

Teil 3.

Moderationsmacht und soziale Plattformen

Die Grenzen privater Normsetzung durch soziale Netzwerke: Zur Drittwirkung der Unionsgrundrechte bei Community Standards am Beispiel von Hate Speech

Amelie Röbling und Johannes Weil*

I. Einleitung

Die Community Standards sozialer Netzwerke haben in jüngerer Zeit immer mehr Aufmerksamkeit erlangt – zuletzt etwa im Kontext der Sperrung diverser Social Media Accounts von Donald Trump.¹ Insbesondere legen die Diensteanbieter darin einseitig fest, welche Inhalte auf ihren Plattformen nicht veröffentlicht werden dürfen. Sie fungieren in dieser Hinsicht als „private Gesetzgeber“, interpretieren die betreffenden Regelungen aber auch und setzen sie selbstständig durch.² Bisweilen überschreiten sie dabei Grenzen, die dem Staat bei vergleichbaren Beschränkungen der Meinungsfreiheit zu setzen wären.³ Angesichts der massiven Konzentrationsprozesse, denen die digitalen Meinungsmärkte unterliegen, kommen so speziell die marktbeherrschenden Netzwerke in eine Machtposition, die ihnen eine weitgehende Beeinflussung der öffentlichen Meinungsbildung erlaubt.⁴

Für die deutsche Rechtsordnung wird daher schon länger diskutiert, ob die Diensteanbieter im Wege der mittelbaren Drittwirkung stärker an die

* Der Beitrag wurde vom Bundesministerium für Bildung und Forschung (BMBF 01IS18026B) durch die Förderung des Zentrums für skalierbare Datenanalyse und Künstliche Intelligenz ScaDS.AI Dresden/Leipzig unterstützt

1 Twitter Inc., Permanent suspension of @realDonaldTrump, Twitter-Blog 2021, https://blog.twitter.com/en_us/topics/company/2020/suspension; Clegg, In Response to Oversight Board, Trump Suspended for Two Years, Facebook-Newsroom 2021, <https://about.fb.com/news/2021/06/facebook-response-to-oversight-board-recommendations-trump>.

2 Pille, Meinungsmacht sozialer Netzwerke, 2016, S. 335.

3 Siehe etwa zur aktuellen Twitter-Kampagne gegen Desinformation zu COVID-19 Reinhardt/Yazicioglu, Grundrechtsbindung und Transparenzpflichten sozialer Netzwerke, DSRITB 2020, 819 (819 f.).

4 Vgl. Dolata, Plattform-Regulierung, Koordination von Märkten und Kuratierung von Sozialität im Internet, Berlin Journal für Soziologie 2019, 179 (183 f.); siehe dazu auch das Impact Assessment der EU-Kommission zum Entwurf des *Digital Services Act*, SWD (2020) 348 final Rn. 85 ff.

Grundrechte ihrer Nutzer zu binden sind, als dies üblicherweise im grundsätzlich gleichgeordneten Verhältnis von Privaten gilt.⁵ Mit ihrem vielbeachteten Entwurf eines *Digital Services Act* (DSA-E)⁶ rückte die EU-Kommission Ende 2020 die Grundrechtsbeziehungen zwischen Anbietern und Nutzern sozialer Medien auch auf unionaler Ebene in den Fokus.⁷ So sieht der als „Grundgesetz für das Internet“⁸ gefeierte Verordnungsentwurf, über den sich das Europäische Parlament und der Rat der EU mit einigen Abänderungen am 23. April 2022 vorläufig geeinigt haben,⁹ eine umfassende Regulierung der Online-Plattformen vor, die eine präzedenzlose Verpflichtung der Diensteanbieter zur Beachtung der Grundrechte ihrer Nutzer bei der Content Moderation miteinschließt (Art. 12 Abs. 2 DSA-E). Das EU-Parlament hat die konsentrierte Fassung am 5. Juli 2022 formal verabschiedet,¹⁰ mit der endgültigen Entscheidung des Rats ist noch im September dieses Jahres zu rechnen.¹¹

5 Jüngst hat auch der BGH als erstes oberstes Gericht in Deutschland zu dieser Frage Stellung bezogen, BGH 29.7.2021 – III ZR 179/20, ZUM 2021, 953; BGH 29.7.2021 – III ZR 192/20, ZUM-RD 2021, 612; im Übrigen s. z. B. Holznagel, Put-back-Ansprüche gegen soziale Netzwerke: Quo Vadis?, CR 2019, 518 (519–522); Lüdemann, Grundrechtliche Vorgaben für die Löschung von Beiträgen in sozialen Netzwerken, MMR 2019, 279; Raue, Meinungsfreiheit in sozialen Netzwerken, JZ 2018, 961; Tuchtfeld, Marktplätze, soziale Netzwerke und die BVerfG-Entscheidung zum „III. Weg“, VerfBlog 2019, <https://verfassungsblog.de/marktplaetze-soziale-netzwerke-und-die-bverfg-entscheidung-zum-iii-weg>; Weinzierl, Warum das Bundesverfassungsgericht Fußballstadion sagt und Soziale Plattformen trifft, JuWissBlog 2018, <https://www.juwiss.de/48-2018>.

6 Europäische Kommission, COM(2020) 825 final.

7 Siehe etwa Appelman/Quintais/Fahy, Using Terms and Conditions to apply Fundamental Rights to Content Moderation, VerfBlog v. 1.9.2021, <https://verfassungsblog.de/power-dsa-dma-06>; Denga, Plattformregulierung durch europäische Werte: Zur Bindung von Meinungsplattformen an EU-Grundrechte, EuR 2021, 569.

8 Ueberbach, Strengere EU-Regeln für Tech-Konzerne, Tagesschau.de 2022, <https://www.tagesschau.de/wirtschaft/weltwirtschaft/eu-digital-service-act-101.html>.

9 Europäisches Parlament, Pressemitteilung v. 23.4.2022, <https://www.europarl.europa.eu/news/de/press-room/20220412IPR27111/digital-services-act-agreement-for-a-transparent-and-safe-online-environment>; zum DSA-E in der Fassung der vorläufigen Vereinbarung von Europäischem Parlament und Rat der EU siehe Ausschuss für Binnenmarkt und Verbraucherschutz des Europäischen Parlaments, PE734.311.

10 Europäisches Parlament, P9_TA(2022)0269.

11 Europäisches Parlament, Pressemitteilung v. 5.7.2022, <https://www.europarl.europa.eu/news/en/press-room/20220701IPR34364/digital-services-landmark-rules-adopted-for-a-safer-open-online-environment>.

Im Anschluss an diese Entwicklung unternimmt der vorliegende Beitrag den Versuch, die Grenzen der Normgebung und -durchsetzung sozialer Netzwerke anhand der Unionsgrundrechte herauszuarbeiten, wobei das virulente Phänomen der Hate Speech im Mittelpunkt der Betrachtung stehen soll.

II. Gemeinschaftsstandards und Instrumente der Durchsetzung

In ihren Nutzungsbedingungen untersagen soziale Netzwerke bestimmte Inhalte und Formen von Äußerungen. Dass sie dabei über das gesetzlich gebotene Maß hinausgehen, dürfte vornehmlich auf die Geschäftsinteressen der Betreiber zurückzuführen sein. So werden die Kommunikationsregeln darauf ausgerichtet, die Plattformen möglichst attraktiv für Nutzer und Werbekunden zu gestalten.¹² Erfasst werden etwa Gewaltaufrufe oder Darstellungen von Gewalt und Nacktheit sowie bestimmte Formen der Desinformation – speziell im Hinblick auf Wahlen und die COVID-19-Pandemie.¹³ Insbesondere verbieten die marktführenden Plattformen jeweils die Veröffentlichung von Inhalten, die von ihnen als Hate Speech qualifiziert werden. Eine Überschneidung gibt es zumindest dahingehend, als darunter direkte Verbalangriffe auf Einzelne und Personengruppen verstanden werden, soweit die Äußerung an deren Zugehörigkeit zu einer geschützten Kategorie anknüpft, wozu etwa Ethnie, Nationalität, Religion, sexuelle Orientierung, Geschlecht und Geschlechtsidentität zählen.¹⁴ Damit greifen die Diensteanbieter – wie auch der Gesetzgeber – auf eine abstrakt-generelle Regelungstechnik und auslegungsbedürftige Begriffe zurück, was der Vielgestaltigkeit der zu regelnden Äußerungen geschuldet ist.¹⁵

Als Instrumentarium zur Durchsetzung ihrer Kommunikationsstandards steht den Anbietern klassischerweise die Löschung der einschlägigen

12 Raue, Meinungsfreiheit in sozialen Netzwerken, JZ 2018, 961 (964) mwN; ähnlich Denga, Plattformregulierung durch europäische Werte: Zur Bindung von Meinungsplattformen an EU-Grundrechte, EuR 2021, 569 (573 f.).

13 Vgl. Facebook, Gemeinschaftsstandards, <https://transparency.fb.com/de-de/policies/community-standards>; Twitter Inc., Regeln und Richtlinien, <https://help.twitter.com/de/rules-and-policies>; s. für weitere Beispiele auch Pille, Meinungsmacht sozialer Netzwerke, 2016, S. 316–320.

14 Siehe Facebook, Hassrede, <https://transparency.fb.com/de-de/policies/community-standards/hate-speech>; Twitter Inc., Richtlinie zu Hass schürendem Verhalten, <https://help.twitter.com/de/rules-and-policies/hateful-conduct-policy>.

15 Vgl. Friehe, Löschen und Sperren in sozialen Netzwerken, NJW 2020, 1697.

Beiträge, aber auch die vorübergehende oder unbefristete Sperrung von Nutzeraccounts zur Verfügung.¹⁶ In jüngerer Zeit sind jedoch weitere Moderationstechniken hinzugekommen. So reduzieren die Plattformen bisweilen die Reichweite unerwünschter Inhalte. Der Beitrag bleibt also erhalten, wird im News Feed anderer Nutzer aber seltener oder überhaupt nicht mehr angezeigt. Eine weitere Methode stellt das Labeling dar, d. h., der Beitrag wird mit einer entsprechenden Kennzeichnung oder inhaltlichen Einordnung versehen.¹⁷

Um die schiere Masse an potenziell einschlägigen Äußerungen zu bewältigen, setzen die Diensteanbieter neben Meldungen durch andere Nutzer inzwischen auch auf KI-Systeme, die unerwünschte Inhalte automatisiert detektieren. Diese werden anschließend durch Menschen evaluiert oder sogar ausschließlich maschinell moderiert.¹⁸ In der Regel kommen dabei Methoden des maschinellen Lernens zum Einsatz. Stark vereinfacht gesprochen entwickeln diese Systeme anhand festgelegter Trainingsdaten und Lernregeln ein Modell, das über die gegebene Datenbasis hinaus die gewünschte Klassifizierung von Inhalten möglich machen soll.¹⁹ Ausgewertet werden zum einen die in den Beiträgen enthaltenen Texte, Bilder und Videos, die zunächst durch Verfahren der Text- und Bildanalyse für das maschinelle Lernen aufbereitet werden.²⁰ Daneben spielen aber auch

16 Friehe, Löschen und Sperren in sozialen Netzwerken, NJW 2020, 1697 (1700).

17 Hierzu Möller/Hameleers/Ferreau, Typen von Desinformation und Misinformation/Ferreau, 2020, S. 44 (52 f.). Eine weitergehende Darstellung der unterschiedlichen Moderationspraktiken sozialer Netzwerke findet sich im vorliegenden Band, Gsenger/Kübler/Wagner, Content governance on social networking sites: Battling disinformation and upholding values, S. 181.

18 Vgl. Enquete-Kommission Künstliche Intelligenz BT-Drs. 19/23700, S. 479–482; siehe z. B. Bradford u. a., Report of the Facebook Data Transparency Advisory Group, 2019, S. 11 f.

