



# Bedarfsanalyse zum Phänomen des digitalen Hasses/der digitalen Hetze mit Schwerpunkt Rassismus in Niedersachsen: eine qualitative Interviewstudie

Ergebnisbericht für den Landespräventionsrat Niedersachsen

Carlotta Alpers, Anna Matthes, Johanna Brüser und Samuel Tomczyk

Universität Greifswald



gefördert durch

**Landesprogramm**  
für Demokratie  
und Menschenrechte



## Inhaltsverzeichnis

<b>1</b>	<b>Hintergrund.....</b>	<b>1</b>
1.1	<i>Begriffsdefinitionen.....</i>	<i>1</i>
1.2	<i>Betroffene von Hass im Netz.....</i>	<i>2</i>
1.3	<i>Auswirkung von „Hass im Netz“ auf Betroffene.....</i>	<i>4</i>
1.4	<i>Unterstützende, schützende und präventive Ansätze.....</i>	<i>5</i>
<b>2</b>	<b>Methoden.....</b>	<b>10</b>
2.1	<i>Zielgruppe und Rekrutierung.....</i>	<i>10</i>
2.1.1	<i>Rekrutierung.....</i>	<i>11</i>
2.1.2	<i>Stichprobe.....</i>	<i>13</i>
2.2	<i>Datenerhebung.....</i>	<i>14</i>
2.2.1	<i>Fragebogen für Demografische Variablen.....</i>	<i>15</i>
2.2.2	<i>Narrative Interviews.....</i>	<i>16</i>
2.2.3	<i>Durchführung der Interviews.....</i>	<i>17</i>
2.3	<i>Datenauswertung.....</i>	<i>17</i>
2.3.1	<i>Transkription und Anonymisierung.....</i>	<i>18</i>
2.3.2	<i>Datenanalyse.....</i>	<i>18</i>
<b>3</b>	<b>Ergebnisse der Bedarfsanalyse.....</b>	<b>20</b>
3.1	<i>Auswirkungen von digitalem Hass und Hetze.....</i>	<i>21</i>
3.2	<i>Umgang mit digitalem Hass und Hetze.....</i>	<i>24</i>
3.2.1	<i>Persönliches Coping und Aufsuchen sozialer Unterstützung.....</i>	<i>24</i>
3.2.2	<i>Direkte Reaktion auf den digitalen Hass und Hetze.....</i>	<i>25</i>
3.2.3	<i>Erhaltene (professionelle) Unterstützung.....</i>	<i>27</i>
3.3	<i>Erwünschte Unterstützung.....</i>	<i>28</i>
3.4	<i>Empfehlungen für weitere Präventionsmaßnahmen.....</i>	<i>29</i>
3.5	<i>Einschätzungen zu bestehenden Maßnahmen.....</i>	<i>32</i>
3.5.1	<i>Notfallbuttons.....</i>	<i>32</i>
3.5.2	<i>Einsatz von Künstlicher Intelligenz (KI).....</i>	<i>33</i>
3.5.3	<i>Workshops und strukturierte Bildungsangebote.....</i>	<i>34</i>
<b>4</b>	<b>Diskussion und Zusammenfassung der Ergebnisse.....</b>	<b>36</b>
<b>5</b>	<b>Literaturverzeichnis.....</b>	<b>42</b>

# 1 Hintergrund

Diese Bedarfsanalyse wurde als interviewbasierte Studie gestaltet, die sich mit der Perspektive Betroffener von Hass im Netz vertieft auseinandergesetzt hat, darunter besonders Personen, die von Rassismus im Netz betroffen sind. Es sollte durch die Analyse herausgearbeitet werden, welche Ausprägung von Hass im Netz die Gruppe häufig betreffen, welche Auswirkungen erlebt und welche Unterstützungsangebote, Präventions- und Interventionsmaßnahmen als hilfreich wahrgenommen werden. Die Interviews wurden in der Vorbereitung, Durchführung und Analyse durch einen Beirat von Erfahrungs- und Fachexpert\*innen begleitet und durch Diskussionen fortlaufend validiert und mit aktuellen Forschungsergebnissen zusammengeführt. Die Bedarferhebung ist über vier Schritte umgesetzt worden, (1) die Leitfadententwicklung und Zielgruppendefinition, (2) eine konsensuelle Validierung des Vorgehens und der Methoden, (3) die Durchführung von narrativen, leitfadengestützten Interviews mit den Zielgruppen und (4) eine qualitative Analyse der erhobenen Daten sowie deren diskursive Validierung. Dieser Ergebnisbericht gibt einen Einblick in die Hintergründe und Methoden der Bedarfsanalyse sowie die zentralen Ergebnisse und Erkenntnisse der Analyse und die daraus abgeleiteten Implikationen.

## 1.1 Begriffsdefinitionen

Digitaler Hass und digitale Hetze sind Phänomene, die durch die zunehmende Verlagerung von Kommunikation in den digitalen Raum an Bedeutung gewonnen haben und verstärkt in den Fokus gesellschaftlicher Debatten rücken. Sie haben nicht nur schwerwiegende Folgen für die direkt Betroffenen, sondern wirken sich auch auf gesamtgesellschaftlicher Ebene negativ aus. Das Kompetenzzentrum „Hass im Netz“ verwendet Hass im Netz als Sammelbegriff, unter dem eine Vielzahl verwandter Phänomene zusammengefasst werden, darunter Cyberhass und Cyberbullying. Das Kompetenzzentrum definiert Hass im Netz daher breit als „eine Vielzahl unterschiedlicher u. a. *abwertender, entwürdigender, auf Einschüchterung zielender* oder

verhetzender Online-Phänomene gegenüber Personen oder bestimmten Personengruppen. Gemeint sind damit sowohl entsprechende Inhalte als auch Handlungen.“ (*Was ist Hass im Netz?*, 2024). Hass im Netz ist explizit kein juristisch definierter Begriff, umfasst aber Phänomene mit strafrechtlicher Relevanz wie Volksverhetzung oder Beleidigung. Mechanismen wie die Verstärkung durch Algorithmen, die sogenannte Online-Disinhibition (Online-Enthemmung) und verstärktes Moral Disengagement tragen maßgeblich zur Verbreitung und Eskalation dieser Phänomene bei (z. B. Lo Cricchio et al., 2021; Nocera et al., 2022; Runions & Bak, 2015; Silvia Casale et al., 2015; Wachs et al., 2022).

## 1.2 Betroffene von Hass im Netz

Ein besonderer Fokus der Forschung (so auch im Rahmen dieser Bedarfsanalyse) liegt auf vulnerablen Gruppen, die vermehrt Opfer von Hass im Netz werden können. Dazu zählen Menschen, mit rassifizierten Merkmalen (d. h. Merkmale wie Hautfarbe, Kleidung, Sprache, aufgrund derer Personen einer Gruppe zugewiesen werden, die sich biologisch reproduziert, sodass die Merkmale als inhärent und distinkt für diese Gruppe gesehen werden), sowie Personen, die aufgrund ihrer sexuellen Orientierung oder Geschlechtsidentität (z. B. Personen, die sich LGBTQIA\* oder FLINTA\* zugehörig fühlen<sup>1</sup>) Zielscheibe von Hass im Netz werden. Besonders betroffen sind auch Personen des öffentlichen Lebens, deren große Sichtbarkeit und Reichweite das Risiko für digitalen Hass und Hetze erhöhen kann, da sie von vielen Personen wahrgenommen werden. Dabei ist oft eine intersektionale Betroffenheit zu beobachten, wenn

---

<sup>1</sup> LGBTQIA\* steht für lesbian, gay, bisexual, transgender/transsexual, queer/questioning, intersex, asexual. (lesbisch, schwul, bisexuell, transgender/transsexuell, queer/fragend, intersexuell, asexuell).

FLINTA\* steht für Frauen, lesbische, inter, nicht-binäre, trans und agender Personen. Das \* (manchmal auch +) dient in beiden Fällen als Platzhalter für weitere Geschlechtsidentitäten. Damit gibt eine gewisse Schnittmenge der beiden Begriffe, FLINTA\*, manchmal auch FINTA\*, richtet sich allerdings primär an Frauen, da sie als marginalisierte Gruppe, auch innerhalb von LGBTQIA\*-Gruppen angesehen werden und daher eigene Räume erhalten sollen, um sich austauschen und entfalten zu können und nicht mit den von ihnen identifizierten patriarchalen Machtstrukturen konfrontiert sein zu müssen.

In dieser Bedarfsanalyse werden beide Begriffe verwendet, FLINTA\*, wenn primär das Merkmal als Frau als Geschlechtsidentität zentral erlebt wurde, LGBTQIA\*, wenn eine sexuelle Orientierung (z. B. Homosexualität) als primär relevant angesehen wurde.

Personen mehrere Merkmale sozialer Ungleichheit aufweisen und eine multiple Diskriminierung oder Anfeindung vorliegt (z. B. eine LGBTQIA\*-Aktivistin, die aufgrund hoher Reichweite in sozialen Medien als Person des öffentlichen Lebens gelten kann) (Tomczyk, 2024). Die Relevanz der Forschung zu digitalem Hass und Hetze zeigt sich auch in aktuellen gesellschaftlichen Entwicklungen und Debatten, wie der gestiegenen Zahl von Migrationsbewegungen und assoziierten Diskursen über gelungene und gescheiterte Integration und Sicherheit. In diesem Zusammenhang kommt dem Erleben von Rassismus eine besondere Bedeutung zu, da Personen aufgrund ihrer Herkunft oder ihrer körperlichen Erscheinung (z. B. Haar- oder Hautfarbe), und damit wahrnehmbarer und nicht wahrnehmbarer Merkmale, rassifiziert kodiert und abgewertet werden können.

Im digitalen Raum kann dies v. a. anhand von sprachlichen und nicht-sprachlichen Codes erfolgen, wie die Verwendung unterschiedlicher Sprach- und Zeichensysteme, die nicht der Mehrheitssprache entsprechen, das Verwenden von Vor- und Nachnamen oder Pseudonymen, die in Verbindung mit einer nicht-deutschen Herkunft stehen können (z. B. Namen arabischen Ursprungs) und die Nutzung von Bildern und Symbolen, die entsprechende Merkmale aufweisen (z. B. Flaggen unterschiedlicher Länder) (vgl. Matamoros-Fernández & Farkas, 2021). Diese Codes können Zeichen der Identitätsbildung und Verbundenheit für Individuen als Teil der relevanten sozialen Gruppe oder Ausdruck individueller Präferenzen sein, und damit z. B. auf Teilhabe, soziale Eingebundenheit oder Solidarität abzielen, aber sie können ebenso für Personen, die Hass ausüben, als Merkmal von Gruppen identifiziert werden, zu denen eine Abgrenzung erfolgen soll, was bspw. durch Ablehnung, Ausschluss aus Online-Diskursen oder auch Bedrohung, Beleidigung und Hass erfolgen kann (z. B. Lingren et al., 2024). Da in sehr vielen Kontexten eine multiple Betroffenheit vorliegt, da Personen z. B. rassifiziert kodiert werden und zugleich aufgrund ihrer geschlechtlichen Identität oder Zuschreibung kategorisiert und diskriminiert werden können, ist es zentral, eine gemeinsame

Betrachtung vorzunehmen. Ziel dieser Bedarfsanalyse ist, Erscheinungsformen des gruppenbezogenen Hasses im Netz – auch mit Blick auf Intersektionalität – zu erfassen und auf dieser Basis Hinweise für konkrete Schutzmaßnahmen abzuleiten, die auf die Bedarfe der betroffenen Zielgruppen zugeschnitten sind. Daher sollen nachfolgend wesentliche Auswirkungen von Hass im Netz sowie relevante Maßnahmen zu Schutz und Prävention zusammengefasst werden, die eine Grundlage für die Analyse und Interpretation der Bedarfsanalyse sind.

### **1.3 Auswirkung von „Hass im Netz“ auf Betroffene**

Das Erleben von Hass im Netz (hier definiert als Cyberviktimsierung) kann tiefgreifende Auswirkungen für die Betroffenen haben. Cyberviktimsierung wird häufig mit psychischen Belastungen wie Stress, Trauer und Wut in Verbindung gebracht, die das Wohlbefinden erheblich beeinträchtigen können (Buelga et al., 2022; Jadambaa et al., 2019; John et al., 2018; Li et al., 2022). Darüber können auch somatische Beschwerden wie Kopfschmerzen und Schlafstörungen auftreten, die die körperliche Gesundheit der Betroffenen beeinflussen können (Alhaboby et al., 2019). In den Folgen der Viktimsierung zeigen sich auch geschlechtsspezifische Unterschiede. Frauen und Mädchen berichten bspw. deutlich häufiger von psychosozialen und psychosomatischen Belastungen wie Angstzuständen und Depressionen (Wright & Wachs, 2020). Diese Unterschiede verdeutlichen die Notwendigkeit, präventive und unterstützende Maßnahmen geschlechtsspezifisch auszurichten, um gezielt auf die unterschiedlichen Bedarfe einzugehen und eine effektive Unterstützung zu gewährleisten. Die individuellen Auswirkungen von Hass im Netz stehen jedoch nicht isoliert, sondern sind eingebettet in soziale und gesellschaftliche Dynamiken. Auf gesellschaftlicher Ebene führt die Normalisierung abwertender Sprache zu einer zunehmenden Akzeptanz von diskriminierenden Haltungen und einer stärkeren Ablehnung von Minderheiten wie Immigrant\*innen sowie religiösen, ethnischen und sexuellen Minderheiten (Bilewicz & Soral, 2020). Diese Entwicklungen verstärken Ausgrenzung und Marginalisierung und schaffen ein Umfeld, das

weiteren Hass fördert. Die Wechselwirkungen zwischen individuellen und gesellschaftlichen Folgen sind hierbei zentral. Gesellschaftliche Entwicklungen, wie die Verrohung des Dialogs, haben einen direkten Einfluss auf die Betroffenen, da sie Diskriminierung verstärken und die psychosozialen Belastungen erhöhen. Gleichzeitig tragen die individuellen Auswirkungen dazu bei, dass sich Betroffene aus gesellschaftlichen Diskursen zurückziehen, was wiederum Ausgrenzung verstärken und eine Art Spiralwirkung entfalten kann (Wright & Wachs, 2020).

