

tungen der freiwilligen Selbstkontrolle. Allerdings zeigt sich immer deutlich, dass sich effektive und effiziente Regulierung digitaler Inhalte nicht automatisch aus den Steuerungserfahrungen der analogen Welt – nicht nur aufgrund der Globalität der Informationsflüsse im Internet – fortschreiben lässt. Vielmehr sind nicht selten spezifische Ansätze für die digitale Sphäre notwendig. Es scheint daher nur folgerichtig, wenn Küllmer und Breindl (2019: 335) das Internet als ein »Testgebiet für neue Arten von Regulationsmethoden« erkennen. Auch wenn dabei im globalen Internet die supranationale Ebene per definitionem eine wichtige Rolle zukommt, wird dennoch die »traditionell [...] zentrale Rolle [des Staates] bei der Regulierung von Inhalten« nicht verdrängt (ebd.). Auch bei der digitalen Inhaltsregulierung stellt die nationalstaatliche Ebene weiterhin einen wichtigen Ankerpunkt dar.

Allein aus historisch gewachsenen kulturellen Unterschieden oder aufgrund verschiedener politischer Ausrichtungen ist ein koordiniertes, einheitliches europäisches oder gar globales Vorgehen nicht immer einfach. Beispielhaft wird dies am Umgang mit Volksverhetzung und Meinungsfreiheit in Deutschland, verglichen mit der Redefreiheit [*freedom of speech*] in den USA (vgl. etwa Menkes 2022: 37). Genauso aber auch bei der, im Konflikt um Art. 17<sup>10</sup> der EU-Urheberrechtsrichtlinie 2019/790 [*Directive on Copyright in the Digital Single Market; DSM-RL*] deutlich gewordenen (vgl. Europäisches Parlament 2019), unterschiedlichen Schwerpunktsetzung aus französischer und deutscher Perspektive in der Abwägung zwischen den Interessen von Urheber:innen und Verwerter:innen auf der einen und einer möglichen Erschwerung neuer Geschäftsmodelle von Internetplattformen auf der anderen Seite.<sup>11</sup> Dementsprechend bieten EU-Richtlinien oftmals einen Umsetzungs- und Interpretationsspielraum, der den Nationalstaaten eine eigene Ausgestaltung, unter anderem aus Steuerungsperspektive, ermöglicht – etwa bei der Wahl des als geeignet wahrgenommenen Steuerungsinstruments. Nicht zuletzt können auch nationale Gesetze den entscheidenden Anreiz für Regulierungsbemühungen auf europäischer Ebene liefern oder als Blaupause für Steuerungskonzepte dienen. Dies zeigt sich auch bei der Plattformregulierung, die im Folgenden anhand des deutschen Netzwerkdurchsetzungsgesetzes (NetzDG) näher beleuchtet wird.

### V.3.1 Policy I: Netzwerkdurchsetzungsgesetz (NetzDG)

Im Kern zielt das NetzDG auf eine schnelle Entfernung (Löschung) oder Sperrung rechtswidriger Inhalte aus sozialen Netzwerken. Somit geht es um eine große Bandbreite an unterschiedlichen Aspekten, die bei der Frage nach Inhaltsregulierung eine Rolle spielen. Terroristische und kriminelle Inhalte bedrohen die nationale Sicherheit, gewaltverherrlichende und sexuelle Inhalte das Kindeswohl, Urheberrechtsverletzungen ökonomische Interessen.

10 Der am Richtlinienentwurf geführten Konflikt referenziert aufgrund der ursprünglichen Nummerierung zumeist Artikel 13.

11 Für die Debatten um Artikel 13 beziehungsweise 17 – die sich unter anderem um die Gefahr von Zensur durch Uploadfilter (siehe dazu auch Kapitel V.3.3.2) und eine weitere Privatisierung der Rechtsdurchsetzung (siehe Kapitel V.3.3.1) drehten – siehe etwa Gnauer (2019) und Askani (2021: 77ff.).

mische Interessen sowie Hassbotschaften und Verleumdung die Menschenwürde (vgl. Küllmer/Breindl 2019: 336).

