ISD | Powering solutions to extremism and polarisation

gefördert durch:
Auswärtiges Amt

Digital Policy Lab '20 Begleitpapiere

- Transparenz, Datenzugang und „Online Harms“
- Nationale & Internationale Modelle zur Online-Regulierung
- Zusammenfassung des EU Digital Services Act & UK Online Safety Bill
- Das freiheitlich-demokratische Internet – Fünf Modelle für eine digitale Zukunft
- Überlegungen zur Zukunft der Online-Regulierung

Chloe Colliver, Milo Comerford,
Jennie King, Alex Krasodomski-Jones,
Christian Schwieter & Henry Tuck

Über das Digital Policy Lab

Das Digital Policy Lab (DPL) ist eine von dem Institute for Strategic Dialogue (London/Berlin) organisierte Initiative zur transatlantischen Koordinierung digitalpolitischer Regulierungsanstrengungen. Gefördert wird das Projekt vom Auswärtigen Amt.

Das DPL ist als eine zwischenstaatliche Arbeitsgruppe konzipiert, die sich darauf konzentriert, den regulatorischen und politischen Weg zur Bekämpfung von Desinformation, Hassrede, Extremismus und Terrorismus im Internet aufzuzeigen. Sie besteht aus einer Kerngruppe hochrangiger Vertreter der zuständigen Ministerien und Regulierungsbehörden ausgewählter liberal-demokratischer Länder.

Die geäußerten Ansichten sind die der Autoren und spiegeln nicht unbedingt die Ansichten der Teilnehmer oder Förderer des DPL wider. Die herausgeberische Verantwortung hat Huberta von Voss-Wittig, Executive Director ISD Germany.

ISD | Powering solutions to extremism and polarisation

Beirut | Berlin | London | Paris | Washington DC

www.isdglobal.org

Inhaltsangabe

ISD | Powering solutions to extremism and polarisation

Digital Policy Lab
Diskussionspapier

Transparenz, Datenzugang und „Online Harms“

Über dieses Diskussionspapier

Dieses Diskussionspapier gibt einen Überblick über die internationale politische Debatte zu Transparenzpflichten für Social-Media Unternehmen im Kampf gegen Desinformation, Hassrede, Extremismus und Terrorismus im Internet. Es stellt die Hauptthemen des ersten Digital Policy Lab vor, das am 12. und 13. November 2020 stattfand, und bezieht die Diskussionen der Veranstaltung mit ein. Die in diesem Papier geäußerten Ansichten sind die der Autoren und spiegeln nicht unbedingt die Meinungen der Teilnehmer am Digital Policy Lab oder der Regierungen wider. Aufgezeichnet wurde der Inhalt von dem auf Extremismusbekämpfung und die Analyse des Internets spezialisierten Think Tank Institute for Strategic Dialogue gGmbH (Berlin) in Zusammenarbeit mit dem Londoner Headquarter des Instituts.

Zur besseren Lesbarkeit wurde auf die gleichzeitige Verwendung männlicher und weiblicher Sprachformen verzichtet. Es wurde das generische Maskulinum verwendet, wobei alle Geschlechter gleichermaßen gemeint sind.

ISD | Powering solutions to extremism and polarisation

Beirut | Berlin | London | Paris | Washington DC

www.isdglobal.org

Warum ist Transparenz wichtig?

Das Internet hat mächtige neue Kräfte geschaffen, die unser Leben entscheidend prägen. Online-Plattformen haben die Art und Weise, wie wir kommunizieren, Entscheidungen treffen und Ansichten vertreten, radikal verändert. Gleichzeitig beklagen Regierungen auf der ganzen Welt das Entstehen sogenannter „Online Harms“.[1] Die explosionsartige Verbreitung der Covid-19-„Infodemie“ im Jahr 2020 hat die Bandbreite der „Online Harms“, denen liberal-demokratische Gesellschaften ausgesetzt sind – von Desinformation über Hassrede bis hin zu Extremismus und Terrorismus – , ebenso deutlich gemacht wie die Dringlichkeit einer internationalen Koordination zur Eindämmung dieser Bedrohungen.

Ohne ein fundiertes Verständnis von Ausmaß und Art dieser Herausforderungen ist es jedoch nahezu unmöglich einzuschätzen, welche potenziellen Lösungen sowohl effektiv als auch verhältnismäßig wären. Es ist von zentraler Bedeutung, dass Regierungen, Regulierungsbehörden, die Zivilgesellschaft und die breite Öffentlichkeit besser verstehen, auf welche Weise das Internet die Gesellschaft und die Demokratie beeinflusst, um erfolgreich seine positiven Auswirkungen fördern und negative Auswirkungen begrenzen zu können.

Regierungen und Amtsträger benötigen genauere Informationen, um ihren Verpflichtungen gegenüber den Bürgern nachzukommen. Zu diesen Verpflichtungen gehören: Die Aufrechterhaltung und Durchsetzung bestehender Gesetze, demokratische Aufsicht, der Schutz der nationalen Sicherheit, die Vertretung der Interessen jener Wähler, die Opfer von „Online Harms“ geworden sind, und das Eintreten für Veränderungen im Namen der Geschädigten. Damit die Regulierungsbehörden effektiv arbeiten können, benötigen sie ausführliche Informationen über Unternehmensrichtlinien, -verfahren und -entscheidungen sowie über die zugrundeliegende Technologie.

Wissenschaft und Medien würden von einem besseren Zugang zu Daten profitieren und könnten ihre Aufgaben im öffentlichen Interesse besser erfüllen, die nicht zuletzt darin bestehen, die oft undurchsichtige Online-Welt besser zu beleuchten. Auch die Zivilgesellschaft würde davon profitieren, da man gleichzeitig eine bessere Beweisgrundlage hätte, um Täter verfolgen und Ursachen und Auswirkungen von „Online Harms“ aufspüren zu können. Zudem würde so eine unabhängige Kontrolle samt Beratung und Unterstützung für gefährdete Gruppen oder Minderheiten ermöglicht. Letztendlich ist Transparenz auch für normale Internet-Nutzer von entscheidender Bedeutung – nicht nur, damit sie ihre Rechte und Pflichten im Internet wahrnehmen können, sondern auch, um die Beziehungen, die sie mit Online-Plattformen eingehen, und das Umfeld, in dem sie ihre Informationen erhalten, besser zu verstehen. Um eine wirksame Aufsicht zu gewährleisten und eine nachhaltige Politik, Gesetzgebung und Regulierung für die Online-Welt entwickeln zu können, brauchen wir also eine solidere Beweisgrundlage.

Transparenz trägt entscheidend dazu bei, dass die entsprechenden Beweise gesammelt werden können: Indem wir die Transparenz in Online-Räumen erhöhen, so die Argumentation, haben wir eine bessere Chance, ein breites Spektrum an illegalen und legalen „Online Harms“ zu erkennen und darauf zu reagieren. Transparenz wird weithin als ein Schlüsselprinzip für „Good Governance“ der öffentlichen Verwaltung benannt und akzeptiert, unter anderem vom Europarat, von der OSZE und von der Europäischen Kommission. Kurz gesagt, es wird davon ausgegangen, dass für eine faire und effiziente Regierungsführung eine unabhängige Aufsicht und Möglichkeiten der öffentlichen Kontrolle notwendig sind. In einem demokratischen System muss es transparente Prozesse geben, die sicherstellen, dass öffentlichen Akteure zur Rechenschaft gezogen werden können.

1 Der Begriff „Online Harms“ geht auf das Online Harms White Paper der britischen Regierung zurück, welches im April 2019 veröffentlicht wurde. „Online Harms“ steht sowohl für konkrete als auch potenzielle Schäden (sprich Gefahren) die z.B. durch das Verbreiten von problematischen Inhalten in sozialen Medien entstehen. Da der englische Begriff die deutschen Begriffe Schäden, Gefahren und Risiken vereint, wird in diesem Arbeitspapier der englische Originalbegriff verwendet.

ISD verfügt im Hinblick auf eine Reihe von Online-Herausforderungen (von Hassrede, Desinformation bis hin zum Terrorismus) über jahrzehntelange Erfahrung in der Zusammenarbeit mit dem privaten Sektor, politischen Entscheidungsträgern und der Zivilgesellschaft. Diese Erfahrung zeigt, dass jeder wirksame regulatorische oder nicht-regulatorische Ansatz zur Bekämpfung der gesamten Bandbreite von „Online Harms“ auf Transparenz beruhen sollte. Transparenz ist kein Selbstzweck, sondern eine Voraussetzung für das öffentliche Vertrauen, die Ausübung der demokratischen Rechenschaftspflicht, und die Zusammenarbeit zwischen Technologieunternehmen, Regierung und der Öffentlichkeit. Die Anforderungen und Erwartungen an diese Transparenz sind jedoch oft unzureichend formuliert oder variieren stark zwischen Regeln für Online-Plattformen und Offline-Rechtsordnungen. Es gibt vielversprechende Rahmenwerke und Modelle für Transparenz, sowohl im digitalen Kontext als auch in anderen verwandten oder vergleichbaren Bereichen. Diese sollten im Hinblick auf das gesamte Online-Ökosystem als „Best Practice“ genutzt werden und als Orientierung für zukünftige Erwartungen an Transparenz im Internet dienen.

Abwägen von Transparenz mit Sicherheits- und Datenschutzbelangen

Schädliche Online-Aktivitäten auf der einen und Initiativen zu ihrer Bekämpfung auf der anderen Seite wurden als Wettrüsten zwischen illegitimen Akteuren und denjenigen, die ihnen entgegenwirken wollen, beschrieben. In diesem Kontext muss ein wirksames Gleichgewicht gefunden werden, da nachvollziehbare Einblicke in die Durchsetzungsmethoden sich auch als kontraproduktiv erweisen können. In Anbetracht eines komplexen, sich konstant verändernden Informationsumfelds ist es jedoch wichtig, dass ein Höchstmaß an Transparenz angestrebt wird, um Vertrauen aufzubauen, Sorgfalt zu gewährleisten und Missbrauch vorzubeugen. Bedenken, dass illegitime Akteure die genauen Methoden zum Aufspüren und zur Mäßigung bestimmter Inhalte ausfindig machen können, sollten ernst genommen werden. Dies sollte jedoch die Entwicklung von Strukturen zur Rechenschaftspflicht nicht ausschließen, um konkrete Schwachstellen zu beseitigen. Darüber hinaus ist zu beachten, dass Transparenz das Recht auf Datenschutz ergänzen muss und nicht aushöhlen darf. Ein gutes Modell für Transparenz schützt die Privatsphäre des Einzelnen und ermöglicht gleichzeitig ein Makroverständnis von Art und Umfang der Arbeitsprozesse von Online-Plattformen sowie aller potenziellen Rechtsverletzungen, die sich aus der Nutzung der Plattform ergeben.

Herausforderungen durch mangelnde Transparenz

Grenzen der aktuellen öffentlichen Forschung zu „Online Harms“

Jüngste ISD-Forschungen zur Untersuchung der von Facebook ergriffenen Maßnahmen gegen „coordinated inauthentic activity“ – definiert als organisierte verdeckte, trügerische und absichtlich irreführende Aktivitäten auf der Plattform – zeigen die derzeitigen Grenzen der von Technologieunternehmen veröffentlichten Transparenzberichte auf. Dies gild insbesondere im Hinblick auf das tatsächliche Ausmaß und die Art der Herausforderung durch schädigende Inhalte und entsprechendes Verhalten auf der Plattform. Die Untersuchung hat deutlich gemacht, dass die von Facebook freigegebenen Daten nur einen Teil der Wirklichkeit abbilden. Die verfügbaren Informationen zeigen das beträchtliche Ausmaß an betrügerischen Aktivitäten, u. a. von Nationalstaaten, PR-Unternehmen und ideologisch motivierten Hassgruppen, die über Facebook versuchen, Wähler auf der ganzen Welt zu beeinflussen. Was die veröffentlichten Daten nicht zeigen können, ist das tatsächliche Ausmaß dieser Aktivitäten, einschließlich derer, die vom Unternehmen nicht entdeckt oder gemeldet werden. Unabhängige Forscher, unter anderen vom ISD, finden nach wie vor Beispiele für großangelegtes „coordinated inauthentic activity“ auf Facebook und Twitter, obwohl sie oft nur minimalen Zugang zu diesen Daten haben. Beispielsweise identifizierten das ISD und Partnerorganisationen im Vorfeld der Wahlen zum Europäischen Parlament neunzehn dieser koordinierten inauthentischen Netzwerke auf Facebook und Twitter, und zwar durch Untersuchungen, die sich lediglich auf sechs EU-Mitgliedsstaaten konzentrierten. Die Beweise deuten darauf hin, dass die von Facebook in den letzten zwei Jahren publizierten Fälle nur einen Bruchteil des wahren Ausmaßes solcher Aktivitäten auf der Plattform darstellen. Dies hat sich im Zuge der „Covid-19-Infodemie“ noch weiter verschärft. Eine Untersuchung von ISD und BBC ergab, dass eine Reihe von Webseiten, die Desinformationen rund um COVID-19 verbreiteten, während der Gesundheitskrise über 80 Millionen Aufrufe auf Facebook erhielten - sechsmal mehr als das US Center for Disease Control and Prevention (CDC) und die WHO zusammen.

Eine kürzlich durchgeführte ISD-Studie zum Ausmaß von digitalen Einschüchterungsversuchen, die sich gegen eine Reihe von Politikern richten, ergab, dass Frauen und Kandidaten, die einer ethnischen Minderheit angehören, überproportional häufig Beleidigungen online ausgesetzt sind. Die Studie hat gezeigt, dass Social-Media-Plattformen im Hinblick auf Belästigungen und Bedrohungen mehr Transparenz über ihre Richtlinien zur Inhaltsmoderation sowie über interne Abläufe und die Ergebnisse von Durchsetzungsmaßnahmen bieten müssen. Hierzu sollten Informationen über Art und Inhalte, die unter die entsprechenden Richtlinien fallen, sowie Ressourcen die für die Inhaltsmoderation inklusive Kulturkenntnissen der Teams bereitgestellt werden, gehören. Ferner muss die Transparenz bei den Einspruchs- und Rechtsbehelfsverfahren für unrechtmäßige Löschungen von Inhalten und das Blockieren von Accounts verbessert werden.

In jüngster Zeit stehen insbesondere schnell wachsende jugendorientierte Social-Media-Plattformen wie TikTok in der Kritik. Besonders im Fokus stehen dabei potenzielle Gefahren wie Desinformation zur öffentlichen Gesundheit während der Covid-19-Pandemie. Die Plattform hat sich zwar verpflichtet, gegenüber ihrer Community Rechenschaft abzulegen, indem sie Informationen über das Löschen von Inhalten, einschließlich Hassrede und Fehlinformationen, herausgibt. Doch das Fehlen eines offiziellen Datenzugangs durch eine API (application programming interface) sowie undurchsichtige Suchfunktionen und Merkmale wie das Mobile-First-Design der Plattform, machen es Forschern schwer, die Datenerfassung zu automatisieren und Trends im großen Maßstab zu beobachten.

Es gab zwar Versuche des privaten Sektors, Wissenschaftlern einen besseren Zugang zu gewähren. Jedoch bringt die Entwicklung sicherer und transparenter Prozesse für die gemeinsame Nutzung anonymisierter Daten in dieser Größenordnung besondere Herausforderungen mit sich. Initiativen wie Social Science One, die darauf abzielen, ausgewählten Akademikern Zugang zu großen anonymisierten Datensätzen zu gewähren, um Desinformation auf Facebook zu untersuchen, zeigen, dass die Zusammenarbeit zwischen Unternehmen und unabhängigen Forschern erst am Anfang steht und dringend ausgeweitet werden muss.

Enthüllungen von Tech-Unternehmen, die den Mangel an Transparenz offenbaren

Neben den Erkenntnissen aus den Transparenzberichten der Unternehmen und den unterschiedlichen von Online-Plattformen bereitgestellten Zugangsmöglichkeiten für Forscher zeigen Daten, die von Unternehmen durch demokratische Kontrolle extrahiert (oder den Medien zugespielt) wurden, das Ausmaß des Informationsdefizits. Dies betrifft sowohl die Online-Räume, in denen „Online Harms“ gedeihen können, wie auch die Entscheidungen privater Unternehmen, die diese Räume gestalten. Einblicke von Insidern der Tech-Unternehmen machen klar, wie begrenzt aktuelle Transparenzbemühungen und groß die Herausforderungen sind. Zahlreiche Informationen, die den Plattformen zur Verfügung stehen werden aufgrund begrenzter Pflichten zur Transparenzberichterstattung weder an Regierungen weitergegeben, noch gelangen sie an die Öffentlichkeit.

Beispielsweise schrieb die kürzlich entlassene Facebook-Mitarbeiterin Sophie Zhang im September 2020 ein Memo, in dem sie Facebook dafür kritisierte, nicht effektiv auf die globalen, inauthentischen und koordinierten politischen Aktivitäten auf der Plattform zu reagieren. Sie sprach Bedenken an, die Forscher und politische Entscheidungsträger schon seit einiger Zeit geäußert hatten – nämlich, dass Facebook es politischen Akteuren auf der ganzen Welt möglich macht, umfangreiche betrügerische Aktivitäten im Hinblick auf Wahlen durchzuführen, weil die Hürden für die Einrichtung entsprechender Seiten sehr niedrig sind.