#### **1.4 Unterstützende, schützende und präventive Ansätze**

Präventive und schützende Maßnahmen gegen Hass im Netz können auf unterschiedlichen Ebenen ansetzen, eine ausführlichere Darstellung findet sich im Gutachten von Tomczyk (2024). Dazu gehören Maßnahmen auf individueller oder sozialer und damit verhaltenspräventiver Ebene (z. B. bei den Betroffenen selbst, bei den Aus\*übenden von Hass im Netz, bei Aktiven in sozialen Netzwerken, etwa als Moderator\*innen) sowie institutioneller und politischer und damit verhältnispräventiver Ebene (etwa in Form von Gesetzgebung). Auf individueller Ebene sind zunächst Bewältigungsstrategien (sog. Coping) zu nennen, mit denen auf Hass im Netz als ein potenziell stressauslösendes Ereignis (sog. Stressor) reagiert wird. Die Forschung zu Coping ist weitläufig und v. a. in der Verarbeitung von Stress, Gesundheit und Krankheit erforscht, im Wesentlichen können dabei Ansätze unterschieden werden, die eine eher emotionsorientierte oder problemorientierte Herangehensweise beschreiben (vgl. Lazarus & Folkman, 1991). *Emotionsorientiertes Coping* bezieht sich auf die Bewältigung der durch den Stressor ausgelösten Gedanken und Emotionen, wie z. B. Ablenkung, Schaffung positiver Emotionen durch Wechsel auf angenehme Aktivitäten, Austausch mit anderen, oder aktive Reduktion von als unangenehm erlebten Emotionen wie Angst sowie Stressreduktion, z. B. durch Bewegung, Atemübungen oder Achtsamkeit. Auch Strategien der Rationalisierung (z. B. Normalisierung des Geschehens, Zuschreibung auf äußere Umstände zur Entlastung der eigenen Person) können darunter fallen. Dies schafft in der Regel kurzfristig Erleichterung, hat

aber keinen Einfluss auf die Situation bzw. das Geschehen. *Problemorientiertes Coping* meint Ansätze, die sich mit der Situation auseinandersetzen und aktiv an Lösungen arbeiten. Im Kontext von Hass im Netz kann dies etwa konkrete Gegenrede bedeuten, aber auch die Nutzung von Meldestellen oder die Moderation von Diskussionen. Da dies unmittelbare Auswirkungen auf das Kommunikationsgeschehen haben kann, z. B. bei Ausschluss/Blockieren von bestimmten Nutzer\*innen, ist zu erwarten, dass es längerfristige Effekte hat als emotionsorientiertes Coping. Die beiden Herangehensweisen treten im Alltag häufig ergänzend auf, da es nicht in jeder Situation möglich ist, z. B. Gegenrede zu gestalten oder komplexe Diskussionen zu moderieren, aber eine kurzfristige Erleichterung oder Ablenkung dennoch als hilfreich erlebt wird (Lazarus & Folkman, 1991). Viele verhaltensbasierte Maßnahmen zielen daher darauf ab, hilfreiche Copingstrategien zu entwickeln bzw. zu unterstützen (z. B. durch Workshops, Trainings oder Coaching) und Raum für deren Umsetzung zu schaffen (z. B. Meldefunktionen oder Notfallbuttons, (KI-gestützte) Hinweise zur Gestaltung von Gegenrede, Training von Kommunikationsfertigkeiten oder Stressbewältigung). Ferner existiert eine Vielzahl von Bildungs- und Empowerment-Programmen, die auf die Stärkung individueller Kompetenzen (wie Stressbewältigung), aber auch die Förderung einer digitalen Zivilcourage und Beteiligung abzielen (z. B. Gegenrede gestalten, diskriminierungs- und stigmafremde Sprache in digitalen Räumen nutzen). Diese Maßnahmen werden weitreichend umgesetzt, sind bislang aber zumeist unzureichend hinsichtlich Akzeptanz und Wirksamkeit evaluiert.

Verhältnispräventive Maßnahmen zielen nicht auf das Verhalten selbst, sondern dessen Rahmenbedingungen ab, daher fallen darunter etwa gesetzliche Regelungen, Funktionen von sozialen Medien/Online-Plattformen, der Einsatz Künstlicher Intelligenz zur Erkennung und Moderation von Hassinhalten oder die Bereitstellung von Tools (z. B. KI) zum Umgang mit Hass im Netz – der Kompetenzaufbau zur Nutzung solcher Tools wäre dann eine ergänzende Maßnahme der Verhaltensprävention. Verhältnispräventive Maßnahmen zielen vielfach darauf

ab, die Rechte von Nutzer\*innen sozialer Medien zu stärken, indem sie ein transparentes Beschwerdemanagement, Auskunfts- und Informationspflichten der Anbieter\*innen sowie Sanktionen wie Bußgelder vorschreiben und mit Meldepflicht schwerer Straftaten verbinden, um Strafverfolgung zu erleichtern und digitale Hasskriminalität abzuschrecken. Dabei werden etwa Anbieter\*innen von sozialen Medien wie Instagram, Twitter/X oder YouTube gezielt angesprochen, da sie entsprechende Funktionen vorhalten und Nachverfolgungen ermöglichen müssen.

Die Durchsetzung dieser Regelungen kann zudem durch Künstliche Intelligenz (KI) unterstützt werden. Diese bietet vielversprechende Möglichkeiten zur Identifikation und Moderation von digitalem Hass, beispielsweise durch (teil-)automatisierte Systeme wie die Nutzung von Lexika wie HurlLex (Poletto et al., 2021). Diese Tools ermöglichen es, hasserfüllte Inhalte schneller zu erkennen und zu bearbeiten. In der Praxis stößt der Einsatz solcher Systeme jedoch auf Herausforderungen, wie die Entwicklung unverzerrter und zuverlässiger Algorithmen, die zuverlässige Erkennung und Analyse sprachlicher Muster und nichttextlicher Inhalte wie Emojis oder Videos (Abarna et al., 2022; Khurana et al., 2023). Und darüber hinaus bleiben kulturelle Unterschiede, Sarkasmus und idiomatische Ausdrücke schwer erfassbar (Sarsam et al., 2020), algorithmenbasierte Ansätze können zudem auch zur Gestaltung von Hassrede genutzt werden (sog. *Dual Use-Problematik*), sodass Hassbotschaften reproduziert, vervielfältigt und disseminiert werden können, was Online-Gemeinschaften schnell überfordern, bedrohen oder zum Rückzug bewegen kann (z. B. mittels Bot-Einsatz) (Garland et al., 2022). In einigen Social-Media-Plattformen selbst sind verschiedene KI-Ansätze bereits integriert und mit Handlungsansätzen verknüpft, um auf Hass im Netz zu reagieren. Dazu gehören technische Maßnahmen wie Algorithmen zur automatisierten Erkennung und Entfernung von Hassinhalten (z. B. über verbotene Worte/Aussagen in Inhalten, die dadurch in ihrer Reichweite eingeschränkt werden), aber auch nutzer\*innen-orientierte Tools wie Meldesysteme, mit denen hasserfüllte Inhalte gemeldet und überprüft werden

können. Plattformen implementieren zudem Richtlinien bzw. Standards (eng. Community-Standards) für die Nutzung ihrer Plattformen (bspw. Meta, 2024). Basierend darauf sollen Inhalte, die von Nutzenden gemeldet wurden von Community-Moderator\*innen oder spezialisierten Teams durchgesetzt werden.

Personen, die Hass verbreiten möchten, reagieren zunehmend auf diese Moderation und nutzen vermehrt internen Code, sog. Dogwhistles. Ziel ist es damit Hass verbreiten zu können, ohne seine Ansichten bzw. Affiliationen zu verbotenen Organisationen für die Mehrheit erkennbar zu kennzeichnen. Die Posts sind ausgezeichnet durch das Nutzen von Begriffen und Emojis, die ohne Hintergrundwissen nicht als radikal, kontrovers oder ausgesprochen verhetzend wahrgenommen werden (Lindgren et al., 2024). Die Folgen sind mangelnde Moderation und oft für außenstehende unerkannte Attacken auf bestimmte Personen oder Personengruppen. Diese Herausforderungen bei der Moderation von Hassinhalten auf Social-Media-Plattformen zeigen, dass technische und algorithmische Ansätze allein nicht ausreichen, um das Problem effektiv zu bekämpfen. Ergänzend dazu spielen pädagogische und beratende Maßnahmen eine wichtige Rolle, die direkt bei den Betroffenen ansetzen, aber auch die Gesellschaft im Umgang mit digitalem Hass sensibilisieren.

Während verhältnispräventive Maßnahmen eine breite Wirkung entfalten, da sie z. B. für alle Nutzenden einer Plattform gelten, sind viele der verhaltensbasierten Maßnahmen spezifisch ausgerichtet, so richtet sich ein Großteil der von Tomczyk (2024) identifizierten Programme direkt oder indirekt an Kinder, Jugendliche und junge Erwachsene. Ältere Erwachsene und hochalte Menschen wurden bislang seltener betrachtet, zudem wurde die Evaluation von Maßnahmen häufig nur unzureichend durchgeführt, sodass die Schlussfolgerungen hinsichtlich der Wirksamkeit von Maßnahmen mit Vorsicht zu betrachten sind. Dennoch können einige Trends identifiziert werden. Als besonders relevant gelten demnach settingbasierte, strukturierte Programme, die eine Mindestdauer von acht Stunden aufweisen und einen Methodenmix aus Präsenz- und digitalen Elementen kombinieren. Beispiele hierfür sind der

Einsatz von virtueller Realität, webbasierte Anwendungen oder Nachrichtendienste (Doty et al., 2021; Mishna et al., 2009; Polanin et al., 2022). Inhaltlich sind die Förderung affektiver Empathie, sozio-emotionaler Kompetenzen, Kommunikationsfähigkeiten sowie die Vermittlung von Digital Citizenship prägende Elemente. Letzteres beschreibt die Befähigung zur gesellschaftlichen Teilhabe durch digitale Medien. Dazu gehören sowohl die Aufklärung über digitale Phänomene wie Filterblasen oder Echokammern als auch die aktive Nutzung digitaler Räume für Partizipation, etwa durch Crowdfunding, digitale Petitionen oder Selbstverteidigung (Choi, 2016; Hintz et al., 2017).

Zugleich ist deutlich, dass bestimmte Zielgruppen bislang kaum im Fokus bestehender Maßnahmen und Untersuchungen stehen. Dazu gehören vor allem Erwachsene sowie Personen, die besonders häufig von Hass im Netz betroffen sind, wie beispielsweise Mitglieder marginalisierter Gruppen oder Vertreter\*innen aus Politik und anderen Bereichen des öffentlichen Lebens (Tomczyk, 2024). Diese Lücke verdeutlicht, dass es an umfassenden, zielgruppenspezifischen Erkenntnissen fehlt, die als Grundlage für die Entwicklung passgenauer Unterstützungs- und Schutzmaßnahmen dienen könnten. Vor diesem Hintergrund wurde die vorliegende Bedarfsanalyse durchgeführt, um die Perspektiven von Betroffenen genauer zu explorieren und für die Entwicklung neuer Präventions- und Unterstützungsangebote aufzubereiten, die derzeit im Auftrag des Landespräventionsrats Niedersachsen entwickelt und erprobt werden.

## 2 Methodik

Im Rahmen der Bedarfsanalyse wurden stratifiziert Personen für Interviews rekrutiert, die sich hinsichtlich relevanter Merkmale (Rassismuserfahrung, LGBTQIA\*, FLINTA\*) und ihrer Kombination unterschieden, um eine Repräsentation dieser Gruppen zu ermöglichen (als sog. *stratified purposeful sampling*) (Palinkas et al., 2015). So wird eine Synthese ermöglicht, die die unterschiedlichen Perspektiven und Erfahrungen einordnet und, ausgehend von Rassismus, weitere Betroffenheit in den Blick nimmt und dafür Empfehlungen ableiten kann (siehe Tabelle 1). Als Merkmale wurden von Rassismus betroffene Personen, hier näherungsweise über Migrationshintergrund und damit verbundene digitale Diskriminierungserfahrung abgebildet, Zugehörigkeit zu LGBTQIA\* und/oder FLINTA\* gewählt, da diese laut Forschung besonders betroffene und vulnerable Gruppen in Bezug auf digitalen Hass sind.

Das Forschungsteam bestand aus einer männlich und drei weiblich gelesenen Personen, den Autor\*innen dieses Berichts, sowie weiteren Hilfskräften mit Qualifikationen in den Fächern Psychologie, Soziologie und Kommunikationswissenschaft (Bachelor of Arts/Science, Master of Art/Science, Promotion) und langjähriger Erfahrung in der ehrenamtlichen Arbeit mit Menschen mit Fluchterfahrung und diverser geschlechtlicher Identität und sexueller Orientierung sowie der qualitativen Sozialforschung. Insgesamt wurden 16 Interviews mit Betroffenen durchgeführt und mittels qualitativer Inhaltsanalyse nach Kuckartz (2018) ausgewertet. Im Folgenden wird die Rekrutierung der Zielgruppe sowie die Durchführung und Auswertung der Interviews genauer beschrieben.

