Right-Wing Studies Working Paper Series* (Juli).
- Fortuna P. und S. Nunes. 2018. A survey on automatic detection of hate speech in text. *ACM Computing Surveys* 51(4): 85:1-30.
- Fox J. und Tang W.Y. 2017. Women's experiences with general and sexual harassment in online video games: Rumination, organizational responsiveness, withdrawal, and coping strategies. *New Media & Society* 19(8): 1290-1307.
- Gagliardone I. et al. 2016. Mechachal: Online debates and elections in Ethiopia – from hate speech to engagement in Social Media. *SSRN Electronic Journal* 2831369.
- Gagliardone I. et al. 2015. *Countering online hate speech: UNESCO Series on Internet Freedom*. United Nations Educational, Scientific, and Cultural Organization, Paris.
- Gatewood C. und I. Boyer. 2019. *Building digital citizenship in France. Lessons from the Sens Critique Project building digital citizenship in France*. Institute for Strategic Dialogue, London.
- Gatewood C. et al. 2020. Cartographie de la haine en ligne: Tour d'horizon du discours haineux en France. Institute of Strategic Dialogue.
- George C. 2015. Hate speech law and policy. In: *The International Encyclopedia of Digital Communication and Society, First Edition*, John Wiley & Sons, Inc. (Hrsg.): 1-10.
- Geschke D. et al. 2019. #Hass Im Netz: Der schleichende Angriff auf unsere Demokratie. Eine Bundesweite Repräsentative Untersuchung. Campact.
- Guhl J., J. Ebner und J. Rau. 2020. *Das Online-Ökosystem rechtsextremer Akteure*. Institute for Strategic Dialogue, London.
- Hawdon J., A. Oksanen und P. Räsänen. 2017. Exposure to online hate in four nations: A cross-national consideration. *Deviant behavior* 38(3): 254-266.
- Haymoz S. et al. 2019. *L'estremismo politico tra i giovani in Svizzera: entità e fattori esplicativi: Rapporto di valutazione per il Canton Ticino*. ZHAW und HETS-FR (Januar).
- Hayne S. C. und R. E. Rice. 1997. Attribution accuracy when using anonymity in group support systems. *International Journal of Human-Computer Studies* 47(3): 429-452.