Ironischerweise wurden die Möglichkeiten von Forschern, solche Aktivitäten aufzudecken, erschwert, nachdem Facebook den API-Zugang von Drittanbietern nach dem Cambridge Analytica-Skandal erheblich eingeschränkt hat. Dies sollte den Missbrauch durch kommerzielle oder politische Akteure verhindern, hat aber auch die Bemühungen zur Erforschung von „Online Harms“ behindert. Beispielsweise hat das Unternehmen den Datenzugriff eines technischen Tools der New York University geblockt, das die Transparenz des Targeting für politische Werbung erhöhen sollte. Demgegenüber veranschaulichten interne Dokumente, die von einem Untersuchungsausschuss des britischen Parlaments eingeholt wurden, dass Facebook den Datenzugang für einige ausgewählte Privatunternehmen im Rahmen spezieller „Whitelisting“-Vereinbarungen zuvor erweiterte.

Auch was analytische Fähigkeiten und die Daten angeht, die den Unternehmen intern zur Verfügung stehen, ist das Allgemeinverständnis begrenzt – abgesehen von den einfachen Kennzahlen rund um die Nutzung von Inhalten, die in Transparenzberichten präsentiert werden. So konnten beispielsweise die von Wissenschaftlern und zivilgesellschaftlichen Forschern seit Jahren geäußerten Bedenken hinsichtlich der potenziell schädlichen Auswirkungen von Empfehlungsalgorithmen aufgrund von Einschränkungen beim Datenzugriff nicht aussagekräftig mit Fakten untermauert werden. Ein kürzlich durchgesickerter interner Facebook-Bericht, der den eigenen Führungskräften im Jahr 2018 vorgelegt wurde, ergab jedoch, dass sich das Unternehmen sehr wohl darüber bewusst war, dass sein Produkt, insbesondere sein Empfehlungsalgorithmus, Zwietracht und Polarisierung schürt. Und bereits 2016 ergab ein anderer interner Bericht, dass 64 % der Personen, die einer extremistischen Gruppe auf Facebook beitraten, dies nur taten, weil der Algorithmus der Plattform sie ihnen empfohlen hatte, wie das Wall Street Journal berichtete.

Andere Studien haben ähnliche Kritik am Empfehlungsalgorithmus von YouTube und seiner Rolle bei der Vertiefung politischer Polarisierung und Radikalisierung geübt. Der fehlende Zugang zu Daten für unabhängige Forscher hat es jedoch auch fast unmöglich gemacht, die Behauptungen von YouTube zu verifizieren, dass es die Empfehlungen von „grenzwertigen Inhalten und schädlichen Fehlinformationen“ um 50% reduziert habe. Mozilla hat eine Reihe von Empfehlungen ausgesprochen, um aussagekräftige Daten für Forscher zu sammeln. Dazu gehören Forderungen nach umfassenderen Kennzahlen für Impressions und Engagement, Bereitstellung von historischen Videoarchiven und Simulationswerkzeugen, die es möglich machen, die Wege des Empfehlungsalgorithmus besser nachzuvollziehen.

In Anbetracht des potenziellen Schadens, den Empfehlungssysteme anrichten können, hat Facebook im Vorfeld der US-Wahl 2020 diskret Empfehlungs-Algorithmen für politische Gruppen ausgesetzt, zumal auch bekannt wurde, dass das Unternehmen eine internen Kontrollmechanismus zur Überwachung von „Gewalt- und Aufwiegelungstrends“ eingerichtet hatte. Dieses Tool, das das Gefährdungspotenzial auf der Grundlage von Hashtags und Suchbegriffen bemisst – und einen 45%igen Anstieg von Gewalt und Aufwiegelung im Wahlzeitraum festgestellt hat – zeigt den erheblichen Mehrwert von Echtzeitdaten, die privaten Unternehmen, nicht aber Regierungen, Regulierungsbehörden oder Forschern zugänglich sind. Es bleibt unklar, wie sich solche Tools auf die Änderungen und Durchsetzung der Gemeinschaftsrichtlinien von Facebook beziehen oder diese auslösen, oder ob ein bestimmtes Maß an „Gewalt und Aufwiegelung“ als akzeptabel angesehen wird.

Schlüsselbereiche, die mehr Transparenz erfordern

In einem Papier aus dem Jahr 2019 hat das ISD vier Schlüsselbereiche der Funktionen und Dienste von Technologieunternehmen und Plattformen dargelegt, die eine verbesserte Transparenz erfordern, um „Online Harms“ zu bekämpfen und Nutzerrechte besser zu wahren.

Inhalt & Moderation

Plattformen, die zu öffentlichen Räumen geworden sind, müssen selbige so allgemeinverständlich wie möglich gestalten. Da Webplattformen und ihre Nutzer eine wachsende Rolle bei der Gestaltung unserer Kultur einnehmen, unsere politischen Entscheidungen beeinflussen und den gesellschaftlichen Wandel vorantreiben, sollten die Aktivitäten, die in diesen Räumen stattfinden, einsehbar sein. Transparenz erfordert hier, dass sowohl Forscher als auch Nutzer systematisch Zugang zu öffentlichen Inhalten erhalten sowie klare Informationen darüber, wie Plattformen diese Inhalte moderieren.

Ersteres kann durch eine leicht navigierbare API erreicht werden, die der Öffentlichkeit ein Mittel zur Abfrage von Live- und historischen Inhalten jenseits der Beschränkungen des Standard-“News Feed“ bietet. Neben der offiziellen API, die von Twitter und anderen Diensten wie der zu Facebook gehörenden Software CrowdTangle angeboten wird, haben einige Forscher ihre eigenen unabhängigen Monitoring- und Suchfunktionen entwickelt, um die Sichtbarkeit der Plattformaktivitäten zu verbessern.

Über den API-Zugang hinaus betonen die jüngsten regulatorischen Initiativen in Deutschland, Frankreich und Australien die Notwendigkeit einer größeren Transparenz bei den Moderationsaktivitäten der Plattformen, besonders im Hinblick auf ein schnelleres Vorgehen der Plattformen gegen illegale und hasserfüllte Inhalte. Viele dieser Initiativen verlangen von Social-Media-Unternehmen regelmäßige Transparenzberichte, in denen eingegangene Beschwerden und getroffene Entscheidungen gegen Hassrede oder „coordinated inauthentic behaviour“ (CIB) auf ihren Plattformen dokumentiert werden. Darüber hinaus verlangen viele Interessengemeinschaften von Social-Media-Unternehmen, dass sie Informationen oder Anfragen zur Sperrung von Inhalten durch Strafverfolgungsbehörden veröffentlichen.

Derartige Transparenzberichte können ein wichtiges Instrument sein, um Forscher, die Zivilgesellschaft, Regulierungsbehörden und politische Entscheidungsträger in ihrer Arbeit zu unterstützen. Beispielsweise wurde für eine kürzlich durchgeführte ISD-Untersuchung das öffentlich zugängliche Archiv der Facebook-Transparenzberichte genutzt, um nachzuvollziehen, wie das Unternehmen inauthentische Aktivitäten auf seinen Plattformen identifiziert und bekämpft. Dabei wurden auch beträchtliche Einnahmen entdeckt, die Facebook in Form von Anzeigenverkäufen über die Accounts generieren konnte. Außerdem können diese Berichte Klarheit über die sich ständig verändernden Nutzungsbedingungen und deren Durchsetzung durch die Plattformen schaffen.

Beschwerden & Rechtsmittel

In der Öffentlichkeit klafft eine erhebliche Informationslücke bezüglich der Fähigkeit der Plattformen, Gewalt fördernde Inhalte abzumildern und entsprechend darauf zu reagieren. Die Sichtbarkeit von Beschwerden, die an Plattformen gerichtet werden, ist wesentlich für die Rechenschaftspflicht, die Unterstützung der Opfer von „Online Harms“, die Sensibilisierung für Herausforderungen, mit denen Nutzer online konfrontiert sind, und die Bereitstellung von Beweisen zwecks Wiedergutmachung. Wie oben beschrieben, haben (teils gesetzlich vorgeschriebene) Transparenzberichte versucht, diese Lücke zu schließen.

Die Transparenzberichte geben jedoch oft keinen aussagekräftigen Einblick in die Moderationsprozesse privater Unternehmen und schränkt damit die Möglichkeiten der Nutzer ein, Einspruch zu erheben und Entscheidungen anzufechten. Tatsächlich wurde auf Grund dessen das erste Bußgeld nach dem deutschen NetzDG-Gesetz gegen Facebook verhängt, und zwar wegen der Bereitstellung unvollständiger Daten in seinem ersten Transparenzbericht 2018. Als Reaktion auf den wachsenden Druck kündigte Facebook 2018 die Einrichtung eines „unabhängigen Aufsichtsgremiums“ an. Die 40 Mitglieder des Gremiums wurden im Mai 2020 vorgestellt. Ende 2020 hat das Gremium mit der Überprüfung begonnen. Ziel

ist es, Nutzern zu ermöglichen, Einspruch gegen die von Facebook getroffenen Inhaltsentscheidungen zu erheben, indem sie an die unabhängigen Mitglieder des Gremiums gerichtet werden können. Die Entscheidungen des Gremiums sind bindend sollen die Moderationsrichtlinien von Facebook in Zukunft prägen.

Eine ähnliche Aufsichtsfunktion nimmt derzeit der Verein Freiwillige Selbstkontrolle Multimedia-Diensteanbieter e.V. (FSM) wahr. Weil es sich hier um eine Selbstkontrolleinrichtung nach dem NetzDG handelt, können Social-Media-Unternehmen entscheiden, schwierige Inhaltsentscheidungen an ein von der FSM einberufenes Expertengremium zu delegieren. Die FSM verfügt ebenfalls über Verfahren, um Beschwerden von Nutzern gegen Entscheidungen zur Inhaltsmoderation im Namen des Social-Media-Unternehmens zu prüfen. Seit April 2020 sind zehn solcher Nutzerbeschwerden eingegangen. Sechs Inhalte wurden als legal angesehen wurden und mussten daher von den Social-Media-Unternehmen wieder eingestellt werden (Stand Dezember 2020).

Werbung

Werbung – insbesondere gezielte politische Werbung – ist eines der wichtigsten Produkte von Online-Plattformen. Sie ermöglicht es Werbetreibenden und Wahlkämpfern, Inhalte direkt an ein ausgewähltes Publikum zu vermitteln. Es ist im öffentlichen Interesse, dass Internetnutzer verstehen wie und warum sie online gezielt angesprochen werden, und dass die Regulierungsbehörden in der Lage sind, Missstände zu erkennen und zu beseitigen.

Viele Interessengemeinschaften auf der ganzen Welt haben beschlossen, dass die Transparenzstandards für politische Werbung über die Anforderungen an unbezahlte öffentliche Inhalte und Kommunikation hinausgehen sollten. Regulierungsinitiativen, wie etwa in Frankreich, Irland, Australien, Kanada, den USA, Großbritannien und der EU, haben vorgeschlagen, die bestehenden Genehmigungsanforderungen für politische Offline-Werbung auf den Online-Bereich auszuweiten. Dies beinhaltet nicht nur die Forderung nach einer eindeutigen Kennzeichnung von bezahlten Inhalten mit dem Hinweis, wer die Werbung autorisiert hat, sondern auch danach, die Nutzer darüber zu informieren, warum sie angesprochen wurden. Um diesen Anforderungen gerecht zu werden, haben Facebook und Twitter ein öffentliches Archiv von Werbeanzeigen eingeführt, das von jedermann erforscht und abgefragt werden kann. Der Mangel an bereitgestellten Details und die Unzuverlässigkeit des Angebots während wichtiger Wahlphasen haben jedoch gezeigt, wie wichtig zivilgesellschaftliche Initiativen wie etwa die in Großbritannien ansässige Organisation Who Targets Me oder das bereits erwähnte NYU Ad Observatory sind, um die Transparenz von Online-Werbung zu verbessern. Viele dieser Initiativen sind älter als die offiziellen Anzeigenarchive der Plattformen und nutzen Browser-Plug-ins, um per Crowdsourcing Informationen darüber zu sammeln, wo und wann Werbung in den Newsfeeds der Nutzer auftaucht.

Plattform-Architektur & -Design

Es gibt nach wie vor erhebliche Bedenken, dass die Architektur von Plattformen der öffentlichen Ordnung schadet. Diese reichen von Hinweisen auf diskriminierende Suchmaschinenergebnisse bis hin zu Befürchtungen, dass das Design von Social-Media-Plattformen die Verbreitung von spalterischen oder irreführenden Inhalten fördert. Im Mittelpunkt dieser Bedenken stehen unbeabsichtigte Konsequenzen durch das Design und die Algorithmen einer Plattform, welche das Nutzererlebnis und den „User Journey“ diktieren. Für diejenigen, die mit den internen Abläufen von Social-Media-Unternehmen nicht vertraut sind, bleiben sie schwer zu überprüfen und zu bewerten. Beispielsweise wurden Empfehlungssysteme dafür kritisiert, dass sie Nutzer dazu bringen, immer extremere Inhalte zu konsumieren, was möglicherweise eine politische Radikalisierung begünstigt. YouTube, dessen Unternehmensvertreter behaupten, dass Nutzer ihre Zeit zu 70 % mit empfohlenen statt aktiv gesuchten Videos verbringen, ist hier besonders erwähnenswert.

Zwar gibt es einige Versuche von Social Media-Unternehmen, in Bezug auf ihre Empfehlungsalgorithmen transparenter zu werden. Doch sind die Erklärungen für den Durchschnittsnutzer oft unverständlich. Für Forscher, die sich ein objektives Bild von der Reichweite von Inhalten machen möchten, stellen Empfehlungssysteme eine Erschwernis dar, weil sich das System an jeden Nutzer anpasst. Jeder einzelne Newsfeed ist einzigartig und damit auch der jeweilige Inhalt, der den Nutzer erreicht. Die Kuratierungsentscheidungen und Nutzersignale, die bestimmen, welche Inhalte präsentiert werden, bleiben für Außenstehende nicht nachvollziehbar. Facebook beispielsweise hat kürzlich Behauptungen widerlegt, dass konservative Kommentatoren auf der Plattform eine größere Reichweite haben als traditionelle Medienunternehmen. Journalisten hatten die öffentlich zugänglichen Daten von Facebook verwendet, um Reaktionen und Kommentare pro Tag und Woche zusammenzufassen und zu zeigen, dass an einem durchschnittlichen Tag die beliebtesten Inhalte von konservativen Experten dominiert werden. In einem Blog-Post erklärte Facebook, dass Engagement-Zahlen und Reichweite zwei verschiedene Metriken sind. Genau die Reichweite ist für unabhängige Forscher oder die breite Öffentlichkeit jedoch nicht messbar oder ersichtlich.

Ähnliche Fragen wurden vom französischen Conseil supérieur del'audiovisuel (CSA) aufgeworfen. Die mangelnde Transparenz der Plattformen in Bezug auf ihre Kuratierungsalgorithmen schränken die Möglichkeiten einer angemessenen Aufsicht erheblich ein. Im Hinblick auf weitere gesetzliche Regelungen hat ein aktueller Bericht des Ada Lovelace Institute eine detaillierte Anleitung veröffentlicht, wie verschiedene Arten von Algorithmus-Audits und Folgenabschätzungen genutzt werden können, um die Transparenz der Plattformen zu verbessern und letztlich die Rechenschaftspflicht der Unternehmen in diesem Bereich zu verschärfen.

Ansätze zum Erreichen von mehr Transparenz

Maßnahmen müssen an den oben genannten vier Kernbereichen ansetzen: 1. Inhalt & Moderation, 2. Beschwerden & Rechtsmittel, 3. Werbung, 4. Plattform-Architektur und -Design. Hier ist zusätzliche Transparenz erforderlich in Bezug auf die Richtlinien und Prozesse die für jeden dieser Bereiche gelten, sowie auf die Ergebnisse und ihre Auswirkungen. Die Richtlinien der Unternehmen, die in ihren Nutzungsbedingungen definiert sind, bestimmen die Regeln, die auf ihren Plattformen gelten. Daher ist es wichtig zu verstehen, wie sie festgelegt wurden und von wem. Die Prozesse der Unternehmen bestimmen dann, wie diese Richtlinien in der Praxis umgesetzt und durchgesetzt werden, einschließlich der menschlichen und automatisierten Systeme, die Entscheidungen auf der Grundlage der geltenden Richtlinien treffen. Auch hier ist Transparenz unerlässlich, um besser zu verstehen, wie und von wem diese Prozesse gestaltet und welche Sicherheitsvorkehrungen getroffen wurden, um Konsistenz zu gewährleisten und Verzerrungen zu verhindern. Schließlich ist weitere Transparenz erforderlich, um die Ergebnisse und Auswirkungen dieser Richtlinien und Prozesse besser zu verstehen und festzustellen, ob diese eventuell von kommerziellen Interessen der Technologieunternehmen ausgehebelt werden.

Als Reaktion auf das Auftreten von „Online Harms“ in Zusammenhang mit Social-Media-Plattformen und digitalen Diensten im Allgemeinen wurden verschiedene Modelle zur Verschärfung der Rechenschaftspflicht durch Transparenz entwickelt. Diese Modelle unterscheiden sich hinsichtlich ihres Umfangs, wobei sich einige auf bestimmte Funktionen und Dienste von Plattformen und/oder die Richtlinien, Prozesse und Ergebnisse in diesen Bereichen konzentrieren, während andere einen eher ganzheitlichen Ansatz verfolgen. Im Folgenden wird ein kurzer Überblick über einige dieser Modelle gegeben.