### 2.1 Zielgruppe und Rekrutierung

Die Auswahl der Zielgruppen für die Studie basierte auf den Erkenntnissen aus dem „*Gutachten zum Phänomen des digitalen Hasses und der digitalen Hetze in Niedersachsen*“ (Tomczyk, 2024) und weiterer Fachliteratur. Zudem wurden Expert\*innengespräche (etwa mit der Fachgruppe Digitale Gefahren des Landespräventionsrats

Niedersachsen, dem Landesrat für Kriminalvorbeugung M-V, und Expert\*innen aus der Forschung und Praxis zu Hass im Netz) genutzt, um die Auswahl zu diskutieren, die Relevanz und Betroffenheit abzubilden und auch mit Blick auf Forschungs- und Entwicklungsbedarfe zu priorisieren (z. B. fehlende Erkenntnisse zu Präferenzen spezifischen Gruppen).

Bekannt ist, dass gruppenbezogene Merkmale wie Migrationsgeschichte und rassifizierte Merkmale, Geschlecht oder sexuelle Orientierung häufige Angriffspunkte für digitalen Hass darstellen können. Zusätzlich wurde die Perspektive von Personen des öffentlichen Lebens als Dimension berücksichtigt, da sich u. a. im Zuge der Rekrutierung herausstellte, dass viele der Personen, die bereit waren, zu dem Thema Interviews zu führen, auch über eine Bekanntheit und Reichweite, z. B. in sozialen Medien, verfügten, da sie oftmals auch politisch und in Interessensverbänden aktiv waren und damit eine gewisse Öffentlichkeit erreichten.

### **2.1.1 Rekrutierung**

Für die Rekrutierung der Teilnehmenden wurde entsprechend der gewählten Zielgruppen ein Stichprobenplan erstellt (siehe Tabelle 1). Um diese Zielgruppe zu erreichen, wurden geeignete Kontakte recherchiert, dabei wurden unter anderem Einzelpersonen, Parteien, Institutionen (Integrationsbeiräte) und Betroffenenvertretungen wie das *Queere Netzwerk Niedersachsen* berücksichtigt. Um möglichst viele betroffene Personen aus Niedersachsen zu erreichen, wurden in den ersten Rekrutierungswellen Organisationen kontaktiert, die vorrangig in Niedersachsen aktiv sind (da es sich bei den relevanten Gruppen vielfach auch um Online-Gruppen bzw. online aktive Gruppen handelt, sind die Mitwirkenden nicht immer auf die Landesgrenzen beschränkt). Da sich die Rekrutierung in mehreren Wellen als schwierig erwies, wurden in späteren Wellen auch Organisationen außerhalb Niedersachsens berücksichtigt. Alle Kontakte wurden in Listen zusammengetragen und die Ergebnisse der Kontaktaufnahme dokumentiert. Die Kontaktaufnahme wurde per E-Mail, Telefon und Instagram bzw. über soziale Medien umgesetzt, wenn sie entsprechend als Kontaktadressen angegeben wurden. Die

Organisationen wurden mit der Bitte kontaktiert, den Aufruf zur Studienteilnahme zu teilen, während Einzelpersonen zudem direkt um eine Teilnahme gebeten wurden.

**Tabelle 1**

*Stichprobenplan für die Rekrutierung von Teilnehmenden für die Bedarfsanalyse.*

	Cis-Männer		FLINTA*	
LGBTQIA*	1	2	2	3
Nicht LGBTQIA*	1	2	2	3

*Anmerkung:* Anzahl an Teilnehmenden gesamt: 16; grau hinterlegte Felder beziehen sich auf das Vorhandensein von Merkmalen, die mit der Betroffenheit von Rassismus assoziiert wurden (z. B. Migrationshintergrund, ethnische Zugehörigkeit).

Personen, die an einer Teilnahme interessiert waren, konnten über einen SoSci Survey-Link (Leiner, 2019) Zugang zu einem kurzen Online-Fragebogen bekommen. Der Fragebogen erfasste grundlegende demografische Informationen wie Alter und Geschlecht sowie eine erste Einschätzung, inwiefern die Person von Hass im Netz betroffen war. Nur Personen, die über 18 Jahre alt waren und bestätigten, dass sie bereits Erfahrungen mit digitalem Hass gemacht hatten, wurden für die Teilnahme an der Studie berücksichtigt. Auf diese Weise wurde die unmittelbare Erfahrung von Hass im Netz als gemeinsames Merkmal definiert und eine Altersgruppe angesprochen, die die bisherige Forschung im Kindes- und Jugendalter ergänzt.

Der Fragebogen diente nicht nur der Erhebung dieser demografischen Daten, sondern enthielt auch eine Einwilligungserklärung zur Teilnahme an der Studie. Nach Abschluss des Fragebogens hatten die Teilnehmenden die Möglichkeit, einen passenden Termin für das Interview aus einem digitalen Kalender auszuwählen. Im Anschluss erhielten sie per E-Mail eine Bestätigung des Termins sowie den Link zum Interview.

### 2.1.2 Stichprobe

Die finale Stichprobe umfasste 16 Interviews, zu gleichen Anteilen geführt von Carlotta Alpers und Anna Matthes, mit Personen aus verschiedenen Bundesländern. Die meisten stammen aus Niedersachsen (5 Personen), gefolgt von Nordrhein-Westfalen (3 Personen), Hessen (2 Personen), Sachsen (2 Personen) und jeweils einer Person aus Sachsen-Anhalt, dem Saarland, Mecklenburg-Vorpommern und Berlin. In Bezug auf die Zugehörigkeit zu einer von Rassismus betroffenen Gruppe, gaben sieben Personen an, sich keiner der genannten Gruppen zugehörig zu fühlen. Zwei Personen identifizierten sich als asiatisch, zwei als muslimisch, zwei als osteuropäisch, und je eine Person als jüdisch und eine als schwarz. Eine Person gab mehrere Zugehörigkeiten an (asiatisch, muslimisch, osteuropäisch). Die Selbstbeschreibung der ethnischen Zugehörigkeit zeigt eine stärkere Differenzierung. So beschrieben viele Personen zusätzlich zu den vorgeschlagenen auch geografische Zuordnungen (z. B. „mitteleuropäisch“, „Vietnamesisch in Ethnie und Kultur mit deutschen Anteilen aufgrund von Geburt und Sozialisierung in Deutschland“, „Spagat zwischen deutschen und russisch/ukrainischen Vorfahren“). In der Durchführung der Interviews wurde die entsprechende Selbstzuschreibung der Personen berücksichtigt, indem spezifisch Nachfragen, z. B. zu Unterschieden und Gemeinsamkeiten im Erleben in verschiedenen Kulturräumen, gestellt wurden.

Geschlecht und geschlechtliche Identifikation der interviewten Personen wurden ebenfalls differenziert erhoben. Die Mehrheit der Teilnehmenden (8 Personen) identifiziert sich als weiblich, während fünf sich als männlich und je eine Person als nicht-binär oder agender/kein Geschlecht bezeichneten. Eine Person gab ihr eine Mehrfachidentifikation an (männlich, Inter\*, queer). Das biologische Geschlecht stimmt in 13 Fällen mit der geschlechtlichen Identifikation überein; in drei Fällen gibt es Unterschiede. Die sexuelle Orientierung zeigt eine breite Vielfalt: sieben Personen identifizieren sich als bi- oder pansexuell, fünf als heterosexuell, drei als homosexuell, und eine als asexuell.

Beim Bildungsstand haben die meisten Teilnehmenden (14 Personen) einen schulischen Abschluss auf Abiturniveau, zwei Personen verfügen über eine Fachhochschulreife bzw. Fachoberschule. Beruflich haben neun Teilnehmenden einen Abschluss einer Universität oder Hochschule, eine Person hat eine Lehre abgeschlossen, und fünf befinden sich noch in Ausbildung oder haben keinen beruflichen Abschluss. Eine Person machte keine Angabe.

Damit haben die Personen im Vergleich zum Bundesdurchschnitt einen sehr hohen Bildungsgrad, was eine Einschränkung der Generalisierbarkeit der Ergebnisse bedeutet. Da die Auseinandersetzung mit dem Thema Hass im Netz und die Differenzierung verschiedener personen- und gruppenbezogener Merkmale allerdings eine Voraussetzung für die Teilnahme und den Studieneinschluss darstellte, ist dies erwartbar.

Das Alter der Teilnehmenden liegt zwischen 19 und 63 Jahren. Der Altersdurchschnitt beträgt 30,5 Jahre. Eine Person machte keine Angaben zu ihrem Alter. Zudem wurden die Personen basierend auf ihren Aussagen im Interview als Person des öffentlichen Lebens kategorisiert, wenn die Person außerhalb ihres persönlichen und beruflichen Umfelds bekannt ist oder war, beispielsweise durch politisches, künstlerisches oder aktivistisches Engagement. Die Einordnung wurde bewusst nach Selbstzuschreibung gewählt, da die Auswahl belastbarer Außenkriterien bei der Vielfalt an Teilnehmenden und der relevanten Öffentlichkeiten nicht möglich erschien. Nach diesem Kriterium wurden acht Teilnehmende (AM01, AM02, AM04, CA01, CA02, CA03, CA04, CA05) als Person des öffentlichen Lebens klassifiziert.

## **2.2 Datenerhebung**

Im Rahmen dieser Studie wurden qualitative Interviews und ein Online-Fragebogen eingesetzt. Jede teilnehmende Person wurde gebeten, vor dem 60- bis 90-minütigen narrativen Interview, den online bereitgestellten Fragebogen auszufüllen. Die Interviews ermöglichen es, tiefgehende Einblicke in die persönlichen Erfahrungen und Perspektiven der Betroffenen von digitalem Hass zu gewinnen. Der ergänzende Fragebogen diente dazu, demografische und

vergleichbare Informationen zu sammeln. Im Folgenden werden die Methode der narrativen Interviews sowie die Erstellung des Fragebogens und des Leitfadens vorgestellt.

### **2.2.1 Fragebogen für Demografische Variablen**

Im Rahmen der Studie wurde ein Fragebogen eingesetzt, um soziodemografische Variablen der Teilnehmenden systematisch zu erfassen. Dieser diente dazu, grundlegende Merkmale der Teilnehmenden zu dokumentieren und Unterschiede zwischen den Teilnehmendengruppen zu analysieren. Vor Beginn der Erhebung stimmten die Teilnehmenden der Nutzung ihrer Daten zu, indem sie eine Einwilligungserklärung unterzeichneten, die sowohl die Teilnahme an der Studie als auch die Verarbeitung ihrer Daten umfasste.

Anschließend wurde die Betroffenheit von Hass im Netz erfragt. Hierzu wurde die Definition des Kompetenznetzwerkes gegen Hass im Netz zitiert und konnte mit den Antwortmöglichkeiten „ja“ und „nein“ beantwortet werden. Es wurde darauf verwiesen, dass nur bei Zustimmung die Befragung fortgeführt werden konnte.

Die nächsten Fragen dienten der Erfassung der soziodemographischen Daten der Teilnehmenden. Hierzu wurden zunächst das Alter, der höchste schulische Abschluss, der höchste berufliche Schulabschluss, das aktuelle Bundesland indem die Teilnehmenden leben erfragt. Orientiert an den Fragen von Horstmann et al., (2022) wurde das biologische Geschlecht und die Geschlechtsidentifikation erfragt. Die Zugehörigkeit zu einer rassifizierten Gruppe wurde mit den Fragen von Ruland et al. (2022) erhoben. Die Fragen zur Sexualität orientierten sich an Lenzner et al. (2022). Anschließend wurden die Teilnehmenden gebeten ein Codewort zu erstellen, bestehend aus den letzten beiden Buchstaben ihres Vornamens, dem Geburtstag (zweistellig) und den letzten beiden Ziffern der Postleitzahl. Dies diente der späteren Zuordnung zu den entsprechenden Interviews. Zum Schluss wurden die Teilnehmenden auf eine externe Webseite (DFN – Terminplaner) weitergeleitet, auf der Sie einen Termin für das Interview vereinbaren konnten.

## 2.2.2 Narrative Interviews

Das narrative Interview ist eine Methode von Schütze (1983), die für die Biografie-forschung entwickelt wurde und darauf abzielt durch eine größtmögliche Offenheit, frei einen gesamt Eindruck der Perspektive und den Gedanken einer Erzählperson zu erfassen. Dabei werden verschiedene Phasen berücksichtigt. Nachdem das das Gespräch eingeleitet wurde, die Aufnahme gestartet und Zustimmung zum Interview und der Verarbeitung der Daten gegeben wurde, beginnt die interviewende Person mit einem Erzählstimulus, der die erste Phase, die offene Erzählung einleitet. In dieser „Stehgreiferzählung“ kann (und sollte) die Interviewte ihre gesamten Erfahrungen berichten, ohne dabei von der Interviewenden unterbrochen zu werden. Wenn die Interviewte Person andeutet keine weiteren Angaben machen zu wollen oder zu können („das war jetzt alles“, „mehr fällt mir jetzt auch nicht ein“), geht die Interviewende in die nächste Phase über. Hier werden gezielt Fragen gestellt, die sich auf das Gesagte beziehen (sog. Immanente Fragen), die zum Verständnis beitragen und ggf. Lücken in der Erzählung füllen. Zuletzt werden noch exmanente Fragen gestellt. Diese entstehen nicht direkt aus dem Gesagten, sondern richten sich vielmehr auf das spezifische Forschungsinteresse der Forschenden. Dazu gehören etwa Fragen nach Erklärungen, Meinungen, Interpretationen oder Bewertungen. Danach wird das Interview final abgeschlossen („Gibt es nun am Ende des Interviews noch etwas, was Sie gerne zum Thema erzählen möchten?“).

