

Hass im Netz

—

Wahrnehmung, Verbreitung und Zivilisierung von Hate Speech

MARTIN EMMER

Geboren 1969 in München, Professor für Publizistik- und Kommunikationswissenschaft an der Freien Universität Berlin, Gründungsdirektor des Weizenbaum-Instituts für die vernetzte Gesellschaft.

Die vielfältigen Kommunikationsmöglichkeiten, die das Internet bereitstellt, wurden lange Zeit vor allem als eine Chance diskutiert, Bürgern mehr Teilhabechancen an öffentlichen Debatten und am politischen Prozess zu ermöglichen. Inzwischen hat der Enthusiasmus deutlich abgenommen. Skeptische

Fragen werden gestellt. Vielfach herrscht die Einschätzung, dass in den von Nutzerbeiträgen geprägten Bereichen des Internets – Kommentarspalten von Medien, Foren, Sozialen Medien wie Facebook oder YouTube – manipulative, beleidigende oder sogar hasserfüllte Inhalte dominieren und ein rationaler Diskurs über politische Streitfragen kaum noch möglich ist.

Selbst wenn aufgrund der beinahe grenzenlosen Menge von Internetinhalten kaum wirklich verlässliche Zahlen über den Umfang dieses Phänomens

gewonnen werden können, so deuten Datenerhebungen in begrenzten Bereichen des Internets auf eine gravierende Zunahme des Problems in den letzten Jahren hin: So ermittelte etwa eine Untersuchung im Zeitraum von Februar 2017 bis Februar 2018 innerhalb eines Jahres einen Anstieg von Beiträgen rechtsextremen, hasserfüllten Inhalts um mehr als das Zehnfache (Kreißel et al., 2018).

Ein wichtiger Kristallisationspunkt von Hassbeiträgen sind alle Themen, in denen es um Ausgrenzung und Abwertung von Minderheiten oder bestimmten Personengruppen geht. Vor allem die Diskussion um die Aufnahme von Flüchtlingen seit Sommer 2015 hat sich hierfür als eine Art Brandbeschleuniger erwiesen. Die Ausprägungen dieser Debatte haben Befürchtungen genährt, dass hasserfüllte Beiträge im Netz eine Bedrohung für den gesellschaftlichen Zusammenhalt darstellen, indem sie zum Rückgang des Vertrauens in politische Institutionen, zum Aufstieg rechtsextremer Parteien und nicht zuletzt auch zur Zunahme von Gewaltstraftaten gegen Angehörige angegriffener Personengruppen führen können.

Aus diesem Grund beschäftigen sich mittlerweile viele Wissenschaftlerinnen und Wissenschaftler, etwa aus der Rechtswissenschaft, den Sprachwissenschaften, den Sozialwissenschaften oder auch der Informatik, mit diesem Problemfeld. Aus Sicht der Medien- und Kommunikationswissenschaft steht vor allem die Frage im Mittelpunkt, wie sich solche Hassbeiträge in digitalen Medien verbreiten und welche Folgen sie dort haben können. Solche Inhalte sind schließlich ein Phänomen, das stark mit dem digitalen Medienwandel verknüpft ist: Sie existieren im Wesentlichen erst, seit der klassische Journalismus seine sogenannte „Gatekeeper“-Funktion für die Öffentlichkeit verloren hat und mit Internetmedien im Prinzip jede Einzelperson an den Massenmedien vorbei Öffentlichkeiten für ihre Botschaften finden kann.

WAS IST HASSREDE?

Die Frage, was Hassrede (englisch „hate speech“) genau ist, ist nicht einfach zu beantworten. Die engste Definition ist in der Regel eine rechtswissenschaftliche, in der sich strafrechtlich relevante Aussagen, etwa Aufrufe zur Gewalt oder „Aufstachelung zum Rassenhass“ (Paragraph 130 Absatz 1 Strafgesetzbuch [StGB]), abgrenzen lassen von allen anderen Aussagen, die – auch wenn sie im Einzelfall sehr unschön sind – durch das in Deutschland relativ weit ausgelegte Recht auf Meinungsfreiheit gedeckt sind. Die Zahlen zur Löschung „offensichtlich rechtswidriger Inhalte“, wie sie kürzlich von Facebook im zweiten Transparenzbericht nach dem Netzwerkdurchsetzungsgesetz veröffentlicht wurden (Facebook, 2019), halten sich jedoch in sehr engen Grenzen und erreichen keinesfalls die Umfänge, die von externen Studien wie der oben zitierten ermittelt wurden. Es gibt also aus sprach- und

gesellschaftswissenschaftlicher Perspektive einen großen Graubereich mutmaßlich legaler, trotzdem aber hasserfüllter und für Betroffene und die Öffentlichkeit schädlicher Inhalte.