- Hine G. E. et al. 2017. Kek, cucks, and god emperor Trump: A measurement study of 4chan's politically incorrect forum and its effect on the web. *Eleventh International AAAI Conference on Web and Social Media (ICWSM), Oktober 2017*.
- Hsueh M. et al. 2015. «Leave your comment below»: can biased online comments influence our own prejudicial attitudes and behaviors? *Human Communication Research* 41(4): 557-576.
- Isbister T. et al. 2018. Monitoring targeted hate in online environments. arXiv preprint arXiv: 1803.04757.
- Jakubowicz A. et al. 2017. How cyber users experience and respond to racism: Evidence from an online survey. In: *Cyber Racism and Community Resilience*, Palgrave Hate Studies, Springer: 65-94.
- Jeong S.-H., H. Cho und Y. Hwang. 2012. Media literacy interventions: A meta-analytic review. *Journal of Communication* 62(3): 454-472.
- Jones D. und Benesch S. 2019. *Combating hate speech through counterspeech*. Berkman Klein Center (9. August).
- Jourová V. 2019a. *Code of conduct on countering illegal hate speech online. Fourth evaluation confirms self-regulation works*. European Commission (12. Februar).
- Jourová V. 2019b. *How the Code of Conduct helped countering illegal hate speech online*. European Commission (Februar).
- Kaspar K., 2017. Hassreden im Internet – Ein besonderes Phänomen computervermittelter Kommunikation? In: K. Kaspar, L. Grässer und A. Riffi. *Online Hate Speech: Perspektiven auf eine neue Form des Hasses*. kopaed verlagsgmbh, Düsseldorf/München.
- Keen A. 2007. *The Cult of the Amateur*. Currency, New York.
- Kenski K., K. Coe und S. A. Rains. 2017. Perceptions of uncivil discourse online: An examination of types and predictors. *Communication Research* 1-20.
- Klein A. 2017. Hate speech in the information age. In: *Fanaticism, Racism, and Rage Online*. Klein A. (Hrsg.): 25-39. Palgrave Macmillan, Cham.
- Kreissel P. et al. 2019. *Hass auf Knopfdruck: Rechtsextreme Trollfabriken und das Ökosystem koordinierter Hasskampagnen im Netz*. Institute for Strategic Dialogue, London.
- Krieger N. 1990. Racial and gender discrimination: Risk factors for high blood pressure? *Social Science & Medicine*, 30, 1273-1281.
- Landesanstalt für Medien NRW. 2018. *Ergebnisbericht Hassrede*. Forsa.
- Lanzke A. et al. 2013. *Viraler Hass: Rechtsextreme Kommunikationsstrategien im Web 2.0*. Amadeu Antonio Stiftung, Heidelberg.
- Lapidot-Lefler N. und A. Barak. 2012. Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. *Computers in Human Behavior* 28(2): 434-443.
- Latzer M., Büchi M. und N. Festic. 2019a. Internetverbreitung und digitale Bruchlinien in der Schweiz 2019: *Themenbericht aus dem World Internet Project – Switzerland 2019*. Zürich: Universität Zürich (Oktober).
- Latzer M., Büchi M., und N. Festic. 2019b. Internetanwendungen und deren Nutzung in der Schweiz 2019. *Themenbericht aus dem World Internet Project – Switzerland 2019*. Zürich: Universität Zürich (Oktober).
- Laubenstein S. und A. Urban. 2018. Fallbeispiele: Welche Arten von Kampagnen gegen Hass und Extremismus im Internet funktionieren, welche funktionieren nicht und warum? In: *Hassrede und Radikalisierung im Netz. Der OCCI-Forschungsbericht*. Baldauf J. et al. (Hrsg.): 55-63. Institute for Strategic Dialogue, London.
- Ledwich M. und A. Zaitsev. 2019. Algorithmic extremism: Examining YouTube's rabbit hole of radicalization. arXiv preprint arXiv: 1912.11211.
- Lim M. 2017. Freedom to hate: Social Media, algorithmic enclaves, and the rise of tribal nationalism in Indonesia. *Critical Asian Studies* 49(3): 411-427.
- Lingiardi V. et al. 2019. Mapping Twitter hate speech towards social and sexual minorities: A lexicon-based approach to semantic content analysis. *Behaviour and Information Technology*. DOI: 10.1080/0144929X.2019.1607903.
- Lowry P. B. et al. 2016. Why do adults engage in cyberbullying on social media? An integration of online disinhibition and deindividuation effects with the social structure and social learning model. *Information Systems Research* 27(4): 962-986.
- Lucas B. 2014. *Methods for monitoring and mapping online hate speech*. GSDRC Applied Knowledge Services.
- Marwick, A. und R. Lewis. 2017. *Media manipulation and disinformation online*. Data & Society Research Institute, New York.
- Marwick A. E. 2015. Instafame: Luxury selfies in the attention economy. *Public Culture* 27(1/75): 137-160.
- Matamoros-Fernández A. 2017. Platformed racism: The mediation and circulation of an Australian race-based controversy on Twitter, Facebook and YouTube. *Information, Communication & Society* 20(6): 930-946.
- Mathew B. et al. 2019. Spread of hate speech in online social media. In: *11th ACM Conference on Web Science (WebSci '19)*, 30. Juni-3. Juli 2019. Boston, MA, USA.
- Mittos A. et al. 2019. And we will fight for our race! A measurement study of genetic testing conversations on Reddit and 4chan. arXiv preprint arXiv:1901.09735.
- Mondal M. et al. 2018. Characterizing usage of explicit hate expressions in social media. *New Review of Hypermedia and Multimedia* 24(2): 110-130.
- Müller K. und C. Schwarz. 2019. Fanning the flames of hate: Social media and hate crime. *SSRN Scholarly Paper*.