Verfahrenstechnische Rechenschaftspflicht

Das im Mai 2019 veröffentlichte „French framework to make social media platforms more accountable“ betont die Notwendigkeit einer engen Zusammenarbeit zwischen privaten Unternehmen, einer unabhängigen Regulierungsbehörde und der Regierung. Ausgangspunkt ist, dass sich das Vertrauensverhältnis zwischen Social-Media-Unternehmen, Regulierungsbehörden und Nutzern in den letzten Jahren verschlechtert hat. Der Kern dieses Vertrauensmangels ist eine Informationsasymmetrie, die nur durch einen kontinuierlichen offenen Dialog zwischen Unternehmen, Regulierungsbehörden und der Zivilgesellschaft verbessert werden kann. Transparenz ist somit ein notwendiger erster Schritt für die gesamte Gesellschaft, um das Problem der „Online Harms“ angehen zu können.

Im Gegensatz zur Ex-post-Regulierung über Notice-and-Takedown-Regelungen fordert das französische Rahmenwerk eine verbesserte verfahrenstechnische Rechenschaftspflicht, d.h. „die Auferlegung starker Verpflichtungen zur Transparenz von Schlüsselsystemen, die von außen nicht einsehbar sind, d.h. das Moderationssystem (und Verfahren zur Entwicklung und Aktualisierung der ihm zugrunde liegenden Nutzungsbedingungen) sowie die Verwendung von Algorithmen zur gezielten und personalisierten Darstellung von Inhalten“. In Anlehnung an das Modell der Finanzprüfung im Bankensektor wird die Umsetzung von Präventiv- und Korrekturmaßnahmen durch “systemrelevante” Unternehmen unter Aufsicht einer Regulierungsbehörde vorgeschlagen. Die Regulierungsbehörde würde nicht versuchen, spezifische Fälle zu regulieren, in denen ein Schaden eintreten könnte oder eingetreten ist, sondern vielmehr jenen Social-Media-Unternehmen, die als „systemische“ Akteure gelten, eine „Sorgfaltspflicht“ („duty of care“ Ansatz) auferlegen.

Transparenz als Standard

Der Verhaltenskodex für altersgerechtes Design des britischen Information Commissioner's Office (ICO) verfolgt einen anderen Ansatz in Bezug auf Transparenz und Rechenschaftspflicht und konzentriert sich neben den unternehmensinternen Prozessen auch auf die Nutzerseite. Im Mittelpunkt des Kodex steht die Annahme, dass es bei Transparenz in erster Linie darum geht, den Nutzern auf „klare, offene und ehrliche“ Weise zu kommunizieren, „was sie erwarten können, wenn sie auf Online-Dienste zugreifen“. Neben klaren Beschreibungen, wie persönliche Daten verwendet werden dürfen, fordert der Kodex transparente und leicht verständliche Nutzungsbedingungen und Community-Standards. Darüber hinaus wird „Nudging“ explizit als ein Werkzeug erwähnt, das proaktiv eingesetzt werden kann, um Nutzer zu einem datenschutzbewussten Verhalten zu ermutigen, etwa durch das Abmelden von Empfehlungssystemen.

Das Fehlen einer nutzerfreundlichen Gestaltung wird auch in der jüngsten unabhängigen Überprüfung des deutschen NetzDG-Gesetzes kritisiert und wurde bereits wiederholt von Jugendschutz.net angesprochen. Beide fordern transparentere und leicht zugängliche Meldeverfahren seitens der Social-Media-Unternehmen. Dazu gehört ein gut sichtbarer Meldebutton, der es Nutzern, auch Kindern und Jugendlichen, ermöglicht, unangemessene Inhalte, die sie auf Plattformen finden, zu melden. Der Kerngedanke dieses „Safety-by-Design“-Ansatzes ist, dass „Online Harms“ nicht nur durch Inhalte entstehen, sondern auch durch Online-Interaktionen zwischen Nutzern, die von Social-Media-Unternehmen vermittelt werden.

Während sowohl der ICO-Kodex als auch die Arbeit von Jugenschutz.net in erster Linie darauf abzielt, Minderjährige vor „Online Harms“ zu schützen, kann verstärkte Nutzererfahrungsforschung zur Förderung eines datenschutzbewussten und sicheren Online-Verhaltens als Modell verwendet werden, um Anbieter digitaler Dienste in einer Vielzahl von Umgebungen zur Verantwortung zu ziehen. Dies betrifft auch die Nutzung sozialer Medien durch Erwachsene.

Ein menschenrechtlicher Rahmen für digitale Politikgestaltung

David Kaye, der ehemalige UN-Sonderberichterstatter zur Meinungsfreiheit, hat sich neben anderen für einen rechtebasierten Ansatz bei der Online-Regulierung und insbesondere bei der Moderation von Inhalten eingesetzt. Seine zentrale Forderung ist, dass sich die Normen für die Moderation von Inhalten an Menschenrechtsstandards, wie sie in der Allgemeinen Erklärung der Menschenrechte festgelegt sind, orientieren sollten. Dies bedeutet nicht, dass jede Form der Meinungsäußerung in den sozialen Medien erlaubt sein sollte. Vielmehr geht es darum, dass in den Geschäftsbedingungen des Unternehmens und in den entsprechenden staatlichen Vorschriften klar zum Ausdruck kommt, wann und warum eine Einschränkung des Rechts auf freie Meinungsäußerung notwendig und verhältnismäßig ist.

Der Versuch, verhältnismäßig und nur dort zu regulieren, wo es notwendig ist, liegt dem risikobasierten Ansatz zugrunde, den das britische „Online Harms White Paper“ vorsieht, oder dem Fokus auf systemische Akteure im französischen Rahmenwerk. Kaye und das Forum for Democracy & Information gehen jedoch noch weiter, indem sie fordern, dass sich Unternehmen oder Regulierungsbehörden bei ihren Entscheidungen auf die internationale Menschenrechtsrechtsrechtsprechung beziehen sollten. Dies erfordert sowohl „Regelwerkstransparenz“ als auch „Entscheidungstransparenz“, die erreicht wird, indem man den Entscheidungsprozess, der hinter einer Plattformaktion steht, klar darlegt. Diese Transparenz kann die Grundlage für die Rechenschaftspflicht von Unternehmen und Regierungen bilden, weil die Öffentlichkeit die getroffenen Entscheidungen überprüfen und Einspruch dagegen erheben kann.

Basierend auf diesem Ansatz hat der jüngste Bericht des Forum for Information & Democracy eine einzige Ex-ante-Anforderung an Plattformen in Bezug auf Transparenz vorgeschlagen, nämlich eine sogenannte menschenrechtliche Folgenabschätzung ihrer Dienste und aller vorgeschlagenen Änderungen daran, einschließlich der Praktiken zur Moderation von Inhalten.

Anhang: Transparenzrahmen

Das folgende Rahmenkonzept zum Verständnis von Bereichen, in denen mehr Transparenz erforderlich sein könnte, wurde vom ISD auf der DPL-Veranstaltung im November 2020 vorgestellt. Es kombiniert die verschiedenen Kategorien, die in diesem Briefing skizziert wurden, gibt Beispiele für aktuelle Fragen zu „Online Harms“ und nennt einige bestehende Initiativen (sowohl aus dem öffentlichen als auch dem privaten Sektor).

Beispiele für „Online Harms“	Richtlinien	Prozesse	Ergebnisse
Desinformation Studie 1 Studie 2 Studie 3 Studie 4	Was ist eine unerlaubte Koordinierung? Wie unterscheiden Unternehmen zwischen gefälschten und koordinierten Accounts?	Wie groß ist der sprachliche und kulturelle Umfang dieser Untersuchungen und Teams? Sind der Iran und Russland die meistzitierten Quellen, weil die Unternehmen sich in ihren Untersuchungen auf diese Länder konzentrieren?	Wie groß ist der Einflussbereich von koordinierten Netzwerken auf reale Nutzerunterhaltungen in den sozialen Medien? Wie viele Nutzer beschäftigten sich mit Inhalten, die von inauthentischen Accounts verbreitet wurden?
Verschwörungs-theorien & Rekrutierung von Extremisten Studie 1 Studie 2 Studie 3 (deutsch) NBC-Bericht	Warum dürfen einige extremistische Gruppen aktiv online bleiben und andere werden entfernt? Wer trifft diese Entscheidungen?	Wie fördern Empfeh-lungsalgorithmen extremistische Gruppen oder stufen sie zurück? Wie hat sich das interne Ziel der Plattform, „Gruppen“ proaktiv zu vergrößern, darauf ausgewirkt, wie viele Nutzer QAnon-Gruppen beigetreten sind?	Wie hoch ist die Anzahl von Mitgliedern öffentlicher und privater extremistischer Gruppen? Wie hoch sind die Gewinne, die mit von extremistischen Gruppen gekaufter Werbung erzielt werden?
Belästigung und Beleidigung Studie 1	Welche Inhalte (Text, Bild, Video) fallen unter relevante Belästigungs-richtlinien, welche nicht? Welche Unternehmensrichtlinien gibt es für die Aufbewahrung von Beweisen für Belästigung und Morddrohungen, damit die Opfer rechtliche Schritte einleiten können?	Welche Unternehmensressourcen werden für Inhaltsmoderation eingesetzt? Wie ist es um die sprachliche und kulturelle Kontextexpertise der entsprechenden Teams bestellt? Wie ist die Balance zwischen KI und menschlicher Moderation bei der Erkennung von Belästigung und Beleidigung? Welche Daten werden verwendet werden, um diese Systeme zu trainieren?	

Beispiele für aktuelle Initiativen

Problembereiche	Richtlinien	Prozesse	Ergebnisse
Inhalt & Moderation	Tech Against Terrorism Mentoring für die Nutzungsbedingungen kleinerer Tech-Plattformen	FSM Social Science One	Gesetzlich vorgeschriebene NetzDG-Transparenzberichte
Beschwerden & Rechtsmittel		Facebook Oversight Board FSM als die nach NetzDG anerkannte Einrichtung der regulierten Selbstregulierung	Öffentliche Berichte über Entscheidungen, die im Rahmen des FSM-Beschwerdeverfahrens getroffen wurden
Werbung			Politische Anzeigenarchive von Facebook und Google
Plattform-Architektur & Design	Algorithmische Kriterien von TikTok		

Potenzielle zukünftige Transparenzanforderungen

Problembereiche	Richtlinien	Prozesse	Ergebnisse
Inhalt & Moderation		Geregeltes Reporting über die Anzahl der Content-Moderatoren, Sprachen, Expertise etc. Freiwilliges Transparenz-Protokoll der OECD	
Beschwerden & Rechtsmittel	Einschätzung menschenrechtlicher Folgen		Transparenzberichterstattung über Umfang und Entscheidungen zu Einsprüchen von Nutzern
Werbung	Regulierte Werbetransparenz-anforderungen für alle Anzeigen		
Plattform-Architektur & Design		Interviews der Regulierungsbehörde mit Datenwissenschaftlern zur Anpassung des algorithmischen Designs	Algorithmische Audits – Kontrolltests Datenübertragbarkeit

ISD | Powering solutions to extremism and polarisation

Digital Policy Lab
Diskussionspapier

Nationale & internationale Modelle zur Online-Regulierung

Über dieses Diskussionspapier

Dieses Diskussionspapier ist eine Momentaufnahme der internationalen politischen Debatte über Ansätze zur Bekämpfung von Desinformation, Hassrede, Extremismus und Terrorismus im Internet. Es fasst die politischen Initiativen zusammen, die während des zweiten, vom Auswärtigen Amt geförderten Digital Policy Lab (DPL) am 10. und 11. Dezember 2020 vorgestellt wurden, und bietet darüber hinaus einen Überblick über zusätzliche nicht-regulatorische und regulatorische politische Initiativen der Five Eyes- und ausgewählter EU-Länder. Diese bilden einen Rahmen für die Diskussion über Vorteile, Fallstricke und Einschränkungen der bestehenden Bemühungen in diesem Bereich. Die in diesem Papier geäußerten Ansichten sind die der Autoren und spiegeln nicht unbedingt die Ansichten der Teilnehmer am Digital Policy Lab oder der Regierungen wider. Aufgezeichnet wurde der Inhalt von dem auf Extremismusbekämpfung und die Analyse des Internets spezialisierten Think Tank Institute for Strategic Dialogue gGmbh (Berlin) in Zusammenarbeit mit den Londoner Headquarter des Instituts.

Zur besseren Lesbarkeit wurde auf die gleichzeitige Verwendung männlicher und weiblicher Sprachformen verzichtet. Es wurde das generische Maskulinum verwendet, wobei alle Geschlechter gleichermaßen gemeint sind.

ISD | Powering solutions to extremism and polarisation

Beirut | Berlin | London | Paris | Washington DC

www.isdglobal.org

Einführung

Extremisten aus dem gesamten ideologischen Spektrum wissen mittlerweile, wie sie Online-Produkte, Mediensysteme und Plattformen manipulieren können, um das Publikum zu täuschen, den Informationsfluss zu verzerren und illegale Aktivitäten durchzuführen. Ausländische, inländische und transnationale Akteure setzen Taktiken ein, um die soziale und politische Polarisierung zu fördern, die Verfügbarkeit genauer und transparenter Informationen zu erschweren, demokratische Prozesse zu untergraben sowie um ausgrenzende und extreme politische Agenden zu verbreiten. Diese neuen Praktiken der Kommunikation und politischen Mobilisierung entwickeln sich viel schneller als der gesetzliche Rahmen, in dem auf sie reagiert werden kann.

In Wahlkampfzeiten werden die Normen, die bisher als Richtschnur dafür dienten, was legitime und illegitime demokratische politische Kampagnen sind, durch eine Vielzahl neuer Technologien in Frage gestellt. Das Online-Ökosystem aus sozialen Medien und Kommunikationswerkzeugen bietet staatlichen und nichtstaatlichen Akteuren neue digitale Wege, ihren Einfluss geltend zu machen und dabei neue Formen der Anonymität und Spurenverwischung zu nutzen. Ausländische Staaten und transnationale extremistische Netzwerke haben Kampagnen gestartet, die falsche Informationen über nahezu alle entscheidenden Zukunftsthemen (wie z.B. den Klimawandel, Migration oder Impfungen), sowie über Politiker, Regierungen, Aktivisten und Minderheiten verbreiten, dazu gibt es ausreichend Beispiele in den USA, Deutschland, Schweden, Italien, Frankreich, Mexiko, Brasilien und andernorts. Wissenschaftler, Regierungen und Technologieunternehmen haben zahlreiche Beweise für die raffinierte und konzertierte Irreführung und Manipulation des Publikums, oft verbunden mit dem Ziel, Intoleranz, Empörung und sogar Gewalt zu fördern. Es besteht kein Zweifel daran, dass die Kombination von Hassrede, Extremismus, Desinformation und Verschwörungstheorien eine Gefahr für die Demokratie darstellt. Es besteht akuter Handlungsbedarf, damit die Gemeinschaft demokratischer Staaten effektiv und angemessen auf diese komplexen Herausforderungen reagieren kann.

Jenseits von Wahlen sind drängende globale Zukunftsthemen wie Klimawandel, Migration und öffentliche Gesundheit ins Visier extremistischer Akteure geraten. Wir beobachten eine wachsende Intersektionalität bei Informationsoperationen, die darauf angelegt sind, das Vertrauen in demokratische Institutionen und die Wissenschaft zu erschüttern. Im aktuellen Kontext von COVID-19 standen beispielsweise Klimawandelleugner oft an der Spitze der Bemühungen, die Schwere der Pandemie mit Hilfe von Online-Desinformation herunterzuspielen. Diese Kräfte sind nicht neu, aber eine Reihe hybrider Bedrohungen wurde durch die digitale Technologie exponentiell aufgeladen. Noch nie zuvor konnten einige wenige Engagierte mit ihren Ideen so schnell so viele Menschen erreichen, begünstigt durch die algorithmische Verstärkung von sensationslüsternen und extremen Botschaften in den sozialen Medien.

Regierungen, Technologieunternehmen, Wahlausschüsse und zivilgesellschaftliche Gruppen versuchen, mit den Auswirkungen dieser neuen Entwicklungen umzugehen. Regulierungsbemühungen konnten bisher nicht mit dem rasanten technologischen Fortschritt im Hinblick auf Werbung, der künstlichen Verstärkung von Narrativen und Publikumssegmentierung Schritt halten. Definitionen bleiben umstritten. Fragen nach der Absicht und dem Einfluss sind nach wie vor von großer Bedeutung, wenn es darum geht, die illegitime Nutzung von Technologien von der legitimen abzugrenzen, wobei diese Unterschiede oft am schwierigsten zu treffen und politisch am stärksten aufgeladen sind.

Demokratien haben sich mit der Frage auseinandergesetzt, wie die Digitalpolitik den sozialen Zusammenhalt und die öffentliche Sicherheit effektiv sichern und gleichzeitig die Rechte der Internetnutzer schützen kann. Der Umgang mit Sprache und Informationen in einer liberalen Gesellschaft ist eine aufwändige Übung in langsamer Regulierung. Sorgfalt und Vorsicht sowie Zaghaftigkeit und Geduld werden in der schnelllebigen Welt der Technologie leicht ausgenutzt. Historisch gesehen war das Internet ein mächtiges Werkzeug für die Projektion liberaler Werte. Die Gleichheit des Rechts auf Information, der freie Zugang zu den Medien und die Meinungsfreiheit sind in den Protokollen und in der Infrastruktur, auf der das Internet aufbaut, festgeschrieben. Diese Möglichkeiten zu schützen und gleichzeitig ihren Missbrauch zu verhindern, ist die zentrale Herausforderung.