19 Zu den verschiedenen Methoden der Hate Speech-Detektion im Überblick etwa Fortuna/Nunes, A Survey on Automatic Detection of Hate Speech in Text, ACM Computing Surveys 2018, 51(4), 1 (16–22); auf den Einsatz von maschinellem Lernen verweisen z. B. Meta AI, How AI is getting better at detecting hate speech, 2020, <https://ai.facebook.com/blog/how-ai-is-getting-better-at-detecting-hate-speech> zu Facebook; Twitter Developer Platform, HateLab, <https://developer.twitter.com/en/community/success-stories/hatelab>; Google, Transparenzbericht, 2022, <https://transparencyreport.google.com/netzdg/youtube> zu YouTube; siehe auch Enquete-Kommission Künstliche Intelligenz BT-Drs. 19/23700, S. 479–481 für einen Überblick zu den großen Plattformen.

20 Dazu Ebers u. a., Künstliche Intelligenz und Robotik/Niederée/Nejdl, 2020, § 2 Rn. 103–110.

externe Faktoren wie die Identität des Verfassers und Interaktionen, z. B. beigefügte Kommentare oder „Likes“ und „Shares“, eine Rolle.²¹

III. Berücksichtigung der EGMR-Judikatur im unionalen Kontext

Im Hinblick auf die Kommunikations- und Medienfreiheiten, die im Verhältnis von Anbietern und Nutzern sozialer Netzwerke von Relevanz sind, fällt die Rechtsprechung des EuGH nicht besonders ergiebig aus, was speziell auch für die grundrechtliche Behandlung von Hate Speech gilt.²² Bei der Auslegung der Unionsgrundrechte kann aber grundsätzlich auf die weitreichendere Rechtsprechung des EGMR zurückgegriffen werden. Die EMRK diente bei der Erarbeitung der deutlich jüngeren GRC in großen Teilen als Leitbild.²³ Insbesondere sieht Art. 52 Abs. 3 S. 1 GRC vor, dass die Grundrechte der Charta die „gleiche Bedeutung und Tragweite“ haben, wie sie die entsprechenden Grundrechte der Konvention aufweisen.²⁴ Dies bezieht sich sowohl auf den Schutzzumfang als auch auf die Schrankenebene.²⁵ Zu beachten ist dabei nicht nur die textliche Fassung der EMRK, sondern auch die dazu ergangene Rechtsprechung des EGMR.²⁶

IV. Wirkung der Unionsgrundrechte zwischen Privaten

Bei den Anbietern sozialer Netzwerke handelt es sich typischerweise um Privatunternehmen, die als solche nicht zu den klassischen Grundrechts-

21 Siehe z. B. Bradford u. a., Report of the Facebook Data Transparency Advisory Group, 2019, S. 11.

22 Dazu Struth, Hassrede und Freiheit der Meinungsäußerung, 2019, S. 109 ff.

23 Stern/Sachs, Grundrechtecharta/Krämer, 2016, Art. 52 Rn. 65.

24 „Entsprechen“ ist dabei iSv Entsprechens auf tatsächlicher und nicht nur auf rechtlicher Ebene zu verstehen, siehe Spiekermann, Die Folgen des Beitritts der EU zur EMRK für das Verhältnis des EUGH zum EGMR und den damit einhergehenden Individualrechtsschutz, 2013, S. 61 ff.; speziell zur Kongruenz der unions- und konventionsrechtlichen Kommunikations- und Medienfreiheiten EuGH 14.2.2019 – C-345/17, ECLI:EU:C:2019:122 Rn. 65 = EuZW 2019, 299 – Buivids.

25 Meyer/Hölscheidt, Charta der Grundrechte der Europäischen Union/Schwerdtfeger, 5. Aufl. 2019, Art. 52 Rn. 59.

26 EuGH 30.6.2016 – C-205/15, ECLI:EU:C:2016:499 Rn. 41 = BeckRS 2016, 81416 – Toma; EuGH 22.12.2010 – C-279/09, ECLI:EU:C:2010:811 Rn. 35 = BeckRS 2010, 91492 – DEB.

verpflichteten zählen. Um aus den Unionsgrundrechten Grenzen für die normative Gestaltungsmacht der Diensteanbieter abzuleiten, müssten jene aber an die GRC gebunden sein. Insofern soll zunächst ein Blick auf die – noch nicht abschließend geklärte – Frage nach der Wirkung der Unionsgrundrechte im horizontalen Verhältnis geworfen werden.

1. Problematik der unmittelbaren Drittwirkung

Den Anwendungsbereich der Charta definiert Art. 51 Abs. 1 S. 1 GRC. Danach gilt diese für die Unionsorgane umfänglich und für die Mitgliedstaaten nur bei der Durchführung von Unionsrecht.²⁷ Die Vorschrift legt somit nahe, dass jene als alleinige Verpflichtungsadressaten der Unionsgrundrechte gesehen werden könnten.²⁸ Neben dem Wortlaut spricht gegen eine unmittelbare Bindung von Privaten an die Unionsgrundrechte, dass die Anforderungen der Charta an die Rechtfertigung von Grundrechtseingriffen nicht auf privatautonomes Verhalten ausgerichtet sind.²⁹ So ist unklar, wie Private dem in Art. 52 Abs. 1 GRC vorgesehenen Gesetzesvorbehalt entsprechen sollten.³⁰ Trotzdem hat der EuGH in seiner jüngeren Rechtsprechung unter bestimmten Voraussetzungen eine unmittelbare Drittwirkung angenommen.³¹ Die Judikatur beschränkt sich aber insoweit auf Gleichheitsrechte und soziale Rechte.³² Offengeblieben ist, in

27 Ausf. zur „Durchführung des Rechts der Union“ durch die Mitgliedstaaten Meyer/Hölscheidt, Charta der Grundrechte der Europäischen Union/Schwerdtfeger, 5. Aufl. 2019, Art. 51 Rn. 36 ff.; zur Umstr. Bindung der Mitgliedstaaten bei nicht vollständig unionsrechtlich determiniertem Recht siehe Haratsch/Koenig/Pechstein, Europarecht, 12. Aufl. 2020, Kap. 3 Rn. 703 ff. mwN.

28 So etwa Meyer/Hölscheidt, Charta der Grundrechte der Europäischen Union/Schwerdtfeger, 5. Aufl. 2019, Art. 51 Rn. 57; Ehlers, Europäische Grundrechte und Grundfreiheiten, 4. Aufl. 2014, § 14 Rn. 81.

29 Haratsch/Koenig/Pechstein, Europarecht, 12. Aufl. 2020, Rn. 723.

30 Ehlers, Europäische Grundrechte und Grundfreiheiten, 4. Aufl. 2014, § 14 Rn. 81; Haratsch/Koenig/Pechstein, Europarecht, 12. Aufl. 2020, Rn. 723.

31 EuGH 6.11.2018 – C-569/16 u. a., ECLI:EU:C:2018:871 Rn. 87 = NZA 2018, 1467 – Bauer u. Willmeroth; ausf. dazu Classen, Zuviel des Guten? Unionsrechtliche Neuaufwertungen beim Grundrechtsschutz, JZ 2019, 1057 (1062 ff.).

32 Siehe für die Diskriminierungsverbote des Art. 21 Abs. 1 GRC etwa EuGH 17.04.2018 – C-414/16, ECLI:EU:C:2018:257 Rn. 76 = NZA 2018, 569 – Egenberger; EuGH 6.11.2018 – C-684/16, ECLI:EU:C:2018:874 Rn. 78 = NZA 2018, 1474 – Max-Planck-Gesellschaft; zu Art. 31 Abs. 2 GRC EuGH 6.11.2018 – C-569/16 u. a., ECLI:EU:C:2018:871 Rn. 87 f. = NZA 2018, 1467 – Bauer u. Willmeroth; zu

wieweit sich diese Rechtsprechung auch auf die übrigen Grundrechte der Charta übertragen lässt.³³

2. Mittelbare Drittwirkung der Unionsgrundrechte

Jedoch erscheint ein Rückgriff auf die Figur der unmittelbaren Drittwirkung in vielen Fällen schon gar nicht vonnöten. Die von Privaten ausgehenden Gefahren können auch hinreichend durch eine mittelbare Wirkung der Unionsgrundrechte im Horizontalverhältnis kompensiert werden.³⁴ Eine solche wird sowohl vom EuGH³⁵ als auch von weiten Teilen der Literatur anerkannt.³⁶ Die mittelbare Drittwirkung wird dadurch hergestellt, dass nicht Private, dafür aber die nationalen und europäischen Gerichte bei der Rechtsanwendung und -auslegung unmittelbar an die Unionsgrundrechte gebunden sind. Dies wirkt insoweit auf Private zurück, als diese horizontal durch Normen des Zivilrechts berechtigt und verpflichtet werden, die ihrerseits im Lichte der Grundrechte durch die Judikative auszulegen sind.³⁷ Kommt es dabei zu einer Kollision von Grundrechten, ist eine Abwägung zwischen den widerstreitenden Rechtspositionen vorzunehmen.³⁸

Art. 157 Abs. 1 AEUV EuGH, 8.4.1976 – C-43/75, ECLI:EU:C:1976:56 Rn. 31 u. 39 = BeckRS 2004, 71181 – Defrenne.

33 Haratsch/Koenig/Pechstein, Europarecht, 12. Aufl. 2020, Rn. 723.

34 Heselhaus/Nowak, Handbuch der Europäischen Grundrechte/Nowak, 2. Aufl. 2020, § 9 Rn. 47; Haratsch/Koenig/Pechstein, Europarecht, 12. Aufl. 2020, Rn. 723; ähnlich Jarass, Die Bedeutung der Unionsgrundrechte unter Privaten, ZEuP 2017, 310 (332).

35 EuGH 24.11.2011 – C-70/10, ECLI:EU:C:2011:771 Rn. 41 ff. = MMR 2012, 174 – Scarlet; EuGH 16.2.2012 – C-360/10, ECLI:EU:C:2012:85 Rn. 52 = BeckRS 2012, 80348 – SABAM; EuGH 18.7.2013 – C-426/11, ECLI:EU:C:2013:521 Rn. 30 = EuZW 2013, 747 – Alemo Herron; EuGH 13.5.2014 – C-131/12, ECLI:EU:C:2014:317 Rn. 68 f. u. 74 = NJW 2014, 2257 – Google Spain; EuGH 16.7.2015 – C-580/13, ECLI:EU:C:2015:485 Rn. 34 = BeckRS 2015, 80946 – Coty.

36 Pechstein/Nowak/Häde, Frankfurter Kommentar/Pache, 2017, Art. 51 Rn. 38; Meyer/Hölscheidt, Charta der Grundrechte der Europäischen Union/Schwerdtfeger, 5. Aufl. 2019, Art. 51 Rn. 57 ff.; aA Callies/Ruffert, EUV/AEUV/Kingreen, 6. Aufl. 2022, Art. 51 GRC Rn. 26.

37 Haratsch/Koenig/Pechstein, Europarecht, 12. Aufl. 2020, Rn. 723.

38 Jarass, Charta der Grundrechte der EU, 4. Aufl. 2021, Art. 51 Rn. 43.

3. Grundrechtsbindung der Plattformen über Art. 12 DSA-E

Fehlt im Unionsrecht momentan noch ein klarer normativer Anknüpfungspunkt, um die bei der Content-Moderation betroffenen Grundrechte im Horizontalverhältnis zur Wirkung zu bringen, zeichnet sich mit dem Entwurf des DSA eine fulminante Neuerung ab.³⁹ So verpflichtet Art. 12 Abs. 2 DSA-E die Plattformen explizit dazu, bei der Anwendung und Durchsetzung ihrer Geschäftsbedingungen verhältnismäßig vorzugehen und dabei Meinungs- und Medienfreiheit, den Medienpluralismus und andere Grundrechte der GRC von allen Beteiligten und speziell ihren Nutzern zu berücksichtigen. Hierin liegt zwar in der Sache eine unmittelbare Grundrechtsbindung der Plattformen in Bezug auf ihre Kommunikationsregeln. Wegen der Ausgestaltung des DSA-E als Sekundärrechtsakt gilt dies aber ohnehin nur in dem von Art. 12 Abs. 2 DSA-E vorgezeichneten Rahmen. Die Netzwerkanbieter werden also nur partiell und sekundärrechtlich vermittelt an die Charta gebunden. Lässt sich diese Regelungstechnik zwar weder eindeutig dem klassischen Konzept der unmittelbaren noch der mittelbaren Drittwirkung zuordnen, ist sie dem Unionsrecht doch nicht unbekannt. So bringt etwa die DS-GVO in einzelnen Vorschriften die Grundrechte direkt in das Privatrechtsverhältnis ein.⁴⁰

Eine schwächere Grundrechtsbindung der Plattformen scheint der Verordnungsgeber dagegen bei der Normsetzung vorzusehen. Während Art. 12 Abs. 1 DSA-E in einer früheren Version des Europäischen Parlaments die Anbieter auch beim Aufstellen der Kommunikationsregeln ausdrücklich auf die Unionsgrundrechte verpflichtete,⁴¹ wurde diese Rege-

39 Appelman/Quintais/Fahy, *Using Terms and Conditions to apply Fundamental Rights to Content Moderation*, *VerfBlog* 2021, <https://verfassungsblog.de/power-dsa-dma-06>.