### 2.2.2.1 Leitfadenerstellung

Der Leitfaden für die narrativen Interviews wurde vom Forschungsteam auf Grundlage der Forschungsfrage und der im Gutachten von Tomczyk (2024) gewonnenen Erkenntnisse erstellt. Dieser erste Entwurf wurde daraufhin konsensuell mit Expert\*innen validiert. Dazu wurden Personen aus verschiedenen Bereichen einbezogen, neben Forschenden wurden auch Personen aus der Praxis und Betroffene von Hass im Netz integriert. In sieben Einzelinterviews mit Expert\*innen (davon vier Forschende) wurde der Leitfaden gemeinsam durchgesprochen, die Definitionen von Hass im Netz und die Erfassung der relevanten Merkmale, wie im Fragebogen

abgebildet, diskutiert und Hinweise für die Durchführung der Interviews abgeleitet. Die Interviews dauerten jeweils zwischen 30 und 60 Minuten. Der Leitfaden wurde entsprechend den Anmerkungen angepasst und dann in Probeinterviews mit studentischen Hilfskräften, die mindestens eines der untersuchten Merkmale aufwiesen, final getestet, um Verständlichkeit, Zumutbarkeit und Akzeptanz sicherzustellen.

### **2.2.3 Durchführung der Interviews**

Die Interviews wurden anschließend über eine Videokonferenzplattform durchgeführt und mit der aufgezeichnet. Zu Beginn jedes Interviews wurden die Teilnehmenden erneut über die Verwendung der Daten und die Aufzeichnung informiert und gaben ihr Einverständnis zur Aufnahme. Die Interviews selbst wurden anhand des Leitfadens geführt, wobei die Reihenfolge der Fragen an den Gesprächsverlauf angepasst werden konnte. Die Gesprächsdauer variierte zwischen 60 und 90 Minuten. Nachdem die Aufnahme gestoppt wurde, wurde noch einmal ein kurzes Debriefing durchgeführt, um sicherzustellen, dass die Teilnehmenden sicher in ihren Alltag zurückkehren konnten. Die Aufzeichnungen wurden in der Folge transkribiert und inhaltsanalytisch ausgewertet. Im Folgenden wird das Vorgehen in Anlehnung an relevante Fachliteratur genauer beschrieben (Kuckartz & Rädiker, 2022, 2024).

## **2.3 Datenauswertung**

Die Datenauswertung umfasste mehrere Schritte, die sicherstellten, dass die erhobenen Informationen systematisch verarbeitet und analysiert werden konnten. Der Fokus lag dabei auf der sorgfältigen Transkription und Anonymisierung der Interviews sowie auf der detaillierten Analyse mithilfe der Methode der inhaltlich strukturierenden qualitativen Inhaltsanalyse nach Kuckartz (2018).

### **2.3.1 Transkription und Anonymisierung**

Die Audioaufzeichnungen der Interviews wurden mittels einer datenschutzgeprüften, KI-gestützten Software an der Universität Greifswald transkribiert und somit in Textform überführt, die Aufnahmen anschließend gelöscht. Die Transkriptionsregeln wurden von Kuckartz (2018) übernommen. Vor der Analyse wurden die Transkripte gemeinsam durch das Forschungsteam anonymisiert, die Anonymisierung orientierte sich an den Empfehlungen von Meyermann et al. (2014).

### **2.3.2 Datenanalyse**

Die Analyse der erhobenen Daten auf Basis der anonymisierten Transkripte erfolgte mittels der Methode der inhaltlich strukturierenden qualitativen Inhaltsanalyse nach Kuckartz und Rädiker (2024). Zu Beginn der Analyse wurde ein deduktives Kategoriensystem entwickelt, das aus den Forschungsfragen sowie dem Leitfaden der Interviews abgeleitet wurde. Diese Kategorien bildeten die Grundlage für den Codierprozess und ermöglichten eine erste Strukturierung der relevanten Themenbereiche. Im ersten Codierprozess wurde eine Auswahl von Daten unabhängig voneinander von mehreren Codierenden analysiert. Dieses konsensuelle Codieren stellte sicher, dass alle Beteiligten ein einheitliches Verständnis der Kategorien hatten. Nach der individuellen Codierung wurden die Ergebnisse verglichen, Abweichungen diskutiert und ein Konsens erzielt. Basierend auf dieser Validierung wurden Anpassungen vorgenommen, darunter die Überarbeitung des Umfangs codierter Abschnitte sowie die Differenzierung und Anpassung von Kategorien. Zusätzliche Kategorien wie „Umgang mit der Situation/Coping“ und „Bewertung der Reaktionen (positiv/negativ)“ wurden hinzugefügt, um die Daten besser zu erfassen. Die Kodierregeln, die im Analyseprozess angewendet wurden, können Tabelle 2 entnommen werden.

## **Tabelle 2**

*Verwendete Kodierregeln nach Kuckartz und Rädiker (2024)*

Regel	Beschreibung
1. Sinneinheiten codieren	Üblicherweise wird mindestens ein vollständiger Satz codiert.
2. Zusammenhängende Segmente codieren	Wenn eine Sinneinheit mehrere Sätze oder Absätze umfasst, werden diese als ein zusammenhängendes Segment codiert.
3. Interviewfragen einbeziehen	Einleitende oder zwischengeschobene Fragen der Interviewenden werden mitcodiert, wenn sie für das Verständnis der Sinneinheit notwendig sind.
4. Angemessene Textumgebung codieren	Beim Codieren wird darauf geachtet, dass die Textstelle ohne den sie umgebenden Kontext für sich verständlich bleibt.

Nach der konsensuellen Validierung wurden zwei weitere Interviews analysiert, um die Interkoder-Übereinstimmung zu prüfen. Aufgrund einer zu dem Zeitpunkt niedrigen Übereinstimmung wurden die Kodierregeln erneut angepasst und die Kategorien weiter differenziert, um die Zuordnung zu erleichtern. Die verbleibenden Interviews wurden dann erneut kodiert, wobei eine zufriedenstellende Interkoderübereinstimmung erzielt werden konnte. Im nächsten Schritt, der induktiven Subkategorienbildung, wurden die Hauptkategorien dann durch eine erneute Analyse des Materials differenziert. Subkategorien sind dabei direkt aus dem Datenmaterial entstanden. Diese Subkategorien wurden in einem zweiten Kodierprozess auf das gesamte Material angewendet werden, wodurch die Differenzierung der Analyse geschärft werden konnte. Im Folgenden werden die Ergebnisse des Kodierprozesses zusammenfassend berichtet und sind anhand der Angaben zu Erfahrungen, Auswirkungen, Präventionsmaßnahmen und Empfehlungen zum Umgang mit Hass im Netz gegliedert.

### 3 Ergebnisse der Bedarfsanalyse

Die im Rahmen der qualitativen Interviews gesammelten Daten geben Einblicke in die vielfältigen und oft belastenden Erfahrungen der Teilnehmenden mit digitalem Hass und Hetze. Dabei wurden sowohl direkte persönliche Angriffe als auch indirekte Formen von Hass geschildert. Die Erlebnisse reichten von sehr gezielten Angriffen (AM05, CA02, CA05, CA07), *„In den Hochzeiten, die sich manchmal über Wochen und Monate gezogen haben, gab es manchmal täglich drei vier, von verschiedenen Akteuren drei, vier Morddrohungen.“* (CA05, Pos. 6) bis hin zu ständigem Hass, der häufig als Reaktion auf geteilte Inhalte und Meinungen auftrat (CA03, CA04, CA05, AM01) *„...im Zuge meiner politischen Arbeit habe ich da immer wieder Erfahrungen gemacht mit Hasskommentaren.“* (CA03, Pos. 10). Einige Teilnehmende berichteten außerdem vom Erleben indirekten Hasses, etwa durch die Beobachtung von Beleidigungen oder Hassbotschaften gegen andere Personen oder Gruppen (AM06, AM03, CA06). *„Und dann habe ich mir sie [eine Musikerin] im Internet angeschaut, auf YouTube und dann ein paar Kommentare gelesen und ich war sehr schockiert darüber, was manche Leute dann da irgendwie schreiben. (CA06, Pos. 32).*

Die Plattformen, auf denen diese Angriffe stattfanden, waren breit gestreut. Genannt wurden unter anderem TikTok, Instagram, X (ehemals Twitter), Threads, Facebook, YouTube, E-Mails und die Kommentarspalten von Zeitungsartikeln (z. B. AM01, CA04). Die gemeldeten Vorfälle reichten von Morddrohungen (z. B. CA05, CA02, CA01) über direkte Beleidigungen bis hin zu hasserfüllten Kommentaren unter Beiträgen. Die Mehrheit der berichteten Vorfälle wurde von den Teilnehmenden als queerfeindlich, rassistisch oder nationalistisch eingeordnet. Personen, die sich durch sprachliche und nicht-sprachliche Codes öffentlich (z. B. über Statusmeldungen in sozialen Medien) als Teil von Gruppen identifiziert haben (z. B. FLINTA\*, Person mit Migrationshintergrund, s. Tab. 1), berichteten dabei besonders von Anfeindungen, die sie ohne öffentliche Bekanntgabe nicht erfahren haben. Das ist vor dem Hintergrund der Sichtbarkeit und Verfügbarkeit dieser Informationen nachvollziehbar.

Gleichsam wird deutlich, dass Personen des öffentlichen Lebens, die sich in bestimmten Bereichen engagieren (z. B. für geschlechtliche oder kulturelle Vielfalt), einem besonderen Risiko von Hass im Netz ausgesetzt sind. Dies findet sich auch in den Aussagen der interviewten Personen wieder. Die Betroffenheit durch Rassismus wurde häufig anhand entsprechender Merkmale (wie Name, Hautfarbe, Migrationsgeschichte) festgestellt, da Hass im Netz in diesen Fällen häufig gezielt durch sprachliche und nichtsprachliche, audiovisuelle Botschaften ausgedrückt wurde, die z. B. Beleidigungen oder inkorrekte Aussprache des Namens von Personen, inkorrekte Darstellungen des Herkunftslandes oder stereotype und stilisierte Darstellung von Kleidung von Personen beinhalteten. Die Bezugnahme auf solche Merkmale zur Rassifizierung und der Erstellung von Hassbotschaften wurde bereits in der bisherigen Forschung herausgestellt (z. B. Matamoros-Fernández & Farkas, 2021). Die Interviewstudie stellt dabei eine zunehmende Bandbreite in der Nutzung unterschiedlicher Plattformen (z. B. TikTok-Videos, Instagram-Stories) fest, die sich in Beständigkeit, Format und Ansprache von bisher vorrangig untersuchten Botschaften (zumeist Beiträge auf Twitter/X und Facebook) unterscheiden. So wurden vermehrt bild- und videobasierte Botschaften wahrgenommen und häufig indirekte und gegen Gruppen gerichtete Kommunikationsinhalte. Dies knüpft an die Darstellung von Matamoros-Fernández und Farkas (2021) an und bietet weitere empirische Belege für die zunehmende Verbreitung solcher Botschaften. Diese Erfahrungen von Hass im Netz sind mit verschiedenen Auswirkungen verbunden.

### **3.1 Auswirkungen von digitalem Hass und Hetze**

Die Auswirkungen von digitalem Hass und Hetze auf die Betroffenen wurden vielschichtig beschrieben und betreffen sowohl das Verhalten in sozialen Medien als auch den Alltag, die psychische und physische Gesundheit. Damit werden Aspekte der Bewältigung bzw. des Copings berührt, aber auch Konsequenzen im emotionalen Erleben. Einige der Teilnehmenden berichteten von einem Rückzug aus sozialen Medien als direkte Reaktion auf ihre Erfahrungen

mit digitalem Hass (AM06, AM04). „*Ich habe halt viele Voreinstellungen jetzt getroffen, dass es [homophobe Kommentare] aktuell nicht mehr ist. Also ich habe zwei Profile ganz deaktiviert, ein Profil nur für befreundete Personen freigeschaltet bei Instagram.* (AM04, Pos. 25)“

Für viele beeinflussten diese Erfahrungen auch ihr Sicherheitsempfinden im Alltag, was sich in verstärkten Sicherheitsvorkehrungen oder einem erhöhten Bewusstsein für potenzielle Bedrohungen – und damit einer eingeschränkten Teilhabe – äußerte (z. B. CA05, CA07, CA02). „*Ich hatte einen Tag danach [einer direkten Drohung] auch eine Veranstaltung an der Uni einer westdeutschen Stadt, wo ich meine Arbeit halt vorstellen wollte. Die musste ich dann auch auf online verlegen. (.) Also das schränkt mich schon ein und ja, das ist kein gutes Gefühl, vor allem mit dieser Ohnmacht.* (CA05, Pos. 30).“

Die psychischen Belastungen waren dabei besonders hervorzuheben. Mehrere Teilnehmende schilderten, dass sie die Erfahrungen nicht loslassen konnten oder sich Selbstvorwürfe machten (AM04). „*...ich habe mir dann halt Vorwürfe gemacht, warum ich diesen Beitrag halt erstellt habe und ob ich das hätte irgendwie umgehen können, dass Leute jetzt da auch so offen angegriffen werden von dieser Person.*“ (AM04, Pos. 47). Der Hass beschäftigte sie auch außerhalb der sozialen Medien und hatte Auswirkungen auf ihren Alltag (z. B. CA02, CA05). „*Und genau, also das macht sich dann schon konkret im Alltag bemerkbar, aber das sind auch Phasen, die, also ich würde sagen, das dauert dann immer so zwei bis drei Wochen an, (.) wo man natürlich dann schon irgendwie, keine Ahnung, wenn man so vier, fünf Kilo abgenommen hat, dann merkt man das ja auch schon so physisch, dass es gerade nicht so cool ist.*“ (CA02, Pos. 24) Bei einigen Personen verstärkte der Hass bereits bestehende psychische Probleme, wie beispielsweise Zwangsgedanken (AM04). Emotionen wie Wut oder Machtlosigkeit wurden ebenfalls häufig genannt (AM05, CA01). „*...aber gleichzeitig habe ich bei eigenen Hasskommentaren eher das Gefühl von Machtlosigkeit oder Wehrlosigkeit, was mich dann ja irgendwie fertig macht.*“ (CA01, Pos. 36).