Ansätze einer digitalen Politik

Vor mehr als zwei Jahrzehnten trat der Communications Decency Act in den USA in Kraft und die E-Commerce-Richtlinie wurde von der Europäischen Union verabschiedet. Damit waren die grundlegenden Haftungsregelungen geschaffen, auf denen wesentliche Teile des öffentlichen Internets aufbauen sollten. Im Mittelpunkt dieser Maßnahmen stand die Überzeugung, dass die Meinungsfreiheit und eine florierende Internetwirtschaft nur dann gewährleistet werden können, wenn die Anbieter von Online-Diensten (Intermediäre) nicht für die von den Nutzern erstellten Inhalte auf ihren Plattformen haftbar gemacht werden können. Seitdem haben die zunehmenden Bedrohungen durch Desinformation, Hassrede und Extremismus in den sozialen Medien die Grenzen dieses Ansatzes aufgezeigt.

Zunächst entstanden zahlreiche **Initiativen zur Selbst- oder Ko-Regulierung**. Diese versuchten, Online-Plattformen zur Zusammenarbeit bei der Bekämpfung sowohl illegaler Aktivitäten wie Terrorismus oder Kindesmissbrauch als auch „legaler aber schädigender Aktivitäten" wie Desinformation oder das Veröffentlichen von Inhalten, die Selbstschädigung fördern, zu ermutigen. Daneben wurden viele andere Ansätze zum Umgang mit Herausforderungen wie Online-Hassrede, Extremismus, Terrorismus und Desinformation umgesetzt – darunter Strategien zur Gegenkommunikation (counter speech), Programme zur Förderung digitaler Medienkompetenzen sowie Kampagnen zur Sensibilisierung der Öffentlichkeit für die Gefahren des Internets.

Obwohl informelle, freiwillige oder industriegeführte Ansätze in bestimmten Bereichen für Verbesserungen gesorgt haben, sahen sich viele Regierungen gezwungen, die Debatte über die Regulierung der digitalen Sphäre neu zu eröffnen, um diese Herausforderungen effektiver anzugehen. Dieser aufkommende Trend zu einer neuen Online-Regulierung zeigt sich in zwei Arten von Ansätzen:

- **Inhaltsbezogene Ansätze**, die oft auf einen bestimmten „Online Harm"[1] abzielen, wie Hassrede oder Desinformation bei Wahlen. Diese Ansätze konzentrieren sich ggf. auf die effektive und rechtzeitige Entfernung dieser Inhalte.

- **Systemische Ansätze**, bei denen Online-Plattformen nachweisen müssen, dass sie bei der Konzipierung und Umsetzung ihrer Richtlinien, Prozesse und Systeme potenziell negative Folgen bedacht haben. Auch wenn dies nicht in den Rahmen dieses Papiers fällt, sei darauf verwiesen, dass die Europäische Union und eine wachsende Zahl demokratischer Länder versuchen, marktbeherrschende Anbieter von Onlinediensten durch Wettbewerbs- und Datenschutzgesetze zu regulieren. Dies könnte mitunter eine positive Wirkung auf den Rückgang von „Online Harms" haben.

[1]Der Begriff „Online Harms" geht auf das Online Harms White Paper der britischen Regierung zurück, welches im April 2019 veröffentlicht wurde. „Online Harms" steht sowohl für konkrete als auch potenzielle Schäden (sprich Gefahren) die z.B. durch das Verbreiten von problematischen Inhalten in sozialen Medien entstehen. Da der englische Begriff die deutschen Begriffe Schäden, Gefahren und Risiken vereint, wird in diesem Arbeitspapier der englische Originalbegriff verwendet.

Terrorismus & gewalttätiger Extremismus

Auf terroristische und extremistische Aktivitäten auf Online-Plattformen reagieren zu können, ist Regierungen seit über einem Jahrzehnt ein wichtiges Anliegen. Große Aufmerksamkeit erlangte das Thema jedoch ab 2014 mit der großflächigen Verbreitung von ISIS-Propagandamaterial und Rekrutierungsnetzwerken in den sozialen Medien, die den Vormarsch der Terrorgruppe im Irak und in Syrien begleiteten. Seit dem Terroranschlag von Christchurch im März 2019 rücken auch rechtsextreme Inhalte und Netzwerke im Internet immer mehr in den Fokus. Sowohl Regierungen als auch Plattformen haben mit dem Ausmaß der Inhalte und der Widerstandsfähigkeit von Online-Netzwerken zu kämpfen. Hinzu kommt, dass komplexe juristische Erwägungen und menschenrechtliche Implikationen von terroristischen Inhalten die Entfernung von Accounts erschweren.

Was die **Selbst- bzw. Ko-Regulierung** anbelangt, hat die **Europäische Kommission** im Dezember 2015 das EU-Internetforum ins Leben gerufen. Das Forum brachte die EU-Innenminister, hochrangige Vertreter großer Internetunternehmen, Europol, den EU-Koordinator für Terrorismusbekämpfung und das Europäische Parlament mit dem Ziel zusammen, einen gemeinsamen, freiwilligen Ansatz auf der Grundlage einer öffentlich-privaten Partnerschaft zu etablieren, um terroristisches Material im Internet erkennen und bekämpfen zu können.

Im Jahr 2017 folgte die Gründung des Global Internet Forum to Counter Terrorism (GIFCT), einer branchenübergreifenden Organisation, die verhindern sollte, dass Terroristen und gewalttätige Extremisten digitale Plattformen ausnutzen. Gefördert wird das GIFCT vom **EU-Internetforum** und vom **britischen Innenministerium**. Dazu gehört die Einrichtung einer plattformübergreifenden Hash-Sharing-Datenbank für gewalttätige terroristische Bilder und Propaganda, die inzwischen über 200.000 eindeutige digitale „Fingerabdrücke" enthält. 2019 wurde das Content Incident Protocol (CIP) von der GIFCT entwickelt, um Plattformen die Möglichkeit zu geben, in Echtzeit auf terroristische Ereignisse zu reagieren, die Koordination zu unterstützen und das Teilen von aufkommenden terroristischen Inhalten zu verhindern. Dieses Protokoll wurde erstmals am 9. Oktober 2019 nach dem Anschlag in Halle aktiviert, und zwar nachdem der Angreifer seinen Angriff gefilmt und den Livestream beim Streaming-Dienst Twitch hochgeladen hatte. Dabei ist anzumerken, dass diese Plattform zu dem Zeitpunkt kein GIFCT-Mitglied war.

Mittlerweile sind einige weitere bemerkenswerte internationale Initiativen zur Eindämmung terroristischer Inhalte im Internet entstanden, darunter der Christchurch Call, der von der **neuseeländischen Premierministerin Jacinda Ardern** und dem **französischen Präsidenten Emmanuel Macron** nach dem Terroranschlag von Christchurch am 15. März 2019 initiiert wurde. Er enthält einen Aktionsplan für kollektive Reaktionen von Tech-Unternehmen, Regierungen und internationalen Organisationen mit dem Ziel, gewalttätige extremistische Online-Inhalte zeitnah zu beseitigen. Es handelt sich dabei um die erste große internationale Initiative, die rechtsextreme terroristische Inhalte auf die internationale Agenda gesetzt hat und damit eine notwendige Debatte erweiterte, die zuvor vom Kampf gegen islamistischen Terrorismus dominiert wurde.

Parallel dazu haben sich weitere internationale Dialoge zur Bekämpfung der terroristischen Nutzung des Internets entwickelt, etwa der vom **jordanischen König Abdullah II.** ins Leben gerufene Aqaba-Prozess und die Zürich-London-Empfehlungen des Global Counterterrorism Forum zur Verhinderung und Bekämpfung von gewalttätigem Extremismus und Terrorismus im Internet. In Anerkennung der Kapazitätslücke bei kleineren digitalen Plattformen arbeiten die Initiative Tech Against Terrorism des **UN Counter Terrorism Executive Directorate (UN CTED)** und das Projekt der **OECD** zur freiwilligen Transparenzberichterstattung mit Unternehmen des Privatsektors zusammen, um den terroristischen Missbrauch des Internets zu bekämpfen und gleichzeitig die Menschenrechte zu respektieren.

Andere nicht-regulatorische Ansätze umfassen **kommunikationsorientierte Bemühungen** zur Bekämpfung terroristischer und extremistischer Propaganda, die von nationalen Regierungen und auf internationaler Ebene, oft in Zusammenarbeit mit dem Privatsektor, durchgeführt werden. Initiativen wie das Global Engagement Center im US-Außenministerium und die Counter Daesh Communication Cell, die

vom britischen Außenministerium geleitet wird und 82 internationale Partner einschließt, haben strategische Kommunikations- und Counter-Speech-Ansätze entwickelt, die auf die „Nachfrageseite" von terroristischen Online-Inhalten abzielen. Es handelt sich hier sowohl um vorgelagerte präventive Kommunikationsansätze, die darauf abzielen, eine gewisse Widerstandsfähigkeit gegen gewalttätige extremistische oder terroristische Narrative breiterer Zielgruppen aufzubauen, als auch um nachgelagerte Strategien, die darauf abzielen, die Narrative gewalttätiger extremistischer oder terroristischer Gruppen direkt zu widerlegen oder zu kontern.

Im regulatorischen Bereich beruhen die meisten Bemühungen zur Bekämpfung von gewalttätigem extremistischen oder terroristischen Material bisher auf der Regulierung von Benutzerinhalten durch ein Notice-and-Takedown-Modell, das dem etablierten Urheberrecht entlehnt ist, insbesondere dem US Digital Millennium Act (DMCA) von 1996. Die Begründung für diese **inhaltsbasierten Ansätze** ist: „Was offline illegal ist, ist auch online illegal." Social-Media-Unternehmen müssen für die Inhalte auf ihren Plattformen verantwortlich gemacht werden, sobald sie darüber informiert oder entsprechend gewarnt wurden.

In **Großbritannien** wurde 2010 die Counter Terrorism Internet Referral Unit (CTIRU) eingerichtet, die sich auf Grundlage der bestehenden Gesetzgebung um die Entfernung von rechtswidrigen Inhalten mit terroristischem Inhalt aus dem Internet bemüht. Inhalte, die zu terroristischen Handlungen auffordern oder diese verherrlichen, können in Großbritannien gemäß Abschnitt 3 des Terrorism Act (2006) entfernt werden. Die von der Metropolitan Police betriebene CTIRU stellt eine Liste von URLs für Material zusammen, das außerhalb Großbritanniens gehostet wird und in öffentlichen Netzwerken gesperrt ist, und meldet die Inhalte zur Entfernung an Internetunternehmen. In einer durchschnittlichen Woche entfernt die CTIRU über 1.000 Inhalte, die gegen die Terrorismusgesetzgebung in Großbritannien verstoßen.

In Anerkennung der Notwendigkeit einer internationalen Koordinierung wurde 2015 die Internet Referral Unit (EU-IRU) von **Europol** eingerichtet, um terroristische und gewalttätige extremistische Online-Inhalte zu markieren, sie mit den zuständigen Regierungspartnern zu teilen und an Unternehmen weiterzuleiten, die diese Inhalte hosten, damit diese sie entfernen können. Während die EU-IRU keine rechtliche Befugnis hat, Unternehmen zur Entfernung von Inhalten zu zwingen, wurden in Frankreich, den Niederlanden, Belgien, Deutschland und Italien parallele Verweisungsstellen und -mechanismen entwickelt.

Im September 2018 legte die **Europäische Kommission** einen Vorschlag vor, der Social-Media-Unternehmen dazu zwingen soll, terroristische Inhalte innerhalb einer Stunde nach Erhalt einer Benachrichtigung durch nationale Behörden zu entfernen. Wenn die Verordnung in ihrer jetzigen Form vom EU-Parlament und Ministerrat angenommen wird, würde sie über ein Notice-and-Takedown-Modell hinausgehen, indem sie einen proaktiven, **systemischen Ansatz** basierend auf einer „Sorgfaltspflicht" (duty of care) verankert. Die Unternehmen müssten also sicherstellen, dass ihre Plattformen nicht zur Verbreitung terroristischer Inhalte genutzt werden. In ähnlicher Weise wird ein spezifischer Verhaltenskodex für die Bekämpfung von Terrorismus im Internet in die bevorstehende britische Gesetzgebung zu „Online Harms" aufgenommen. Dieser soll 2021 im Parlament eingebracht werden (siehe Seite 11).

Hassrede

In den letzten fünf Jahren gab es eine Reihe von staatlichen Initiativen, vor allem der Europäische Kommission und des deutschen Ministeriums für Justiz und Verbraucherschutz, um Social-Media-Unternehmen zu verpflichten, Hassrede auf ihren Plattformen durch Selbstregulierung zu bekämpfen. Im Juni 2016 hat die Europäische Kommission einen Verhaltenskodex zur Bekämpfung illegaler Hassrede im Internet eingeführt und Social-Media-Plattformen zu dessen Unterzeichnung aufgefordert. Die teilnehmenden Unternehmen haben sich freiwillig verpflichtet, schneller auf illegale Hassrede zu reagieren, ihre Mitarbeiter besser zu schulen und verstärkt mit dem zivilgesellschaftlichen Sektor zusammenzuarbeiten. Der EU-Kodex folgt den freiwilligen Selbstverpflichtungen, die Branchenvertreter bereits im Dezember 2015 im Rahmen der von Heiko Maas (damals Bundesjustizminister) geleiteten deutschen Task Force „Hass im Netz" abgegeben haben.

Während die Bewertung des eigenen Verhaltenskodexes durch die Europäische Kommission positiv ausfiel, stellte Jugendschutz.net gravierende Mängel bei der Beseitigung von Hassrede im Rahmen dieser Selbstregulierungsansätze fest. Als Reaktion darauf verfolgte das deutsche Justizministerium einen **inhaltsbasierten Regulierungsansatz** und legte im März 2017 einen Entwurf für ein sogenanntes Netzwerkdurchsetzungsgesetz (NetzDG) vor. Dieser Gesetzentwurf wurde u. a. vom UN-Sonderberichterstatter für Meinungsfreiheit kritisiert. Er äußerte Bedenken hinsichtlich eines möglichen „Overblocking"[1] von Nutzerinhalten durch Social-Media-Unternehmen. Das Gesetz trat dennoch mit geringfügigen Änderungen im Juni 2017 in Kraft.

Seit Januar 2018 verpflichtet das NetzDG Social-Media-Unternehmen mit mehr als 2 Millionen deutschen Nutzern, zeitnah Maßnahmen gegen illegale, auf ihren Plattformen geteilte Inhalte zu ergreifen, nachdem diese von Nutzern markiert wurden. In der Regel haben Social-Media-Unternehmen sieben Tage Zeit, um auf gemeldete Inhalte zu reagieren. Diese Frist verkürzt sich auf 24 Stunden, wenn der gemeldete Inhalt offensichtlich rechtswidrig ist.

Bei Nichtbeachtung drohen Geldstrafen von bis zu 50 Millionen Euro. Das Gesetz verlangt von den Unternehmen auch, dass sie alle sechs Monate einen Bericht über Bemühungen zur Reduzierung von Hassrede auf ihrer Plattform vorlegen. Dies schließt Zahlen zu den von Nutzern erhaltenen Meldungen und den jeweils ergriffenen Maßnahmen ein. Nutzer, die von Hassrede betroffen sind, können sich an ein Gericht wenden, mit dessen Zustimmung die jeweilige Plattform Daten zu ihrem konkreten Fall herausgeben muss, anhand derer eine Identifikation des Täters möglich ist. Social-Media-Unternehmen können in schwierigen Fällen die Entscheidung bezüglich gemeldeter Inhalte auch an einen „regulierten Selbstregulierer", wie die FSM, delegieren.

Im Juli 2019 wurde vom Bundesamt für Justiz ein Bußgeldbescheid in Höhe von zwei Millionen Euro gegen Facebook erlassen, weil die Plattform in seinen halbjährlichen Berichten keine genauen Angaben zu Kennzeichnungs- und Löschstatistiken gemacht hatte. Im September 2020 stellte eine vom Justizministerium in Auftrag gegebene unabhängige gerichtliche Überprüfung fest, dass das Gesetz im Großen und Ganzen wirksam zur Reduzierung von Hassrede auf Plattformen beiträgt, empfahl aber nutzerfreundlichere Kennzeichnungsprozesse, verbesserte Transparenzanforderungen für Social-Media-Unternehmen und bessere Möglichkeiten für Nutzer, sie betreffende Entscheidungen anzufechten. Im Juni 2020 wurde eine Gesetzesänderung vom Bundestag verabschiedet, die Plattformen dazu verpflichtet, bestimmte Arten von gekennzeichneten kriminellen Inhalten direkt an die Strafverfolgungsbehörden zu melden. Allerdings wurde die Gesetzesänderung auf Grund von verfassungsrechtlichen Bedenken bisher noch nicht vom Bundespräsidenten unterzeichnet (Stand Januar 2021).

1 „Overblocking" bezeichnet die systematische, maßlose Löschung auch rechtmäßiger Inhalte auf Grund von gesetzlichen Löschpflichten

Frankreich hat ähnliche inhaltsbezogene Anstrengungen unternommen, und zwar mit einem Gesetzentwurf „zur Bekämpfung von Online-Hass", der nach seiner Verfasserin Laetitia Avia als „Avia-Gesetz" bezeichnet wird. Zusätzlich zu der 24-Stunden-Frist für die Entfernung von Inhalten, ähnlich wie beim NetzDG, sah das Gesetz eine 1-Stunden-Frist für die Entfernung von terroristischen Inhalten oder Kindesmissbrauchsmaterial sowie Geldstrafen von bis zu 1,25 Millionen Euro für Versäumnisse vor. Anders als beim NetzDG wäre der Vollzug des Gesetzes von der Medienaufsichtsbehörde, dem Conseil supérieur de l'audiovisuel (CSA) überwacht worden. Im Jahr 2020 wurde der Großteil des Gesetzes jedoch vom französischen Verfassungsrat als Verstoß gegen die Meinungsfreiheit verworfen. Es wurden Bedenken hinsichtlich der möglichen Unverhältnismäßigkeit und des Potenzials für ein „Overblocking" von Inhalten geäußert.