- Munger K. 2017. Tweetment effects on the tweeted: Experimentally reducing racist harassment. *Political Behavior* 39: 629-649.
- Murthy D. und S. Sharma. 2019. Visualizing YouTube's comment space: Online hostility as a networked phenomena. *New Media & Society* 21(1): 191-213.
- Musial J. 2017. «My Muslim sister, indeed you are a mujahidah» - Narratives in the propaganda of the Islamic State to address and radicalize Western Women. An Exemplary analysis of the online magazine Dabiq. *Journal for Deradicalization* (9): 39-100.
- Naguib T. 2014. *Begrifflichkeiten zum Thema Rassismus im nationalen und internationalen Verständnis: Eine Auslegeordnung unter Berücksichtigung des Völker- und Verfassungsrechts*. Expertise im Auftrag der Fachstelle für Rassismusbekämpfung FRB, Eidgenössisches Departement des Innern EDI. Zürcher Hochschule für Angewandte Wissenschaften (ZHAW), Winterthur / Bern (27. August).
- Neue deutsche Medienmacher & No Hate Speech Movement Deutschland. 2019. *Wetterfest durch den Shitstorm. Leitfaden für Medienschaffende zum Umgang mit Hass im Netz*. Berlin.
- O'Callaghan D. et al. 2012. An analysis of interactions within and between extreme right communities in social media. In: *Ubiquitous social media analysis*, Atzmueller M. (Hrsg.): 88-107. Springer, Berlin/Heidelberg.
- Olteanu A. et al. 2018. The effect of extremist violence on hateful speech online. In: *Proceedings of the Twelfth International AAAI Conference on Web and Social Media (ICWSM)*: 221-230. AAAI Press, Palo Alto, California.
- Ortiz S. M. 2019. The meanings of racist and sexist trash talk for men of color: A cultural sociological approach to studying gaming culture. *New Media & Society* 21(4): 879-894.
- Pariser E. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. Penguin, UK.
- Perry B. und P. Olsson. 2009. Cyberhate: The globalization of hate. *Information & Communications Technology Law* 18(2): 185-199.
- Quent M. 2018. Zivilgesellschaft: Das globale Dorf verteidigen: Strategien gegen den kulturellen Backlash in sozialen Medien. In: *Hassrede und Radikalisierung im Netz. Der OCCI-Forschungsbericht*. Baldauf J. et al. (Hrsg.): 48-54. Institute for Strategic Dialogue, London.
- Rafael S. und A. Ritzmann. 2018. Hintergrund: Das ABC des Problemkomplexes Hassrede, Extremismus und NetzDG. In: *Hassrede und Radikalisierung im Netz. Der OCCI-Forschungsbericht*. Baldauf J. et al. (Hrsg.): 11-19. Institute for Strategic Dialogue, London.
- Refaeil N. und A. Wiecken. 2018. *Rassismus Im Netz – Schweizer Recht*. Fachstelle für Rassismusbekämpfung. Präsentation, Bern (Dezember).
- Reichelmann A. et al. 2020. Hate knows no boundaries: Online hate in six nations. *Deviant Behavior*: 1-12.
- Republik Österreich: Parlamentsdirektion. 2016. [Digitale Courage](#) (November).
- Reynolds L. und H. Tuck. 2016. *The counter-narrative monitoring & evaluation handbook*. Institute for Strategic Dialogue, London.
- Ribeiro M. H. et al. 2018. Characterizing and detecting hateful users on Twitter. In: *Proceedings of the Twelfth International AAAI Conference on Web and Social Media (ICWSM)*: 676-679. AAAI Press, Palo Alto, California.
- Rost K., L. Stahel und B. S. Frey. 2016. Digital social norm enforcement: Online firestorms in social media. *PLoS One* 11(6): e0155923.
- Salminen J. et al. 2018. Online hate interpretation varies by country, but more by individual: A statistical analysis using crowdsourced ratings. In: *Fifth International Conference on Social Networks Analysis, Management and Security (SNAMS), 15.-18. Oktober 2018*: 88-94.
- Salminen J. et al. 2019. Online hate ratings vary by extremes: A statistical analysis. *Proceedings of the 2019 Conference on Human Information Interaction and Retrieval (CHIIR), 10.-14. März 2019*: 213-217. Glasgow, United Kingdom.
- Schabas, W. A. 2001. Hate-Speech in Ruanda: The road to genocide. *McGill Law Journal*: 46, 301-315.
- Schieb C. und M. Preuss. 2016. Governing hate speech by means of counterspeech on Facebook. In: *66th Annual Conference of the International Communication Association (ICA), 9.-13. Juni 2016*. Fukuoka, Japan: 1-23.
- Schmidt A. und M. Wiegand. 2017. A survey on hate speech detection using natural language processing. In: *Proceedings of the Fifth International Workshop on Natural Language Processing for Social Media*. Valencia, Spain: 1-10.
- Sellers A. 2016. Defining hate speech. *Berkman Klein Center Research Publication No. 2016(20)*; *Boston Univ. School of Law, Public Law Research Paper No. 16(48)*
- Shapiro A. 2019. Predictive policing for reform? Indeterminacy and intervention in big data policing. *Surveillance & Society* 17(3/4): 456-472.
- Siegel A. 2020. Online hate speech. In: *Social Media and Democracy: The State of the Field*. Tucker J. und Persily N. (Hrsg.), Cambridge University Press.
- Siegel A. et al. 2019. *Trumping hate on Twitter? Online hate in the 2016 US election and its aftermath*. Working Paper.
- Silverman T. et al. 2016. *The impact of counter-narratives*. Institute for Strategic Dialogue, London.
- Smahel D. et al. 2020. *EU kids online 2020: Survey results from 19 countries*. London School of Economics and Political Science, London, UK.
- Soral W., B. Michal und M. Winiewski. 2018. Exposure to hate speech increases prejudice through desensitization. *Aggressive Behavior* 44(2): 136-146.