Einige Teilnehmende erlebten die Belastung indirekt durch zusätzliche Arbeit, etwa im Umgang mit hasserfüllten Nachrichten oder bei der Moderation von Inhalten (AM01). *„Ich glaube, wenn man so dieses Ausmaß [an Hasskommentaren] hört, ich glaube, dann macht das schon eher was mit einem, weil man schon das Gefühl hat, okay, man muss schon auch vorsorgen. Also bevor ich jetzt [...] noch mehr in der Öffentlichkeit stand als vorher, war dann schon auch klar die Ansage, okay... muss schon gucken, was zum Beispiel, was du irgendwie im Internet von dir hast, was hast du jemals auf Twitter geliked oder was hast du auf Instagram geliked oder irgendwie noch auf deinem Profil.“* (AM01, Pos. 45).

Körperliche Symptome wie Schwindel (AM03) und Schlaflosigkeit (AM01, AM05, CA04) wurden ebenfalls beschrieben. *„Man ärgert sich maßlos und regt sich drüber auf und schläft vielleicht mal zwei Nächte nicht, weil man es immer wieder hin und her kaut.“* (AM05, Pos. 59). Einige Personen haben die Auswirkungen jedoch als weniger gravierend geschildert. Sie beschrieben einen Gewöhnungseffekt (und ein „Abstumpfen“), wonach die Belastungen durch Hass mit der Zeit abgenommen haben, während sie bei den ersten großen Hasserfahrungen noch deutlich stärker waren (AM01, CA01, CA07). *„Ja, also ich glaube, je öfter man sowas irgendwie erlebt, desto mehr ist man demgegenüber abgehärtet. Also so einfache Kommentare oder einfache Nachrichten, wo jetzt irgendwie Hass geschrieben wird, sind mir eigentlich jetzt relativ egal. Oder nicht egal, aber... Sie haben keine große Auswirkung auf mich.“* (AM01, Pos. 61). *„Man gewöhnt sich dran, das ist ein Gewöhnungseffekt. Und ich bin abgestumpft.“* (CA07, Pos. 56)

Die Erfahrungen verdeutlichen, dass die Auswirkungen von digitalem Hass individuell unterschiedlich sind und je nach Kontext und Intensität variieren. Personen, die von Rassismus betroffen sind, berichteten ebenfalls „Gewöhnungseffekte“, die Tendenz zu einer verstärkten Kontrolle des eigenen Verhaltens, der Einschränkung der Teilhabe sowie der längerfristigen Auseinandersetzung mit möglichen Reaktionen wurde vor allem von Personen des öffentlichen Lebens berichtet.

## 3.2 Umgang mit digitalem Hass und Hetze

Der Umgang mit digitalem Hass und Hetze variierte stark, je nach den individuell verfügbaren und bekannten Strategien der Betroffenen. Die Ansätze umfassten persönliche Bewältigungsversuche (sog. Coping) durch Rückzug und Vermeidung des digitalen Raumes, das Aufsuchen angenehmer Aktivitäten außerhalb des digitalen Raumes, aber auch die problemorientierte Bewältigung, die direkte Reaktionen auf die Hass-Inhalte in den sozialen Medien oder auch die Kommunikation mit anderen (z. B. Bystandern) umfasste (vgl. Lazarus & Folkman, 1991).

### 3.2.1 Persönliches Coping und Aufsuchen sozialer Unterstützung

Für einige Teilnehmende war der Rückzug aus den sozialen Medien nicht nur eine Auswirkung, sondern sie berichteten diese als Strategie, um mit digitalem Hass umzugehen (z. B. AM04, CA02). *„Und so ganz persönlich versuche ich dann, auch wenn mir das nicht immer gelingt, das einfach nicht anzuschauen. Also man kann sich ja selbst entscheiden, will ich jetzt auf YouTube gucken?“* (CA02, Pos. 24). Andere gaben an, dass sie nur noch Teile von sich selbst preisgaben, um sich vor weiteren Angriffen zu schützen (AM01). Eine weitere Strategie war das gezielte Aussuchen von Freund\*innen und Familie, die als Unterstützungsquelle dienen konnten (z. B. AM01, AM08, CA02). Insbesondere der Rückhalt durch enge Vertraute (auch durch Kolleg\*innen und andere Betroffene) schien für viele eine wichtige Quelle der emotionalen Unterstützung zu sein (CA02, CA01, AM08). *„(...) also bei Freunden halt besonders die, die selbst dann auch von solchen Nachrichten betroffen wären oder die im gleichen Verband sind wie ich und auch schon Erfahrungen einfach mit solchen Kommentaren haben.“* (AM08, Pos. 61). Darüber hinaus berichteten einige, dass sie Ablenkung durch Arbeiten oder durch andere Aktivitäten suchten, um sich von den belastenden Erfahrungen zu distanzieren. *„[Dann] gehe ich halt zum Betrieb und flex da mal acht Stunden. Und dann geht es mir auch wieder gut. Einfach, weil ich extrem fokussiert auf andere Dinge bin. Ansonsten*

*hilft rausgehen mit Dingen, wo man nicht so viel Zeit hat zu grübeln. Also jetzt so spazieren gehen, nicht so optimal. Segeln gehen, perfekt.“* (CA02, Pos. 40) Auch die Interaktion mit Haustieren wurde als hilfreiche Strategie genannt, um emotionale Entlastung zu finden (CA02).

### **3.2.2 Direkte Reaktion auf den digitalen Hass und Hetze**

Einige Betroffene entschieden sich, den digitalen Hass zu melden. Dabei äußerten sie jedoch Frustration über die langsame Reaktionszeit der Plattformen oder darüber, dass Hass nicht immer als solcher anerkannt wurde (z. B. AM04, CA02, CA05). „...*Meiner Erfahrung nach, bei Facebook muss man das halt immer wieder tun, bis die überhaupt mal reagieren, weil ich habe halt sehr oft auch die Erfahrung gemacht, dass dann kam, nö, XY-Kommentar ist nicht das und das, das ist noch im Rahmen unserer Richtlinie, ...*“ (AM04, Pos. 55) Manche wendeten sich auch an die Polizei, obwohl sie teilweise Unbehagen dabei empfanden, als marginalisierte Person mit den Behörden in Kontakt zu treten (AM04). „*Polizei, da bin ich immer so... da traue ich mich manchmal nicht so wirklich als queere Person, weil ich da auch immer bestimmte Ängste habe. Wie reagieren PolizistInnen einfach auf mich und was kommt da jetzt auf mich zu, wenn ich da als queere Person so hingehere und das offen melde.*“ (AM04, Pos. 57). Zudem berichteten einige, dass sie wenig Vertrauen in den Nutzen solcher Maßnahmen hatten und daher die Anzeige von Hass-Inhalten als wenig vielversprechend ansahen (AM07, AM06). „*Nee, Anzeige erstattet habe ich nicht, weil ich finde auch, dass das wahrscheinlich zu kompliziert wird und dass am Ende eh keiner angeklagt wird. Und meistens kommen die dann auch noch frei. Das ist dann auch noch irgendwie ziemlich erfolglos. Es verspricht halt keinen Erfolg.*“ (AM07, Pos. 59). Dennoch war das Anzeigen bei der Polizei für andere eine Maßnahme, die sie trotz wenig Erfolg als wichtig erachteten (CA02, CA05, CA07). „*Und dann ist das auch für mich insofern eine Genugtuung, dass ich dann sage, okay, ich habe jetzt diese Anzeige erstattet.*“ (CA05, Pos. 50)

Andere Teilnehmende suchten Hilfe bei speziellen Meldestellen für Hass im Netz, wobei einige eine Personen schnelle Reaktion/Hilfe erhielten (AM04/05), während andere angaben, nicht ausreichend informiert zu sein, um die Meldestellen effektiv zu nutzen (AM06). Auch wenn die Meldeverfahren zunehmend vereinfacht wurden, bestanden dennoch Bedenken oder Hürden in deren Nutzung (z. B. Unsicherheit bezüglich des Datenschutzes, der anzugebenden Daten). Dies waren in den Interviews Einzelfälle, sie werfen aber Fragen an mögliche Barrieren in der breiten Bevölkerung auf. Die selbstbestimmte Nutzung von Angeboten wie Meldestellen setzt gewisse digitale Kompetenzen voraus, da bekannt sein muss, wann und wie diese Stellen erreichbar sind, welche Informationen benötigt werden (und gespeichert werden sollten) usw. Auch das Blockieren von Nutzenden, die Hass verbreiten, wurde von einigen Personen genutzt, um sich vor weiteren Angriffen zu schützen. Dies war insbesondere bei persönlichen Kontakten oder Plattformen mit Direktkontakt (u. a. Instagram, WhatsApp) hilfreich und wurde auch als selbstwirksam wahrgenommen, da Personen sich somit selbst aktiv schützen konnten. Das stellte einen Mehrwert gegenüber KI-basierten Methoden dar, die automatisierte Filter anboten, deren Algorithmen aber nicht für alle verständlich oder zugänglich waren. Somit wurde den Nutzenden weniger Autonomie ermöglicht und es bestand eine Unsicherheit hinsichtlich der Arbeitsweise der KI-basierten Filter. Zudem wurde hier eine Ambivalenz für Personen deutlich, die eine Gewöhnung an Hass und ein Abstumpfen im Interview berichteten (darunter Personen mit Rassismuserfahrung und genderqueere Personen). Da sie häufige Betroffenheit berichteten und die Phänomene als Teil ihres Alltags einordneten, sahen sie mit der Zeit weniger Notwendigkeit, etwa im Sinne von Meldeverfahren tätig zu werden. Damit wurden derartige Angebote seltener in Anspruch genommen. Die wahrgenommene Präsenz von Hass im Netz und eine erlebte Gewöhnung sowie möglicherweise erfolglose Versuche der Meldung oder Beseitigung von Hassbotschaften könnten auch Anzeichen für eine wahrgenommene Hilflosigkeit von Betroffenen sowie die Normalisierung von Hassrede sein, die dazu beiträgt, dass systemische Diskriminierung verstärkt wird (Bilewicz & Soral, 2020).

### 3.2.3 Erhaltene (professionelle) Unterstützung

Die Inanspruchnahme externer, professioneller Unterstützung im Umgang mit Hass im Netz wurde selten berichtet. Einige Teilnehmende suchten Unterstützung bei spezialisierten Organisationen wie HateAid oder anderen Beratungs- und Informationsangeboten (AM04, CA04). *„Oder auch Beratungsstellen, die dann halt, die ich dann halt erreichen konnte zum Teil. Also mehr auch offline, wo ich dann halt mal reden konnte.“* (AM04, Pos. 61). Auch Beratungsstellen gegen rechte Gewalt wurden als Anlaufstelle genannt (z. B. CA02, CA05, CA07). *„Und was immens geholfen hat, ist das (gemeinnütziger Verein für Bildung, Beratung und lokale Erprobung in Bundesland), das sage ich ganz einfach. Also die Leute sind super ausgebildet, die können dir vom Kopf her, ähm, dich auffangen.“* (CA07, Pos. 60).

Die meisten Personen gaben an, das Thema selbständig, ohne externe Unterstützung zu bearbeiten, was für ihre Selbstwirksamkeit im Umgang mit Hassbotschaften förderlich erscheint. Eine weitere Hürde für die Nutzung professioneller Unterstützung war allerdings die Wahrnehmung, dass die eigene Erfahrung nicht schwerwiegend genug sei, um Hilfe in Anspruch zu nehmen. Einige berichteten, dass sie sich, z. B. beim Erfahren von indirektem Hass, nicht zur Zielgruppe der Beratungsstellen gehörig fühlten und daher auf externe Unterstützung verzichteten. *„Ja, ich glaube einmal würde mir jetzt da nicht so direkt eine Betroffenen-Stelle einfallen, wo ich mich melden, also wo ich das Gefühl habe, da ist jetzt so die Zielgruppe, da bin ich unter der Zielgruppe, dass ich mich da so, (...) niedrigschwellig melden könnte und darüber zu reden oder dadurch vielleicht [Entlastung] zu bekommen.“* (CA06, Pos. 54). Es ist daher im Einzelfall unklar, welche Schwelle der Betroffenheit als ausreichend für eine Meldung gesehen wird und wie die Relevanz einzuschätzen ist, insbesondere bei indirekter Betroffenheit, wenn Hass im Netz z. B. als Reaktion auf eigene Beiträge anderer Personen entgegengebracht wird (etwa eigenen „Followern“ im Fall von Personen des öffentlichen Lebens, die soziale Medien als Kommunikationsmedium nutzen). In diesen Fällen wurde oftmals Unsicherheit berichtet und Unwissen hinsichtlich möglicher

Beratungs- und Unterstützungsangebote (für die Person selbst sowie als Hilfestellung/leistung für die Betroffenen) deutlich. Analoge Prozesse werden in der Literatur bspw. für die Inanspruchnahme professioneller Hilfe bei psychischer Belastung beschrieben (z. B. Tomczyk et al., 2020). Dort spielen ebenfalls die wahrgenommene Symptomschwere, Fertigkeiten und Wissen, angemessene Hilfe zu erhalten sowie die Verfügbarkeit solcher Angebote eine zentrale Rolle für die tatsächliche Nutzung. Im Kontext von Hass im Netz lassen sich vergleichbare Faktoren beschreiben, da z. B. Melde- oder Beratungsstellen vorhanden und bekannt sein müssen, Personen sich kompetent im Umgang mit diesen fühlen und sie der Meinung sein müssen, dass ihre aktuelle Situation schwerwiegend genug für eine Meldung oder Beratung ist. In dieser Bedarfsanalyse lassen sich diese Prozesse skizzieren, anhand empirischer Erhebungen könnte darüber hinaus geprüft werden ist, ob sie sich auch quantitativ bestätigen und ob vergleichbare Barrieren (bspw. Stigmatisierung, geringe Selbstwirksamkeit) existieren.

