



## Mit Mathematik gegen Hass und Fliegenfallen

### Erkenntnisse aus der Analyse von vernetzten sozialen Online-Netzwerken

Ereignisse aus der jüngsten Vergangenheit – wie die Ermordung des Kasseler Regierungspräsidenten Walter Lübcke am 2. Juni 2019 [1] oder die Beurteilung von 22 beleidigenden Äußerungen gegenüber der Politikerin Renate Künast durch das Berliner Landgericht (am 9. September 2019) als „zulässige Meinungsäußerungen“ (und nicht als strafbare Beleidigungen; [9]) – machen klar, dass wir Hass-Reden ernsthafter entgegenzutreten müssen. Erkenntnisse aus den Sozialwissenschaften zeigen dies zudem überdeutlich (Spitzer, in diesem Heft). Aber was kann man tun? Antworten kommen aus einer wissenschaftlichen Disziplin, an die man eher zunächst nicht denkt, wenn man über diese Frage nachdenkt: Mathematik.

Man kann Strukturen und Prozesse auf verschiedenen Ebenen betrachten, auf kleinsten molekularer Ebene einerseits oder auf größtmöglicher „ganzheitlicher“ Ebene andererseits, wobei die Wahrheit (im Sinne einer Betrachtungsebene, die interessante neue Erkenntnisse liefert) nicht selten irgendwo in der Mitte auf der Ebene von Systemen oder Netzwerken liegt. Um Aphasien zu verstehen, untersucht man am besten das Zusammenspiel zwischen den an der Produktion bzw. der Dekodierung von Sprache beteiligten Strukturen, bei denen es sich jeweils um Netzwerke von Neuronen handelt. Netzwerke von Netzwerken sind mithin in der Systemneurobiologie (im englischen spricht man von „systems neuroscience“) nichts Ungewöhnliches. Die Systemeigenschaften solcher dynamisch interagierenden Netzwerke (die sich oft erst durch das Zusammenspiel ergeben, und nur dann) lassen sich mathematisch beschreiben und vor allem auch modellieren, so dass man solche Modelle analysieren und dadurch sogar Einsichten in mögliche Änderungen der Dynamik durch äußere Eingriffe gewinnen kann.

Um nichts weniger geht es in einer kürzlich im Fachblatt *Nature* publizierten Arbeit zur

„globalen Online-Hass-Ökologie“<sup>1</sup>, wie die Autoren auf verschiedenen sozialen Online-Plattformen untereinander verbundene, global agierende Hassgruppen verschiedener Länder und Sprachen bezeichnen [7]. Es geht um Facebook, WhatsApp, Snapchat, Instagram, Telegram, VKontakte, WeChat oder Weibo. Und es geht um rechtsradikalen, islamistischen oder wie auch immer motivierten Hass, den es zu bekämpfen gilt, indem man seine Dynamik versteht [5].

Die Autoren untersuchten hierzu öffentlich zugängliche Daten von den genannten Plattformen aus verschiedenen Ländern in verschiedenen Sprachen. Ohne sich auf die Ebene einzelner Individuen zu begeben fanden sie durch die Analyse von Webseiten, Chatgruppen und Blog-Gemeinden zunächst einmal, dass sich Hassgruppen durch Prozesse der Selbstorganisation bilden und erstaunlich stabil sind. Für die Nutzer solcher als „Cluster“ bezeichneter (je nach Plattform unterschiedlich formeller oder informeller) Gruppierungen ergeben sich Vorteile, da in ihrem Cluster sogenannte Trolle, Bots oder (menschliche) politische Gegner unterdrückt werden, die sie damit weniger belästigen.