- Stahel L. und C. Schoen. 2019. Female journalists under attack? Explaining gender differences in reactions to audiences' attacks. *New Media & Society*: 1461444819885333.
- Sticca F. und S. Perren. 2013. Is cyberbullying worse than traditional bullying? Examining the differential roles of medium, publicity, and anonymity for the perceived severity of bullying. *Journal of Youth and Adolescence* 42: 739-750.
- Stiftung gegen Rassismus und Antisemitismus (GRA) und Schweizerischer Israelitischer Gemeindeverband (SIG). 2019. *Antisemitismusbericht*. Zürich.
- Suler J. 2004. The online disinhibition effect. *Cyberpsychology & Behavior* 7(3): 321-326.
- Sunstein C. R. 2000. Deliberative trouble? Why groups go to extremes. *The Yale Law Journal* 110(1): 71-119.
- Tuck H. und T. Silverman. 2016. *The counter-narrative handbook*. Institute for Strategic Dialogue, London.
- Tynes B., J. Torro und F. Lozada. 2019. An unwelcomed digital visitor in the classroom: The longitudinal impact of online racial discrimination on school achievement motivation. *School Psychology Review* 44(4): 407-424.
- Tynes B. M. und S. L. Markoe. 2010. The role of color-blind racial attitudes in reactions to racial discrimination on social network sites. *Journal of Diversity in Higher Education* 3(1): 1-13.
- Tynes B. M. et al. 2008. Online racial discrimination and psychological adjustment among adolescents. *Journal of Adolescent Health* 43(6): 565-569.
- Vosoughi S., D. Roy und S. Aral. 2018. The spread of true and false news online. *Science* 359(6380): 1146-1151.
- Wachs S. et al. 2019. Associations between witnessing and perpetrating online hate in eight countries: The buffering effects of problem-focused coping. *International Journal of Environmental Research and Public Health* 16(20): 3992.
- Waqas A. et al. 2019. Mapping online hate: A scientometric analysis on research trends and hotspots in research on online hate. *PloS one* 14(9).
- Winter A. 2019. Online hate: From the far-right to the 'Alt-Right' and from the margins to the mainstream. In: *Online Othering, Palgrave Studies in Cybercrime and Cybersecurity*, Maras M. und T.J. Holt (Hrsg.): 39-63. Palgrave Macmillan.
- Wolfe C. 2019. Online trolls, journalism and the freedom of speech: Are the bullies taking over? *Ethical Space: The International Journal of Communication Ethics* 16(1): 11-21.
- Yanagizawa-Drott D. 2014. Propaganda and conflict: Evidence from the Rwandan genocide. *The Quarterly Journal of Economics* 129(4): 1947-1994.
- Ziegele M. et al. 2019. *Aufräumen im Trollhaus: Zum Einfluss von Community-Managern und Aktionsgruppen in Kommentarspalten*. Institute for Internet and Democracy, Düsseldorf.
- Ziegele M., C. Koehler und M. Weber. 2018. Socially destructive? Effects of negative and hateful user comments on readers' donation behavior toward refugees and homeless persons. *Journal of Broadcasting & Electronic Media* 62(4): 636-653.