Im September 2020 hat die **österreichische** Regierung unter der Ägide des Kanzleramtes und Justizministeriums ein ähnliches Gesetz mit dem Namen „Kommunikationsplattformgesetz" vorgeschlagen, das auf Plattformen mit mehr als 100.000 inländischen Nutzern oder einem Jahresumsatz von mehr als 500.000 € abzielt. Zusätzlich sieht der **österreichische** Entwurf ein spezielles Beschwerdeverfahren für Nutzer vor, um ihre Inhalte wiederherzustellen, wenn die Entfernung durch die Plattform als ungerechtfertigt angesehen wird (etwa durch die Aufsichtsbehörde KommAustria). Neben möglichen Bußgeldern von bis zu 10 Mio. € sieht das Gesetz indirekten finanziellen Druck vor, der bei Nichteinhaltung ausgeübt werden könnte, etwa die Sperrung der Zahlung von Werbeeinnahmen an Plattformen. Die österreichische Regierung plant allerdings nicht, Meldepflichten an die Strafverfolgungsbehörden einzuführen, anders als die jüngste **deutsche** Gesetzesnovelle.

Desinformation & ausländische Einmischung

Anders als die Bemühungen, Hassrede sowie gewalttätiges, extremistisches oder terroristisches Material zu bekämpfen, konzentrierten sich Regulierungs- und Selbstregulierungsinitiativen rund um Desinformation hauptsächlich auf Wahlen. Diese Initiativen hatten eine Erhöhung der Transparenzanforderungen zur Folge, die politische Kampagnen auf sozialen Medien, Suchmaschinen und anderen offenen Plattformen betreffen.

In jüngerer Zeit haben die COVID-19-Pandemie und die Verbreitung von Gesundheitsfehlinformationen im Internet den Umfang der (Selbst-) Regulierungsbemühungen erweitert, was bisher hauptsächlich zu Änderungen der Community-Richtlinien der Social-Media-Unternehmen geführt hat– mit mäßigem Erfolg wie das ISD herausfand. Die Regulierung oder Gesetzgebung in Bezug auf die manipulative Nutzung von Informationen ist eine Herausforderung, die mit potenziellen Fallstricken, kontraproduktiven Nebeneffekten und besorgniserregenden Risiken für die Meinungsfreiheit behaftet ist. Doch die Art der Bedrohung durch Desinformation erfordert eine schnelle und umfassende Reaktion der politischen Entscheidungsträger, um demokratische Prozesse und Gesellschaften vor Angriffen mit gefährlichen gezielten, vorsätzlichen Unwahrheiten zu schützen, die das Potenzial haben, Einzelpersonen und Gesellschaften zu schaden.

Selbstregulierungsansätze gehörten zu den ersten Maßnahmen, die zur Bekämpfung von Desinformation eingeführt wurden. Im September 2018 veröffentlichte die Europäische Kommission ihren Verhaltenskodex für Desinformation, der einen Opt-in-, also selbstregulierenden Rahmen zur Bekämpfung der Verbreitung falscher oder irreführender Inhalte im Internet bietet. Der Kodex fordert mehr Anzeigentransparenz, aktive Maßnahmen gegen Fake-Accounts, eine verbesserte Sichtbarkeit vertrauenswürdiger Quellen und die Zusammenarbeit mit der Wissenschaft, um einen besseren Datenzugang zu ermöglichen. Dieser Kodex wurde inzwischen von Facebook, Google, Twitter, Mozilla, Microsoft, TikTok und Vertretern der Werbebranche unterzeichnet. Die Umsetzung des Kodex wird regelmäßig sowohl von den Unterzeichnern als auch von der Kommission überwacht, obwohl der Fortschritt bisher laut Kommission unbeständig war. Dazu hat das Institute for Strategic Dialogue (ISD) eine eigenen Bewertung abgegeben. Es gibt zahlreiche Probleme mit der Einhaltung und Wirksamkeit der Maßnahmen. Trotz verbesserter öffentlicher Kontrolle hat der Kodex die Grenzen der Selbstregulierung deutlich gemacht. Insbesondere die anhaltende Trennung von Plattformen, Forschern und der Zivilgesellschaft hat sich zur Bekämpfung von Desinformation als äußerst hinderlich erwiesen.

Neben Ansätzen zur Selbstregulierung sind **digitale Kompetenzen und Medienkompetenz** die vielleicht häufigsten nicht-regulatorischen Maßnahmen zur Stärkung der Resilienz einer breiten Öffentlichkeit gegenüber Desinformation. Die meisten Länder haben diesbezügliche Bemühungen eingeleitet oder vorgeschlagen, auch wenn das Angebot innerhalb und über die Grenzen hinweg noch sehr unterschiedlich ist. Auf internationaler Ebene hat die neu gegründete Europäische Beobachtungsstelle für digitale Medien die Aufgabe, Ansätze zur Bekämpfung von Desinformation sowohl aus der Forschungs- als auch aus der Bildungsperspektive zu bündeln. Dazu gehören die Koordinierung akademischer Bemühungen, die Verbesserung von Werkzeugen für die Datenanalyse, die Schulung von Forschern und Journalisten, die Ausweitung von Faktencheck-Initiativen und der Aufbau einer Datenbank mit Materialien und bewährten Verfahren für Pädagogen.

Über freiwillige Initiativen hinaus haben verschiedene Länder ihre Wahlgesetze geändert, um Online-Desinformation durch **inhaltsbasierte Ansätze** zu bekämpfen. **Kanada** verabschiedete im Dezember 2018 den Election Modernization Act, der vorsieht, dass Unternehmen, die absichtlich Werbeflächen an „Ausländer, die Wähler unangemessen beeinflussen“ verkaufen, mit Geld- oder sogar Haftstrafen belegt werden können. Das Gesetz verpflichtet Social-Media-Unternehmen außerdem, ein digitales Register aller Anzeigen zu führen, die in Zusammenhang mit Bundeswahlen stehen. Auch müssen die Namen der Personen genannt werden, die die Anzeigen seitens der Werbetreiber autorisiert haben. Darüber hinaus ist es eine Straftat, „falsche Aussagen über einen Kandidaten zu machen, um das Ergebnis einer Wahl zu beeinflussen“.

Australien hat im März 2018 ähnliche Transparenzregeln verabschiedet. Sie verlangen detaillierte Genehmigungserklärungen für bezahlte Werbung, die während eines Wahlkampfs in sozialen Medien verbreitet wird. Der im Dezember 2018 eingeführte Foreign Influence Transparency Scheme Act weitet die Offenlegungspflichten weiter aus und bezieht auch Personen oder Unternehmen ein, die im Auftrag eines ausländischen Auftraggebers Informationen erstellen und öffentlich verbreiten. Der überparteiliche Honest Ads Act, der im Oktober 2017 in den **US-Senat** eingebracht wurde, spiegelt im Großen und Ganzen die in Kanada und Australien umgesetzten Prinzipien und Maßnahmen wider, wurde aber noch nicht verabschiedet.

Im Dezember 2018 führte **Frankreichs Gesetz** gegen die „Manipulation von Informationen" die weitreichendsten Transparenzanforderungen für Online-Plattformen ein. Es verlangt von Social-Media-Unternehmen, Details über die Geldgeber und die für politische Werbung ausgegebenen Gelder bereitzustellen und Statistiken darüber zu veröffentlichen, wie Plattform-Algorithmen Inhalte im Zusammenhang mit „einer Debatte von nationalem Interesse" fördern. Dies schließt die Rolle personenbezogener Daten bei der gezielten Verbreitung von Inhalten ein. Darüber hinaus müssen die Unternehmen den Nutzern die Möglichkeit geben, Inhalte zu melden, die sie für irreführend oder gefälscht halten. Diese Anforderung baut auf dem französischen Wahlgesetzbuch auf, das die Verbreitung von Falschnachrichten explizit verbietet und mit Geld- und in einigen Fällen auch mit Haftstrafen belegt. Das Gesetz von 2018 ermächtigt Richter außerdem, jede „verhältnismäßige und notwendige Maßnahme" anzuordnen, um die Verbreitung von gefälschten oder irreführenden Informationen im Internet zu unterbinden. Klagen können von „jeder Person, die daran interessiert ist, zu handeln" (*personne ayant intérêt à agir*) eingereicht werden und Entscheidungen müssen innerhalb von 48 Stunden getroffen werden.

Hin zu systemischen Ansätzen[2]

Viele der „Online Harms" und Online-Aktivitäten, die von extremistischen Gruppen oder ausländischen Staaten koordiniert werden und gegen die die oben beschriebenen Regelungen vorgehen wollen, überschreiten bestehende rechtliche Schwellenwerte. Einige werden neu reguliert, etwa die Verbreitung von Desinformationen, wie sie während Wahlperioden in Frankreich identifiziert wurden. Andere sorgen weiterhin für heftige gesetzgeberische Diskussionen, wie das in Irland in geltende Verbot von Bots zur Förderung politischer Kampagnen. Aber die meisten liegen immer noch in einer Grauzone der Akzeptanz, im Spannungsfeld zwischen den Nutzungsbedingungen von Technologieunternehmen und nationalen Gesetzen, die bisher noch nicht mit der sich entwickelnden Bedrohung Schritt halten. Darüber hinaus macht das grenzübergreifende Internet die Zuordnung von Verantwortung nicht einfach. Die Grenze zwischen staatlichen Akteuren und nicht-staatlichen Netzwerken ist immer schwieriger zu ziehen. Ebenso schwierig ist es, die Urheber illegaler Hassrede, extremistischer Inhalte oder Desinformationen zu ermitteln.

Diese Probleme haben einige Gesetzgeber veranlasst, über die inhaltsbasierte Regulierung hinauszugehen und eine schadensübergreifende Perspektive für die Online-Regulierung einzunehmen. Das Ziel dieser eher „systemischen" Ansätze ist es, einen Aufsichtsrahmen zu entwickeln, mit dem eine Vielzahl von „Online Harms" bekämpft werden kann, die von Hass, Extremismus und Terrorismus bis hin zu spezifischen Gefahren für Kinder und Jugendliche, Cyber-Mobbing und Desinformation reichen.

Großbritannien hat im „Online Harms White Paper" (veröffentlicht im April 2019) verschiedene gesetzgeberische Maßnahmen vorgestellt, die zur Regulierung einer breiten Palette von Themen eingesetzt werden könnten, mit dem Ziel, Großbritannien zum „sichersten Ort der Welt im Internet" zu machen. Das Herzstück des vorgeschlagenen Rahmens ist eine gesetzliche Sorgfaltspflicht, die von einer zentralen Medien-Regulierungsbehörde (Ofcom) überwacht wird. Dies würde Online-Service-Provider gesetzlich verpflichten, „ihre Nutzer vor Schaden zu bewahren". Im starken Gegensatz zu Regulierungsansätzen, die versuchen, bestehende Gesetze im Internet durchzusetzen, etwa das deutsche NetzDG, würde die britische Regulierung auch Inhalte umfassen, die zwar nicht illegal sind, aber dennoch als schädlich angesehen werden (sogenannte „legale Schäden"). Ähnlich wie der Vorschlag der Europäischen Kommission zu terroristischem Material würde dieser Ansatz Unternehmen zwingen, proaktive Maßnahmen zu ergreifen (und nachzuweisen), die verhindern, dass schädliche Inhalte erscheinen oder bekannt werden, anstatt lediglich auf Anfragen Dritter zu reagieren, die bestimmte Inhalte melden und entfernen wollen.

In der **Europäischen Union betont** der Ende 2020 veröffentlichte Europäische Aktionsplan für Demokratie die Notwendigkeit einer Kombination aus regulatorischen und nicht-regulatorischen Initiativen zum Schutz von Wahlen, zur Sicherung des Medienpluralismus und zur Bekämpfung von Desinformation auf europäischer Ebene. All das erfordert eine verstärkte Zusammenarbeit und in einigen Fällen eine Ko-Regulierung zwischen dem öffentlichen und dem privaten Sektor. Außerdem wird seitens der EU die E-Commerce-Richtlinie in Form des Digital Services Act (DSA) überarbeitet. In ihrer ersten Folgenabschätzung hat die Europäische Kommission eine breite Palette von Problemen identifiziert, die bei der Regulierung digitaler Dienste in der EU angegangen werden müssen, darunter gesellschaftliche Schäden, illegale Aktivitäten und unzureichender Schutz der Grundrechte. Als Hauptmängel des derzeitigen Systems wurden vor allem eine ineffektive Aufsicht und eine unzureichende Verwaltungszusammenarbeit identifiziert, was auf Pläne hindeutet, über eine Notice-and-Takedown-Regelung hinaus eine systematischere Aufsicht über Social-Media-Plattformen durch europäische Regulierungsbehörden einzuführen.

[2]Für Details bitte das separate Briefing-Papier zum EU Digital Services Act und der UK Online Safety Bill konsultieren.

Desweiteren gibt es einen Vorschlag der deutschen Bundesregierung zur Überarbeitung des inhaltsorientierten **deutschen** NetzDG, in dem Aufsichtsmandat für das Bundesamt für Justiz vorgesehen ist. Alle diese Entwicklungen deuten darauf hin, dass mit dem 20. Jahrestag der E-Commerce-Richtlinie die neue Generation von regulatorischen und nicht-regulatorischen Initiativen zur Bekämpfung von Hass, Extremismus und Terrorismus zunehmend die internen Prozesse von Social-Media-Unternehmen in den Mittelpunkt stellt. Neben anhaltenden Debatten darüber, was schädliche Online-Inhalte sind, wird der Schwerpunkt darauf liegen, sicherzustellen, dass den Regulierungsbehörden die richtigen Instrumente zur Verfügung stehen, um ihre Aufsichtsfunktion zu erfüllen und gleichzeitig ihre operative und funktionale Unabhängigkeit zu wahren. Dies ist eine markante Entwicklung der digitalen Regulierung in einem Bereich, der bisher von einem (oft freiwilligen) Notice-and-Takedown-Modell dominiert wurde.

ISD | Powering solutions to extremism and polarisation

Digital Policy Lab
Zusammenfassung

EU Digital Services Act & UK Online Safety Bill

Über diese Papier

Dieses Papier bietet eine Zusammenfassung des Digital Services Act (DSA) und der Online Safety Bill (OSB). Diese digitalpolitischen Gesetzesinitiativen der Europäischen Union (DSA) und Großbritanniens (OSB) wurden beide im Dezember 2020 veröffentlicht. Dieses Papier bildete die Diskussionsgrundlage für die Sitzung des Digital Policy Lab vom 27. Januar 2021. Sie konzentriert sich auf Kernelemente der beiden umfangreichen Vorlagen in Bezug auf ihre Relevanz für die Bekämpfung von Terrorismus, Extremismus, Hassrede und Desinformation im Internet. Die in diesem Dokument geäußerten Ansichten sind die der Autoren und spiegeln nicht unbedingt die Ansichten der Teilnehmer oder vertretenden Regierungen wider.

Zur besseren Lesbarkeit wurde auf die gleichzeitige Verwendung männlicher und weiblicher Sprachformen verzichtet. Es wurde das generische Maskulinum verwendet, wobei alle Geschlechter gleichermaßen gemeint sind.

ISD | Powering solutions to extremism and polarisation

Beirut | Berlin | London | Paris | Washington DC

www.isdglobal.org

Der Digital Services Act (DSA) der Europäischen Union

Am 15. Dezember 2020 hat die Europäische Kommission ein neues zweigliedriges Regelwerk für alle digitalen Dienste, einschließlich sozialer Medien, Online-Marktplätze und anderer Online-Plattformen, die in der Europäischen Union tätig sind, vorgestellt: den Digital Services Act und den Digital Markets Act.

Das **Gesetz über digitale Dienste (DSA)**[1] regelt die Pflichten digitaler Dienste, die als Vermittler zwischen Verbrauchern und Waren, Dienstleistungen und Inhalten fungieren. Sie sollen dafür sorgen, dass „die Menschen weniger mit illegalen Aktivitäten und gefährlichen Gütern zu tun bekommen und sie werden den Schutz der Grundrechte gewährleisten, sodass ein sichereres Online-Umfeld entsteht, damit die Bürgerinnen und Bürger frei ihre Ideen äußern, miteinander kommunizieren und online einkaufen können“.[2] Des Weiteren aktualisiert der DSA die E-Commerce-Richtlinie, die seit 2000 in Kraft ist.