40 Die Rechtmäßigkeit einer Datenverarbeitung wird von der Abwägung mit den Grundrechten der betroffenen Person (Art. 6 Abs. 1 S. 1 lit. f DS-GVO), das „Recht auf Vergessenwerden“ von der Erforderlichkeit der Verarbeitung zur Ausübung des Rechts auf freie Meinungsäußerung und Information (Art. 17 Abs. 3 lit. a DS-GVO) abhängig gemacht. Darauf hinweisend auch Denga, *Plattformregulierung durch europäische Werte: Zur Bindung von Meinungsplattformen an EU-Grundrechte*, *EuR* 2021, 569 (594), der allerdings Art. 12 DSA-E als ein rechtstechnisches Novum betrachtet, weil dieser im Gegensatz zu den Normen der DS-GVO eine einseitige Grundrechtsbindung Privater statuiere. Darin scheint er eine sekundärrechtlich unzulässige Ausgestaltung der primärrechtlich verankerten Grundrechtsfunktionen zu erkennen. Art. 12 DSA-E lässt jedoch zumindest im Wege der grundrechtskonformen Auslegung eine Abwägung zu → V. 2.

41 Siehe Europäisches Parlament P9_TA(2022)0014, Abänderung 513.

lung in der aktuellen Fassung in die Erwägungsgründe verschoben und hat damit eine normative Abwertung erfahren. Nunmehr weist lediglich ErwG 38b S. 1 DSA-E darauf hin, dass die Diensteanbieter bei der Konzeption ihrer Kommunikationsregeln die Grundrechte der Rezipienten des Netzwerks beachten sollten. ErwG 38b S. 2 DSA-E hebt beispielhaft sehr große Online-Plattformen hervor – d.h. solche mit einer durchschnittlichen Anzahl an Nutzern von mindestens 45 Millionen und einer entsprechenden Einstufung durch die Kommission (Art. 25 Abs. 1 DSA-E) –, welche Meinungs- und Informationsfreiheit in besonderem Maße berücksichtigen sollten. Diese grundrechtlichen Wertungen können mittelbar im Rahmen von Art. 12 Abs. 1 DSA-E einfließen, der den Netzwerkanbietern vorschreibt, jegliche Beschränkungen der Nutzung ihrer Plattformen verbindlich in den allgemeinen Geschäftsbedingungen festzuschreiben.

V. Betroffene Grundrechtspositionen bei der Content-Moderation

Um die Grenzen der Normsetzung und -durchsetzung zu bestimmen, ist in einem weiteren Schritt herauszuarbeiten, welche Grundrechte in den normativen Rahmen von Art. 12 DSA-E einzubeziehen sind.

1. Grundrechte des Urhebers eines Beitrags

Ausdrücklich adressiert werden von Art. 12 Abs. 2 DSA-E und ErwG 38b die Rechte der Nutzer und die Meinungs- und Medienfreiheit. Dies liegt auch nahe, führen die Community Standards und deren Durchsetzung doch offensichtlich zu einer Beschränkung der Äußerungsmöglichkeiten von Nutzern, was insbesondere die Urheber eines Beitrags einschließt.

a) Genereller Schutz durch die Kommunikationsfreiheiten des Art. 11 Abs. 1 GRC

Aufseiten der Nutzer spielt das in Art. 11 Abs. 1 GRC garantierte Recht der freien Meinungsäußerung eine zentrale Rolle. Erfasst wird sowohl die Meinungsfreiheit als auch die Freiheit, Informationen und Ideen zu empfangen und weiterzugeben. Die Umschreibung des Regelungsbereichs mit der Trias aus Meinungen, Informationen und Ideen deutet bereits auf einen umfassenden Schutz hin, der sich nicht nur auf Werturteile, sondern auch

auf Tatsachenbehauptungen erstreckt.⁴² Neben der Äußerung der eigenen Meinung werden auch jegliche Formen der Verbreitung geschützt, was die Online-Kommunikation in sozialen Netzwerken umfasst.⁴³ Auch das Verbreiten und Teilen fremder Inhalte – etwa durch *Reposten* oder *Hyperlinks* – stellt ein von Art. 11 Abs. 1 GRC geschütztes Verhalten dar.⁴⁴ Inhalt und Qualität der Äußerung sind dabei unerheblich,⁴⁵ sodass auch verletzend, schockierende oder beunruhigende Äußerungen in den Schutzbereich fallen.⁴⁶

Bestimmte Formen von Hassrede könnten jedoch unter Anwendung des Missbrauchsverbots in Art. 54 GRC vom Schutzbereich des Art. 11 Abs. 1 GRC auszunehmen sein. Zwar fehlt zu dieser Frage bisher eine Stellungnahme des EuGH.⁴⁷ Mit Verweis auf das konventionsrechtliche Pendant in Art. 17 EMRK hat aber der EGMR klargestellt, dass antisemitische und antimuslimische Äußerungen, die Leugnung des Holocausts sowie die Rechtfertigung von Kriegsverbrechen keinen Schutz durch die Meinungsfreiheit genießen.⁴⁸ Der Schutz soll jedoch nur ausnahmsweise und in extremen Fällen versagt werden.⁴⁹

Diese Vorgehensweise des EGMR wird aber zu Recht dahingehend kritisiert, dass letztlich unklar bleibt, welche konkreten Anforderungen erfüllt sein müssen, um einer Äußerung den Schutz durch die Meinungsfrei-

42 Von der Groeben/Schwarze/Hatje, Europäisches Unionsrecht/Augsberg, 7. Aufl. 2015, Art. 11 GRC Rn. 6; GA Trstenjak 24.11.2010 – C-316/09, ECLI:EU:C:2010:712 Rn. 81 = BeckRS 2010, 91345 – Sharp; so auch EGMR 8.7.1986 (Pl.) – 9815/82 Rn. 46 – Lingens/Österreich.

43 Vgl. EuGH 6.11.2003 – C-101/01, ECLI:EU:C:2003:596 Rn. 86 = EuZW 2004, 245 – Lindqvist; siehe Heselhaus/Nowak, Handbuch der europäischen Grundrechte/Kühling, 2. Aufl. 2020, § 27 Rn. 18; Meyer/Hölscheidt, Charta der Grundrechte der Europäischen Union/Bernsdorff, 5. Aufl. 2019, Art. 11 Rn. 12.

44 GA Trstenjak 24.11.2012 – C-316/09, ECLI:EU:C:2010:712 Rn. 81 = BeckRS 2010, 91345 – Sharp; speziell zum Setzen von Hyperlinks EuGH 8.9.2016 – C-160/15, ECLI:EU:C:2016:644 Rn. 45 – GS Media.

45 Calliess/Ruffert, EUV/AEUV/Calliess, 6. Aufl. 2022, Art. 11 GRC Rn. 5; vgl. auch die Rechtsprechung des EGMR 7.12.1976 (Pl.) – 5493/72, BeckRS 1976, 107942 Rn. 49 – Handyside/Vereinigtes Königreich; EGMR 26.04.1979 (Pl.), 6538/74, BeckRS 1979, 108526 Rn. 65 – Sunday Times/Vereinigtes Königreich.

46 EuGH 6.3.2001 – C-274/99, ECLI:EU:C:2001:127 = EuR 2001, 542 – Connolly im Anschluss an EGMR 7.12.1976 (Pl.) – 5493/72, BeckRS 1976, 107942 Rn. 49 – Handyside/Vereinigtes Königreich.

47 Siehe Struth, Hassrede und Freiheit der Meinungsäußerung, 2019, S. 109 ff.

48 Instrukтив EGMR 6.1.2011 (GK) – 34932/04, NVwZ 2011, 1307 (1308 f.) mwN aus der Rspr – Paksas/Litauen.

49 EGMR 6.1.2011 (GK) – 34932/04, NVwZ 2011, 1307 (1308.) – Paksas/Litauen; EGMR 15.10.2015 (GK) – 27510/08, NJW 2016, 3353 (3354) – Perincek/Schweiz.

heit zu verwehren.⁵⁰ Vorzugswürdig erscheint es, die Missbräuchlichkeit der Äußerung erst im Rahmen der Rechtfertigung zu berücksichtigen. Die Frage der Schutzwürdigkeit einer Äußerung ist keine des Schutzbereichs, sondern auf der Schrankenebene anzusiedeln.⁵¹ Die Versagung des Schutzes würde außerdem dazu führen, dass ein Eingriff auch im Hinblick auf Art und Schwere nicht mehr auf seine Verhältnismäßigkeit hin überprüft würde.⁵² Demnach bleibt zu hoffen, dass sich der EuGH insoweit nicht dem EGMR anschließt, entsprechende Äußerungen also zunächst in den Schutzbereich des Art. 11 Abs. 1 GRC einbezieht.

b) *Beeinträchtigung der Medienfreiheit durch Sperrung publizistisch aktiver Nutzer*

Nicht abschließend geklärt ist, inwieweit sich die Nutzer sozialer Medien auch auf die in Art. 11 Abs. 2 GRC garantierte Medienfreiheit berufen können. Neben den klassischen Medien wie Presse, Rundfunk und Film werden von jener auch moderne digitale Formen der Massenkommunikation erfasst.⁵³ Demzufolge muss sich der Nutzer mit seinen Beiträgen an die Allgemeinheit richten.⁵⁴ Voraussetzung dafür ist, dass er seine Inhalte öffentlich zugänglich macht, sodass sie grundsätzlich einem unbestimmten Personenkreis, also zumindest allen registrierten Nutzern, offenstehen.⁵⁵

50 So etwa Meyer-Ladewig/Nettesheim/von Raumer, EMRK/Neidhardt, 4. Aufl. 2017, Art. 17 Rn. 2; Struth, Hassrede und Freiheit der Meinungsäußerung, 2019, S. 105; ähnlich Pabel/Schmahl, Internationaler Kommentar zur Europäischen Menschenrechtskonvention/Schiedermaier, 28. EL 2020, Art. 10 Rn. 29.

51 Pabel/Schmahl, Internationaler Kommentar zur Europäischen Menschenrechtskonvention/Schiedermaier, 28. EL 2020, Art. 10 Rn. 29; Ehlers, Europäische Grundrechte und Grundfreiheiten/Pünder, 4. Aufl. 2014, § 17 Rn. 10.

52 Cannie/Voorhoof, The Abuse Clause and Freedom of Expression in the European Human Rights Convention: an Added Value for Democracy and Human Rights Protection?, Netherlands Quarterly of Human Rights 2011, 54 (70).

53 Ehlers, Europäische Grundrechte und Grundfreiheiten/Pünder, 4. Aufl. 2014, § 17 Rn. 15; Grabenwarter, Europäischer Grundrechtsschutz/Walter, 2014, § 12 Rn. 18.