Nachrichtenartikel

- Berners-Lee T. 2017. [I invented the web. Here are three things we need to change to save it.](#) *The Guardian* (12. März 2017).
- Fisher M. 2018. [Inside Facebook's secret rules for global political speech.](#) *New York Times* (27. Dezember 2018).
- Germann M. 2018. [7000 gesperrte Kommentare.](#) *WOZ* (20. September 2018).
- Gunaratna S. 2016. [Neo-Nazis tag \(\(\(Jews\)\)\) on Twitter as hate speech, politics collide.](#) *CBS News* (10. Juni 2016).
- Klepper D. 2020. [Facebook removes nearly 200 accounts tied to hate groups.](#) *ABC News* (6. Juni 2020).
- Lizza, R. 2016. [Twitter's anti-semitism problem.](#) *New Yorker* (19. Oktober 2016).
- Newton C. 2019. [The Trauma Floor.](#) *The Verge* (25. Februar 2019).
- Priebe M. 2020. [Wie Rassisten das Coronavirus für sich nutzen.](#) *FAZ* (3. Februar 2020).
- Serafini S. 2015. [Edelweiss-Streit: Auf Whatsapp blüht der jugendliche Patriotismus.](#) *Aargauer Zeitung* (21. Dezember 2015).
- Spiegel.de. 2016. [Freundlicher Frosch wird Hasssymbol](#) (28. September 2016).
- Schweiz aktuell. 2020. [Ja zum Transitplatz für Fahrende in Wileroltigen.](#) *SRF* (9. Februar 2020).
- Stahel L. 2019. [Gehässige Leserreaktionen können die Qualität der Berichterstattung auch erhöhen.](#) *NZZ* (7. Mai 2019).
- 20min. 2019. [Primarschüler machen sich über Dunkelhäutige lustig](#) (30. Dezember 2019).
- 20min. 2019. [«Rechtsextreme Chats gibt es an jeder Schule»](#) (19. März 2019).

Sonstige Internetquellen

(reine Namen von Webseiten, die schon im Text genannt wurden und insbesondere die Hyperlinks aus Tabelle 1-8 sind hier nicht mehr aufgelistet)