Laut der Europäischen Kommission verfolgt der DSA „ein ausgewogenes Gleichgewicht in Bezug auf die Haftung von Vermittlern und sieht wirksame Maßnahmen zur Bekämpfung illegaler Inhalte und zur Minderung gesellschaftlicher Risiken im Internet vor“. Die Kommission erklärt, dass diese neuen Regeln „ein wichtiger Schritt zur Verteidigung europäischer Werte im Online-Raum“ sind mit dem Ziel neue Maßstäbe zu setzen, „an denen sich die Regulierungsansätze für Online-Vermittler auch auf weltweiter Ebene messen lassen“.[3]

Der DSA wurde zusammen mit dem **Digital Markets Act (DMA)** vorgestellt. Dieses Gesetz legt neue Regeln für sogenannte „Gatekeeper“-Unternehmen fest. Das sind Plattformen, die als wichtiges Tor dienen, über das andere Unternehmen ihre Kunden erreichen. Dazu zählen z.B. Google und Facebook. Diese Unternehmen kontrollieren mindestens einen sogenannten „zentralen Plattformdienst“ (wie Suchmaschinen, soziale Netzwerkdienste, bestimmte Messenger-Dienste, Betriebssysteme und Online-Vermittlungsdienste) und haben eine dauerhafte, große Nutzerbasis in mehreren Ländern der EU. Unter dem Digital Markets Act müssen Unternehmen, die als Gatekeeper identifiziert wurden, proaktiv bestimmte Verhaltensweisen umsetzen und unfaire Verhaltensweisen unterlassen. Das DMA zielt darauf ab, diese Gatekeeper daran zu hindern, Unternehmen und Verbrauchern unfaire Bedingungen aufzuerlegen, und sie zu verpflichten, die Transparenz wichtiger digitaler Dienste zu gewährleisten. Es gibt zahlreiche Beispiele für die unlauteren Bedingungen, die Gatekeeper anderen Stakeholdern auferlegen. Dazu gehört der restriktive oder unmögliche Zugang zu den eigenen Unternehmensdaten oder Situationen, in denen Nutzer an einen bestimmten Dienst gebunden sind und nur begrenzte Möglichkeiten haben, zu alternativen Diensten zu wechseln.

Die Vorschläge der Europäischen Kommission bilden den Ausgangspunkt für das Europäische Parlament und die Mitgliedstaaten, um Rechtsvorschriften auf Ebene der Europäischen Union zu verabschieden. Als Mitgesetzgeber werden sie die Vorschläge zunächst entlang ihrer Präferenzen abändern, bevor sie sich auf einen Kompromisstext einigen. Dieses Verfahren dauert bis zu drei Jahre und wird voraussichtlich frühestens 2023 abgeschlossen sein. In den folgenden Abschnitten werden die wichtigsten Inhalte des Gesetzes über digitale Dienste vorgestellt.

Neue horizontale Regeln für alle Kategorien von Inhalten vs. branchenspezifische Regeln

Der DSA-Vorschlag aktualisiert die horizontalen Vorschriften, die die Zuständigkeiten und Pflichten von Anbietern digitaler Dienste (insbesondere von Online-Plattformen) in der Europäischen Union festlegen. Diese Regeln gelten in der EU ausnahmslos auch für die Online-Vermittler, die außerhalb der Europäischen Union niedergelassen sind, ihre Dienste aber in der EU anbieten. Damit führt er einen horizontalen Rahmen für alle Kategorien von Inhalten, Produkten, Dienstleistungen und Aktivitäten auf Vermittlungsdiensten ein.[4] Im Rahmen dieses Briefings ist es sinnvoll zu klären, wie der Vorschlag mit drei oft diskutierten Kategorien von Inhalten umgeht. Dabei geht es um schädliche Inhalte, illegale Inhalte und offensichtlich illegale Inhalte (siehe unten).

Der Vorschlag erkennt an, dass die zunehmende Nutzung bestimmter Dienste „bei der Vermittlung und Verbreitung rechtswidriger oder anderweitig schädlicher Informationen und Tätigkeiten eine immer wichtigere Rolle spielen" (§ 5). Er definiert **schädliche Informationen und Aktivitäten** jedoch nicht und argumentiert, es gebe unter den Interessenträgern ein:

> *„allgemeines Einvernehmen darüber, dass „schädliche" (aber nicht oder zumindest nicht unbedingt illegale) Inhalte im Gesetz über digitale Dienste nicht definiert werden sollten und dass sie keiner Pflicht zur Entfernung unterliegen sollten, da es sich hierbei um einen heiklen Bereich handele, der schwerwiegende Auswirkungen auf den Schutz der Meinungsfreiheit habe". (S.9)*

Stattdessen schafft die Kommission Sorgfaltspflichten für Plattformen, etwa in Bezug auf „die Tätigkeiten der Anbieter von Vermittlungsdiensten, mit denen illegale Inhalte oder Informationen, die von Nutzer bereitgestellt werden und mit den allgemeinen Geschäftsbedingungen des Anbieters unvereinbar sind, erkannt, festgestellt und bekämpft werden sollen" (Artikel 2.p). Was beispielsweise in der britischen Online Safety Bill als schädliche Inhalte angesehen wird, wird hauptsächlich von dem Verweis auf „Informationen, die mit den Geschäftsbedingungen einer Plattform unvereinbar sind" abgedeckt. Die Kommission stellt in ihren Begleitmaterialien ausdrücklich fest, dass „[s]chädliche Inhalte [...], soweit sie nicht illegal sind, nicht genauso behandelt werden [sollten] wie illegale Inhalte".[5] Regeln zur Entfernung oder Förderung der Entfernung von Inhalten (siehe unten) sollen nur für illegale Inhalte gelten, unter voller Wahrung der Meinungsfreiheit.

Illegale Inhalte sind im weitesten Sinne definiert als alle „Informationen im Zusammenhang mit illegale Inhalten, Produkten, Dienstleistungen oder Tätigkeiten", wie sie durch Unionsrecht oder ein nationales Gesetz eines Mitgliedstaats definiert sind (Artikel 2g, §12). Daher verändert der DSA nicht die nationalen oder EU-Gesetze, die festlegen, was illegale Inhalte sind. Der DSA deckt Informationen ab, die an sich illegal sind, wie Hassrede, terroristische Inhalte, diskriminierende Inhalte, aber auch Informationen, die sich auf Aktivitäten beziehen, die illegal sind, etwa die Weitergabe von Darstellungen sexuellen Missbrauchs von Kindern (Child Sexual Abuse Material/CSAM) der rechtswidrigen Weitergabe privater Bilder ohne Zustimmung (z.B. sogenannte „Rachepornos") oder Online-Stalking bzw. Online-Belästigung.

Offensichtlich rechtswidrige Inhalte sind Inhalte, bei denen es für einen Laien ohne inhaltliche Analyse offensichtlich ist, dass der Inhalt rechtswidrig ist (§ 47). Diese Definition ist nur in Zusammenhang mit Artikel 20 relevant (siehe unten, Abschnitt 4.3), der besagt, dass Online-Plattformen Nutzer sperren können, die solche Inhalte häufig bereitstellen.

Die vorgeschlagene Verordnung ergänzt bestehende, sektorspezifische Rechtsvorschriften wie die Richtlinie über audiovisuelle Mediendienste (AVMD), die Urheberrechtsrichtlinie, den gemeinschaftlichen Besitzstand im Bereich des Verbraucherschutzes oder die künftige Richtlinie über terroristische Inhalte.[6] Sie gelten als lex specialis (§ 9), was bedeutet, dass die früheren sektorspezifischen Verpflichtungen dem DSA übergeordnet sind, sofern dieser das gleiche Thema weniger detailliert regelt. Ein Beispiel: Spezifische Verpflichtungen, die Video-Sharing-Plattformen wie YouTube von der AVMD auferlegt wurden, um Hassrede zu bekämpfen, gelten auch weiterhin. Der DSA würde für diese Video-Sharing-Anbieter nur dann gelten, sofern die AVMD-Richtlinie diese Themen nicht – oder nicht vollständig – regelt (DSA, S.4).

Aktualisierter Rahmen für die bedingte Haftungsbefreiung von Anbietern von Vermittlungsdienstleistungen

Kapitel II des DSA enthält Bestimmungen, die Anbieter von Vermittlungsdiensten von der Haftung befreit. Genauer gesagt enthält es die Bedingungen, unter denen Anbieter von reinen Durchleitung- (Artikel 3), Caching- (Artikel 4) und Hosting-Diensten (Artikel 5) von der Haftung für die von ihnen übermittelten und gespeicherten Informationen Dritter befreit sind. Die Definitionen dieser Kategorien von Vermittlungsdiensten sind gleichgeblieben (siehe Artikel 2f). Die Artikel 3 und 4 sind Kopien der Artikel 12 und 13 der E-Commerce-Richtlinie. Diese Regeln gelten für jeden Anbieter eines Vermittlungsdienstes „ungeachtet des Orts der Niederlassung des Anbieters dieser Dienste" (Artikel 1.3), sofern er Dienstleistungen in der EU erbringt, die durch eine „wesentliche Verbindung" zur EU nachgewiesen werden. Eine wesentliche Verbindung könnte aus spezifischen faktischen Kriterien abgeleitet werden, etwa der Anzahl der Nutzer in der EU oder der Ausrichtung der Aktivitäten auf einen oder mehrere Mitgliedstaaten (Artikel 2d, §7-8).

In diesem Vorschlag werden nun zum ersten Mal **„Online-Plattformen" als Unterkategorie von Hosting-Diensten** eingeführt (siehe Artikel 2f, § 13). Diese werden als Anbieter von Hosting-Diensten angesehen, die nicht nur Informationen speichern, die von den Empfängern des Dienstes auf deren Wunsch bereitgestellt werden, sondern diese Informationen auch in der Öffentlichkeit verbreiten, wiederum auf deren Wunsch (Artikel 2h). Diese neue Kategorie ist aus haftungsrechtlicher Sicht zwar nicht von Bedeutung, wohl aber wichtig für die Bestimmung der Kategorien von Anbietern, für die die neuen Sorgfaltspflichten gelten (siehe unten, Abschnitt 4). Zwischenmenschliche Kommunikationsdienste wie Telegram oder WhatsApp fallen nicht unter diese Definition, ebenso wenig wie „Vermittlungsdienste" wie „IT-Dienstleistungen auf Distanz oder Transport-, Beherbergungs- oder Lieferdienste" (§6).

Der Vorschlag **behält die Haftungsregeln für Anbieter von Vermittlungsdiensten bei, wie sie in den letzten zwei Jahrzehnten vom europäischem Gerichtshof ausgelegt wurden** (§16). Die Faustregel lautet nach wie vor: Wenn ein Hosting-Dienst tatsächliche Kenntnis von illegalen Inhalten erlangt, muss er unverzüglich handeln, um diese Inhalte zu entfernen oder den Zugang zu ihnen zu sperren (Artikel 5.1, §22).[7] Die Ausnahmen von der Haftung gelten nicht für Anbieter, die „dahingehend eine aktive Rolle [einnehmen], dass [sie] Wissen oder Kontrolle über" die von einem Nutzer bereitgestellten Informationen erlangen (§18).

Ein neues Element des DSA ist die Einführung einer sogenannten „**Guter Samariter"-Klausel**. Der neue Artikel 6 soll Anreize für freiwillige, proaktive Untersuchungen seitens der Anbieter um nationalen Rechtsbestimmungen nachzukommen, schaffen. Er stellt klar, dass Untersuchungen, die darauf abzielen, illegale Inhalte aufzuspüren und zu entfernen, eine Plattform nicht untauglich für eine Haftungsbefreiung machen (Artikel 6, § 25).

Artikel 7 enthält ein **Verbot allgemeiner Überwachungs-** oder aktiver Ermittlungspflichten für diese Anbieter (Artikel 7) – ähnlich wie Artikel 15 der E-Commerce-Richtlinie. Schließlich erlegt dieser Abschnitt des DSA den Anbietern von Vermittlungsdiensten die Verpflichtung auf, die Anordnungen nationaler Justiz- oder Verwaltungsbehörden im Hinblick auf die Ahndung illegaler Inhalte (Artikel 8) der Bereitstellung entsprechender Informationen (Artikel 9) angemessen umzusetzen.

Abgestufte Struktur der Sorgfaltspflichten für verschiedene Arten von Vermittlungsdienstleistungen

Abgesehen davon, dass der DSA einen Rahmen für die bedingte Freistellung der Erbringung von Vermittlungsdienstleistungen im Binnenmarkt schafft, führt der Entwurf auch neue Sorgfaltspflichten ein, die an die Art der betreffenden Vermittlungsdienstleistung angepasst sind. Es gibt eine Reihe von grundlegenden Sorgfaltspflichten, die für alle Anbieter von Vermittlungsdiensten gelten. Sie werden dann durch zusätzliche Pflichten für Anbieter von Hosting-Diensten und Online-Plattformen ergänzt.

Für sehr große Online-Plattformen (VLOPs) legt der Vorschlag asymmetrische Sorgfaltspflichten in Abhängigkeit von der Art ihrer Dienstleistungen

und ihrer Größe fest. Der Vorschlag sieht einen „beaufsichtigten Risikomanagementansatz“ (DSA, S.11) vor, der bestimmte materielle Verpflichtungen auf VLOPs beschränkt, die aufgrund ihrer Reichweite eine zentrale, systemische Rolle für die Erleichterung der öffentlichen Debatte und wirtschaftliche Transaktionen erlangt haben (§53). Mit diesem Ansatz werden bestimmte identifizierte Probleme nur dort angegangen, wo sie auftreten, während Anbieter, die von diesen Problemen nicht betroffen sind, nicht übermäßig durch zusätzliche Bürokratie belastet werden. Daraus ergeben sich in Bezug auf vier verschiedene Kategorien von Vermittlungsdiensten folgende Verpflichtungen:

<table>
<tr><th>Vermittlungsdienste</th><th>Hosting-Dienste</th><th>Online-Plattformen</th><th>Sehr große Plattformen (VLOPs)</th></tr>
<tr><td colspan="4">Berichterstattung zu Transparenz
Berücksichtigung der Grundrechte in den Nutzungsbedingungen
Zusammenarbeit mit nationalen Behörden bei Anordnungen
Kontaktstellen und gegebenenfalls gesetzliche Vertretung</td></tr>
<tr><td rowspan="4"></td><td colspan="3">Meldung und Abhilfe sowie Pflicht zur Unterrichtung der Nutzer</td></tr>
<tr><td colspan="2"></td><td>Beschwerde- und Rechtsbehelfsmechanismus sowie außergerichtliche Streitbeilegung
Vertrauenswürdige Hinweisgeber („trusted flaggers“)
Maßnahmen gegen missbräuchliche Meldungen sowie Gegendarstellungen
Sicherheitsüberprüfung von Drittanbietern („KYBC“ – „know your business customer“)
Transparenz von Online-Werbung gegenüber Nutzer
Meldung von Straftaten</td></tr>
<tr><td colspan="3"></td><td>Risikomanagement-Pflichten und Compliance-Beauftragte
Externe Risikoprüfung und öffentliche Rechenschaftspflicht
Transparenz der Empfehlungssysteme und Wahlmöglichkeiten für Nutzer beim Zugriff auf Informationen
Datenaustausch mit Behörden und Forschern
Verhaltenskodizes
Zusammenarbeit im Krisenfall</td></tr>
</table>

Verpflichtungen für alle Anbieter von Vermittlungsdiensten

Die Verpflichtungen für alle Anbieter von Vermittlungsdiensten sind im ersten Abschnitt von Kapitel 3 des DSA dargelegt und umfassen die folgenden vier Hauptpflichten:

- Die **Verpflichtung, eine operative zentrale Anlaufstelle einzurichten**, um die direkte Kommunikation mit den Behörden der Mitgliedstaaten, der Europäischen Kommission und dem Europäischen Gremium für digitale Dienste („das Gremium", siehe unten) zu erleichtern (Artikel 10, § 36).

- Die Verpflichtung für Dienstleistungserbringer einen **gesetzlichen Vertreter** in der Union zu benennen, sofern sie nicht in einem Mitgliedstaat niedergelassen sind, aber ihre Dienstleistungen in der Union anbieten (Artikel 11, § 37). Der benannte gesetzliche Vertreter kann für die Nichteinhaltung der Verpflichtungen aus dieser Verordnung haftbar gemacht werden.

- Eine Transparenzverpflichtung: (1) **in ihren Allgemeinen Geschäftsbedingungen alle Beschränkungen darzulegen, die sie für die Nutzung ihrer Dienste auferlegen können**. Diese Informationen umfassen Angaben zu allen Strategien, Verfahren, Maßnahmen und Werkzeugen, die zum Zweck der Inhaltsmoderation eingesetzt werden, einschließlich algorithmischer Entscheidungsfindung und menschlicher Überprüfung (Artikel 12.1, § 38), und (2) die Verpflichtung bei der Anwendung und Durchsetzung dieser Beschränkungen verantwortungsvoll zu handeln (Artikel 12.2).

- Alle Anbieter von Vermittlungsdiensten sollten **jährlich über die von ihnen vorgenommene Moderation von Inhalten berichten**, unabhängig davon, ob es sich um illegale Inhalte handelt oder ob sie gegen die Geschäftsbedingungen des Anbieters verstoßen (Artikel 13, § 39). „Sehr große Plattformen" (VLOPs, siehe unten) sind verpflichtet alle sechs Monate Bericht zu erstatten. Dieser soll hauptsächlich aggregierte Daten enthalten über:
  - Die Anzahl der von den Behörden der Mitgliedsstaaten erhaltenen Take-Down-Anordnungen (13.1a);
  - Die Anzahl der eingereichten Meldungen durch Nutzer (13.1b);
  - Die Anzahl der eingegangenen Beschwerden über die Entscheidungen des Anbieters zur Inhaltsmoderation (Artikel 13.1.d);
  - Informationen über „die auf Eigeninitiative des Anbieters durchgeführte Moderation von Inhalten, einschließlich der Anzahl und Art der ergriffenen Maßnahmen, die sich auf die Verfügbarkeit, Sichtbarkeit und Zugänglichkeit der von den Nutzer bereitgestellten Informationen auswirken, und der Möglichkeiten der Nutzer, solche Informationen bereitzustellen, aufgeschlüsselt nach der Art des Grundes und der Grundlage für das Ergreifen dieser Maßnahmen" (Artikel 13.1.c).