54 Meyer/Hölscheidt, Charta der Grundrechte der Europäischen Union/Bernsdorff, 5. Aufl. 2019, Art. 11 Rn. 18; siehe auch EuGH 14.2.2019 – C-345/17, ECLI:EU:C:2019:122 Rn. 53 = EuZW 2019, 299 – Buivids zum datenschutzrechtlichen Medienprivileg (Art. 9 DS-RL, nunmehr Art. 85 Abs. 2 DS-GVO).

55 Vgl. Giere, Grundrechtliche Einordnung sozialer Netzwerke vor dem Hintergrund des Netzwerkdurchsetzungsgesetz (NetzDG), 2021, S. 82.

In Abgrenzung zur Meinungsfreiheit ist die Medienfreiheit weiterhin nur betroffen, soweit es um mit Presse und Rundfunk vergleichbare medien-spezifische Vermittlungsleistungen geht.⁵⁶ Steht hingegen der Inhalt eines einzelnen Beitrags im Vordergrund, ist die Meinungsfreiheit anzuwenden.⁵⁷ Das Verbot von Hate Speech und daran anschließende Durchsetzungsmechanismen wie das Labeling, die Sichtbarkeitsreduzierung und Löschung von Beiträgen zielen klar auf den Äußerungsinhalt ab. Anders verhält es sich dagegen bei der Sperrung von Nutzerkonten. Diese knüpft zwar in der Regel auch an einen Verstoß gegen die Kommunikationsregeln an und weist damit eine inhaltliche Dimension auf. Durch den Ausschluss aus dem Netzwerk wird dem Nutzer aber insbesondere die von der Medienfreiheit geschützte Infrastruktur entzogen, die für die Publikation seiner Inhalte erforderlich ist.⁵⁸

Letztlich dürften sich aber neben klassischen Medienschaffenden nur solche Nutzer auf die Medienfreiheit berufen können, welche die Plattform für eine dem Journalismus ähnelnde Tätigkeit nutzen, was etwa Formen des Laienjournalismus wie Blogging und andere publizistische Betätigungen umfassen kann.⁵⁹

2. Grundrechte der Netzwerkbetreiber

Wenngleich Art. 12 Abs. 2 DSA-E die Plattformen auf den ersten Blick einseitig auf die Grundrechte der Nutzer zu verpflichten scheint, spricht die Norm doch in abgeschwächter Form nur von „berücksichtigen“ (Abs. 2). Insbesondere das Erfordernis der Verhältnismäßigkeit legt nahe, dass die Vorschrift Raum für eine Abwägung lassen soll, über die auch die Interes-

56 Jarass, Charta der Grundrechte der EU, 4. Aufl. 2021, Art. 11 Rn. 17; instruktiv unter Hinweis auf die Abgrenzungsschwierigkeiten zur individuellen Meinungsäußerungsfreiheit Grabenwarter, Europäischer Grundrechtsschutz/Walter, 2014, § 12 Rn. 20.

57 Grabenwarter, Europäischer Grundrechtsschutz/Walter, 2014, § 12 Rn. 20.

58 Vgl. Grabenwarter, Europäischer Grundrechtsschutz/Walter, 2014, § 12 Rn. 20.

59 Vgl. Stern/Sachs, Europäische Grundrechtecharta/von Coelln, 2016, Art. 11 Rn. 38; siehe auch EuGH 14.2.2019 – C-345/17, ECLI:EU:C:2019:122 Rn. 52 = EuZW 2019, 299 – Buivids, der für das datenschutzrechtliche Medienprivileg (Art. 9 DS-RL, nunmehr Art. 85 Abs. 2 DS-GVO) klargestellt hat, dass jenem nicht nur Medienunternehmen unterfallen, sondern jede journalistisch tätige Person.

sen der Netzbetreiber zur Geltung kommen können.⁶⁰ Da die Netzanbieter selbst Träger von Grundrechten sind, ist dies jedenfalls im Wege einer grundrechtskonformen Auslegung auch geboten.

a) Grundrechtsfähigkeit der Plattformen

Als privatrechtlich ausgestaltete Organisationen können sich soziale Netzwerke auf jene Grundrechte der Charta berufen, die begrifflich nicht an einen Menschen als Rechtsträger anknüpfen.⁶¹ Dass sich das Aktionszentrum der Plattformen dabei regelmäßig in den USA befindet, steht deren Grundrechtsfähigkeit prinzipiell nicht entgegen. Die GRC beschränkt diese – anders als Art. 19 Abs. 3 GG – nicht auf juristische Personen des Privatrechts mit Sitz im europäischen Inland.⁶² In der Folge zählen auch private Vereinigungen aus Drittstaaten zu den Grundrechtsberechtigten, sofern sie durch das Unionsrecht in vergleichbarer Weise betroffen sind wie innereuropäische Unternehmen.⁶³

b) Vollautomatisierte Moderation als grundrechtsrelevantes Verhalten

Bei der vollautomatisierten Moderation fehlt es im konkreten Fall an einer unmittelbaren menschlichen Leistung, sodass sich bereits die Frage stellt, ob diese dem Netzbetreiber als grundrechtlich relevantes Verhalten zugerechnet werden kann.⁶⁴ Der EuGH stellt bei automatisierten Vorgängen darauf ab, ob der Anwender das technische Mittel in Kenntnis seiner

60 Anders Denga, Plattformregulierung durch europäische Werte: Zur Bindung von Meinungsplattformen an EU-Grundrechte, EuR 2021, 569 (594), der von einer „einseitigen Grundrechtsbindung von Meinungsplattformen“ ausgeht.

61 EuGH 22.12.2010 – C-279/09, ECLI:EU:C:2010:811 Rn. 39 = BeckRS 2010, 91492 – DEB.

62 Eifert/Gostomzyk, Netzwerkrecht/Müller-Terpitz, 2018, S. 45 (51).

63 Heselhaus/Nowak, Handbuch der Europäischen Grundrechte/Nowak, 2. Aufl. 2020, § 2 Rn. 18; vgl. etwa EuGH 31.1.2019 – C-225/17 P, ECLI:EU:C:2019:82 Rn. 100 ff. = BeckRS 2019, 637 – Islamic Republic of Iran Shipping Lines.

64 Vgl. Milstein/Lippold, Suchmaschinenenergebnisse im Lichte der Meinungsfreiheit der nationalen und europäischen Grund- und Menschenrechte, NVwZ 2013, 182 (184 f.); siehe auch Schiedermaier/Weil, Online-Intermediäre als Träger der Meinungsfreiheit – Eine Betrachtung der deutschen, europäischen und amerikanischen Judikatur zum Content Management von Webplattformen, DÖV 2022, 305 (308 f.).

Folgewirkung einsetzt⁶⁵ oder der Vorgang seiner Kontrolle unterliegt.⁶⁶ Für den Einsatz der besagten KI-Systeme durch soziale Netzwerke lässt sich dies zweifellos bejahen. Schließlich werden diese von den Plattformen zur Durchsetzung ihrer Kommunikationsstandards implementiert und zu diesem Zweck kontinuierlich überwacht und angepasst.⁶⁷

c) *Schutz durch die unternehmerische Freiheit, Art. 16 GRC*

In jedem Fall ist aufseiten der Plattformen die von Art. 16 GRC geschützte unternehmerische Freiheit einzustellen. Diese umfasst die Freiheit, eine wirtschaftliche oder geschäftliche Tätigkeit auszuüben, den freien Wettbewerb und die Vertragsfreiheit.⁶⁸ Letztere schließt auch die freie Gestaltung von Vertragsinhalten ein.⁶⁹ Da es sich bei den Kommunikationsregeln um einen Bestandteil der Nutzungsverträge handelt, fällt das Aufstellen und die Durchsetzung der Community Standards bereits aus diesem Grund unter die unternehmerische Freiheit. Darüber hinaus basiert das Geschäftsmodell der Online-Dienste maßgeblich auf Einnahmen durch Werbung, die regelmäßig mithilfe der erhobenen Nutzerdaten direkt an die jeweilige Zielgruppe gerichtet wird (sog. Targeting).⁷⁰ Deswegen liegt es im wirtschaftlichen Interesse der Anbieter, möglichst viele aktive Nutzer zu akquirieren. Vor diesem Hintergrund verfolgen die Kommunikationsregeln den Zweck, das Netzwerk attraktiv für Nutzer und Werbekunden zu gestalten.

65 Vgl. EuGH 17.6.2021 – C-597/19, ECLI:EU:C:2021:492 Rn. 48 f. = GRUR 2021, 1067 – M.I.C.M. zur Zugänglichmachung iSv Art. 3 Abs. 1 u. 2 RL 2001/29/EG.

66 Dazu Milstein/Lippold, Suchmaschinenergebnisse im Lichte der Meinungsfreiheit der nationalen und europäischen Grund- und Menschenrechte, NVwZ 2013, 182 (186) unter Bezug auf EuGH 23.3.2010 – C-236/08 u. a., ECLI:EU:C:2010:159 Rn. 51–57 = GRUR 2010, 445 – Google France.

67 Vgl. Bradford u. a., Report Of The Facebook Data Transparency Advisory Group, 2019, S. 11 f.

68 EuGH 22.1.2013 – C-283/11, ECLI:EU:C:2013:28, Rn. 42 = EuZW 2013, 347 – Sky Österreich; EuGH 17.10.2013 – C-101/12, ECLI:EU:C:2013:661 Rn. 25 = BeckRS 2013, 81980 – Schaible; EuGH 30.6.2016 – C-134/15; ECLI:EU:C:2016:498 Rn. 28 = BeckRS 2016, 81408 – Lidl.

69 Jarass, Charta der Grundrechte der EU, 4. Aufl. 2021, Art. 16 Rn. 10 mN aus der Rspr.

70 Tief, Kommunikation auf Facebook, Twitter & Youtube, 2020, S. 31.

So zielt etwa das Verbot von Hate Speech darauf ab, einem verrohten Umgangston vorzubeugen, der abschreckend auf jene wirken könnte.⁷¹

d) *Schutz durch die Meinungsfreiheit des Art. 11 Abs. 1 GRC*

Ob sich die Diensteanbieter auch auf die Kommunikationsfreiheiten berufen können, lässt sich indes weniger einfach beantworten.⁷² Im Hinblick auf die Verpflichtung eines sozialen Netzwerks, einen Content-Filter zur Verhinderung von Urheberrechtsverletzungen zu implementieren, hat der EuGH jedenfalls nur auf die unternehmerische Freiheit der Plattform abgestellt, ohne sich mit der Meinungsfreiheit derselben auseinanderzusetzen.⁷³ Der EGMR lässt Online-Diensten dagegen generell den Schutz von Art. 10 Abs. 1 EMRK zukommen, wenn und weil sie Mittel für andere zur Verfügung stellen, Meinungen und Informationen zu teilen und zu empfangen.⁷⁴

Allein wegen der von Art. 11 GRC vorgesehenen Aufgliederung in Meinungs- und Informationsfreiheit auf der einen und einem speziellen Mediengrundrecht auf der anderen Seite erscheint im Rahmen der Grundrechtecharta aber eine differenzierte Betrachtungsweise geboten.⁷⁵ Zunächst lassen sich die Kommunikationsstandards als Meinungsäußerungen iSv Art. 11 Abs. 1 GRC einordnen, da sie in wertender Weise zum Ausdruck bringen, welche Inhalte oder Äußerungsweisen die Diensteanbieter auf ihren Plattformen missbilligen.⁷⁶ Dass die Regeln zumeist aus einem kom-

71 Zu allem BGH 29.7.2021 – III ZR 179/20, ZUM 2021, 953 Rn. 73; BGH 29.7.2021 – III ZR 192/20, ZUM-RD 2021, 612 Rn. 85 in Bezug auf Art. 5 Abs. 1 S. 1 Alt. 1 GG.

72 Ausf. dazu Schiedermaier/Weil, Online-Intermediäre als Träger der Meinungsfreiheit – Eine Betrachtung der deutschen, europäischen und amerikanischen Judikatur zum Content Management von Webplattformen, DÖV 2022, 305.