### **3.3 Erwünschte Unterstützung**

Die Teilnehmenden wurden außerdem dazu befragt, welche Unterstützung für Betroffene von digitalem Hass sie sich wünschen würden, ausgehend von ihren eigenen Erfahrungen und Erwartungen, aber auch mit Blick auf die für sie relevanten Bezugsgruppen. Ein häufig genannter Punkt war der Wunsch nach besser geschultem Personal bei der Polizei. Dies wurde insbesondere in Hinblick auf den Umgang mit marginalisierten Gruppen betont, da einige Betroffene das Gefühl hatten, nicht ausreichend ernst genommen zu werden (AM04). *„dass die halt irgendwie auch mehr geschult werden, da einfach sensibler für bestimmte Themen zu sein.“* (AM04, Pos. 75). Diese Wahrnehmung könnte, aus Sicht der Betroffenen, auch den Unwillen verstärken, im Fall von Hass im Netz Meldestellen zu nutzen und z. B. Strafverfolgung in Erwägung zu ziehen. Wenngleich diese nur eine von vielen verhaltensbasierten Strategien darstellt, weist sie darauf hin, dass gerade Personen in Schnittstellenpositionen angemessen auf das Thema Hass im Netz vorbereitet sein müssen.

Auch die Forderung nach schnellerem Handeln von Verantwortlichen von Social-Media-Plattformen wurde mehrfach genannt (AM06, AM07). Dazu wurde auch erwähnt, dass auf weniger bekannten Plattformen und Foren, in denen bislang keine Moderation stattfindet und Hass weitgehend unbeachtet bleibt, dringend Handlungsbedarf besteht (AM04, AM07).

*„Oder ich sag mal, wenn gemeldet wird, dass das dann schneller runter geht oder so, aber... Ich finde, dass das an sich nicht viel bringt, weil ich sag mal, klar, dann kriegt man die Leute vielleicht von Facebook oder Instagram runter, die diese Posts machen, aber dann tauchen die halt irgendwo anders im Internet wieder auf.“* (AM07, Pos. 75)

Darüber hinaus wurde der Wunsch nach mehr Unterstützung und Vernetzung geäußert, um Betroffenen das Gefühl zu nehmen, mit ihren Erfahrungen allein zu sein (AM06). Wenngleich es hier bereits einige Initiativen und Verbände gibt (vgl. Kompetenznetzwerk gegen Hass im Netz), scheinen sie noch nicht alle Zielgruppen angemessen erreicht zu haben.

Ein weiteres Anliegen war die Bereitstellung niederschwelliger Unterstützung, beispielsweise bei der Meldung, Blockierung oder Anzeige von Hassinhalten (AM05, AM06, CA06, CA07). *„Oder generell auch, glaube ich, dass die Polizei da eher agieren würde oder sowas dann. Und dass man da eher Anlaufstellen kennen würde, beziehungsweise auch wüsste, dass das irgendwie erfolgreich sein könnte, wenn man da irgendwas meldet.“* (AM06, Pos. 81)

Auch dieses Zitat verdeutlicht einen Bedarf an Informationen, um Handlungssicherheit zu gewinnen, mehr Transparenz hinsichtlich der Ermittlungsarbeit im Bereich von Hass im Netz herzustellen und Personen darin zu bestärken, bestehende Angebote gut nutzen zu können. Die erwünschten Handlungsfelder stimmen mit den als wirksam erwiesenen Maßnahmen überein, aber auch hier zeigen sich deutlich mögliche Barrieren für die Umsetzung.

### **3.4 Empfehlungen für weitere Präventionsmaßnahmen**

Die Teilnehmenden äußerten zudem unterschiedliche Gedanken und Ansätze, wie Hass im Netz darüber hinaus präventiv adressiert werden könnte. Dabei wurde auch hervorgehoben,

dass es schwierig sei, konkrete Präventionsmaßnahmen zu formulieren. Manche Betroffene sahen das Phänomen des digitalen Hasses als nahezu unaufhaltsam oder unvermeidlich an (AM05, CA07). I: *...was denken Sie braucht es, um präventiv gegen Hass im Netz vorzugehen?* (CA07, Pos. 81) B: *Ich glaube, das werden wir nicht mehr aufhalten können. (...)* (CA07, Pos. 82). Dies bringt eine gewisse Resignation zum Ausdruck und lässt sich auch mit den Erfahrungen von Gewöhnung und Abstumpfen als Zeichen möglicher Hilflosigkeit interpretieren. Da Prävention aber nicht nur die Verhinderung des Auftretens umfasst, sondern z. B. auch die Verhinderung der Eskalation/Verschlimmerung bedeuten kann, konnten einige andere Ansatzpunkte identifiziert werden. Ein zentraler Vorschlag war eine stärkere Bildung und Aufklärung in Schulen, um ein Bewusstsein für die Problematik schon frühzeitig zu schaffen. *„Ich glaube, dass es vor allem hilfreich ist, zum Beispiel an Schulen oder auch in, naja Freizeitaktivitäten ist es auch sehr weit gefasst, aber dass eine gewisse Medienkompetenz beigebracht wird“* (AM03, Pos. 63) Gleichzeitig wurde betont, dass auch bei Erwachsenen mehr Aufklärungsarbeit durch Kampagnen oder Initiativen am Arbeitsplatz notwendig sei (z. B. AM03, AM04, AM06, CA07). *„Ja oder Workshops für Unternehmen anzubieten [zur Prävention], weil Mobbing findet ja auch oft am Arbeitsplatz leider statt. In Radios vielleicht darauf aufmerksam zu machen.“* (AM04, Pos. 89)

Eine weitere Forderung war die konsequentere Verfolgung von Hass im Netz, um die Ausübenden von Hass im Netz zur Verantwortung zu ziehen. (CA05, CA03). *„Also ich glaube, was, was (..), ähm, Leute würden mehr melden und auch mehr anzeigen, wenn die Erfolgsaussichten da höher wären, denke ich. Ähm. Also und da müsste quasi die Moderation der Plattform (.) ausgebaut werden und gleichzeitig (.) müsste die Polizei zumindest ein Minimum an Energie reinstecken in die Ermittlungen der Täterinnen oder Täter.“* (CA03, Pos. 78). Zudem wurde eine bessere und aktivere Moderation in sozialen Medien als entscheidender Faktor genannt (AM07, CA03). *„Also ich glaube, dass am meisten würde ich mir von den Plattformen wünschen, da so, ähm, je nach Plattform sind die Moderationstools sehr*

*frustrierend, also vor allem TikTok finde ich, ähm, finde ich es sehr also es ist einfach quasi von der Art, wie die Plattform aufgebaut ist, ähm, sehr schwer eine große Menge an Kommentaren zu moderieren, ...“ (CA03, Pos. 74). Die Diskussion um Präventionsmaßnahmen ging in einigen Fällen auch über technische Lösungen hinaus. Einige Teilnehmende betonten, dass eine gesamtgesellschaftliche Veränderung notwendig sei, um die Ursachen von Hass anzugehen (AM07, CA02, AM01, CA01).*

*„Ich würde sagen, es ist ein kollektives Problem, das wir haben und nicht ein individuelles Problem. Eine betroffene Person von Hasskommentaren ist nicht die einzige, die oder ist nicht verantwortlich oder schuld an diesen Hasskommentaren und sollte folglich auch nicht in der alleinigen Verantwortung stehen, dagegen vorgehen zu müssen. (..) Das wäre das Erste. Das heißt, es müsste eine kollektive Antwort sein“. (CA01, Pos. 58)*

Hierbei sahen sie die Politik in der Verantwortung, etwa durch die Förderung von interkulturellem Austausch, um Vorurteile und Hass abzubauen (AM07). *„... zum Beispiel Projekte macht, wo man aus vielen Kultur- und religiösen Kreisen die Leute irgendwie einlädt, dann irgendwas zu machen. Das kann alles Mögliche sein, aber dass man halt sozusagen diese Freundschaften, Kontakte knüpft, damit man so die, ich sag mal, den Hass abbauen kann, ...“ (AM07, Pos. 93)*

Gleichzeitig wurde jedoch auch die Sorge geäußert, dass Maßnahmen zur Eindämmung von Hassrede zu weitreichenden Einschränkungen oder Zensur führen könnten (AM06, CA01).

*„Es wird ja oft gefordert, dass man die Anonymität aufhebt. Ich finde aber, dass die Anonymität andere Vorteile mit sich bringt, die ich eigentlich verteidigen möchte. Also ich finde es eigentlich ganz gut, dass Menschen sich irgendwo halbwegs anonym auch äußern können“ (CA01, Pos. 56)*

In der Zusammenschau wurden die auch aus der Fachliteratur ableitbaren Empfehlungen wie Stärkung von Medienkompetenz, Empathie und positiver sozialer Identität (etwa durch Stärkung interkultureller Begegnung und Kooperation), Moderation in sozialen Medien und

technisch gestützte, niedrighschwellige Lösungen (z. B. Melde- und Steuerungsfunktionen in sozialen Medien, einfach verständliche Hinweise und Handlungsanleitungen für die Nutzung von Online-Meldestellen) von den Betroffenen genannt (vgl. Tomczyk, 2024). Zugleich wurden auch Herausforderungen benannt, wie der Schutz der Anonymität, die Berücksichtigung Vielfalt (z. B. das Vorhalten von vielsprachigen und barrierearmen Dokumenten und Informationen) und die Stärkung des Vertrauens und der (Kommunikations-)Kompetenzen in Schnittstellenpositionen (z. B. Polizei, Beratungs- oder Meldestelle). Diese Punkte gelten insbesondere für von Rassismus betroffene Personen, da sie vielfach entsprechende barrierearme Kommunikation vermissen oder von Enttäuschung und Vertrauensverlust berichten, z. B. in der Kontaktaufnahme mit der Polizei oder mit Beratungsstellen.

### **3.5 Einschätzungen zu bestehenden Maßnahmen**

Teilnehmenden wurden darüber hinaus zu drei konkreten Maßnahmen befragt, die im Einklang mit den zu entwickelnden und zu prüfenden Angeboten im Land Niedersachsen stehen: Einem Notfallbutton für Hass im Netz, wie er in Frankreich bereits Pflicht ist, einer App, die mittels künstlicher Intelligenz (KI) die Moderation von Hasskommentaren unterstützt und Workshops für Betroffene, der Empowerment und Autonomie fördern soll und z. B. zur Umsetzung von Gegenrede und dem Aufbau positiv wirksamer Online-Gemeinschaften befähigen soll.

#### **3.5.1 Notfallbuttons**

Ein Notfallbutton wurde von den meisten Teilnehmenden grundsätzlich positiv aufgenommen (z. B. CA05, CA04, CA03, AM04). „*Fände ich super. Ja. Wann kommt es [Notfallbutton]?*“ (CA04, Pos. 88). Gleichzeitig wurde jedoch betont, dass ein solches Instrument nur dann wirksam sei, wenn es gut besetzt und mit ausreichend Ressourcen ausgestattet ist (AM02), um schnelle und verlässliche Hilfe zu leisten. Darüber hinaus wurde angemerkt, dass der Notfallbutton sensibel für verschiedene Themenbereiche gestaltet sein

müsste, um den individuellen Bedürfnissen der Betroffenen gerecht zu werden (AM01). „*Ich kann mir gut vorstellen, dass es [der Kontakt zu Unterstützungsangeboten] schwierig wäre, das genau genug zu spezifizieren, dass man dann überhaupt versteht, was dann der genaue Fall ist. Aber wenn man das schaffen würde, dann finde ich es gut.*“ (AM01, Pos. 107)

### **3.5.2 Einsatz von Künstlicher Intelligenz (KI)**

Der Einsatz von Künstlicher Intelligenz im Umgang mit digitalem Hass wurde von vielen Teilnehmenden mit Skepsis betrachtet. Dabei wurden Fragen aufgeworfen, beispielsweise in Bezug auf die Datengrundlage, mit der KI-Systeme trainiert werden (AM01, CA02).

„*KI ist ja auch immer noch biased und hat ja auch zum Beispiel rassistischen Bias oder einen misogynen Bias oder whatever und ich glaube, da würde man das Problem einfach nur auslagern und man hätte dann nicht mehr so ein, also man könnte dann ja nicht mehr irgendwie auf den Finger zeigen und sagen, ja, eigentlich ist es Aufgabe der Plattformen, sowas zu klären, sondern dann ist halt die KI schuld*“ (AM01, Pos. 109)

Auch die Zweifel an der Effektivität solcher Systeme wurden deutlich, da einige unsicher waren, wie gut KI in der Lage sei, Hassrede präzise zu erkennen und zu klassifizieren (z. B. AM06, CA01) „*Also mein Eindruck ist, dass bisher ein bisschen überschätzt wird, wie (..) welche Feinheiten diese KI also diese künstlichen Intelligenzen verarbeiten können.*“ (CA01, Pos. 66)

Unter der Voraussetzung ausreichender Kontrollmöglichkeiten äußerten einige jedoch Akzeptanz für den Einsatz von KI. Besonders in strafrechtlichen Kontexten schien die Nutzung solcher Technologien für viele sinnvoll (AM02, AM04). „*Was schon interessant ist, ist natürlich, dass man dann die darauf sucht und dann direkt eine Strafanzeige stellt und so.*“ (AM02, Pos. 86)

Diese Bedenken und Potenziale, die in der Nutzung von KI gesehen werden, spiegeln ebenfalls die Befunde der aktuellen Forschung wider, die einerseits eine Teilautomatisierung

der Erkennung von Hass im Netz und damit verbundene Handlungsempfehlungen herausstellt (z. B. Meldefunktionen, Formulierungsvorschläge für Gegenrede) und andererseits Herausforderungen in der KI-basierten Erkennung von Hass im Netz sieht, wenn z. B. Soziolekt oder Jugendsprache als Hasssprache fehlklassifiziert werden (vgl. Tomczyk, 2024).