»Das Internet ist kein rechtsfreier Raum, hier gelten die Gesetze genauso wie in der analogen Welt« – diese Aussage von Bundesjustizministerin Lambrecht (2020) ist alles andere als der übliche Standard. Immer wieder kolportieren Politiker:innen einen rechtsfreien Raum im Internet (wobei sie dabei eigentlich die auf diesem aufsetzenden Dienste meinen)<sup>12</sup> und verbindet damit Forderungen nach mehr oder weniger starken staatlichen Eingriffen. Allzu oft hapert es jedoch nicht an fehlenden rechtlichen Eingriffsmöglichkeiten, sondern vielmehr an deren konsequenter Anwendung.<sup>13</sup> Um Letzteres geht es beim NetzDG, was auch daran zu erkennen ist, dass das Gesetz nicht die Rechtsetzung, sondern die Rechtsdurchsetzung im Titel trägt.

Übergreifend lässt sich das Gesetz dabei in zwei grundlegende Debattenkontexte im digitalen Zeitalter einordnen: Zum einen ist die Herausforderung im Umgang sowohl mit *Hate-Speech* [Hassbotschaften] und *Fake News* (siehe Kapitel II.2.3) als auch mit dokumentiertem Kindesmissbrauch im Internet und zum anderen die Regulierung monopolartiger Internetplattformen zu nennen (siehe Kapitel II.2.2). So nannte auch der Gesetzesentwurf der Bundesregierung (2017b: 1) die »Debattenkultur im Netz«, die sich durch »Hasskriminalität und andere strafbare Inhalte« massiv verändert habe, sowie die »Bekämpfung von strafbaren Falschnachrichten« als Begründung für die Gesetzesinitiative. Beide Problemlagen kumulieren im konkreten Fall in den sozialen Netzwerken (wobei bereits deren Definition im Rahmen des Gesetzgebungsprozesses kontroversen Diskussionen unterlag).

Dass Hassbotschaften in sozialen Netzwerken, die bis hin zu Mordaufrufen reichen, für viele eine Alltagserfahrung im Internet darstellen, zeigt den dringenden Handlungsbedarf – nicht zuletzt auch aufgrund der damit verbundenen Debatten um den schwindenden gesellschaftlichen Zusammenhalt während der Coronapandemie. In der seit 2018 jährlich erhobenen Hate Speech Studie von forsa (2022: 2) gaben 2022 (wie bereits 2018) nur 22 Prozent der Internetnutzer:innen an, »noch nie« persönlich Hasskommentare im Internet gesehen zu haben, 36 Prozent dagegen »häufig« oder »sehr häufig«.<sup>14</sup> Der Anteil derjenigen, die einen Hasskommentar nicht nur gesehen, sondern auch der Plattform gemeldet haben, stieg zwischen 2018 und 2022 von 26 auf

12 So etwa die bayerischen Innenminister Günther Beckstein im Jahr 2006 (vgl. Lischka 2009) und Joachim Herrmann 2012 (vgl. Der Spiegel 2012).

13 Daher stellt sich regelmäßig die Frage, inwieweit es sich eher um symbolische Politik handelt. Symbolische Politik in dem Sinne, dass diese wie verschärfte Gesetzgebung nach »spektakulären Gewaltverbrechen [...] zur Beruhigung der Wähler« und nicht zur Verbesserung der staatlichen Kontrolle diene, denn »in den meisten Fällen [würde] die Anwendung der vorhandenen Bestimmung vollauf genügen« Reinhard (2007: 28).

14 Abweichend davon gab nur 1 % der Internetnutzer:innen an, dass nach ihrer eigenen Einschätzung Nutzer:innen im Internet Hasskommentare »überhaupt nicht« begegnen; 60 % schätzen dagegen »häufig« oder »sehr häufig«. Allerdings konnten 11 % dies nach eigener Aussage gar nicht einschätzen (vgl. forsa 2022: 16).