Unter Rückgriff auf Zugänge der Sprach- und Sozialwissenschaften hat sich für Hassrede ein Verständnis herausgebildet, das im Wesentlichen auf drei Aspekten basiert: *erstens* auf dem Aspekt der „Rede“, das heißt, es geht um Äußerungen in einer Öffentlichkeit, auch wenn diese eher klein ist, etwa in einer Gaststätte oder einem Online-Forum, wo zufälliges Publikum anwesend sein kann. Persönliche Beleidigungen, die in einem privaten Gespräch, zum Beispiel aus einer situativen Verärgerung heraus, ausgesprochen werden, sind also nicht Gegenstand des Problems und seiner Analyse.

Zweitens definiert sich Hassrede durch eine Herabwürdigung einzelner Personen oder ganzer Gruppen; sprachlich kann dies viele verschiedene Formen annehmen, etwa emotional oder sachlich, über Beleidigungen, Bedrohungen oder auch abwertende Stereotypisierungen (zum Beispiel „terroristische Muslime“, „stehende Polen“). Ein *drittes* Merkmal ist schließlich der Rückgriff der Herabwürdigung auf pauschale Gruppenmerkmale; eine Person wird aufgrund bestimmter Eigenschaften, auf die sie meist gar keinen Einfluss hat, abgewertet (etwa die Herkunft oder Hautfarbe, im Fall von sogenannter „Misogynie“/Frauenfeindlichkeit auch das Geschlecht). Die Beleidigung einer Person aufgrund einer konkreten Handlung, etwa weil sie mir gerade den Parkplatz weggeschnappt hat, würde nicht darunter fallen.

ZWISCHENFORMEN UND ÜBERGANGSBEREICHE

Auf Basis solcher Definitionen versuchen weltweit zurzeit viele interdisziplinäre Projekte, Hassrede genauer zu verstehen und auch deren Entstehungsbedingungen und mögliche Wirkungen zu erklären, so etwa das Projekt NOHATE an der Freien Universität und der Beuth-Hochschule in Berlin (www.nohate.online). Dabei geht es nicht nur darum, eine einfache Ja-Nein-Entscheidung über Hassrede oder Nicht-Hassrede zu treffen, sondern auch Zwischenformen und Übergangsbereiche zu identifizieren, etwa von sogenannter „unziviler“ Kommunikation, die bestimmte Regeln verletzen, ohne bereits Hassrede zu sein. Es gibt mittlerweile einige Studien, die zeigen, dass Diskussionen durch solche minderschweren sprachlichen Entgleisungen auf eine schiefe Ebene geraten können, durch die im weiteren Verlauf das Auftreten schwerwiegenderer Formen von Hassrede begünstigt wird. Lernen wir, diese Mechanismen besser zu verstehen, können wir bessere Strategien entwickeln, um frühzeitig reagieren zu können.

Hassrede ist, wie bereits erwähnt, ein Phänomen des sogenannten „User Generated Content“, also von Inhalten, die von Personen oder Organisationen außerhalb der klassischen journalistisch-politischen Öffentlichkeitsakteure

verbreitet werden. Man sollte sich jedoch vor dem Kurzschluss hüten, jeder Beitrag in diesen frei zugänglichen Foren des Internets würde auch wirklich von einer Einzelperson geschrieben, die ihre persönliche Meinung äußern will. Die oben zitierte Studie von Kreißel et al. zeigt, dass etwa die Hälfte der großen Menge an Hassbeiträgen von nur einer Handvoll Nutzeraccounts stammt. Man überschätzt deshalb in der Regel aufgrund der reinen Masse von Beiträgen die Größe der dahinterstehenden Bevölkerungsgruppen, und mutmaßlich ist genau das auch der Zweck eines solchen massenhaften Verbreitens von Botschaften.

BEZAHLTE AGENTEN UND AUTOMATISIERTE ACCOUNTS

Aus der Kommunikationsforschung wissen wir, dass sich dadurch bei anderen Nutzern der Eindruck einer starken Meinungsmehrheit ergeben kann, durch die viele es dann nicht mehr wagen, eine andere Meinung öffentlich zu äußern. Der Nachweis, dass hinter solchen Hasswellen im Netz vielfach strategisch handelnde politische Akteure stehen, die eventuell bezahlte Agenten oder (halb)automatisierte Accounts (Social Bots) einsetzen, ist allerdings schwer zu führen. Denn im Internet lassen sich Spuren leicht verwischen, und die Tatsache, dass es sich bei Facebook und Co. um Plattformen von Privatunternehmen und dementsprechend bei den Inhalten um deren Firmeneigentum handelt, schränkt die Möglichkeiten der unabhängigen Wissenschaft zur Analyse dieser Inhalte deutlich ein.