Dies ist besonders für das Thema Rassismus relevant, da oftmals Begriffe aus Sprach- und Kulturräumen von Minderheiten z. B. in sozialen Medien aufgegriffen und durch eine zumeist weiße, affluente Mehrheit rekontextualisiert und genutzt werden, sodass sich daraus dominierende Einflüsse auf mediale Diskurse ergeben (vgl. Matamoros-Fernández & Farkas, 2021). Die vorliegende Bedarfsanalyse mag auch nicht frei von solchen Einflüssen sein, da sie von weißen Forschenden durchgeführt wurde, die zwar Überschneidungen mit den Merkmalen der untersuchten Gruppen besitzen (z. B. Migrationshintergrund, LGBTQIA\* und FLINTA\*), aber aus und in einem System agieren, in dem diese Gruppen Minoritäten darstellen. Für die Nutzung von sprachlichen und nichtsprachlichen Codes lassen sich unterschiedliche Nutzungskontexte und Bedeutungsräume zwischen den Minoritäten und der Majorität feststellen, sodass eine Minderheit möglicherweise Bezeichnungen (z. B. queer/genderqueer) als Identitätssymbol und Kennzeichen verwendet, während die Mehrheit sie als Zuschreibung ablehnt. Wenn KI-basierte Lösungen dann aus Sicht der Mehrheit entwickelt werden, kann dies dazu führen, dass Inhalte, die darauf rekurieren, fälschlicherweise als Hass klassifiziert werden, von den betroffenen Gemeinschaften aber gar nicht so erlebt werden. Wenn diese Erkennung dann in der Folge mit Sanktionen (z. B. Ausschluss aus digitalen Gemeinschaften, Meldeverfahren) verbunden wird, kann dies zu einer weiteren Benachteiligung von Minderheiten beitragen.

### **3.5.3 Workshops und strukturierte Bildungsangebote**

Workshops wurden von den Teilnehmenden grundsätzlich als gute Idee bewertet. Allerdings wurde häufig betont, dass die Verantwortung für den Umgang mit Hass im Netz nicht allein bei den Betroffenen liegen sollte (AM01, CA01, CA02).

*„Also ich finde es immer grundsätzlich falsch, wenn Menschen, die von etwas betroffen sind, da noch mehr Zeit reinstecken müssen. Also ich hätte lieber Workshops zu kritischer Männlichkeit als für Betroffene sexualisierte Gewalt so.“ (CA02, Pos. 80)*

Während Workshops, die auf Empowerment und Kompetenzentwicklung ausgerichtet sind und eine positive und stärkende Erfahrung bieten sollen, begrüßt wurden, war es den Teilnehmenden wichtig, dass der Fokus nicht ausschließlich darauf liegt, sich selbst zu schützen oder zu verteidigen (CA02, CA03).

*„Ich fände es gut, wenn er quasi nicht nur die Komponente hätte von, wie gehst du damit um, sondern auch vielleicht aus den TeilnehmerInnen irgendwie ein Netzwerk entsteht, auf das man mal zurückgreifen kann, auch in anderen Kontexten, also wenn quasi noch ein Benefit mit dabei ist, der sich nicht nur damit beschäftigt, sondern der einem sonst im Leben ein bisschen hilft und Nachteile, die in einer patriarchalen Gesellschaft entstehen, ein bisschen mit ausgleicht.“ (CA02, Pos. 82)*

Dabei wurde deutlich, dass vor allem solche Angebote präferiert werden, die allgemeine Kompetenzen stärken (z. B. Sozialkompetenz, Medienkompetenz) und weniger spezifisch z. B. nur auf den Umgang mit Hass im Netz ausgerichtet sind. Das stimmt inhaltlich mit Forschungsergebnissen überein, die bislang weniger belastbar für spezifische Interventionen wie Workshops zur Gestaltung von Gegenrede sind (z. B. Blaya, 2019), aber eine breitere Basis für komplexe Interventionen bieten, die z. B. Empathie, Medienkompetenz und Kommunikationsfertigkeiten stärken sollen (Tomczyk, 2024). Wichtig ist, dass solche Angebote in einen entsprechenden Rahmen eingebettet werden sollten, der neben individuellen auf strukturelle Faktoren berücksichtigt. Das bedeutet, dass z. B. neben Workshops auch entsprechende Anlauf- und Meldestellen vorliegen sollten und technische Möglichkeiten wie Notfallbuttons existieren sollten, die die Gestaltung einer weitgehend hassfreien Kommunikation sowie einen angemessenen Umgang mit Hass im Netz begünstigen.

#### **4 Diskussion und Zusammenfassung der Ergebnisse**

Die vorliegende Bedarfsanalyse zum Phänomen des digitalen Hasses/der digitalen Hetze mit Schwerpunkt Rassismus hat mittels narrativer Interviews mit Personen, die aufgrund ihrer sexuellen Orientierung, Geschlechtsidentität oder potenziell rassifizierter Merkmale Hass im Netz erlebt haben, exploriert, welche Erfahrungen bei ihnen vorliegen, welche Auswirkungen erlebt oder wahrgenommen wurden und welche Strategien im Umgang mit Hass im Netz eingesetzt sowie als hilfreich empfunden wurden. Zudem wurden Aspekte intersektionaler Betroffenheit sowie aktuelle Entwicklungen im Bereich der Prävention und Intervention bei Hass im Netz mit den Teilnehmenden diskutiert und von ihnen bewertet. Im Ergebnis lassen sich einige Gemeinsamkeiten der Gruppen sowie Besonderheiten feststellen, die Hinweise für die Gestaltung von Präventionsansätzen bieten.

In der Gruppe der (intersektional) von Rassismus betroffenen Personen sowie bei Personen, die aufgrund ihrer sexuellen Orientierung Anfeindung erlebt haben (LGBTQIA\*), wurden wiederholt negative Erfahrungen mit externen Unterstützungsangeboten und eine wahrgenommene Gewöhnung an Hass im Netz berichtet, die eher mit vermeidendem und emotionsorientierten Coping und weniger problemorientiertem Coping assoziiert waren (Lazarus & Folkman, 1991). Dies kann eine Hürde für die Nutzung von Unterstützungsangeboten und damit ein Faktor anhaltender Belastung für die Betroffenen sein.

Im Zuge der Rekrutierung hat sich überdies die Kategorie der Personen des öffentlichen Lebens als weitere relevante Dimension für die Analyse herausgestellt, da viele der teilnehmenden Personen sich auch durch politische Partizipation und Aktivismus auszeichnen (z. B. für Gleichberechtigung der Geschlechter oder die Integration von Menschen mit Fluchterfahrung), was z. B. in sozialen Medien zu einer gewissen Reichweite und Bekanntheit geführt hat. Da diese Eigenschaft über unterschiedliche Gruppen hinweg festgestellt werden konnte, kann die Dimension auch in einer vergleichenden Analyse berücksichtigt werden.

Alle Zielgruppen berichten von direkten und indirekten Erfahrungen von Hass im Netz, dabei war die Häufigkeit und Frequenz für Personen des öffentlichen Lebens deutlich höher als für Privatpersonen. Art und Umfang richtete sich wesentlich an den Themen und Profilen aus, mit denen die Personen öffentlich erkennbar waren, z. B. Parteizugehörigkeit, sexuelle Orientierung, Nationalität. Personen aus der LGBTQIA\*-Gruppe sowie Personen mit potenziell rassifizierten Merkmalen berichteten dabei über besonders häufiges Erleben von Hass im Netz sowie Alltagsdiskriminierung, bei dem diese beiden Merkmale als ausschlaggebend wahrgenommen wurden, auch wenn weitere der hier untersuchten Merkmale vorlagen. Da im Fall einer intersektionalen Betroffenheit auch ein politischer Aktivismus berichtet wurde, kann dies auch mit der verstärkten Exposition konfundiert sein, die mit der höheren Bekanntheit in Verbindung steht.

Interessant ist, dass alle Gruppen, auch Privatpersonen, eine große Bandbreite an Auswirkungen berichten, dabei wurden somatische, psychische, soziale und Auswirkungen beschrieben, wie sie auch in der Literatur zu finden sind (z. B. Buelga et al., 2022; Jadambaa et al., 2019; John et al., 2018; Li et al., 2022). Die Häufigkeit des Erlebens von Hass im Netz konnte jedoch nicht eindeutig mit der Intensität der Auswirkungen in Verbindung gebracht werden, unter anderem, da bei häufig betroffenen Gruppen Rationalisierung zu beobachten war (d. h. Gewöhnungseffekte, Beschreibung von Hass im Netz als erwartbarer Teil des Alltags für Personen des öffentlichen Lebens). Das kann als eine Bewältigungsstrategie beschrieben werden und findet sich auch im Modell des Moral Disengagement wieder, wo es eine Strategie darstellt, um von einer Verhaltensänderung Abstand zu nehmen (z. B. Lo Cricchio et al., 2021; Nocera et al., 2022; Runions & Bak, 2015). Moral Disengagement sollte daher als motivationaler Faktor nicht nur für die Entstehung und Aufrechterhaltung von Hass im Netz, sondern auch den Rückzug bzw. die Vermeidung und das Ausbleiben von stärker problemorientierter Bewältigung bei Betroffenen mehr Beachtung finden.

FLINTA\*-Personen beschreiben häufiger Beleidigungen bezogen auf das Geschlecht, auch wenn sie keine geschlechtsspezifischen Inhalte teilten. Hier scheint eine Identifizierung als FLINTA\* ausreichend zu sein. Männliche Personen erlebten weniger geschlechtsbezogene Beleidigungen, hatten hingegen häufiger Schwierigkeiten die emotionalen Auswirkungen von Hass im Netz zu beschreiben. Da die daran anschließenden Bewältigungsstrategien weniger spezifisch waren, wäre eine wichtige Anschlussfrage, wie dies in Zusammenhang mit dem Erfolg von Bewältigung steht und ob dies mit der geringeren Inanspruchnahme von Unterstützung in Zusammenhang steht, die von männlich gelesenen Personen berichtet wird. Auch hier scheinen Parallelen zu anderen Kontexten der Hilfesuche oder Inanspruchnahme von unterstützenden Angeboten auf, wie etwa dem Umgang mit psychischen Belastungen (vgl. Tomczyk et al., 2020). Der Zusammenhang von Geschlecht (d. h. Gender) und Cybermobbing und -viktimisierung ist bereits gut erforscht (z. B. Wright & Wachs, 2020), für den Umgang mit Hate Speech, insbesondere im Kontext intersektionaler Betroffenheit, besteht noch weiterer Forschungsbedarf.

In Bezug auf Präventions- und Interventionsmöglichkeiten sind sich die interviewten Personen einig, dass eine Kombination aus kompetenzorientierten, individuellen und gruppenbezogenen Angeboten (z. B. Workshops) sowie verhältnisorientierten Maßnahmen (z. B. Notfallbuttons in sozialen Medien, automatisierte Verweise auf Meldestellen) zielführend erscheinen, um Hass im Netz vorbeugen und besser damit umgehen zu können. Die Umsetzung von Workshops alleine wurde kritisch gesehen, da der Eindruck entstehen könnte, dass es allein Aufgabe der betroffenen Personen sei, sich auch noch um ihren Schutz zu kümmern und damit zusätzlich Arbeit zu leisten. Hier wurde der Wunsch geäußert, zumindest auch Anbieter von sozialen Medien stärker in die Pflicht zu nehmen und Personal an Schnittstellen (z. B. Beratung, Polizei) besser zu schulen, um sensibel mit den Anliegen umgehen zu können, gerade wenn sie aus marginalisierten Gruppen berichtet werden.

In Angeboten, die künstliche Intelligenz nutzen (z. B. Textvorlagen für Gegenrede, (teil-)automatisierte Erkennung und Filterung von Hassbotschaften) wurde zwar Potenziale gesehen, allerdings war die Mehrheit der interviewten Personen skeptisch hinsichtlich der Qualität der Ergebnisse (etwa aufgrund systematischer Verzerrung/Bias bei der sprachbasierten Analyse) und eher zurückhaltend in der Bereitschaft zur Nutzung solcher Anwendungen. Dies stellt eine bekannte Herausforderung in der KI-basierten Innovation dar (vgl. Abarna et al., 2022; Khurana et al., 2023; Poletto et al., 2021; Sarsam et al., 2020), weist aber auch hier auf die Notwendigkeit hin, partizipativ KI-Anwendungen zu gestalten und frühzeitig Akzeptanz und Alltagsnutzung in den Blick zu nehmen. Für Personen, die von Rassismus betroffen sind, kann so beispielsweise gemeinsam daran gearbeitet werden, ihre Alltagskontexte der Nutzung und ihre sprachlichen und nichtsprachlichen Codes (z. B. in der Erkennung von Hassbotschaften) besser zu berücksichtigen, die andernfalls fehlklassifiziert werden oder zu einer geringen Nutzung führen könnten. Dies wurde in der bisherigen Forschung selten umgesetzt, da viele Anwendungen von mehrheitlich weißen Personen in westlich-industrialisierten Ländern entwickelt wurden, die die Erfahrungsexpertise der betroffenen Gruppen nicht teilen (Matamoros-Fernández & Farkas, 2021).

Für die Gruppe der von Rassismus betroffenen Personen lassen sich darüber hinaus weitere Erkenntnisse festhalten. Nicht nur sticht hier ebenfalls der Aspekt der Gewöhnung hervor, aufgrund des gehäuften Erlebens von Hass im Alltag und im Netz, sondern ebenfalls eine geringere Nutzung von Anlaufstellen und Meldestellen von Hass im Netz. Hier wurden eine mangelnde Sensibilität des Personals sowie negative zurückliegende Erfahrungen, z. B. mit der Polizei, als Barrieren ins Feld geführt. Wenngleich es sich um wenige Fälle handelt, da die Analyse qualitativ umgesetzt worden ist, wirft dies die Frage auf, ob das Personal ausreichend sensibilisiert ist, um mit rassistisch geprägtem Hass im Netz gut umgehen zu können. Dies kann etwa durch Schulungen des Personals unterstützt werden.