73 EuGH 16.2.2012 – C-360/10, ECLI:EU:C:2012:85 Rn. 44 = BeckRS 2012, 80348 – SABAM.

74 EGMR 18.12.2012 – 3111/10 Rn. 49 f. – Ahmet Yildirim/Türkei; EGMR 19.2.2013 – 40397/12, GRUR Int 2013, 476 (478) – Neij u. Sunde Kolmisoppi/Schweden; EGMR 19.3.2019 – 43624/14, NJW 2020, 2093 Rn. 68 – Høiness/Norwegen; EGMR 20.1.2020 (GK) – 201/17 Rn. 86–88 – Magyar Kétfarkú Kutya Párt/Ungarn.

75 Ob es sich bei Art. 11 Abs. 2 GRC um ein eigenständiges Grundrecht handelt, ist umstr. Dafür sprechen aber neben der separaten Stellung auch teleologische und entstehungsgeschichtliche Gründe, siehe dazu Stern/Sachs, Europäische Grundrechtecharta/von Coelln, 2016, Art. 11 Rn. 29.

76 Vgl. BGH 29.7.2021 – III ZR 179/20, ZUM 2021, 953 Rn. 74; BGH 29.7.2021 – III ZR 192/20, ZUM-RD 2021, 612 Rn. 86 zu Art. 5 Abs. 1 S. 1 Var. 1 GG.

merziellen Interesse heraus erlassen werden dürften, ist dabei unschädlich, da auch wirtschaftliche Äußerungen dem Schutzbereich unterfallen.⁷⁷ Wird nun ein einzelner Beitrag moderiert, um die Nutzungsbedingungen durchzusetzen, ist darin folgerichtig ebenso ein Werturteil zu sehen.⁷⁸ Selbiges gilt für die Sperrung von Nutzeraccounts, soweit diese als Reaktion auf Verstöße gegen die Kommunikationsregeln erfolgt. Schließlich drückt sich darin die Wertung aus, dass das Kommunikationsverhalten des Nutzers nicht länger geduldet wird.⁷⁹

e) *Schutz durch die Medienfreiheit nach Art. 11 Abs. 2 GRC*

Darüber hinaus erscheint es plausibel, dass sich die Netzwerkbetreiber ebenfalls auf die von Art. 11 Abs. 2 GRC gewährleistete Medienfreiheit berufen können. Zunächst richten sich soziale Netzwerke grundsätzlich an einen unbestimmten Personenkreis, stellen also ein massenkommunikatives Forum zur Verfügung. So können zumindest öffentliche Beiträge jederzeit und von allen registrierten Nutzern parallel eingesehen und kommentiert werden.⁸⁰

Da die Funktion eines sozialen Netzwerks vornehmlich darin besteht, eine Plattform für Inhalte Dritter zur Verfügung zu stellen, könnte dagegen die bereits besprochene medienspezifische Vermittlungsleistung in Frage gestellt werden. Schließlich reicht eine bloß technische Vermittlung dafür nicht aus. Vielmehr ist eine gewisse inhaltliche Zusatzleistung erforderlich, die allerdings nicht das Niveau einer redaktionellen Aufbereitung im klassischen Sinne erreichen muss.⁸¹ Im Fall der Meinungsplattformen wird dies durch die Implementierung und Durchsetzung von Kommunikationsregeln geleistet. Die Anbieter wirken dadurch auf die Kommunikationsweise der Nutzer ein und steuern so einen massenkommunikativen

77 EuGH 17.12.2015 – C-157/14, ECLI:EU:C:2015:823 Rn. 64 = BeckRS 2015, 82002 – Neptune Distribution.

78 Vgl. BGH 29.7.2021 – III ZR 179/20, ZUM 2021, 953 Rn. 74; BGH 29.7.2021 – III ZR 192/20, ZUM-RD 2021, 612 Rn. 86 zu Art. 5 Abs. 1 S. 1 Var. 1 GG.

79 Kellner, Die Regulierung der Meinungsmacht von Internetintermediären, 2019, S. 102.

80 Giere, Grundrechtliche Einordnung sozialer Netzwerke vor dem Hintergrund des Netzwerkdurchsetzungsgesetzes (NetzDG), 2021, S. 82; ausf. dazu und speziell zur automatisierten und personalisierten Vorsortierung von Inhalten durch die Netzwerke S. 77–86.

81 Heselhaus/Nowak, Handbuch der europäischen Grundrechte/Kühling, 2. Aufl. 2020, § 28 Rn. 14 f.

Prozess, womit die Verbreitung der Fremdbeiträge eine inhaltliche Komponente erhält.⁸² Auch der Einsatz KI-basierter Instrumente führt nicht zu einer technischen Verbreitung im obigen Sinne. Die Systeme als solche unterliegen zwar rein technischen Prozessen und Parametern, werden von den Betreibern aber gezielt zur Durchsetzung ihrer Nutzungsbedingungen eingesetzt, die sich wiederum im Ergebnis jedes einzelnen automatisierten Moderationsakts konkretisieren.⁸³

3. Einbeziehung der Grundrechte Dritter

Die „multipolare Gemengelage“ der digitalen Meinungsplattformen erfordert es zudem, die Grundrechte Dritter zu berücksichtigen,⁸⁴ weshalb auch der Verordnungsentwurf explizit die Rechte aller Beteiligten (Art. 12 Abs. 2 DSA-E) und sämtlicher Rezipienten der Plattform (ErwG 38b S. 1 DSA-E) miteinbezieht. Die Community Standards sozialer Netzwerke dienen nicht nur dem Interesse der Anbieter, sondern ebenso dem anderer Nutzer des Netzwerks.⁸⁵ Gleichermaßen können Dritte durch die Content-Moderation negativ betroffen sein.

a) Beeinträchtigung des Privatlebens durch Hate Speech

So kann Hate Speech unter Umständen das Recht auf Achtung des Privatlebens anderer Nutzer beeinträchtigen. Geschützt wird jenes von Art. 7

82 Siehe Kellner, Die Regulierung der Meinungsmacht von Internetintermediären, 2019, S. 138; ähnlich in der Begründung, aber mit Einordnung unter die Meinungsfreiheit iRd deutschen Grundgesetzes BGH 29.7.2021 – III ZR 179/20, ZUM 2021, 953 Rn. 74; BGH 29.7.2021 – III ZR 192/20, ZUM-RD 2021, 612 Rn. 86; für eine allgemeine Einordnung sozialer Netzwerke unter die Medienfreiheit als neuartige „Filtermedien“ Eifert/Gostomzyk, Netzwerkrecht/Ingold, 2018, S. 125 (128 ff.).

83 Vgl. W.D. Okla. 27.5.2003 – CIV-02–1457-M Rn. 6 – Search King, Inc. v. Google Technology, Inc. zum U.S. Const. amend. I.

84 Denga, Plattformregulierung durch europäische Werte: Zur Bindung von Meinungsplattformen an EU-Grundrechte, EuR 2021, 569 (593).

85 Lüdemann, Grundrechtliche Vorgaben für die Löschung von Beiträgen in sozialen Netzwerken, MMR 2019, 279 (282).

Abs. 1 GRC, der inhaltlich grundsätzlich Art. 8 EMRK entspricht.⁸⁶ Letzterer erfasst nach dem EGMR auch die Rechtspositionen der Betroffenen von Hate Speech.⁸⁷ Dabei brachte der Gerichtshof Art. 8 EMRK nicht nur in Stellung, wenn die Äußerung auf eine einzelne Person abzielte, sondern auch dann, wenn sie sich gegen eine ganze Gruppe richtete.⁸⁸ So könne durch die Herabsetzung einer ethnischen Gruppe sowohl deren Identitätsgefühl als auch der Selbstwert ihrer Mitglieder betroffen sein.⁸⁹ Auch mit Blick auf die Leugnung des Genozids an den Armeniern 1915/16 erkannte der Gerichtshof an, dass die durch die Äußerung beeinträchtigten Rechte in den Schutzbereich des Art. 8 EMRK fallen können.⁹⁰

b) *Hate Speech als Angriff auf die Menschenwürde*

Ebenso kann die von Art. 1 GRC geschützte Menschenwürde anderer Nutzer durch Hassrede berührt sein.⁹¹ Eine Konkretisierung des normativen Gehalts der Menschenwürde durch den EuGH fehlt bisher weitestgehend.⁹² Mit unterschiedlicher Begründung wird jedoch davon ausgegangen, dass auch im Rahmen der Charta die in der deutschen Rechtsprechung entwickelte „Objektformel“ angewandt werden kann.⁹³ Meinungs-

86 EuGH 14.2.2019 – C-345/17, ECLI:EU:C:2019:122 Rn. 65 = EuZW 2019, 299 – Buivids; siehe auch Schiedermaier, Der Schutz des Privaten als internationales Grundrecht, 2012, S. 344.

87 EGMR 15.3.2012 (GK) – 4149/04 u. a., NJOZ 2013, 378 Rn. 58 ff. – Aksu/Türkei; EGMR 22.4.2010 – 34050/05 Rn. 78 – Hagenauer/Frankreich.

88 Struth, Hassrede und Freiheit der Meinungsäußerung, 2019, S. 424.

89 EGMR 15.3.2012 (GK) – 4149/04 u. a., NJOZ 2013, 378 Rn. 58 – Aksu/Türkei.

90 EGMR 15.10.2015 (GK) – 27510/08, NJW 2016, 3353 Rn. 155 ff. – Perincek/Schweiz.

91 Art. 1 GRC stellt sowohl ein konstituierendes Bekenntnis der EU zur Menschenwürde als fundamentalen Wert als auch ein eigenständiges Grundrecht dar, siehe Calliess/Ruffert, EUV/AEUV/Calliess, 6. Aufl. 2022, Art. 1 GRC Rn. 2; explizit zur Menschenwürde als Grundrecht EuGH 18.12.2007 – C-341/05, ECLI:EU:C:2007:809 Rn. 94 = NZA 2008, 159 – Laval.

92 In der Rspr. des EuGH lassen sich bisher nur einzelfallbezogene Anforderungen an den Schutz der Menschenwürde finden, siehe zum Asylrecht EuGH 21.12.2011 – C-411/10 u. a., ECLI:EU:C:2011:865 Rn. 86 ff. = ZAR 2012, 115 – N.S. u. a.; zum europäischen Haftbefehl EuGH 5.4.2016 – C-404/15 u. a., ECLI:EU:C:2016:198 Rn. 90 = NJW 2016, 1709 – Aranyosi u. Căldăraru; ausf. Calliess/Ruffert, EUV/AEUV/Calliess, 6. Aufl. 2022, Art. 1 GRC Rn. 27 ff.

93 So etwa Calliess/Ruffert, EUV/AEUV/Calliess, 6. Aufl. 2022, Art. 1 GRC Rn. 36, der mit Verweis auf einen gemeinsamen Kern der verschiedenen Menschenwürde

äußerungen stellen deshalb zumindest dann einen Angriff auf die Menschenwürde dar, wenn sie den Menschen im Kern seiner Persönlichkeit treffen, indem er als minderwertig dargestellt und ihm das Lebensrecht bestritten wird.⁹⁴ Dabei kann auch die unwürdige Behandlung einer Gruppe die Würde einzelner Mitglieder verletzen.⁹⁵

c) *Meinungs- und Medienfreiheit der Inhabeanbieter*

Teilt ein Nutzer einen fremden Inhalt, kann auch die Meinungs- oder Medienfreiheit des dahinterstehenden Urhebers betroffen sein. Die Moderation eines solchen Beitrages durch ein soziales Netzwerk schränkt demnach nicht nur die Freiheit des Äußernden ein, sondern auch die Rechte des Inhabeanbieters aus Art. 11 GRC.⁹⁶

d) *Informationsfreiheit anderer Nutzer*

Die Content-Moderation kann daneben auch die Informationsfreiheit anderer Nutzer und der Öffentlichkeit beeinträchtigen (Art. 11 Abs. 1 GRC). Die Informationsfreiheit schützt die Nutzer in ihrem Recht, Informationen, Ideen und Meinungen zu empfangen,⁹⁷ was die Aufnahme von Informationen über das Internet umfasst.⁹⁸ Durch die Sperrung eines Beitrags oder Nutzers wird der (potenzielle) Zugang zu den damit verbundenen In-

dekonzepte der nationalen Teilrechtsordnungen und unter Heranziehung der EMRK als Rechtserkenntnisquelle die „Objektformel“ als richtungsweisend erachtet; ebenfalls mit Verweis auf die EMRK Pechstein/Nowak/Häde, Frankfurter Kommentar/Frenz, 1. Aufl. 2017, Art. 1 GRC Rn. 29, der zusätzlich die Rechtsprechung des EuGH zur Biopatentrichtlinie als Begründung heranzieht; anders von der Groeben/Schwarze/Hatje, Europäisches Unionsrecht/Augsberg, 7. Aufl. 2015, Art. 1 Rn. 4, der auf die Rechtsprechung des BVerfG Bezug nimmt.