In diesen Clustern überlappen sich Hass-Themen (antisemitisch, anti-Immigration, anti-LGBT<sup>2</sup>), Sprachen, Kulturen und Plattformen, wie die Autoren am Beispiel des Attentats in Christchurch, Neuseeland, vom 15. März 2019 zeigen. Mittels Schusswaffen tötete der 28-jährige terroristische Einzeltäter, der Australier Brenton Tarrant, 51 Menschen und verletzte weitere 50 Menschen, zum Teil schwer. Der ideologische Hintergrund be-

stand in einer Mischung aus rechten, islamfeindlichen, rassistischen und gewaltverherrlichenden Motiven, die auf unterschiedlichen Plattformen verbreitet wurden, wie der Wikipedia-Eintrag zu dieser extremen Gewalttat beispielhaft zeigt:

„Auf Tarrants Waffen und Munitionsmagazinen befanden sich kyrillische und osteuropäische Inschriften mit den Namen von historischen Schlachten, Kämpfern gegen das Osmanische Reich und früheren antimuslimischen Terroristen [...] Darunter waren die Namen Ernst Rüdiger von Starhemberg, der 1683 die Zweite Wiener Türkenbelagerung abgewehrt hatte, der venezianische Offizier Marcantonio Bragadin, der Terrorist Alexandre Bissonnette, der 2017 beim Anschlag auf das Centre culturel islamique de Québec sechs Menschen erschossen hatte. [...] Tarrant hatte im Internet ein Manifest unter dem Titel „Der große Austausch“ veröffentlicht. Er bezog sich damit auf die im Neonazismus und bei der Neuen Rechten beliebte These eines angeblichen geplanten Bevölkerungsaustauschs der weißen Europäer durch Muslime und eine ‚Islamisierung‘ Europas. Er nannte sich im Manifest ‚Rassist‘, ‚Ethnonationalist‘ und ‚Ökofaschist‘. [...] Er erwarte, nach 27 Jahren freizukommen und dann den Friedensnobelpreis zu erhalten, den nach dem Sieg seines Volkes ja auch der Terrorist Nelson Mandela für die gleichen Taten erhalten habe. [...] Auf seinem Facebookprofil hatte er viele Artikel mit Europabezug verlinkt, darunter einen Bericht über rechtsextreme Soldaten in der deutschen Bundeswehr, und geschrieben: ‚Mein Blut ist europäisch‘. [...] Tarrant verbreitete sein Manifest über dieselben Webseiten, die auch die Terrorgruppe Islamischer Staat nutzte. [...] Er platzierte Links in sozialen Medien so, dass sie möglichst viele Leser erreichen“ [2]. Die Autoren halten die in diesem Text näher ausgeführte Mischung für paradigmatisch, beschreiben sie doch selbst den Sachverhalt wie folgt:

1 Der vollständige Titel der Arbeit lautet: „Versteckte Widerstandsfähigkeit und adaptive Dynamik der globalen Online-Hass-Ökologie“.

2 Abkürzung für „lesbian, gay, bisexual und transgender“.

„Das in hohem Maße wechselseitig verbundene Netzwerk von Netzwerken mischt unterschiedliche Hass-Themen (z. B. antisemitisch, anti-immigration, anti-LGBT+), Sprachen, Kulturen und Plattformen. Dieses Online-Mischen geschah beim Attentat in Christchurch: Der mutmaßliche Angreifer war Australier, der Angriff in Neuseeland, und die Schusswaffen trugen mehrere Inschriften in mehreren europäischen Sprachen zu historischen Themen, die auf über ganz Europa verteilt online-Hass-Clustern erwähnt werden [8].<sup>3</sup>

Schon vor 3 Jahren war gezeigt worden, dass 80 % der Fake-News auf Twitter während der US-amerikanischen Präsidentschaftswahlen von nur 0,1 % der Nutzer weitergeleitet worden waren [6]. Auch in der vorliegenden Studie konnte am Beispiel von Ku Klux Klan Hass-Clustern gezeigt werden, dass deren Größe einer abnehmenden Exponential-Funktion entspricht, d. h. wenige sehr große und viele kleine Knoten im Netzwerk vorhanden sind. Man kann dies als Hinweis darauf werten, dass solche vernetzten Cluster durch Prozesse der Selbstorganisation entstehen.<sup>4</sup>

Anhand von Daten aus zwei der genannten sozialen Netzwerke – Facebook und VKontakte<sup>5</sup> – um die Zeit des Christchurch-Attentats herum konnten die Autoren zei-

gen, dass diese unabhängig funktionierenden Plattformen, die im Wettbewerb um Nutzer wirtschaftliche Konkurrenten sind, durch die Aktivität von globalen Hass-Netzwerken miteinander verknüpft werden. Dies geschieht durch Selbstorganisation, also ohne einen Plan dahinter, von allein, aufgrund der Eigendynamik ihrer Funktion. Dadurch widerstehen diese Online-Hass-Netzwerke (die Autoren sprechen von hate-cluster-networks) Versuchen ihrer Bekämpfung mit prinzipiell drei Typen von Reaktion, welche die Autoren als „Spiegelung“, „Reinkarnation“ und „direkte Verlinkung“ bezeichnen.