Weniger einfach zu bearbeiten sind die für diese Gruppe inhärenten Barrieren, die etwa eine Sprachferne mit sich bringen kann, was ebenfalls berichtet wurde. Aufgrund eingeschränkter Sprachkenntnisse ist die Kontaktaufnahme zu Melde- und Anlaufstellen mitunter erschwert, wenn keine adäquaten mehr- oder nichtsprachlichen Angebote zur Verfügung stehen. Einerseits könnten KI-gestützte Angebote gerade für diese Gruppe viele Vorteile bieten, da sie die Kommunikationsbarrieren etwa durch automatisierte Übersetzungen reduzieren können. Andererseits sind sie besonders herausfordernd, da sich eine mögliche Benachteiligung durch KI bei systemischer Diskriminierung schwerlich ausschließen lässt. Hinzu kommen Vorbehalte zur Nutzung solcher Angebote, zumal wenn keine ausreichenden Hilfestellungen zur Nutzung zur Verfügung stehen (z. B. Anleitungen und Erläuterungen), die für die Zielgruppe verständlich und nutzbar sind. Denn die erforderlichen digitalen Kompetenzen müssen ebenfalls mitgedacht werden, um eine Verschärfung der Diskriminierung und eine digitale Spaltung (z. B. fehlender Zugang aufgrund fehlender Kenntnisse) zu vermeiden.

Zuletzt ist für diese Gruppe sowie die LGBTQIA\*-Gruppe eine große Ambivalenz in der Kommunikation festzustellen. Beide Gruppen berichten die häufige Nutzung von Zeichen und Codes, die in Zusammenhang mit ihrer Gruppenzugehörigkeit und sozialen Identität stehen. Dazu gehören z. B. Flaggen (Landesflaggen, Regenbogenflaggen), Frisuren oder Kleidungsstücke, die mit Kulturen oder Praktiken in Verbindung gebracht werden. Für beide Gruppen erfüllt dies wichtige Funktionen neben der sozialen Identität, wie etwa das Erleben eigener Wirksamkeit oder den Ausdruck von Solidarität. Gleichsam sind es diese Zeichen, die von anderen Personen als Anlass von Rassifizierung und gruppenbezogener Diskriminierung (z. B. aufgrund sexueller Orientierung) durch andere gelten können. Sie sind nicht nur für Einzelpersonen leicht zu erkennen, sondern auch für Algorithmen, dank ihrer oftmals öffentlich sichtbaren Position. Dies erleichtert auch die Nutzung von KI-Anwendungen zur Gestaltung und Verbreitung von Hassbotschaften und die Nutzung sog. Dogwhistles, die solche Merkmale als Anlass nehmen können und daran Botschaften knüpfen, die z. B. durch die Verwendung

von Emojis, Ironie oder überzeichneter Sprache nicht eindeutig als Hassbotschaften klassifiziert werden können, von den Adressaten aber klar verstanden werden können (Lindgren et al., 2024). Im Sinne der Dual Use-Problematik bietet der Fortschritt KI-basierter Systeme zur Analyse von Hassbotschaften auch die Möglichkeit, entsprechende Botschaften gestalten und gezielt verbreiten zu können. Die weitere Entwicklung in diesem Feld erfordert daher eine sorgfältige Abwägung möglicher Chancen und Risiken sowie eine stärker partizipative und lebensweltorientierte Entwicklung, um die Akzeptanz und Nutzung zu erhöhen und mögliche Schäden zu minimieren.

## 5 Literaturverzeichnis

- Abarna, S., Sheeba, J. I., Jayasrilakshmi, S., & Devaneyan, S. P. (2022). Identification of cyber harassment and intention of target users on social media platforms. *Engineering Applications of Artificial Intelligence*, *115*, 105283.
- Alhaboby, Z. A., Barnes, J., Evans, H., & Short, E. (2019). Cyber-victimization of people with chronic conditions and disabilities: A systematic review of scope and impact. *Trauma, Violence, & Abuse*, *20*(3), 398–415.
- Bilewicz, M., & Soral, W. (2020). Hate Speech Epidemic. The Dynamic Effects of Derogatory Language on Intergroup Relations and Political Radicalization. *Political Psychology*, *41*(S1), 3–33. <https://doi.org/10.1111/pops.12670>
- Blaya, C. (2019). Cyberhate: A review and content analysis of intervention strategies. *Aggression and Violent Behavior*, *45*, 163-172.
- Buelga, S., Cava, M.-J., Ruiz, D. M., & Ortega-Barón, J. (2022). Cyberbullying and suicidal behavior in adolescent students: A systematic review. *Revista de Educación*, *397*, 43–66.
- Choi, M. (2016). A concept analysis of digital citizenship for democratic citizenship education in the internet age. *Theory & Research in Social Education*, *44*(4), 565–607.
- Doty, J. L., Girón, K., Mehari, K. R., Sharma, D., Smith, S. J., Su, Y.-W., Ma, X., Rijo, D., & Rousso, B. (2021). The dosage, context, and modality of interventions to prevent cyberbullying perpetration and victimization: A systematic review. *Prevention Science*, 1–15.
- Garland, J., Ghazi-Zahedi, K., Young, J.-G., Hébert-Dufresne, L., & Galesic, M. (2022). Impact and dynamics of hate and counter speech online. *EPJ Data Science*, *11*(1), 3.
- Hintz, A., Dencik, L., & Wahl-Jorgensen, K. (2017). Digital citizenship and surveillance| digital citizenship and surveillance society—Introduction. *International Journal of*

*Communication*, 11, 9.

- Horstmann, S., & Bolte, G. (2022). Geschlechtliche Vielfalt erfassen. *Pflegezeitschrift*, 75(12), 52–54. <https://doi.org/10.1007/s41906-022-1960-y>
- Hutson, E., Kelly, S., & Militello, L. K. (2018). Systematic review of cyberbullying interventions for youth and parents with implications for evidence-based practice. *Worldviews on Evidence-Based Nursing*, 15(1), 72–79.
- Jadambaa, A., Thomas, H. J., Scott, J. G., Graves, N., Brain, D., & Pacella, R. (2019). Prevalence of traditional bullying and cyberbullying among children and adolescents in Australia: A systematic review and meta-analysis. *Australian & New Zealand Journal of Psychiatry*, 53(9), 878–888.
- Jia, Y., & Schumann, S. (2023). Tackling Hate Speech Online: The Effect of Counter-speech on Subsequent Bystander Reactions. *Preprint (without Peer Review)*. <https://doi.org/10.33767/osf.io/9jmza>
- John, A., Glendenning, A. C., Marchant, A., Montgomery, P., Stewart, A., Wood, S., Lloyd, K., & Hawton, K. (2018). Self-harm, suicidal behaviours, and cyberbullying in children and young people: Systematic review. *Journal of Medical Internet Research*, 20(4), e9044.
- Khurana, D., Koli, A., Khatter, K., & Singh, S. (2023). Natural language processing: State of the art, current trends and challenges. *Multimedia Tools and Applications*, 82(3), 3713–3744.
- Kuckartz, U. (2018). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung* (4. Auflage). Beltz Juventa.
- Kuckartz, U., & Rädiker, S. (2022). *Qualitative Inhaltsanalyse: Methoden, Praxis, Computerunterstützung: Grundlagentexte Methoden* (5th ed.). Beltz Juventa.
- Kuckartz, U., & Rädiker, S. (2024). *Fokussierte Interviewanalyse mit MAXQDA: Schritt für Schritt*. Springer Fachmedien Wiesbaden. <https://doi.org/10.1007/978-3-658-40212->

- Lazarus, R. S., & Folkman, S. (1991). The concept of coping. In: A. Monat & R. S. Lazarus, Eds. *Stress and coping: An anthology* (pp. 189-206). Columbia University Press.
- Leiner, D. J. (2019). *SoSci Survey* (Version 3.1.06) [Computer software].  
<https://www.soscisurvey.de>
- Lenzner, T., Schick, L., Hadler, P., Behnert, J., Finzer, M., & Neuert, C. (2022). FReDA-W2b – Fragen zu den Themen Fruchtbarkeit von Frauen, affektives Berührungsverhalten, sexuelle Orientierung und transnationale Familien: Kognitiver Online-Pretest. *GESIS Projektbericht*. <https://doi.org/10.17173/PRETEST121>
- Li, C., Wang, P., Martin-Moratinos, M., Bella-Fernández, M., & Blasco-Fontecilla, H. (2022). Traditional bullying and cyberbullying in the digital age and its associated mental health problems in children and adolescents: A meta-analysis. *European Child & Adolescent Psychiatry*, 1–15.
- Lindgren, E., Rönnerstrand, B., Rettenegger, G., Breitholtz, E., Cooper, R., & Sayeed, A. (2024). Coded Appeals and Political Gains: Exploring the Impact of Racial Dogwhistles on Political Support. *Journalism & Mass Communication Quarterly*, 10776990241280373. <https://doi.org/10.1177/10776990241280373>
- Lo Cricchio, M. G., Garcia-Poole, C., te Brinke, L. W., Bianchi, D., & Menesini, E. (2021). Moral disengagement and cyberbullying involvement: A systematic review. *European Journal of Developmental Psychology*, 18(2), 271–311.
- Matamoros-Fernández, A., & Farkas, J. (2021). Racism, Hate Speech, and Social Media: A Systematic Review and Critique. *Television & New Media*, 22(2), 205-224.  
<https://doi.org/10.1177/1527476420982230>
- Meta. (2024). *Hassrede | Transparency Center*. Meta Transparency Center.  
<https://transparency.meta.com/de-de/policies/community-standards/hate-speech/>
- Mishna, F., Cook, C., Saini, M., Wu, M.-J., & MacFadden, R. (2009). Interventions for

- children, youth, and parents to prevent and reduce cyber abuse. *Campbell Systematic Reviews*, 5(1), i–54.
- Nocera, T. R., Dahlen, E. R., Poor, A., Strowd, J., Dortch, A., & Van Overloop, E. C. (2022). Moral disengagement mechanisms predict cyber aggression among emerging adults. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 16(1), 1.
- Polanin, J. R., Espelage, D. L., Grotzinger, J. K., Ingram, K., Michaelson, L., Spinney, E., Valido, A., Sheikh, A. E., Torgal, C., & Robinson, L. (2022). A systematic review and meta-analysis of interventions to decrease cyberbullying perpetration and victimization. *Prevention Science*, 23(3), 439–454.
- Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research*, 42, 533–544.
- Poletto, F., Basile, V., Sanguinetti, M., Bosco, C., & Patti, V. (2021). Resources and benchmark corpora for hate speech detection: A systematic review. *Language Resources and Evaluation*, 55, 477–523.
- Ruland, M., Link, S., Malina, A., & Fahrenholz, M. (2022). *Methodenbericht: Aufbau eines Online-Access-Panels und Erstbefragung im Nationalen Diskriminierungs- und Rassismusmonitor (NaDiRa)* (p. 64). Institut für angewandte Sozialwissenschaft GmbH (infas).  
[https://www.rassismusmonitor.de/fileadmin/user\\_upload/NaDiRa/Rassismus\\_Symptome/Infas\\_Methodenbericht\\_Fragebogen.pdf](https://www.rassismusmonitor.de/fileadmin/user_upload/NaDiRa/Rassismus_Symptome/Infas_Methodenbericht_Fragebogen.pdf)
- Runions, K. C., & Bak, M. (2015). Online moral disengagement, cyberbullying, and cyber-aggression. *Cyberpsychology, Behavior, and Social Networking*, 18(7), 400–405.
- Sarsam, S. M., Al-Samarraie, H., Alzahrani, A. I., & Wright, B. (2020). Sarcasm detection

- using machine learning algorithms in Twitter: A systematic review. *International Journal of Market Research*, 62(5), 578–598.
- Schütze, F. (1983). Biographieforschung und narratives Interview. *Neue Praxis*, 13(3), 283–293.
- Silvia Casale, Guilia Fiovaranti, & Scott Caplan. (2015). Online Disinhibition. *Journal of Media Psychology*, 27(4), 170–177.
- Tomczyk, S. (2024). *Gutachten zum Phänomen des digitalen Hasses und der digitalen Hetze in Niedersachsen: Bedingungen, Erscheinungsformen und Ansätze der Prävention und Intervention*.  
<https://www.mj.niedersachsen.de/startseite/aktuelles/presseinformationen/gemeinsam-gegen-digitalen-hass-230489.html>
- Tomczyk, S., Schomerus, G., Stolzenburg, S., Muehlan, H., & Schmidt, S. (2020). Ready, willing and able? An investigation of the theory of planned behaviour in help-seeking for a community sample with current untreated depressive symptoms. *Prevention Science*, 21(6), 749-760.
- Wachs, S., Bilz, L., Wettstein, A., Wright, M. F., Krause, N., Ballaschk, C., & Kansok-Dusche, J. (2022). The online hate speech cycle of violence: Moderating effects of moral disengagement and empathy in the victim-to-perpetrator relationship. *Cyberpsychology, Behavior, and Social Networking*, 25(4), 223–229.
- Was ist Hass im Netz?* (2024). Kompetenznetzwerk gegen “Hass im Netz.”  
<https://kompetenznetzwerk-hass-im-netz.de/was-ist-hass-im-netz/>
- Wright, M. F., & Wachs, S. (2020). Adolescents’ cyber victimization: The influence of technologies, gender, and gender stereotype traits. *International Journal of Environmental Research and Public Health*, 17(4), 1293.