32 Prozent – gleichzeitig sank der Anteil derjenigen, die auf einen solchen Kommentar kritisch geantwortet haben, von 25 auf 22 Prozent (vgl. ebd.: 4).<sup>15</sup>

Das Problem wurde jedoch bereits früher erkannt. Der Google-Suchtrend zu »Hate Speech« in den vergangenen 20 Jahren stieg Mitte 2016 sprunghaft an.<sup>16</sup> In diesen Zeitraum fällt die vom Bundesministerium für Justiz und Verbraucherschutz (BMJV) unter Heiko Maas ins Leben gerufene Task Force »Umgang mit rechtswidrigen Hassbotschaften im Internet«,<sup>17</sup> um gegen Hate Speech und Fake News in sozialen Netzwerken vorzugehen. In dieser Task Force unter Leitung des Ministeriums waren Vertreter:innen aus der Zivilgesellschaft (eco, FSM, jugendschutz.net, klicksafe.de, Amadeu Antonio Stiftung, Gesicht zeigen!) und Plattformunternehmen (Google, Facebook, Twitter) vertreten. Aus deren Arbeit resultierte eine Selbstverpflichtung der Unternehmen zur Etablierung eines anwenderfreundlichen Melde- und Beschwerdesystems. Nach der Prüfung von Beschwerden »durch ein darauf spezialisiertes Team[]« sollte »die Mehrheit der gemeldeten Inhalte [...] in weniger als 24 Stunden geprüft und, falls erforderlich, entfernt« werden, so das Ergebnispapier der Task Force (BMJV 2015: 2f.).

Das Ministerium setzte also aus Steuerungsperspektive zunächst auf freiwillige Koordination und damit auf *Selbstregulierung*.<sup>18</sup> Eifert (2018: 14) erkannte hierin auch den

- 
- 15 Hierzu passt, dass nur 25 % der Internetnutzer:innen eine Gegenrede für eine wirksame Strategie gegen Hasskommentare und Hate Speech im Internet halten, während das Melden bei der Plattform 69 %, das schnelle Löschen 76 % und die strafrechtliche Verfolgung der Verfasser:innen 79 % für wirksam erachten (vgl. forsa 2022: 18).
- 16 Google Trends vom 17.08.2022 (<https://trends.google.de/trends/?geo=DE>) zum Suchbegriff »Hate Speech«. Verglichen mit der höchsten Anzahl an Suchanfragen im Februar und März 2022 lag diese zwischen 2012 und Mitte 2016 bei 9,7 % und zwischen Mitte 2016 und 2022 bei 50,5 %.
- 17 Auch auf der europäischen Ebene gründete EU-Kommissarin Vera Jourová eine Task Force, deren Arbeit 2016 zu einem Code of Conduct mit Facebook, Microsoft, Twitter und YouTube führte. In diesem verpflichteten sich die Unternehmen unter anderem dazu, »to review the majority of valid notifications for removal of illegal hate speech in less than 24 hours and remove or disable access to such content, if necessary« (Europäische Kommission 2016: 2). Als weitere Unternehmen akzeptierten Instagram, Snapchat und Dailymotion 2018, Jeuxvideo.com 2019, TikTok 2020 und LinkedIn 2021 den Code of Conduct (vgl. Europäische Kommission 2021a).
- 18 An der Beteiligung der Plattformunternehmen zeigen sich auch hier Parallelen zur Etablierung von Gremien und Instrumente der Selbstkontrolle im Medienbereich. So wurde der deutsche Presseerat unter Beteiligung der Unternehmen etabliert, um einer staatlichen Regulierung zu entgehen beziehungsweise dieser zuvorzukommen (vgl. Bastian/Fengler 2016: 214). Selbstregulierung spielt im Umgang mit Fake News weiterhin eine Rolle, etwa bei Meta. Der Konzern kooperiert in diesem Umfeld der Inhaltsregulierung darüber hinaus mit dritten Akteuren. So existiert seit 2016 (damals noch Facebook) weltweit das Meta Journalism Project zur unabhängigen Faktenprüfung bei Facebook, Instagram und WhatsApp (vgl. Meta o.J.). Dabei werden alle möglicherweise Fake News geflaggte Inhalte von (dafür bezahlten) sogenannten Factchecker:innen überprüft und gegebenenfalls für alle Nutzer:innen sichtbar als Desinformation gekennzeichnet oder mit Informationen aus seriösen Quellen kontrastiert (da sich Inhalte mitunter nicht immer eindeutig als Fake News einordnen lassen, sind Faktenchecker:innen durch kontroverse Entscheidungen angreifbar). Hierfür kooperiert Facebook mit Organisationen, in Deutschland etwa mit der dpa und dem geschäftlichen Ableger des Recherchenetzwerks CORRECTIV. Diese Kooperation erfährt durchaus Kritik (vgl. beispielsweise Matzat 2019) und wird auch von den Beteiligten als »Dilemma« gesehen (vgl. CORRECTIV 2018).