So beobachteten die Autoren beispielsweise, dass nach der Blockade des Ku Klux Klan auf Facebook etwa 60 Ku Klux Klan Cluster auf VKontakte zu finden waren, zum Teil in Ukrainischer Sprache. Als die Regierung der Ukraine später VKontakte verbot, wurde zuvor von dort (über Verbindungen zwischen den beiden sozialen Online-Netzwerken in Europa, den USA und Süd-Afrika) das Ku Klux Klan Cluster auf Facebook reinkarniert, mit „Ku Klux Klan“ in kyrillischer Schrift, so dass dies für automatische (in Englisch operierende) Suchalgorithmen im Verborgenen blieb. Damit gab es plötzlich den Ku Klux Klan wieder auf Facebook mit tausenden von Nutzern. Wie kann das sein?

Ganz offensichtlich liegt dies nicht zuletzt daran, dass sich immer der gleiche Hass an unterschiedlichen Inhalten manifestiert, was man mit reiner Logik nicht unbedingt nachvollziehen kann. Die Autoren beschreiben dies wie folgt: „Insbesondere in Europa gibt es eine komplexe Hass-Ökologie, die miteinander verschränkte Narrative enthält, die sowohl Sprachen als auch Begründungszusammenhänge überbrücken – beispielsweise Neonazi-Cluster mit Mitgliedern in Großbritannien, Kanada, den USA, Australien und Neuseeland, die Material über Fußball, den Brexit und Bilder von Skinheads enthalten und zugleich die Musik der Schwarzen promoten. Der Hass mag also rein sein, die Überlegungen dahinter sind es nicht, was nahelegt, dass diese Online-Umgebung wie eine globale Fliegenfalle funktioniert, die rasch neue Mitglieder aus vielen Plattformen, Ländern und Sprachen rekrutieren kann, insbesondere wenn diese noch keinen klaren Fokus

für ihren Hass haben“<sup>6</sup> [7]. Durch weitere mathematische Analysen zu den Verbindungen zwischen Netzwerken konnten die Autoren zeigen, dass manche Interventionen kontraintuitive, unerwünschte Konsequenzen haben können: Wird beispielsweise auf Facebook Zensur ausgeübt und auf VKontakte keine, so kann die Zensur auf nur einem Netzwerk (Facebook) die Verbreitung von Hass über beide Netzwerke unter bestimmten Bedingungen sogar erleichtern. Es können „dark pools“ [8] von Verbindungen, also unbekannte und schwer erkennbare Verbindungen entstehen, welche die Bekämpfung von Online-Hass erschweren.

Insbesondere, wenn Hass-Cluster unter Druck geraten, ändert sich ihre Struktur, wodurch ihre Widerstandsfähigkeit steigt. Die Autoren zeigen dies anhand von Veränderungen der Netzwerkstruktur des Ku Klux Klan Clusters nach dem Schulmassaker von Parkland (Florida, USA) am 14. Februar 2018, bei dem der 19-jährige Nikolas Cruz an seiner ehemaligen Schule 14 Schüler und drei Erwachsene erschossen hatte [2]. Weil in der Presse über Verbindungen des Attentäters zum Ku Klux Klan berichtet wurde, änderte sich dessen Struktur auf VKontakte (► **Abb. 1**): Aus einer losen Anordnung wurde durch Selbstorganisation „von unten“ (bottom-up) eine dichte, klare Struktur.