94 Siehe BVerfG 25.3.2008 – 1 BvR 1753/03, NJW 2008, 2907 (2909).

95 Calliess/Ruffert, EUV/AEUV/Calliess, 6. Aufl. 2022, Art. 1 GRC Rn. 9.

96 Spiecker gen. Döhmann, A new framework for information markets – the ECJ Google Spain decision, Common Market Law Review 2015, 1033 (1046); vgl. dazu auch BVerfG 6.11.2019 – 1 BvR 276/17, NJW 2020, 314 Rn. 108 f. – Recht auf Vergessen II.

97 Stern/Sachs, Europäische Grundrechtecharta/von Coelln, 2016, Art. 11 Rn. 20.

98 Siehe EuGH 24.11.2011 – C-70/10, ECLI:EU:C:2011:771 Rn. 50 = MMR 2012, 174 – Scarlet; EuGH 16.2.2012 – C-360/10, ECLI:EU:C:2012:85 Rn. 48 = BeckRS 2012, 80348 – SABAM; EuGH 24.9.2019 – C-136/17, ECLI:EU:C:2019:773 Rn. 66 = EuZW 2019, 906 – GC u. a.; dazu auch EGMR 1.12.2015, 48226/10 u. a. Rn. 57 –

formationen jedoch versagt,⁹⁹ durch die Sichtbarkeitsreduzierung zumindest erschwert.

VI. Grenzen der Normsetzung und -anwendung im Horizontalverhältnis

Mit Ausnahme der normativen Abstufung zwischen Art. 12 Abs. 2 DSA-E und ErwG 38b (→ IV.3) enthalten die Bestimmungen keine weiteren Vorgaben dazu, welche Tragweite der Pflicht zur Berücksichtigung der Grundrechte seitens der Plattformen zukommen soll.¹⁰⁰ Inwieweit die Netzwerkanbieter in ihren Nutzungsbedingungen Hate Speech verbieten und einzelne Nutzerbeiträge darunter subsumieren dürfen, hängt demnach allein davon ab, wie die skizzierten Grundrechtspositionen zueinander ins Verhältnis zu setzen sind.

1. EGMR-Rechtsprechung zum staatlichen Verbot von Hate Speech

Mangels einer spezifischen Rechtsprechung des EuGH zu dieser Materie bietet sich einmal mehr ein Blick auf den Ansatz des EGMR an. So hat dieser Kriterien zur Bestimmung von Hate Speech entwickelt, bei deren Vorliegen staatlicherseits ein Verbot und grundsätzlich auch schwerwiegendere Maßnahmen gerechtfertigt sein können.¹⁰¹ Unterhalb dieser Schwelle lässt der Gerichtshof jedenfalls bei Äußerungen zu politischen Angelegenheiten und anderen Themen von öffentlichem Interesse wenig Raum für Einschränkungen,¹⁰² wenn und weil sie zur öffentlichen Meinungsbildung

Cengiz u. a./Türkei, der die Sperrung des Zugangs zu YouTube als Eingriff in die Informationsfreiheit wertet.

99 Vgl. zum Löschen von Einträgen bei Suchmaschinen EuGH 13.5.2014 – C-131/12, ECLI:EU:C:2014:317 Rn. 68 f. u. 81 = NJW 2014, 2257 – Google Spain; EuGH 24.9.2019 – C-136/17, ECLI:EU:C:2019:773 Rn. 53, 66 u. 68 = EuZW 2019, 906 – GC u. a.

100 Dazu Appelmann/Quintais/Fahy, Using Terms and Conditions to apply Fundamental Rights to Content Moderation, VerfBlog 2021, <https://verfassungsblog.de/power-dsa-dma-06>.

101 Siehe z. B. EGMR 17.7.2018 – 38004/12, NJOZ 2020, 50 Rn. 227 – Mariya Alekhina u. a./Russland; zu allem im Kontext antireligiöser Rede bereits Claes/Weil, Die Strafbarkeit antireligiöser Rede in der Judikatur des EGMR, EuGRZ 2021, 589 (595).

102 EGMR 15.10.2015 (GK) – 27510/08, NJW 2016, 3353 Rn. 197 – Perincek/Schweiz; EGMR 17.7.2018 – 38004/12, NJOZ 2020, 50 Rn. 227 – Mariya Alekhi-

beitragen und damit von zentraler Bedeutung für eine demokratische Gesellschaft im Sinne der Konvention sind.¹⁰³ Ähnliches gilt für die Kunstfreiheit, die der EGMR ebenfalls unter Art. 10 EMRK fasst.¹⁰⁴ Weniger schutzwürdig sind dagegen Äußerungen, die grundlos beleidigend erscheinen¹⁰⁵ oder rein kommerzieller Natur sind.¹⁰⁶

Unter Hate Speech verstehen die europäischen Richter dabei solche Äußerungen, die mit Elementen von Gewalt verbunden sind oder zu Hass und Intoleranz aufstacheln oder selbiges rechtfertigen.¹⁰⁷ Zur Beurteilung zieht der EGMR verschiedene Faktoren heran, wobei er herausstellt, dass für das Ergebnis in erster Linie das Zusammenspiel derselben entscheidend sei, nicht aber ein einziger Faktor als solcher.¹⁰⁸ Abzustellen ist zunächst darauf, ob die Äußerung vor einem angespannten politischen oder sozialen Hintergrund getätigt wurde. Gemeint sind dabei speziell Konstellationen, welche die innere Sicherheit und Ordnung eines Staates tangieren.¹⁰⁹ Zu beachten ist jedoch, dass Äußerungen in einem derartigen gesellschaftlichen Kontext regelmäßig politischer Natur sein dürften und ihnen damit nach dem skizzierten demokratisch-funktionalen Ansatz des EGMR ein besonderer Schutz zukommt.

In einem weiteren Schritt prüft der EGMR, ob die Äußerung bei fairer Auslegung und unter Beachtung ihres Kontexts als Aufstachelung zu oder Rechtfertigung von Gewalt, Hass oder Intoleranz verstanden werden kann. Ein besonderes Augenmerk legen die europäischen Richter darauf, ob

na u. a./Russland; EGMR 5.12.2019 – 13274/08 Rn. 37 – Tagiyev u. a./Aserbaidshan, jeweils mwN.

103 EGMR 23.5.1991 (Pl.) – 11682/85 Rn. 58 – Oberschlick/Österreich.

104 EGMR 22.10.2007 (GK) – 21279/02 u. a. Rn. 47 – Lindon u. a./Frankreich.

105 Siehe etwa EGMR 2.5.2006 – 50692/99, NVwZ 2007, 316 Rn. 23 – Aydın Tatlav/Türkei.

106 Vgl. EGMR 24.2.1994 – 15450/89 Rn. 50 – Casado Coca/Spanien.

107 EGMR 4.12.2003 – 35071/97 Rn. 51 – Gündüz/Türkei; EGMR 15.10.2015 (GK) – 27510/08, NJW 2016, 3353 Rn. 204 – Perincek/Schweiz; EGMR 17.7.2018 – 38004/12, NJOZ 2020, 50 Rn. 227 – Mariya Alekhina u. a./Russland; EGMR 5.12.2019 – 13274/08 Rn. 38 – Tagiyev u. a./Aserbaidshan; EGMR 6.10.2020 – 16435/10 Rn. 87 – Karastelev u. a./Russland.

108 EGMR 15.10.2015 (GK) – 27510/08, NJW 2016, 3353 Rn. 208 – Perincek/Schweiz.

109 Claes/Weil, Die Strafbarkeit antireligiöser Rede in der Judikatur des EGMR, EuGRZ 2021, 589 (595). So führt die Große Kammer des Gerichtshofs in EGMR 15.10.2015 (GK) – 27510/08, NJW 2016, 3353 Rn. 206 – Perincek/Schweiz Beispiele aus der Rspr. an, die innerstaatliche bewaffnete Konflikte, Gefängnisrevolten mit Todesopfern sowie migrationsbedingte Konflikte betreffen.

pauschal ganze ethnische, religiöse oder andere Gruppen angegriffen oder verunglimpft werden.¹¹⁰

Schlussendlich berücksichtigt der Gerichtshof Natur und Form der Äußerung sowie mögliche Folgeschäden, die direkt oder mittelbar aus der Meinungskundgabe resultieren können. Neben der Bedrohlichkeit der jeweiligen Ausdrucksform kann es dabei auf das Verbreitungspotential ankommen.¹¹¹ In diesem Sinne hat der EGMR für die Online-Kommunikation bereits betont, dass jene ein Gefährdungspotential in sich birgt, weil Inhalte in Sekundenschnelle weltweit verbreitet werden können und teils permanent online verfügbar bleiben. Auf der anderen Seite stelle die Möglichkeit für Nutzer, eigene Online-Inhalte zu generieren, eine präzedenzlose Form der Ausübung der Meinungsfreiheit dar.¹¹²

2. Übertragbarkeit auf das horizontale Grundrechtsverhältnis

Grundsätzlich lässt sich der Standard des EGMR auf den Rechtsrahmen der GRC übertragen, geht doch speziell die Konzeption von Art. 11 GRC klar auf Art. 10 EMRK zurück.¹¹³ Da insoweit das in Art. 10 Abs. 2 EMRK enthaltene Leitbild einer „demokratischen Gesellschaft“ Einzug finden kann, gilt dies ebenfalls für die vom EGMR angelegte demokratisch-funktionale Betrachtungsweise.¹¹⁴ Allerdings bezieht sich die Judikatur des EGMR in erster Linie auf die Beschränkung von Hassrede durch den Staat. Zu klären bleibt somit, inwieweit sich dieser Ansatz auf das Verhältnis

110 EGMR 15.10.2015 (GK) – 27510/08, NJW 2016, 3353 Rn. 206 – Perincek/Schweiz.

111 Claes/Weil, Die Strafbarkeit antireligiöser Rede in der Judikatur des EGMR, EuGRZ 2021, 589 (595); siehe EGMR 15.10.2015 (GK) – 27510/08, NJW 2016, 3353 Rn. 207 mwN – Perincek/Schweiz.

112 EGMR 16.6.2015 (GK) – 64569/09, NJW 2015, 2863 Rn. 110 – Delfi/Estland.

113 EuGH 14.2.2019 – C-345/17, ECLI:EU:C:2019:122 Rn. 65 = EuZW 2019, 299 – Buivids; Heselhaus/Nowak, Handbuch der europäischen Grundrechte/Kühling, 2. Aufl. 2020, § 27 Rn. 11.