Versuch, eine mit unterschiedlichen nationalstaatlichen-rechtlichen Anforderungen einhergehende Segmentierung der Internetregulierung zu vermeiden:

»Der Gesetzgeber hatte durchaus Gespür für diese Problemlage und wählte zunächst den dann letztlich doch für nicht hinreichend effektiv angesehenen Ansatz eines runden Tisches und daraus hervorgehender Selbstverpflichtungen der Anbieter.«

Diese Form der Steuerung zeigte zwar durchaus Wirkung, der Outcome unterschied sich jedoch zwischen den Plattformen erheblich. Dies zeigte das von BMJV und Bundesministerium für Familie, Senioren, Frauen und Jugend (BMFSFJ) finanzierte Monitoringprojekt von jugendschutz.net (2017a), das in einem Test Anfang 2017, mit 540 über die Beschwerdesysteme gemeldeten strafbaren Beiträgen, Löschoroten von 90 Prozent bei YouTube, 39 Prozent bei Facebook und ein Prozent bei Twitter auswies.<sup>19</sup> Diese Daten zog das BMJV auch bei der Begründung des Gesetzesentwurfes heran (vgl. Bundesregierung 2017b: 1), obwohl jugendschutz.net (2017c: 1) selbst in einer Stellungnahme zum Entwurf darauf hinwies, dass ihr Monitoring »nicht mit dem Ziel durchgeführt [wurde], eine empirische Grundlage für gesetzliche Regelungen zu schaffen, sondern um Defizite und Potenziale zur Verbesserung der Beschwerdesysteme zu ermitteln.«

Gleichwohl folgt mit dem am 01.09.2017 vom Deutschen Bundestag verabschiedeten Gesetz zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken, kurz Netzwerkdurchsetzungsgesetz (NetzDG), eine gesetzliche Regelung (vgl. Deutscher Bundestag 2017).<sup>20</sup> Dabei wurde der Geltungsbereich des NetzDG auf folgende Anwendungsbereiche beschränkt (§ 1 Abs. 1):

»Dieses Gesetz gilt für Telemediendiensteanbieter, die mit Gewinnerzielungsabsicht Plattformen im Internet betreiben, die dazu bestimmt sind, dass Nutzern beliebige Inhalte mit anderen Nutzern teilen oder der Öffentlichkeit zugänglich machen (soziale Netzwerke). Plattformen mit journalistisch-redaktionell gestalteten Angeboten, die vom Diensteanbieter selbst verantwortet werden, gelten nicht als soziale Netzwerke im Sinne dieses Gesetzes. Das Gleiche gilt für Plattformen, die zur Individualkommunikation oder zur Verbreitung spezifischer Inhalte bestimmt sind.«

Das NetzDG reguliert somit nur soziale Netzwerke, keine Messenger (daher sind etwa WhatsApp oder Facebook Messenger ausgenommen). Telegram spielt hierbei eine Sonderrolle: Als Messenger gestartet, stellt es heute auch Funktionen eines sozialen Netzwerkes zur Verfügung (öffentliche Kanäle und Gruppen). Darüber hinaus wird es nicht nur als eine der zentralen Plattformen zur Verbreitung von Hate Speech und Fake News

19 Diese Löschoroten erhöhten sich auf 100 % (YouTube, Twitter) beziehungsweise 93 %, sobald über die Beschwerdefunktion für Nutzer:innen der Plattform hinaus eine weitere Meldung mit dem akkreditierte Account von jugendschutz.net und eine direkte E-Mail-Kontaktaufnahme stattfand (vgl. jugendschutz.net 2017a).

20 Damit nahm der Bundesgesetzgeber seine Ausgestaltungs-kompetenz wahr, indem er die sogenannte Hostproviderhaftung konkretisierte, ohne dabei in die bei den Ländern liegende Kompetenz zur Medienregulierung einzugreifen (vgl. Hentsch 2022).

wahrgenommen, sondern zeigt auch die Vollstreckungsprobleme des NetzDG auf (siehe hierzu die Ausführungen weiter unten in Kapitel V.3.3.1).