Für Plattformen und sehr große Plattformen gelten zusätzliche Transparenzanforderungen (siehe unten, Artikel 23 & 33).

Verpflichtungen für Hosting-Dienste

In Kapitel 3, Abschnitt 2 werden Verpflichtungen festgelegt, die zusätzlich zu den Verpflichtungen nach Abschnitt 1 für alle Anbieter von Hosting-Diensten gelten. Dies würde auch Dateispeicher- und -Freigabedienste, Webhosting-Dienste, Werbeserver und „Paste Bins" einschließen, soweit sie als Anbieter von Hosting-Diensten im Sinne dieser Verordnung in Frage kommen. Insbesondere verpflichtet dieser Abschnitt die Anbieter dieser Dienste, benutzerfreundliche Melde- und Aktionsmechanismen einzurichten, die es Dritten ermöglichen, das Vorhandensein mutmaßlich illegaler Inhalte zu melden (Artikel 14, § 40, 41). Artikel 14.2 definiert welche Informationen eine Meldung enthalten muss, um zu gewährleisten, dass die Meldung zu einer tatsächlichen Kenntnisnahme seitens des Anbieters führt.

Wenn ein solcher Anbieter beschließt, den Zugang zu bestimmten Informationen, die von einem Empfänger des Dienstes zur Verfügung gestellt wurden, zu entfernen oder zu sperren, ist er außerdem verpflichtet, diesem Empfänger eine Begründung zu geben und ihn über die verfügbaren Rechtsbehelfsmöglichkeiten zu informieren (Artikel 15, § 42).

Verpflichtungen für alle Online-Plattformen

In Kapitel 3, Abschnitt 3 werden Verpflichtungen festgelegt, die für alle Online-Plattformen gelten, zusätzlich zu den Verpflichtungen aus den Abschnitten 1 und 2. Dieser Abschnitt gilt nicht für Online-Plattformen, die Kleinst- oder Kleinunternehmen sind, „es sei denn ihre Reichweite und Wirkung sind so erheblich, dass sie die Kriterien für eine Einstufung als sehr große Online-Plattformen im Sinne dieser Verordnung erfüllen“ (Artikel 16, § 43). Anbei werden die wesentlichen Punkte aufgelistet:

- Online-Plattformen werden zur **Bereitstellung eines internen Beschwerdesystems verpflichtet, damit Dritte** gegen Entscheidungen zur Entfernung oder Sperrung des Zugangs vorgehen können, zu denen es wegen mutmaßlich illegaler Inhalte oder Verstöße gegen Geschäftsbedingungen gekommen ist (Artikel 17). Dies soll es den Nutzer ermöglichen, bestimmte Entscheidungen zur Inhaltsmoderation einfach und effektiv anzufechten (§ 44).

- Online-Plattformen werden verpflichtet, sich zur Beilegung von Streitigkeiten mit Nutzer ihrer Dienste an zertifizierte **außergerichtliche Streitbeilegungsstellen zu wenden** (Artikel 18, §44/45).

- Online-Plattformen werden verpflichtet, sicherzustellen, dass Meldungen vorrangig behandelt werden, die von Einrichtungen (nicht von Einzelpersonen) eingereicht werden, die vom nationalen Koordinator für digitale Dienste den Status eines **vertrauenswürdigen Hinweisgebers (sog. trusted flaggers) haben** (Artikel 19, §46).

- Festlegung von **Maßnahmen, die Online-Plattformen gegen Missbrauch zu ergreifen haben** (Artikel 20, §47), und zwar für den Fall, dass Nutzer häufig offensichtlich illegale Inhalte bereitstellen oder auch vermehrt offensichtlich unbegründete Hinweise oder Beschwerden einreichen. Unter bestimmten Voraussetzungen sollen Online-Plattformen „ihre einschlägigen Dienste für die an missbräuchlichem Verhalten beteiligte Person vorübergehend aussetzen“ (§47).

- Schaffung einer **Verpflichtung, die zuständigen Strafverfolgungsbehörden zu informieren**, wenn der begründete Verdacht besteht, dass ein Nutzer „eine schwere Straftat begangen hat, begeht oder vermutlich begehen wird, die das Leben oder die Sicherheit von Personen in Gefahr bringt“ (Artikel 21, §48).

- Schaffung einer **Verpflichtung, die Zuverlässigkeit der Gewerbetreibenden, die ihre Dienste nutzen, zu bewerten und spezifische Informationen über sie zu veröffentlichen**, wenn diese Online-Plattformen es Verbrauchern ermöglichen, Fernabsatzverträge mit diesen Gewerbetreibenden abzuschließen (Artikel 22, § 49).

- Schaffung einer **Verpflichtung**, ihre Schnittstelle so zu gestalten, dass die Händler das Verbraucher- und Produktsicherheitsrecht der Union einhalten können (Artikel 22, § 50).

- **Zusätzlich zu ihren Transparenzverpflichtungen in Artikel 13** sind Online-Plattformen auch verpflichtet, in ihren Jahresberichten Daten zu veröffentlichen über:

  – die Anzahl der Streitfälle, die den außergerichtlichen Streitbeilegungsgremien vorgelegt wurden;

  – die Anzahl der gemäß Artikel 20 verhängten Aussetzungen der Dienste;

  – jeden Einsatz automatischer Mittel zum Zweck der Inhaltsmoderation, einschließlich Angabe der genauen Zwecke, der Indikatoren für die Genauigkeit der automatischen Mittel bei der Erfüllung dieser Zwecke und aller angewandten Schutzmaßnahmen (Artikel 23, § 51).

- **Transparenzpflichten für Online-Plattformen, die Werbung anzeigen** (Artikel 24, §52). Die Präambel hebt hervor, wie Online-Anzeigen zu erheblichen Risiken beitragen können, da sie (1) illegale Inhalte enthalten können, (2) finanzielle Anreize für die Veröffentlichung oder Verstärkung von illegalen oder anderweitig schädlichen Inhalten und Aktivitäten im Internet bieten oder (3) diskriminierende Anzeigen mit negativen Auswirkungen auf die Gleichbehandlung und Chancengleichheit der Bürger anzeigen können. Infolgedessen erlegt der DSA diesen Plattformen zusätzliche Transparenzmaßnahmen auf, die es Nutzer ermöglichen sollen, für jede Anzeige in Echtzeit

  – zu erkennen, dass es sich bei den angezeigten Informationen um eine Anzeige handelt.

  – die natürliche oder juristische Person zu identifizieren, in deren Auftrag die Anzeige geschaltet wurde.

  – aussagekräftige Informationen über die wichtigsten Parameter zu bekommen, die verwendet werden, um die Empfänger zu bestimmen, denen die Werbung angezeigt wird, was auch „aussagekräftige Erläuterungen zur zugrunde liegenden Logik [...], einschließlich der Angabe, wann Profiling genutzt wird“ beinhaltet (§52);

  – Der Verhaltenskodex für Online-Werbung in Artikel 36 soll weiter ausgearbeitet werden, um aufzuzeigen wie die Bereitstellung von „aussagekräftigen Informationen“ über die „wichtigsten Parameter“ in der Praxis funktionieren kann (vgl. Artikel 36.2.b).

Risikobewertung, Risikominderung und Prüfungspflichten für VLOPs

Abschnitt 4 enthält zusätzlich zu den in den Abschnitten 1 - 3 festgelegten Verpflichtungen für sehr große Online-Plattformen (gemäß Definition in Artikel 25, §53-55) eine weitere Verpflichtung zum Umgang mit systemischen Risiken. Die **operative Schwelle für Anbieter, die in den Anwendungsbereich dieser Verpflichtungen fallen**, umfasst Online-Plattformen mit einer erheblichen Reichweite in der EU. Diese wird derzeit als mehr als 45 Millionen aktive Nutzer pro Monat festgelegt. Ändert sich die Bevölkerungszahl der EU um einen bestimmten Prozentsatz, passt die Kommission die Zahl der für den Schwellenwert berücksichtigten Empfänger so an, dass sie stets 10 Prozent der Gesamtbevölkerung entspricht.

Sobald eine Plattform diese Schwelle erreicht, können die von ihr ausgehenden systemischen Risiken angesichts ihrer Reichweite und ihrer Fähigkeit, die öffentliche Debatte zu gestalten und Informationen online zu verbreiten, unverhältnismäßig negative Auswirkungen auf unsere Gesellschaften haben. Die Kommission argumentiert, dass die Art und Weise, wie diese Plattformen ihre Dienste gestalten, „im Allgemeinen auf eine Optimierung ihres oft werbegestützten Geschäftsmodells ausgerichtet [ist] und kann Anlass zu gesellschaftlichen Bedenken geben. Besteht keine wirksame Regulierung und Durchsetzung, können die Plattformen die Spielregeln bestimmen, ohne dass dabei die mit ihnen verbundenen Risiken und der dadurch möglicherweise entstehende gesellschaftliche und wirtschaftliche Schaden wirksam ermittelt und gemindert werden kann“ (§ 56).

Daher sind VLOPs **verpflichtet**, mindestens einmal jährlich **Risikobewertungen durchzuführen**, also Bewertungen der systemischen Risiken, die sich aus den Funktionen und aus der Nutzung ihres Dienstes sowie aus potenziellem Missbrauch durch die Empfänger des Dienstes ergeben, und dann geeignete Maßnahmen zu ergreifen, um diese Risiken zu mindern (Artikel 26, § 57). In dem Vorschlag werden folgende systemische Risiken genannt:

- (26a): **die Verbreitung illegaler Inhalte über ihre Dienste**, etwa Material über sexuellen Kindesmissbrauch oder illegale Hassrede, sowie illegale Aktivitäten, wie der Verkauf gefälschter Produkte. Die Kommission hebt die Risiken von „Konten mit einer besonders großen Reichweite“ hervor, die diese Inhalte verbreiten oder sich an dieser Art von Verhalten beteiligen (§ 57).

- (26b): **etwaige nachteilige Auswirkungen des Dienstes auf die Ausübung der Grundrechte**,

auf Achtung des Privat- und Familienlebens, der Meinungs- und Informationsfreiheit, des Diskriminierungsverbots und der Rechte des Kindes, wie sie in den Artikeln 7, 11, 21 bzw. 24 der Charta der Grundrechte der Europäischen Union verankert sind. Dies ist eine wichtige Kategorie, da der Präambel deutlich macht, dass diese Risiken „beispielsweise auf die Gestaltung der Algorithmensysteme sehr großer Online-Plattformen oder auf den Missbrauch ihres Dienstes für die Übermittlung missbräuchlicher Nachrichten oder auf andere Methoden zur Verhinderung der freien Meinungsäußerung oder zur Behinderung des Wettbewerbs zurückzuführen sein“ (§ 57).

- (26c) die **vorsätzliche Manipulation ihres Dienstes**, auch durch nicht authentische Nutzung oder automatisierte Nutzung des Dienstes, mit tatsächlichen oder vorhersehbaren negativen Auswirkungen auf den Schutz der öffentlichen Gesundheit, von Minderjährigen, des zivilgesellschaftlichen Diskurses oder tatsächlichen oder vorhersehbaren Auswirkungen im Zusammenhang mit Wahlprozessen und der öffentlichen Sicherheit. Solche Risiken können zum Beispiel durch die Erstellung von Fake Accounts, die Verwendung von Bots und andere automatisierte oder teilweise automatisierte Verhaltensweisen entstehen, die zu einer schnellen und weiten Verbreitung von Informationen führen können, bei denen es sich um illegale Inhalte handelt oder die mit den Geschäftsbedingungen einer Online-Plattform unvereinbar sind.

Die Kommission stellt fest, dass VLOPs bei der Durchführung von Risikobewertungen insbesondere berücksichtigen müssen, „wie ihre Systeme zur Moderation von Inhalten, ihre Empfehlungssysteme und ihre Systeme zur Auswahl und Anzeige von Werbung die in Absatz 1 genannten systemischen Risiken beeinflussen, sowie die Möglichkeit der raschen und weiten Verbreitung von illegalen Inhalten und von Informationen, die mit ihren allgemeinen Geschäftsbedingungen unvereinbar sind“ (Artikel 26.2).

VLOPs sollten, wo angemessen, ihre Risikobewertungen und ihre Maßnahmen zur Risikominderung gemeinsam mit Vertretern der Dienstleistungsempfänger, Vertretern von Gruppen, die potenziell von ihren Dienstleistungen betroffen sind, unabhängigen Experten und zivilgesellschaftlichen Organisationen, durchführen und entwickeln. (§59).

Nachdem entsprechende Risiken identifiziert wurden, sollten VLOPs „angemessene, verhältnismäßige und wirksame“ **Maßnahmen einsetzen, um diese Risiken zu mindern** (Artikel 27.1, §58). Geeignete Maßnahmen können sein:

- (27.1a) Anpassung ihrer Geschäftsbedingungen sowie der Gestaltung und Funktionsweise ihrer Inhaltsmoderationsprozesse, algorithmischen Empfehlungssysteme, Online-Schnittstellen oder anderer Funktionsmerkmale ihrer Dienste. Dies kann beispielsweise die Verbesserung der Sichtbarkeit von maßgeblichen Informationsquellen beinhalten.

- (27.1b) Gezielte Maßnahmen zur Einschränkung von Werbung, einschließlich des Verzichts auf Werbeeinnahmen für bestimmte Inhalte.

- (27.1.d/e) Initiierung oder Anpassung der Zusammenarbeit mit vertrauenswürdigen Hinweisgebers („trusted flaggers“), auch durch Schulungen und Austausch mit vertrauenswürdigen Meldeorganisationen, und Zusammenarbeit mit anderen Dienstleistern, auch durch die Entwicklung neuer oder bestehender Verhaltenskodizes oder anderen Selbstregulierungsmaßnahmen.

Die Kommission kann allgemeine Leitlinien für die Anwendung dieser spezifischen risikomindernden Maßnahmen erlassen.

In Anbetracht der Tatsache, dass sie eine Überprüfung durch unabhängige Experten gewährleisten müssen, sollten VLOPs **durch unabhängige Audits** Rechenschaft darüber ablegen, dass sie die in dieser Verordnung festgelegten Verpflichtungen ebenso einhalten wie ggf. ergänzende Verpflichtungen, die sich aus Verhaltenskodizes und Krisenprotokollen ergeben. Um dieses Ziel zu erreichen, sollen sie sich auch externen und unabhängigen Audits unterziehen (Artikel 28, §60), die – falls der Prüfungsvermerk nicht positiv ausfällt – zu operativen Empfehlungen über spezifische Maßnahmen zur Einhaltung der Verordnung führen (Artikel 28.3 f). Die Plattformen sollten dem Auditor

Zugang zu allen relevanten Daten gewähren, die für die ordnungsgemäße Durchführung des Audits erforderlich sind. Die Prüfer sollten auch auf andere objektive Informationsquellen zurückgreifen können, etwa auf Studien von zugelassenen Wissenschaftlern. Der Auditbericht wird an den Koordinator für digitale Dienste der Einrichtung und an das EU Gremium geschickt (siehe unten). Sie prüfen, ob die vorgeschlagenen Empfehlungen ordnungsgemäß umgesetzt wurden. Ist dies nicht der Fall, kann die Kommission weitere Untersuchungen durchführen (Artikel 51 – siehe unten) und schließlich eine Geldstrafe oder andere vorläufige Maßnahmen gegen die VLOP verhängen. Der Abschnitt enthält auch **eine besondere Verpflichtung, wenn VLOPs Empfehlungssysteme verwenden** (Artikel 29, §62) **oder Online-Werbung** auf ihrer Online-Schnittstelle **anzeigen** (Artikel 30, §63).

Die Kommission erkennt an, dass **Empfehlungssysteme** „eine wichtige Rolle bei der Verstärkung bestimmter Nachrichten, der viralen Verbreitung von Informationen und der Stimulierung des Online-Verhaltens" spielen können (§ 62). Daher sollten VLOPs sicherstellen, dass die Nutzer angemessen informiert werden und leicht Einfluss darauf nehmen können, welche Informationen ihnen empfohlen werden. Sie sollten die wichtigsten Parameter für solche Empfehlungssysteme auf einfache und verständliche Weise darstellen und so dafür sorgen, dass die Nutzer verstehen, wie die Informationen für sie priorisiert werden. Sie sollten auch sicherstellen, dass den Nutzern alternative Optionen für die Hauptparameter zur Verfügung stehen, etwa solche Optionen, die nicht auf dem Profiling des Nutzers basieren.

Die Kommission erkennt an, dass **Werbesysteme** „besondere Risiken" bergen und „aufgrund ihres Umfangs und ihrer Fähigkeit, Empfänger des Dienstes auf der Grundlage ihres Verhaltens innerhalb und außerhalb der Online-Schnittstelle dieser Plattform gezielt anzusprechen und zu erreichen, eine weitere öffentliche und regulatorische Aufsicht erfordern" (§ 63). Daher sollten VLOPs über ihre Programmierschnittstellen (APIs) öffentlichen Zugang zu einem Anzeigenarchiv gewährleisten, das folgende Informationen enthält: den Inhalt jeder Anzeige; die natürliche oder juristische Person, in deren Namen die Anzeige aufgegeben wurde; den Zeitraum, in dem die Anzeige angezeigt wurde; die wichtigsten Targeting-Parameter und die Gesamtzahl der Empfänger der Anzeige.