Skepsis ist teilweise auch hinsichtlich anderer Wirkungen angebracht: So kann eine unzivile oder gar hasserfüllte Kommunikation zwar Hemmschwellen senken und Menschen zu ähnlichen Äußerungen mobilisieren; einem unmittelbaren Einfluss auf die Einstellungen der Gesamtbevölkerung steht allerdings der soziologische Befund entgegen, dass wir Menschen in der Regel durch ein relativ stabiles System von Werten und Einstellungen geprägt sind, das sich nicht so einfach durch Medieninhalte ändern lässt. Stattdessen wenden sich Menschen lieber Inhalten zu, die ihren schon existierenden Einstellungen entsprechen, und vermeiden solche, die diesen widersprechen. Gerade diese urmenschlichen Verhaltensweisen werden von den meisten Social-Media-Plattformen noch verstärkt, indem uns durch Vorschlags- und Empfehlungssysteme (etwa bei YouTube) oder direkte Filterung von Inhalten (etwa bei Facebook) meist nur solche Inhalte angeboten werden, die wir mutmaßlich gern sehen und die uns zu weiterer Nutzung animieren – um uns weitere bezahlte Werbung anzeigen zu können. Eine Folge dieses Systems aus verhaltensverstärkender Angebotsorganisation ist, dass vor allem solche Nutzer zum Beispiel fremdenfeindliche und herabwürdigende Inhalte suchen und sie auch stärker angezeigt bekommen, die bereits entsprechende Voreinstellungen haben. Folge dieses Systems kann einerseits eine Radikalisierung

und Mobilisierung von Menschen mit rechtsradikaler Einstellung sein, andererseits eine Abwendung und das zunehmende Schweigen von Menschen mit gegensätzlicher Haltung.

Was lässt sich unternehmen, um zu verhindern, dass solche Wirkungen demokratische Werte und Prozesse erodieren lassen? Natürlich muss man zuerst auch gegenüber neuen Anbietern öffentlicher Diskussionsräume wie etwa Facebook auf der Einhaltung des Rechts bestehen. Allerdings kann es dabei nur um die Löschung beziehungsweise Strafverfolgung eindeutig illegaler Inhalte gehen, die aber, wie beschrieben, nur einen kleinen Teil des Problems darstellen.

Wichtiger für eine „Zivilisierung“ der Debatte wären Maßnahmen wie eine Verringerung der oben beschriebenen Filtereffekte, etwa indem man Plattformen in gewissen Grenzen zu einer Organisation von Vielfalt beziehungsweise zu einer Begrenzung der rein auf positive Nutzererfahrung ausgerichteten Inhaltsdarstellung verpflichtet. Darüber hinaus gibt es durchaus positive Erfahrungen mit verschiedenen Formen der Moderation von Internet-Diskussionsforen: Schnelles Intervenieren bei Hassbeiträgen, Sperrung betrügerischer Accounts, Verweis der Nutzer auf Forenregeln und sachliche Argumentation können hitzige Diskussionen versachlichen und beruhigen.

Damit lassen sich sicher nicht alle Probleme lösen – in einer Welt, in der Bürger sich jederzeit frei äußern können, wird man dauerhaft damit leben müssen, auch unangenehmen Beiträgen oder unbelehrbaren „Trollen“ oder Aktivisten zu begegnen. Will man keine „chinesische Lösung“, also eine zentralisierte, automatisierte Vollkontrolle mit umfangreicher Löschung von Nutzerinhalten, muss man Lösungen finden, die mit demokratischen Grundwerten vereinbar sind. Hier ist, neben Plattformregulierung und mehr Investitionen in die Moderation von Inhalten, nicht zuletzt Medien-, Kommunikations- und Bürgerkompetenz gefragt: Wir alle haben in der neuen digitalen Kommunikationswelt, die uns viele neue Freiheiten eröffnet, eben auch eine größere Verantwortung, mit diesen Freiheiten umzugehen. Wir müssen lernen, zu erkennen, welche Qualität bestimmte Inhalte haben (dies gilt auch für andere Arten der Problemmunikation, etwa für das Problem von „Fake News“), wir müssen unangenehme Inhalte ertragen können, und schließlich sollten wir lernen, wie wir selbst durch eigene Beiträge gewinnbringende Diskussionen mit anderen Menschen im Netz führen können.

Literatur

KreiBel, Philip / Ebner, Julia / Urban, Alexander / Guhl, Jakob: Hass auf Knopfdruck: Rechtsextreme Trollfabriken und das Ökosystem koordinierter Hasskampagnen im Netz, Institute for Strategic Dialogue (ISD) & #ichbinhier, London 2019, u. a.: www.isdglobal.org/wp-content/uploads/2018/07/ISD_Ich_Bin_Hier_2.pdf [Zugriff am 05.06.2019].

Facebook: NetzDG-Transparenzbericht, Dublin, Januar 2019, fbnewsroomus.files.wordpress.com/2019/01/facebook_netzdg_januar_2019_deutsch52.pdf [Zugriff am 05.06.2019].