- Alexander N. 2019. [Reconquista Germanica meldet sich ab](#). *Youtube* (30. November 2019).
- Amnesty International. 2019. [Il barometro dell'odio](#) (Datum fehlend, 2019).
- Anglin A. 2016. [A Normie's Guide to the Alt-Right](#). *Daily Stormer* (31. August 2016).
- Banaszczuk Y. 2019. [Toxic Gaming: Rassismus, Sexismus und Hate Speech in der Spieleszene](#). *Bundeszentrale für politische Bildung* (26. Juli 2019).
- Committee on the Elimination of Racial Discrimination CERD. 2013. [Allgemeine Bemerkungen Nr. 35 betreffend «Combating racist hate speech»](#) (26. September 2013).
- Council of Europe. 1997. [Council of Europe's Committee of Ministers' Recommendation 97\(20\)](#) (30. Oktober 1997).
- European Commission against Racism and Intolerance ECRI. 2015. [General Policy Recommendation N°15 on Combating Hate Speech](#) (8. Dezember 2015).
- Fachstelle für Rassismusbekämpfung FRB. [Online-Rechtsratgeber](#).
- Fachstelle für Rassismusbekämpfung FRB. [Medien und Internet](#).
- Garland C. 2008. [Klan's new message of cyber-hate](#). *Cult Education Institute* (27. März 2008).
- Generation D. 2017. [Das Shitposting 1x1](#) (18. Mai 2017).
- Getthetrollsout!. 2018. [«Troll of the month»](#) (5. Dezember 2018).
- Gold extra. [Tools of Subversion](#).
- Heiderich K. 2018. [Extremismusforscherin Julia Ebner: „Hasskampagnen folgen einem klaren Muster“](#). *Fearless Democracy* (26. April 2018).
- Humanrights.ch. 2020. [Rechtsentwicklung: Verstärkung des gesetzlichen Schutzes vor Diskriminierung?](#) (19. März 2020).
- Humanrights.ch. 2017. [Hassaufrufe im Internet - Schweizer Fälle und Politik der Newsportale](#) (23. August 2017).
- Humanrights.ch. 2017. [Hassreden: Die Grenzen der Meinungsäusserungsfreiheit](#) (6. Februar 2017).
- INACH. 2019. [The state of policy on cyber hate in the EU](#) (19. November 2019).
- ISD Global. [Youth Civil Activism Network \(YouthCAN\)](#)
- Jetzt. 2018. [Rechtsextreme „Feministinnen“ stören die Berlinale](#) (20. Februar 2018).
- LIGHT ON. 2014. [Visual database of racist and discriminatory symbols and images](#) (Oktober 2014).
- Lorentzen M. K. 2017. [Don't feed the trolls - Fight them](#). TEDx Talk. *Youtube* (7. Januar 2020).
- Pfeiffer, J. 2020. [Hate Speech in der Incel-Szene](#). *Belltower News* (25. März 2020).
- Reitman J. 2018. [All-American Nazis](#). *Rolling Stones* (2. Mai 2018).
- sCAN. 2020. [Hotspots of Hate: The online responsibility of public figures](#) (Datum fehlend 2020).
- sCAN. 2018a. [Hate Ontology](#) (30. November 2018).
- sCAN. 2018b. [Mapping study: Countering online hate speech with automated monitoring tools](#) (26. November 2018).
- sCAN. 2019. [Monitoring Report 2018-2019](#) (Datum fehlend 2019).
- Southern Poverty Law Center. 2020. [Alt-Right](#).
- Southern Poverty Law Center. 2018. [McInnes, Molyneux, and 4chan: Investigating pathways to the alt-right](#). (19. April 2018).
- Southern Poverty Law Center. 2015. [The Council Of Conservative Citizens: Dylann Roof's gateway into the world of white nationalism](#). (21. Juni 2015).
- Stiftung gegen Rassismus und Antisemitismus (GRA). [Hate Speech / Hassrede im Internet](#).
- Stiftung gegen Rassismus und Antisemitismus (GRA). [Meldetool](#).
- Wikipedia. [Metapedia](#).
- Zivilcourage und Anti-Rassismus-Arbeit ZARA. [Hass im Netz melden](#).

Bildnachweise

Grafik 7: <https://derpicdn.net/img/2013/4/6/289767/full.png>

Grafik 8: <https://me.me/i/nazi-pony-tumblr-912b75e26c2345f8aea216a60a0723c9>

Interviewverzeichnis

Birdwell, Jonathan (Institute of Strategic Dialogue, London)

Bischof, Michael (Integrationsförderung, Stadt Zürich)

Pershan, Claire (Renaissance Numérique, Paris)

Pugatsch, Dominic (GRA, Zürich)

Spiess-Hegglin, Jolanda (#Netzcourage, Zug)

Voigt, Hansi (dasnetz.ch; bajour)