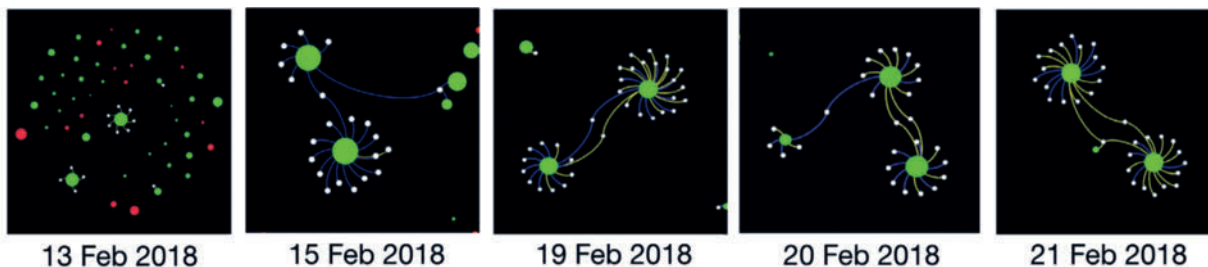
Das gleiche Phänomen konnten die Autoren auch nach dem (missglückten) Anschlag auf den Terroristen und IS-Führer Abu Bakr al-Baghdadi am 18. März 2015 nachweisen: Der djihadistische anti-westliche Hass organisierte sich neu im Sinne einer starken Verdichtung.

3 „The highly interconnected network-of-networks mixes hate narratives across themes (for example, anti-Semitic, anti-immigrant, anti-LGBT+), languages, cultures and platforms. This online mixing manifest itself in the 2019 attack in Christchurch: the presumed shooter was Australian, the attack was in New Zealand, and the guns carried messages in several European languages on historical topics that are mentioned in online hate clusters across continents“ [8].

4 This suggests that the online hate ecology is self-organized, because it would be almost impossible to engineer this distribution using top-down control [8].

5 VKontakte ist ein seit dem Jahr 2006 bestehendes russisches, in 81 Sprachen operierendes soziales Online-Netzwerk mit weltweit etwa 100 Millionen Nutzern. In Russland liegt es im Hinblick auf die Anzahl seiner Nutzer auf Platz drei, in Deutschland auf Platz 9. Die Plattform reagiert auf radikale Inhalte, die hochgeladen werden, praktisch nicht und verbreitet rechtsradikales, rassistisches, neonazistisches und neofaschistisches Gedankengut [3].

6 Im Original: „Europe shows a particularly complex hate ecology, which reflects intertwined narratives that cross languages and declared causes—for example, neo-Nazi clusters with membership drawn from the United Kingdom, Canada, United States, Australia and New Zealand feature material about English football, Brexit and skinhead imagery while also promoting black music genres. So although the hate may be pure, the rationale given is not, which suggests that this online ecology acts like a global fly-trap that can quickly capture new recruits from any platform, country and language, particularly if they do not yet have a clear focus for their hate“ [8].



► **Abb. 1** Dynamik der Struktur des „Ku Klux Klan Ökosystems“ auf der Plattform VKontakte vor und in den Tagen nach dem Schulmassaker von Parkland. In Form weißer Punkte sind nur diejenigen einzelnen Nutzer dargestellt, deren Status sich im nächsten Zeitschritt geändert hat. Rot sind geschlossene Cluster, grün sind offene Cluster dargestellt. Verbindungen sind nach ihrem zeitlichen Verlauf in gelb und blau dargestellt. Gelb: die Verbindung verschwindet zwischen den Zeitpunkten  $t$  und  $t + 1$ ; blau: die Verbindung entsteht am Tag  $t$  (nach Daten aus [7], ► **Fig. 3a**). Wie man sieht, verändert sich das Netzwerk im Verlauf weniger Tage von einer losen Häufung in eine klare Struktur.

Abschließend schlagen die Autoren 4 Interventionsstrategien vor:

- Die Anzahl großer Cluster reduzieren. Hierbei könnte man meinen, dass man am besten mit dem größten anfängt, was sich jedoch aus Sicht der Autoren ungünstig auswirkt. Ihre Analysen zeigen jedoch bereits im Jahr 2016 anhand der empirischen Untersuchung von ISIS-Clustern [7] und in der hier vorliegenden allgemeineren Studie, dass bei einem Wegfall des größten Clusters sofort wieder ein großes durch Selbstorganisation entsteht. Weil mittelgroße und kleinere Cluster häufiger und leichter aufzufinden und auch zu verbieten seien, sei es langfristig eher zielführend, Gegenmaßnahmen bei mittelgroßen Clustern zu treffen (aus denen dann mit geringerer Wahrscheinlichkeit erneut große Cluster entstehen).
- Einen kleinen Bruchteil zufällig ausgewählter Nutzer über viele Hass-Plattformen hinweg aus dem Netz entfernen. Man würde auf diese Weise am wenigsten Gegenwehr produzieren, sodass die Schwächung nachhaltiger sei.
- Anti-Hass Cluster aus Hass-Gegnern könnten unterstützt werden. Diese Hass-Gegner gibt es und man könnte ihre Clusterbildung fördern sowie deren Anzahl und Aktivität steigern.
- Weil sich Hass-Cluster nicht selten gegenseitig widersprechen, könnte man sie gegeneinander ausspielen. Hierzu müsste man neue Cluster mit spezifischem Profil etablieren bzw. stärken, was dazu führen könnte, dass sich ein-

zelne Hass-Cluster gegenseitig neutralisieren.

Mit den heute gegebenen Möglichkeiten der Modellierung halten es die Autoren für möglich, wirksame Strategien gegen Online-Hass zu entwickeln, ihre Auswirkungen zu messen und dann rational im Einzelfall informiert zu entscheiden, wie gegen diese oder jene Form von Hass am besten vorzugehen ist.

Man kann nur hoffen, dass sie Recht behalten und ihre Erkenntnisse zu einer Verminderung des weltweit stärker werdenden Hasses beiträgt, der ansonsten unsere Gesellschaft langfristig gefährdet.

### Interessenkonflikt

Es liegt kein Interessenkonflikt vor.

### Korrespondenzadresse

**Prof. Dr. Dr. Manfred Spitzer**  
Universität Ulm  
Abteilung für Psychiatrie  
Leimgrubenweg 12–14  
89075 Ulm

### Literatur

- [1] Anonymus. Mordfall Walter Lübcke. Wikipedia [https://de.wikipedia.org/wiki/Mordfall\\_Walter\\_Lübcke](https://de.wikipedia.org/wiki/Mordfall_Walter_Lübcke); abgerufen am 26.9.2019

- [2] Anonymus. Terroranschlag auf zwei Moscheen in Christchurch. Wikipedia [https://de.wikipedia.org/wiki/Terroranschlag\\_auf\\_zwei\\_Moscheen\\_in\\_Christchurch](https://de.wikipedia.org/wiki/Terroranschlag_auf_zwei_Moscheen_in_Christchurch); abgerufen am 26.9.2019
- [3] Anonymus. vk.com (VKontakte). Wikipedia <https://de.wikipedia.org/wiki/Vk.com>; abgerufen am 28.9.2019
- [4] Anonymus. Schulmassaker von Parkland. Wikipedia [https://de.wikipedia.org/wiki/Schulmassaker\\_von\\_Parkland](https://de.wikipedia.org/wiki/Schulmassaker_von_Parkland); abgerufen am 28.9.2019
- [5] Derzsy N. The dynamics of online hate. *Nature* 2019; 573: 203–204
- [6] Grinberg N, Joseph K, Friedland L, et al. Fake news on Twitter during the 2016 U. S. presidential election. *Science* 2019; 363: 374–378
- [7] Johnson NF, Zheng M, Vorobyeva Y, et al. New online ecology of adversarial aggregates: ISIS and beyond. *Science* 2016; 352: 1459–1463
- [8] Johnson NF, Leahy R, Johnson Restrepo N, et al. Hidden resilience and adaptive dynamics of the global online hate ecology. *Nature* 2019; 573: 261–265
- [9] Siebert P. Beschimpfungen – Künast will gegen Urteil vorgehen. *Berliner Morgenpost* 18.9.2019 <https://www.morgenpost.de/berlin/article227129109/RenateKuenastuebelbeschimpftKritikanBeschlussvonBerlinerGericht.html>; abgerufen am 26.9.2019
- [10] Spitzer M. Von Hass-Sprache zur Kriminalität. Facebook, Twitter, die AfD und der Brexit. *Nervenheilkunde* 2019; 38: 794–802

### Bibliografie

DOI <https://doi.org/10.1055/a-0952-6806>  
*Nervenheilkunde* 2019; 38: 854–856  
© Georg Thieme Verlag KG Stuttgart · New York  
ISSN 0722-1541