114 Heselhaus/Nowak, Handbuch der europäischen Grundrechte/Kühling, 2. Aufl. 2020, § 27 Rn. 53. In diesem Sinne betont auch der EuGH, dass die Meinungsfreiheit eine wesentliche Grundlage einer Demokratie als in Art. 2 EUV verankertes Grundprinzip der Union darstelle, EuGH 21.12.2016 – C-203/15, ECLI:EU:C:2016:970 Rn. 93 = NJW 2017, 717 – Tele2 Sverige, und im Hinblick auf ihre gesellschaftliche Funktion betrachtet werden müsse, EuGH 12.6.2003 – C-112/00, ECLI:EU:C:2003:333 Rn. 80 = EuZW 2003, 592 – Schmidberger.

zwischen privaten Akteuren auf sozialen Netzwerken im Rahmen der GRC übertragen lässt.

a) *Vereinbarkeit mit dem Prinzip der praktischen Konkordanz*

Als grundlegendes Prinzip für die Güterabwägung im Horizontalverhältnis zieht der EuGH die praktische Konkordanz heran. Zwischen den widerstreitenden Grundrechten ist demnach ein angemessener Ausgleich herzustellen.¹¹⁵ Dabei darf weder eines der Grundrechte in qualifizierter Weise beeinträchtigt noch in seinem Wesensgehalt verletzt werden.¹¹⁶

Diesen Grundsätzen wird der Ansatz des EGMR durchaus gerecht. Durch die Möglichkeit, Gewaltaufrufe und die Aufstachelung zu Hass und Intoleranz zu verbieten, können soziale Netzwerke Äußerungsformen verhindern, die in besonderer Weise abschreckend und damit geschäftsschädigend wirken. Handelt es sich um bloß beleidigende oder wirtschaftlich motivierte Beiträge, sind die Anforderungen sogar herabzusetzen. Gleichmaßen werden die Interessen der von Hate Speech betroffenen Dritten hinreichend berücksichtigt, weil jedenfalls solche Äußerungen untersagt werden können, die einen Angriff auf die Menschenwürde darstellen oder grob ehrverletzend sind. Schließlich wird so auch die Grenze getroffen, an der es nicht unplausibel erscheint, dass die angegriffenen Nutzergruppen selbst wegen der abschreckenden Wirkung in der Ausübung ihrer Meinungsfreiheit gehemmt werden könnten.¹¹⁷

Auf der anderen Seite wird politischen und künstlerischen Äußerungen der gebotene Raum gegeben, sodass auch das vom EuGH aufgegriffene Diktum Straßburgs zur Geltung kommt, wonach nicht nur Meinungen geschützt seien, die Zustimmung erfahren oder als harmlos oder unerheblich betrachtet werden, sondern auch solche, die einen Teil der Bevölkerung beleidigen, aus der Fassung bringen oder stören.¹¹⁸ Die Meinungs-

115 EuGH 12.6.2003 – C-112/00, ECLI:EU:C:2003:333 Rn. 81 = EuZW 2003, 592 – Schmidberger; EuGH 13.5.2014 – C-131/12, ECLI:EU:C:2014:317 Rn. 81 = NJW 2014, 2257 – Google Spain; EuGH 17.4.2018 – C-414/16, ECLI:EU:C:2018:257 Rn. 80 = NZA 2018, 569 – Egenberger.

116 EuGH 16.7.2015 – C-580/13, ECLI:EU:C:2015:485 Rn. 35 = GRUR 2015, 894 – Coty Germany.

117 So für die Religionsfreiheit Temperman, *Blasphemy, Defamation of Religions & Human Rights Law*, Netherlands Quarterly of Human Rights 2008, 517 (535 f).

118 EuGH 6.3.2001 – C-274/99 P, ECLI:EU:C:2001:127 Rn. 39 = EuR 2001, 542 – Connolly.

und Medienfreiheit der Netzwerke muss in diesen Fällen wegen ihrer vornehmlich wirtschaftlichen Zwecksetzung hinter die demokratisch bedeutungssameren Formen der Rede zurücktreten.

b) *Staatsähnliche Bindung der Netzwerke*

Ein Problem könnte sich allerdings daraus ergeben, dass in der unveränderten Übertragung des EGMR-Standards eine staatsgleiche Bindung der Netzwerke an die Meinungsfreiheit der Nutzer liegen würde. Eine stärkere Inanspruchnahme der Diensteanbieter rechtfertigt sich aber über die genannten Gründe hinaus durch die herausgehobene kommunikative Bedeutung und Marktmacht der großen Plattformen. In vielen gesellschaftlichen Gruppen und Kontexten erfolgt ein erheblicher Teil der Kommunikation über soziale Netzwerke. Durch die Unterdrückung der Meinung oder gar den Ausschluss eines Nutzers wird dessen Teilhabe am öffentlichen und privaten Diskurs daher empfindlich beeinträchtigt.¹¹⁹

Zwar käme in Betracht, die betreffenden Nutzer auf die Möglichkeit der Nutzung anderer, ggf. weniger reglementierter Plattformen zu verweisen. Dem stehen aber Netzwerk- und Lock-in-Effekte entgegen. Der Großteil der Social-Media-Nutzer konzentriert sich auf wenige Plattformen, die zwar teils ähnliche Zielgruppen bedienen, sich aber speziell in den zur Verfügung stehenden Kommunikationsmitteln unterscheiden. Der Wechsel auf ein anderes Netzwerk geht damit nicht nur mit dem Verlust der bestehenden Kommunikation und dem geschaffenen Bestand an sozialen Kontakten, sondern auch mit dem der spezifischen Kommunikationsweise einher.¹²⁰

Aus dieser sozialen Abhängigkeit resultiert eine strukturelle Überlegenheit der Plattformen gegenüber ihren Nutzern, die bei der Abwägung Berücksichtigung finden muss.¹²¹ Dies entspricht auch der Normstruktur von

119 Raue, Meinungsfreiheit in sozialen Netzwerken, JZ 2018, 961 (966); ausf. dazu Pille, Meinungsmacht sozialer Netzwerke, 2016, S. 311 f.; siehe auch eurostat, Are you using social networks?, 2019, <https://ec.europa.eu/eurostat/web/products-eurostat-news/-/EDN-20190629-1>, wonach schon 2018 mehr als die Hälfte der Bevölkerung der EU-Mitgliedstaaten in sozialen Netzwerken aktiv war, in der Altersgruppe von 16 bis 24 Jahren sogar beinahe 90 Prozent.

120 Elsaß/Labusga/Tichy, Löschungen und Sperrungen von Beiträgen und Nutzerprofilen durch die Betreiber sozialer Netzwerke, CR 2017, 234 (235 u. 239); ähnlich Raue, Meinungsfreiheit in sozialen Netzwerken, JZ 2018, 961 (966).

121 Vgl. Raue, Meinungsfreiheit in sozialen Netzwerken, JZ 2018, 961 (966). Auch die Rechtsprechung des EuGH legt nahe, dass jener die beherrschende Stel-

Art. 12 DSA-E. Dass die Vorschrift den Akzent auf die Verpflichtung der Plattformen legt, dürfte gerade dem asymmetrischen Verhältnis von Netzwerk und Nutzern geschuldet sein.¹²² Um den Unterschied zwischen den grundrechtsverpflichteten Mitgliedstaaten und den grundrechtsberechtigten Netzwerkbetreibern jedoch nicht „völlig einzuebnen“,¹²³ ist den Plattformen ein weiter Beurteilungsspielraum hinsichtlich der Frage zuzusprechen, worin ein Gewaltaufruf oder eine Aufstachelung zu Hass und Intoleranz gesehen werden kann und ob die dafür heranzuziehenden Kriterien einschlägig sind.¹²⁴ Dies muss erst recht für den Bereich der Normsetzung gelten (Art. 12 Abs. 1 DSA-E), für den die Grundrechtsbindung anders als bei der Normanwendung (Art. 12 Abs. 2 DSA-E) nicht unmittelbar im Gesetzestext angeordnet wird, sondern nur im Rahmen der Erwägungsgründe Erwähnung findet (→ IV.3).

c) *Anwendbarkeit der Kriterien auf Private*

Weiterhin lässt sich zumindest für das Kriterium des angespannten gesellschaftlichen Hintergrunds die Übertragbarkeit auf das grundrechtliche Verhältnis zwischen Privaten bezweifeln. Schließlich speist sich dieser Faktor aus der Gefährdung der inneren Sicherheit eines Staates, welche die betreffende Äußerung verschärfen könnte. Bei der Wahrung der inneren Sicherheit handelt es sich aber um eine originär staatliche Aufgabe.¹²⁵ Dies spiegelt sich auch in der Schrankenbestimmung des Art. 10 Abs. 2 EMRK wider, die den Schutz der nationalen sowie öffentlichen Sicherheit und Ordnung als zulässige staatliche Eingriffsziele vorsieht. Trotzdem kann dieser Faktor auch im Horizontalverhältnis Anwendung finden. Dass auf diese Weise vor allem gewaltvolle und hasserfüllte Äußerungen, die einen Bezug auf eine ganz reale Gefahr aufweisen, unterdrückt werden können, entspricht eindeutig dem Interesse der betroffenen Drittnutzer. Ebenso

lung von Privaten entsprechend berücksichtigt, siehe Denga, Plattformregulierung durch europäische Werte: Zur Bindung von Meinungsplattformen an EU-Grundrechte, EuR 2021, 569 (586–588).

122 Siehe dazu auch das Impact Assessment zum DSA-E, Europäische Kommission SWD(2020) 348 final Rn. 51 ff. u. 77 ff.

123 Friehe, Löschen und Sperren in sozialen Netzwerken, NJW 2020, 1697 (1699).

124 Vgl. Denga, Plattformregulierung durch europäische Werte: Zur Bindung von Meinungsplattformen an EU-Grundrechte, EuR 2021, 569 (593).

125 Dazu etwa Isensee, Das Grundrecht auf Sicherheit, 1983, S. 17.

können sich die Netzwerkbetreiber so besonders abschreckender, weil gefahrenträchtiger Rede entledigen.

d) Verfahrensrechtliche Lösung als Alternative?

Für die deutsche Grundrechtsordnung hat der BGH jüngst eine vornehmlich verfahrensrechtliche Lösung für die vorliegende Konstellation entwickelt. Danach ist ein soziales Netzwerk zwar grundsätzlich frei darin, in seinen Nutzungsbedingungen bestimmte Inhalte zu verbieten.¹²⁶ Eine Sperrung dürfe aber nicht willkürlich erfolgen, sondern müsse auf einem sachlichen Grund und einem objektiv überprüfbaren Tatbestand beruhen.¹²⁷ Insbesondere legt der BGH verfahrensrechtliche Mindestanforderungen fest, die sich auch in den Geschäftsbedingungen widerspiegeln müssen. Im Wesentlichen hat das Netzwerk die Nutzer umgehend über die Maßnahme zu informieren, eine Begründung anzugeben und eine Möglichkeit zur Gegenäußerung mit anschließender Neubescheidung zu gewähren.¹²⁸

Durch diese formellen Anforderungen wird zwar eine grundlegende Absicherung der Meinungsfreiheit der Nutzer gewährleistet. Auch in Kombination mit dem materiell gelagerten Willkürverbot werden jene allein aber der besonderen demokratischen Bedeutung der Meinungsfreiheit nicht gerecht. Den Netzwerkbetreibern bliebe es damit weitestgehend unbenommen, sämtliche unliebsame Meinungen zu unterdrücken, sofern sich dafür ein sachlicher Grund – im Fall der Netzwerke in der Regel ein wirtschaftliches Interesse – anführen lässt. Indem die Diensteanbieter aber Foren des öffentlichen Austausches schaffen, müssen sie in einem gewissen Maß auch die damit verbundenen Kontroversen, Konflikte und unbequemen Meinungsäußerungen hinnehmen.¹²⁹ Letztlich scheint auch der DSA-E von einem stärker materiell angebundenen Ansatz auszugehen. Schließlich regelt der Entwurf das Verfahren in Form eines Melde- und Beschwerdemechanismus bereits in umfassender Weise (vgl. Art. 14 ff. DSA-E), ver-

126 BGH 29.7.2021 – III ZR 179/20, ZUM 2021, 953 Rn. 78; BGH 29.7.2021 – III ZR 192/20, ZUM-RD 2021, 612 Rn. 90.

127 BGH 29.7.2021 – III ZR 179/20, ZUM 2021, 953 Rn. 80 ff.; BGH 29.7.2021 – III ZR 192/20, ZUM-RD 2021, 612 Rn. 92 ff.

128 BGH 29.7.2021 – III ZR 179/20, ZUM 2021, 953 Rn. 83 ff.; BGH 29.7.2021 – III ZR 192/20, ZUM-RD 2021, 612 Rn. 95 ff.

129 Vgl. Raue, Meinungsfreiheit in sozialen Netzwerken, JZ 2018, 961 (967).

pflichtet die Plattformen in Art. 12 Abs. 2 DSA-E aber zusätzlich auf die Grundrechte der Nutzer.