Zudem sollten explizit journalistische Angebote und implizit zivilgesellschaftliche Angebote wie Wikipedia oder Foren von den Regelungen des NetzDG ausgenommen werde.<sup>21</sup> Um darüber hinaus kleinere Anbieter:innen zu entlasten und die Markteintrittsbarriere für neue konkurrierende Unternehmen zu senken, wurde eine Untergrenze von zwei Millionen Nutzer:innen für die Umsetzungsnotwendigen der festgelegten Berichts- und Beschwerdeverfahrenspflichten (§ 2 und § 3) festgelegt (§ 1 Abs. 2).<sup>22</sup>

Die Steuerungsmaßnahmen lassen sich fünf Ebenen zuordnen:

1. Beschwerden: Aufbau eines nutzer:innenfreundlichen Verfahrens, um rechtswidrige Inhalte zu melden, sowie eines wirksamen und transparenten Verfahrens des Beschwerdemanagements (§ 3 Abs. 1).<sup>23</sup>
2. Transparenz: Es besteht eine halbjährliche Berichtspflicht, unter anderem über die Anzahl der Beschwerden sowie der entfernten oder gesperrten Inhalte (§ 2).<sup>24</sup>
3. Fristen: Die Plattformen müssen *offensichtlich rechtswidrige* Inhalte innerhalb von 24 Stunden (§ 3 Abs. 2 Satz 2) sowie jedweden *rechtswidrigen* Inhalt innerhalb von sieben Tagen nach Beschwerdeingang entfernen oder sperren (§ 3 Abs. 2 Satz 3).<sup>25</sup>
4. Bußgelder: Die Umsetzung dieser Maßnahmen soll durch ihre Bußgeldbewehrung sichergestellt werden (§ 4). Vorgesehen sind bis zu fünf Millionen Euro (bei vorsätzlicher und fahrlässiger) nicht (vollständiger) Umsetzung der Maßnahmen.<sup>26</sup>
5. Ansprechpartner:in: Die Plattformen müssen eine:n Zustellbevollmächtigte:n im Inland sowie eine empfangsberechtigte Person für Auskunftersuchen inländischer Strafverfolgungsbehörden benennen (§ 5).

Mit dem NetzDG werden also die Versuche einer freiwilligen Koordination vonseiten der Plattformen durch relativ klare Regelungen abgelöst. Wenn Eifert (2018: 16) davon

- 
- 21 Unterschiedliche Akteure bemängelten jedoch, dass der Anwendungsbereich dennoch nicht ausreichend konkret benannt sei und tendenziell auch andere Plattformen (Bewertungsportale, Diskussionsforen, Messaging-Dienste) mit umfassen würde, die das BMJV im Kern gar nicht adressieren wollte (siehe Tabelle 17).
  - 22 Der DJV (2017: 3) kritisierte bereits beim Gesetzesentwurf, dass journalistische Inhalte, die über Plattformen vertrieben werden, die unter das Gesetz fallen, nicht anders behandelt werden als Inhalte anderer Nutzer:innen der Plattform.
  - 23 Die SRIW (2017: 5) kritisiert »wenige greifbare Details für das vorzuhaltende Beschwerdeverfahren«.
  - 24 Die SRIW (2017: 9) kritisierte, dass die Wirksamkeit ausschließlich quantitativ darüber bemessen werde, wie viele rechtswidrige Inhalte in welcher Zeit gelöscht worden seien (Löschquote), qualitative Aspekte damit aber zu kurz kämen.
  - 25 Zu den rechtswidrigen Inhalten zählen 21 Straftatbestände nach dem Strafgesetzbuch (StGB) wie Beleidigung, üble Nachrede, Verleumdung, Volksverhetzung oder Gewaltdarstellung (vgl. forum privatheit 2018: 4). Während sich jugendschutz.net (2017b) für eine Ausweitung der Tatbestände einsetzte, kritisierte der DJV (2017) die aufgeführte Liste als willkürlich.
  - 26 Die SRIW (2017: 6) kritisierte die Bußgeldvorschriften: Diese »überschreiten klar die Grenze des Bestimmtheitsgebots.«

spricht, dass der Grundansatz des NetzDG »in der Strukturierung der Verantwortungswahrnehmung großer gewerblicher Plattform-Intermediäre« bestehe, wird darin der mit der Gesetzgebung einhergehende Steuerungswandel von der Selbstregulierung zur *regulierten* Selbstregulierung deutlich.