3. Grenzen der Normdurchsetzung

Betrifft der dargestellte Standard des EGMR vor allem die Frage, unter welchen Voraussetzungen ein Verbot und damit verbundene Sanktionen überhaupt möglich sind, was im Rahmen von Art. 12 DSA-E vor allem das Aufstellen und die Anwendung der Kommunikationsregeln tangiert, stellt sich weiterhin die Frage, welche Maßnahme zur Durchsetzung derselben im Einzelfall zulässig ist. Insoweit kommt es vor allem auf die von Art. 12 Abs. 2 DSA-E geforderte Verhältnismäßigkeit an. Danach darf die Maßnahme nicht das überschreiten, was zur Erreichung der verfolgten Ziele geeignet und erforderlich ist, wobei die durch sie verursachten Nachteile nicht außer Verhältnis zu den angestrebten Zielen stehen dürfen.¹³⁰

Die „minimalinvasivste“ Form der anfangs skizzierten Moderationstechniken stellt das Labeling dar. Die Äußerung bleibt als solche unberührt, erfährt aber inhaltlich eine Abwertung. Deutlich eingriffsintensiver fällt dagegen die Sichtbarkeitsreduzierung aus, führt diese doch zu einer objektiven Beeinträchtigung der Meinungsfreiheit des betreffenden Nutzers und der Informationsfreiheit Dritter, indem die Reichweite des Beitrags verringert wird. Die schwerwiegendste inhaltsbezogene Maßnahme bildet die Löschung des Beitrags. Die betreffende Meinungsäußerung wird vollständig unterdrückt und kann auch von anderen Nutzern nicht mehr rezipiert werden.

Als besonders schwerwiegende Beeinträchtigung stellt sich die temporäre oder sogar unbefristete Sperrung eines Nutzerkontos dar. Dem Nutzer wird dadurch vollständig die Möglichkeit genommen, auf der Plattform eigene Meinungen kundzutun und fremde Meinungen zu rezipieren.¹³¹ Umgekehrt sind auch andere Nutzer daran gehindert, mit ihm zu kommunizieren und auf seine Inhalte zuzugreifen. Diese Form der Content-Moderation ist daher nur als *ultima ratio* im Fall von wiederholten oder besonders schwerwiegenden Verstößen gegen die Community Standards einzusetzen.¹³²

130 Siehe z. B. EuGH 15.2.2016 – C-601/15 PPU, ECLI:EU:C:2016:84 Rn. 54 = NVwZ 2016, 1789 – J.N.

131 Pille, Meinungsmacht sozialer Netzwerke, 2016, S. 314.

132 Vgl. Friehe, Löschen und Sperren in sozialen Netzwerken, NJW 2020, 1697 (1700).

Ein spezielles Problem tritt bei dem Einsatz von KI-Systemen zur automatisierten Moderation auf, weil jene die Gefahr des *Overblocking* mit sich bringen. Die gegenwärtigen Filtersysteme weisen Schwächen bei der Berücksichtigung des Kontexts einer Äußerung auf, der gerade für die Einordnung als Hate Speech von besonderer Relevanz ist. In der Folge kann es zu falsch-positiven Ergebnissen kommen, sodass auch an sich zulässige Inhalte entfernt werden.¹³³ Die KI-basierte Content-Moderation kann daher die Kommunikationsfreiheiten der Nutzer breitenwirksam beeinträchtigen.¹³⁴ Allerdings ist zu berücksichtigen, dass die automatisierte Filterung im Interesse der Netzbetreiber und geschädigten Nutzer erfolgt. Schließlich besteht darin eine effiziente und kostensparende Methode zur Überprüfung großer Mengen an Inhalten, die händisch nur schwer zu bewältigen wäre.¹³⁵ Sofern sich die Fehlerquote in einem zumutbaren Rahmen hält, bietet Art. 17 Abs. 5 DSA-E dafür eine interessengerechte Lösung. Danach darf zumindest die Entscheidung der Plattform über die gegen einen Moderationsakt gerichtete Beschwerde nicht allein mit automatisierten Mitteln getroffen werden.

Fällt die Fehlerrate jedoch zu hoch aus, greift die Risikominderungspflicht aus Art. 27 Abs. 1 DSA-E, die sich allerdings auf sehr große Online-Plattformen im Sinne von Art. 25 DSA-E beschränkt. Diese sind zunächst gem. Art. 26 Abs. 1 DSA-E dazu verpflichtet, zumindest einmal jährlich systemische Risiken zu evaluieren, die auf ihre algorithmischen Systeme zurückzuführen sind. Dazu zählen gem. Art. 26 Abs. 1 lit. b DSA-E auch negative Effekte für die Ausübung von Meinungs- und Informationsfreiheit, welche von einem Moderationssystem mit signifikanter Fehlerrate ausgehen können. Daran anschließend schreibt Art. 27 Abs. 1 DSA-E vor, dass die Plattformen verhältnismäßige und effektive Maßnahmen zur Minderung des identifizierten Risikos treffen müssen, was Anpassungen des Moderationsverfahrens (lit. c) und der algorithmischen Systeme (lit. ca) einschließt. Im Fall der automatisierten Moderation kommt dafür natürlich die Reduzierung der Fehlerrate in Betracht, aber auch die Implementierung anderer Schutzvorkehrungen. So könnte ein menschlicher Entschei-

133 Instrukтив Finck, *Artificial Intelligence and Online Hate Speech*, 2019, S. 5 ff. mwN.; Europäische Kommission SWD(2020) 348 final, Rn. 80; vgl. Keller, *Facebook Filters, Fundamental Rights, and the CJEU's Glawischnig-Piesczek Ruling*, GRUR Int. 2020, 616 (622).

134 Denga, *Plattformregulierung durch europäische Werte: Zur Bindung von Meinungsplattformen an EU-Grundrechte*, EuR 2021, 569 (583) mwN.

135 Vgl. Finck, *Artificial Intelligence and Online Hate Speech*, 2019, S. 6.

der bereits zum Zeitpunkt der Erstentscheidung eingebunden werden, um das maschinelle Ergebnis vorab zu überprüfen.

Da die Pflicht zur Risikoeermittlung grundsätzlich nur jährlich besteht und die Schwelle zur Klassifizierung einer Plattform als „sehr groß“ mit durchschnittlich 45 Millionen aktiven Nutzern pro Monat (Art. 25 Abs. 1 DSA-E) sehr hoch liegt, bestehen insoweit jedoch gewisse Schutzlücken. In Bezug auf Uploadfilter für Urheberrechtsverletzungen hat der EuGH bereits betont, dass diese nur mit Art. 11 Abs. 1 GRC vereinbar seien, wenn sichergestellt wird, dass sie hinreichend zwischen zulässigen und unzulässigen Inhalten unterscheiden können.¹³⁶ In diesem Sinne weist auch ErwG 38 DSA-E darauf hin, dass die Diensteanbieter bei der Verwendung automatisierter Instrumente die Fehlerquote so weit wie möglich begrenzen sollen. Ist das System nicht hinreichend akkurat und greifen die Art. 26 f. DSA-E aus den vorgenannten Gründen nicht, erscheint daher aus grundrechtlicher Sicht ein Korrektiv notwendig, das auf Art. 12 Abs. 2 DSA-E gestützt werden und etwa in der Einbindung eines menschlichen Entscheiders bei der Erstentscheidung liegen kann.

VII. Schlussbetrachtungen

Speziell mit Art. 12 Abs. 2 DSA-E kündigt sich eine bedeutsame Rechtsentwicklung an, welche die grundrechtliche Dimension des multipolaren Gefüges sozialer Netzwerke in das unionale Privatrecht inkorporiert. Verspricht die Regelung zwar eine erhebliche Verbesserung für die Rechtspositionen der strukturell unterlegenen Netzwerknutzer, ist sie angesichts ihrer potenziell weitreichenden Folgen doch recht vage formuliert.¹³⁷ Da die Vorschrift die Pflicht zur Berücksichtigung der Grundrechte, die sie den Plattformen auferlegt, nicht näher bestimmt, läge es momentan an der Rechtsprechung, ihren normativen Gehalt zu konkretisieren.

Im Hinblick auf Hate Speech liefert die Judikatur des EGMR insoweit einen guten Ausgangspunkt. Danach können die Plattformen Äußerungen, die mit Elementen von Gewalt verbunden sind oder zu Hass und

136 EuGH 26.4.2022 – C-401/19, ECLI:EU:C:2022:297 Rn. 86 = GRUR 2022, 820; siehe bereits EuGH 24.11.2011 – C-70/10, ECLI:EU:C:2011:771 Rn. 52 = MMR 2012, 174 – Scarlet; EuGH 16.2.2012 – C-360/10, ECLI:EU:C:2012:85 Rn. 50 = BeckRS 2012, 80348 – SABAM.

137 Ausf. dazu Appelman/Quintais/Fahy, Using Terms and Conditions to apply Fundamental Rights to Content Moderation, VerfBlog 2021, <https://verfassungsblog.de/power-dsa-dma-06>.

Intoleranz aufstacheln oder selbiges rechtfertigen, ohne Weiteres verbieten. Im Verhältnis von Privaten ist den grundrechtsberechtigten Netzbetreibern hierfür ein weiter Beurteilungsspielraum zuzusprechen. Abstrakt entsprechen die marktführenden Plattformen diesem Standard bereits jetzt, knüpfen sie doch ähnlich dem EGMR bei ihrer Definition von Hassrede speziell an einen Angriff auf die Zugehörigkeit zu einer bestimmten Gruppe an. Ob das auch für die konkrete Anwendung und Verhältnismäßigkeit der Durchsetzung im Einzelfall gilt, darf bezweifelt werden, wird doch immer wieder Kritik an der Moderationspraxis der Netzwerke laut.¹³⁸ Zuletzt sind dabei auch die zur Detektion eingesetzten KI-Systeme stärker in den Fokus geraten.¹³⁹ Im Einzelnen könnten sich daraus schwierige Abwägungsfragen ergeben, die Rechtsprechung und Rechtswissenschaft noch lange beschäftigen dürften.

-
- 138 Einen Überblick vermitteln etwa die Beispiele bei Electronic Frontier Foundation, <https://www.eff.org/de/tossedout>; Raue, Meinungsfreiheit in sozialen Netzwerken, JZ 2018, 961 (964); Pille, Meinungsmacht sozialer Netzwerke, 2016, S. 315 ff; zum speziellen Problem der unverhältnismäßigen Benachteiligung bestimmter Personengruppen etwa Haimson/Delmonaco/Nie/Wegner, Disproportionate Removals and Differing Content Moderation Experiences for Conservative, Transgender, and Black Social Media Users: Marginalization and Moderation Gray Areas, Proceedings of the ACM on Human-Computer Interaction 2021, 5(CSCW2), 1.
- 139 Zur übermäßigen Detektion von Beiträgen bestimmter Personengruppen etwa Ghaffary, The algorithms that detect hate speech online are biased against black people, Vox 2019, <https://www.vox.com/recode/2019/8/15/20806384/social-media-hate-speech-bias-black-african-american-facebook-twitter>. Auch die Leistungsfähigkeit der KI-Systeme wurde zuletzt durch die Auswertung interner Berichte von Facebook (sog. *Facebook Files*) in Zweifel gezogen, dazu Seetharaman/Horwitz/Scheck, Facebook Says AI Will Clean Up the Platform. Its Own Engineers Have Doubts, The Wall Street Journal 2021, <https://www.wsj.com/articles/facebook-ai-enforce-rules-engineers-doubtful-artificial-intelligence-11634338> 184.