Auf der einen Seite beinhaltet das NetzDG als hierarchische Steuerungsformen zu klassifizierende Elemente. Es enthält Vorgaben und Verpflichtungen, denen Plattformbetreiber innerhalb eines definierten Zeitrahmens nachkommen müssen, wenn sie nicht mit empfindlichen Strafen belegt werden wollen. Hieran zeigt sich auch, dass die immer mal wieder formulierte Annahme der Unregulierbarkeit global agierender Digitalkonzerne auf nationalstaatlicher Ebenen (und auch auf europäischer Ebene, wie beispielsweise die Datenschutzgrundverordnung und der in Kapitel V.3.6 behandelte Digital Services Act zeigen) nicht haltbar ist (vgl. auch Pistor 2021: 21f.).<sup>27</sup> Diese können »durchaus unter Druck gesetzt und zu Regulierung gezwungen werden« (Thiel 2019: 53; auch Pohle/Thiel 2019: 62f.). Auf der anderen Seite überträgt das NetzDG implizit und explizit staatliche Aufgaben auf Private – dieser, von unterschiedlichen Akteuren bereits beim Gesetzesentwurf kritisch angemerkte Aspekt (siehe nachfolgendes Kapitel und Tabelle 17) wird später als Phänomen der Privatisierung der Rechtsdurchsetzung (siehe Kapitel V.3.3) sowie der Privatisierung der Rechtsetzung (siehe Kapitel V.3.4) vertiefend aufgegriffen.

### V.3.2 Polity: Akteure und Akteurskonstellationen

Die Akteurskonstellation im Feld der Inhaltsregulierung auf internetbasierten Plattformen fällt sehr heterogen aus (siehe Tabelle 15). Auf der einen Seite stehen – im Sinne des NetzDG als Steuerungsobjekte – die Plattformbetreiber:innen. Hierunter fallen insbesondere die großen Diensteanbieter:innen wie Alphabet (mit Google und YouTube), Meta (mit Facebook und Instagram), reddit, Telegram, TikTok und Twitter. Auf der anderen Seite befinden sich die Nutzer:innen dieser Dienste, die selbst wiederum sehr heterogen ausfallen. Zum einen sind hier die »normalen« Nutzer:innen zu nennen. Diese übernehmen jedoch als aktive Prosumer:innen nicht nur primär die Content-Generierung auf den Plattformen, sondern beanstanden als Melder:innen auch (vermeintlich) rechtswidrige Inhalte oder stellen solche als Störer:innen online. Zum anderen kommen die professionellen Plattformnutzer:innen hinzu. Hierzu zählen sowohl Wirtschaftsakteure wie Unternehmen, die selbst Inhalte generieren oder Werbung schalten, Werbenetzwerke sowie Medienunternehmen oder (freischaffende) Journalist:innen und Autor:innen. Für die Frage des Umgangs mit Internetinhalten kommen als dritte Gruppe die zivilgesellschaftlichen Organisationen hinzu. Hierbei lassen sich insbesondere die Bereiche der auf Freiheits- und Bürger:innenrechte fokussierten Organisationen, Vereinigungen mit

27 Ähnliches könnte sich auch mit Blick auf die durch Netzwerk- und Skaleneffekte zurückzuführende Monopol- beziehungsweise Oligopolbildung in der Digitalwirtschaft zeigen. So laufen etwa in den USA Bemühungen, Facebook zu zerschlagen – wobei das 20. Jahrhundert mit der Zerschlagung von American Telephone & Telegraph (AT&T) und Standard Oil als Blaupause dient. »Allerdings sprechen manche Argumente dafür, dass die digitale Ökonomie hier eine systematisch höhere Dynamik aufweist und schwerer zu kontrollieren ist« (Mergel 2022: 250).