Darüber hinaus legt der Abschnitt 4 die Bedingungen fest, zu denen VLOPs dem Koordinator für digitale Dienste am Niederlassungsort oder der Kommission über Online-Datenbanken oder APIs **Zugang zu Daten** gewähren müssen, die für die Gewährleistung und Bewertung der Einhaltung dieser Verordnung erforderlich sind (Artikel 31.1, 31.3, § 64). Dazu zählen Daten, „die erforderlich sind, um die mit den Systemen der Plattform verbundenen Risiken und mögliche Schäden zu bewerten, sowie Daten zur Genauigkeit, Funktionsweise und Prüfung von Algorithmensystemen für die Moderation von Inhalten, Empfehlungs- oder Werbesysteme oder Daten zu Verfahren und Ergebnissen der Moderation von Inhalten oder von internen Beschwerdemanagementsystemen" (§64). Wichtig ist auch, dass VLOPs (auf Anfrage des Koordinators für digitale Dienste am Niederlassungsort oder der Kommission) zugelassenen Wissenschaftlern Zugang zu den Daten gewähren müssen, „zum ausschließlichen Zweck der Durchführung von Forschungsarbeiten, die zur Ermittlung und zum Verständnis systemischer Risiken [...] beitragen" (Artikel 31.2). Eine Zulassung erfolgt nur, wenn die Wissenschaftler einer akademischen Einrichtung angehören, unabhängig von kommerziellen Interessen sind und nachweislich über Fachwissen in den Bereichen verfügen, die mit den untersuchten Risiken oder den entsprechenden Forschungsmethoden zusammenhängen. Außerdem müssen sie sich verpflichten und in der Lage sein, die je nach Anfrage spezifischen Anforderungen an die Datensicherheit und Vertraulichkeit zu erfüllen (Artikel 31.4).

Schließlich verpflichten sich die VLOPs, einen oder mehrere Compliance-Beauftragte zu benennen, die für die Einhaltung der in der Verordnung festgelegten Verpflichtungen sorgen (Artikel 32). Artikel 33 listet weitere spezifische Transparenzverpflichtungen auf, die in den Artikeln 26-28 beschrieben sind.

Mittel zur Implementierung & Standardisierung der Sorgfaltspflichten

In Abschnitt 5 werden die Verfahren beschrieben, die von der Kommission unterstützt und gefördert werden, um die wirksame und kohärente Anwendung der Verpflichtungen aus dem DSA zu erleichtern, die möglicherweise eine Umsetzung mit technischen Mitteln erfordern. Die Kommission möchte „freiwillige Branchennormen" wie Verhaltenskodizes fördern, die „bestimmte technische Verfahren umfassen [...] soweit die Industrie dazu beitragen kann, genormte Instrumente für die Einhaltung dieser Verordnung zu entwickeln, z. B. durch die Möglichkeit, Mitteilungen etwa über Anwendungsprogrammierschnittstellen zu übermitteln, oder durch eine bessere Interoperabilität von Werbearchiven." (§ 66).

Die Kommission und der Vorstand fördern die Ausarbeitung von **Verhaltenskodizes**, die zur Anwendung des DSA beitragen (Artikel 35, § 67-69). Wichtig ist, dass diese Verhaltenskodizes Verpflichtungen zum Ergreifen spezifischer Maßnahmen zur Risikominderung enthalten können (Artikel 35.2), die auf der Grundlage von Leistungsindikatoren (KPIs) bewertet werden (Artikel 35.3). Die Kommission kann Plattformen und andere interessierte Parteien auffordern, sich freiwillig an dem Verhaltenskodex zu beteiligen (Artikel 35.2), wobei der Anreiz zur Beteiligung erheblich ist, da „die Beteiligung einer sehr großen Online-Plattform an einem Verhaltenskodex und dessen Einhaltung als geeignete Risikominderungsmaßnahme angesehen werden" (§ 68). Das ist mit der „Koregulierungssicherung" gemeint. Verhaltenskodizes sind freiwillig, aber die Nichtteilnahme an ihnen erhöht das Risiko der Nichteinhaltung des DSA.

Der DSA identifiziert eine Reihe von Bereichen, die für solche Verhaltenskodizes in Betracht kommen, insbesondere (1) Maßnahmen zur Risikominderung in Bezug auf bestimmte illegale Inhalte und (2) negative Auswirkungen systemischer Risiken wie Desinformation oder manipulative und missbräuchliche Aktivitäten auf die Gesellschaft und die Demokratie. Dazu gehören koordinierte Operationen, die darauf abzielen, Informationen zu verstärken, einschließlich Desinformation, etwa der Einsatz von Bots oder gefälschten Accounts zur Erstellung gefälschter oder irreführender Informationen, manchmal mit dem Ziel, einen wirtschaftlichen Gewinn zu erzielen. Diese Art von Desinformation ist besonders schädlich für schutzbedürftige Empfänger von Online-Diensten, wie etwa für Kinder (§ 68).

Als spezifischen künftigen Verhaltenskodex führt die Kommission **einen Verhaltenskodex für Online-Werbung** auf. Dieser Kodex würde über die verpflichtenden Anzeigenarchive (Artikel 30) und die nutzerorientierten Transparenz-Tools für Anzeigen (Artikel 24) hinausgehen. Sein Ziel ist es, verschiedene Akteure der Ad-Tech-Wertschöpfungskette mit Organisationen der Zivilgesellschaft oder den zuständigen Behörden zusammenzubringen, um mehr Transparenz bezüglich der „Übermittlung der relevanten Informationen" in der Ad-Tech-Wertschöpfungskette zu schaffen, von den Herausgebern bis zu den Werbetreibenden (Artikel 36, §70), um ein „wettbewerbsorientiertes, transparentes und faires Umfeld in der Online-Werbung" zu gewährleisten (Artikel 36.2).

Es gibt auch eine Bestimmung zu **Krisenprotokollen, um außergewöhnlichen, die öffentliche Sicherheit oder die öffentliche Gesundheit betreffenden Umständen zu begegnen** (Artikel 37, §71). Dazu gehört „jedes unvorhersehbare Ereignis wie z. B. Erdbeben, Wirbelstürme, Pandemien und andere ernste grenzüberschreitende Bedrohungen für die öffentliche Gesundheit sowie Krieg und terroristische Handlungen [...], bei denen Online-Plattformen z. B. für eine schnelle Verbreitung von illegalen Inhalten oder Desinformation missbraucht werden können oder eine rasche Verbreitung verlässlicher Informationen erforderlich ist" (§71). VLOPs sollten bei der Erstellung und Anwendung spezifischer Krisenprotokolle unterstützt werden. Solche Krisenprotokolle sollten nur für einen begrenzten Zeitraum aktiviert werden, und die ergriffenen Maßnahmen sollten sich auf das beschränken, was unbedingt notwendig ist, um den außergewöhnlichen Umstand zu bewältigen, etwa durch „hervorgehobene Darstellung von Informationen über die Krisensituation, die von den Behörden der Mitgliedstaaten oder auf Unionsebene bereitgestellt werden" (Artikel 37.2.a).

Überwachung & Durchsetzung

Der DSA arbeitet basierend auf drei Faustregeln:

- Die Gewährleistung einer angemessenen Aufsicht und Durchsetzung sollte grundsätzlich Sache der Mitgliedstaaten sein.
- Der Mitgliedstaat, in dem sich die Hauptniederlassung des Anbieters der Vermittlungsdienste befindet, ist für die entsprechenden Sorgfaltspflichten zuständig, was bei VLOPs in der Regel eine irische Regulierungsbehörde wäre.
- Für den Fall, dass unionsweit systemische Risiken durch VLOPs entstehen, sieht die vorgeschlagene Verordnung eine Überwachung und Durchsetzung auf Unionsebene vor - sei es durch das Europäische Gremium für digitale Dienste oder die Europäische Kommission.

Koordinatoren für digitale Dienste (DSCs)	**Europäisches Gremium für digitale Dienste (Gremium)**	**Europäische Kommission (CE)**
• Unabhängige Behörden • Direkte Überwachung und Durchsetzung (standardmäßig) • Koordinierung mit anderen zuständigen nationalen Behörden • Koordination und Kooperation auf EU-Ebene mit dem Gremium, CE und anderen DSCs	• Unabhängige Ad-hoc-Beratergruppe • Zusammengestellt von DSCs • Vorsitz durch CE, keine Abstimmungsrechte für CE • Beratung von DSCs und CE, Empfehlung von Maßnahmen • Keine rechtlich-bindenden Akte, muss aber von DSC und CE berücksichtigt werden • Zusammenarbeit mit anderen EU-Gremien, Agenturen und Ämtern in verwandten Angelegenheiten	• Direkte Durchsetzungsbefugnisse gegenüber sehr großen Online-Plattformen für: • Besondere Verpflichtungen für VLOPs (nach DSC-Überwachung) • Alle anderen Verpflichtungen (bei Untätigkeit der DSC) • Administrative Unterstützung des Vorstands • Berät bei grenzüberschreitenden Streitigkeiten • Greift auf Anfrage der DSC ein

Überwachung auf nationaler Ebene

Jeder Mitgliedstaat sollte „mindestens eine“ bestehende oder neue nationale Behörde benennen, deren Aufgabe es ist, den DSA durchzusetzen, insbesondere hinsichtlich der Haftungsregelung für Vermittlungsdienste. Spezifische Regulierungsbehörden können für bestimmte Aufsichts- oder Durchsetzungsaufgaben zuständig sein (etwa Medienaufsichtsbehörden oder Verbraucherschutzbehörden) (§ 72), aber nur eine Behörde kann „Koordinator für digitale Dienste“ sein und als einheitliche Anlaufstelle fungieren.

Die Koordinatoren für digitale Dienste haben **Ermittlungsbefugnisse**. Dazu gehört die Befugnis, von Anbietern die Herausgabe von Informationen zu verlangen, die sich auf einen möglichen Verstoß gegen den DSA beziehen (Artikel 41a); die Befugnis, Inspektionen vor Ort durchzuführen, um solche Informationen zu beschlagnahmen (Artikel 41b), und die Befugnis, jeden Mitarbeiter aufzufordern, Erklärungen zu diesen Informationen abzugeben (Artikel 41c).

Sie haben auch **Durchsetzungsbefugnisse**, namentlich die Befugnis, die Einstellung des Verstoßes anzuordnen, Geldbußen zu verhängen und einstweilige Maßnahmen anzuordnen. Die Nichteinhaltung der Verpflichtungen aus dem DSA kann mit einer Geldstrafe von bis zu 6 % des Jahreseinkommens oder -umsatzes des Anbieters geahndet werden, während Strafen für die Bereitstellung falscher, unvollständiger oder irreführender Informationen eine Geldstrafe von bis zu 1 % des Jahreseinkommens oder -umsatzes des Anbieters zur Folge haben können (Artikel 42.3). Wenn der Verstoß nach Ausschöpfung all dieser Möglichkeiten immer noch andauert und schwere Straftaten nach sich zieht, kann der Koordinator für digitale Dienste schließlich eine Justizbehörde ersuchen, die vorübergehende Sperrung des Zugangs für Nutzer für mindestens vier Wochen anzuordnen (Artikel 41.3b).

Wichtig ist, dass der Mitgliedstaat, in dem sich die Hauptniederlassung des Anbieters von Vermittlungsdiensten befindet, für die Sorgfaltspflichten der Plattformen **zuständig** ist. In der Regel wäre das eine irische Regulierungsbehörde (Artikel 40.1). Allerdings können Koordinatoren für digitale Dienste aus anderen Rechtsordnungen – oder der Vorstand – den Koordinator für digitale Dienste ersuchen, die erforderlichen Ermittlungs- und Durchsetzungsmaßnahmen zu ergreifen (Artikel 45). Einzelpersonen oder repräsentative Organisationen sollten in der Lage sein, alle Beschwerden, die in Zusammenhang mit der Einhaltung dieser Verordnung eingegangen sind, beim Koordinator für digitale Dienste in dem Land einzureichen, in dem sie den Dienst erhalten haben (Artikel 43, §81). Falls erforderlich, wird der Koordinator für digitale Dienste die Beschwerde an den Koordinator für digitale Dienste in der entsprechenden Niederlassung weiterleiten.

Supranationale Beaufsichtigung

Mit dem DSA wird ein **Europäisches Gremium für digitale Dienste** eingerichtet, das als unabhängiges Beratungsgremium auf EU-Ebene die Kommission unterstützt und ihr hilft, die Maßnahmen des Koordinators für digitale Dienste (DSC) zu koordinieren, u. a. durch Vorschläge für geeignete Untersuchungs- und Durchsetzungsmaßnahmen. Das Gremium wird aus allen DSCs bestehen und an der Ausarbeitung von Verhaltenskodizes mitwirken. Es wird außerdem unverbindliche Stellungnahmen für die DSCs oder andere zuständige nationale Behörden abgeben (Artikel 47, §88-90).

Für den Fall, dass durch VLOPs unionsweit systemische Risiken entstehen, sieht die vorgeschlagene Verordnung eine Aufsicht und Durchsetzung auf Unionsebene vor. Abschnitt 3 betrifft die Beaufsichtigung, Untersuchung, Überwachung von VLOPs, sowie eventuelle Vollstreckungsmaßnahmen. Er sieht eine verstärkte Aufsicht für den Fall vor, dass Plattformen gegen die Bestimmungen aus Abschnitt 4 verstoßen (Artikel 50, § 94).

Sobald ein Verstoß festgestellt wurde, etwa aufgrund von individuellen oder gemeinsamen Untersuchungen, Audits oder Beschwerden, sollte der Koordinator für digitale Dienste (DSC) am Niederlassungsort alle Maßnahmen überwachen, die von der betreffenden VLOP gemäß ihres Aktionsplans ergriffen wurde (Artikel 50.1/50.2). Wenn der DSC Bedenken hat, dass diese Maßnahmen nicht wirksam sein könnten, kann ein weiteres Audit dieser Maßnahmen verlangt werden (Artikel 50.3/95). Wurde der Verstoß nach Ansicht des DSCs vom VLOP nicht eingestellt, kann die Kommission weitere Ermittlungen anstellen (Artikel 51).

Der DSA sieht auch die Möglichkeit vor, dass die Kommission von sich aus oder auf Ersuchen des Boards bei VLOPs interveniert (Artikel 51, §96). In diesen Fällen kann die Kommission Untersuchungen durchführen, u. a. indem sie VLOPs dazu zwingt, alle relevanten Dokumente, Daten und Informationen zur Verfügung zu stellen, die für die Untersuchung notwendig sind, einschließlich Erklärungen zu Datenbanken und Algorithmen (Artikel 52, §99), Befragungen (Artikel 53) und Vor-Ort-Inspektionen (Artikel 54). Bei Nachprüfungen vor Ort können die Kommission und die von ihr benannten Prüfer oder Sachverständigen von der betreffenden VLOP oder einer anderen in Artikel 52 Absatz 1 genannten Person Erklärungen zu ihrer Organisation, ihrer Arbeitsweise, ihrem IT-System, ihren Algorithmen, ihrem Umgang mit Daten und ihrem Geschäftsgebaren verlangen (Artikel 54 Absatz 3).

Die Kommission kann einstweilige Maßnahmen erlassen (Artikel 55), Zusagen von VLOPs verbindlich machen (Artikel 56) sowie deren Einhaltung der Verordnung überwachen (Artikel 57). Im Falle der Nichteinhaltung kann die Kommission Entscheidungen zur Nichteinhaltung (Artikel 58) treffen sowie Geldbußen (Artikel 59) und Zwangsgelder (Artikel 60) für Verstöße gegen die Verordnung sowie für die Erteilung unrichtiger, unvollständiger oder irreführender Auskünfte im Rahmen der Untersuchung verhängen.

Antwort auf die Konsultation zum „Online Harms White Paper“ in Großbritannien[8]

Die Konsultationsphase der britischen Regierung zu den Vorschlägen, „Großbritannien zum sichersten Ort der Welt zu machen, um online zu gehen, und zum besten Ort, um ein digitales Unternehmen aufzubauen und zu gründen“ (1.0)[9], ist nun abgeschlossen[10]. Die Vorschläge zur Regulierung von „Online Harms“ wurden erstmals im Online-Harms-White Paper erwähnt, das im April 2019 veröffentlicht wurde. Ein Zwischenbericht zur Konsultation wurde im Februar 2020 veröffentlicht, der vollständige Konsultationsbericht im Dezember 2020. Der neue, leicht veränderte Rechtsrahmen wird über das bevorstehende „Online Safety Bill“ hinausgehen. Er wird für 2021 erwartet. Der Kern des Vorschlags ist eine neue gesetzliche Sorgfaltspflicht, die von der unabhängigen Regulierungsbehörde Ofcom durchgesetzt wird.

Unternehmen im Geltungsbereich

Der Gesetzesrahmen gilt für Unternehmen, die entweder „**nutzergenerierte Inhalte** hosten, auf die Nutzer im Vereinigten Königreich zugreifen können“ oder „öffentliche oder private Online-Interaktionen zwischen Nutzern von Diensten“ ermöglichen, von denen sich mindestens eine Partei im Vereinigten Königreich befindet (1.1). Es wurde klargestellt, dass **Suchmaschinen** in den Anwendungsbereich fallen (1.3). Die britische Regierung hat betont, dass sie einen risikobasierten Ansatz verfolgen und die Regulierungsmaßnahmen auf Unternehmen konzentrieren wird, deren Dienste das größte Schadensrisiko darstellen (20). Zu den Aktivitäten, die vom Anwendungsbereich ausgeschlossen werden sollen, gehören ISPs, Hosting-Anbieter, App-Stores sowie Business-to-Business-Dienste (1.2).

In der Konsultationsantwort heißt es außerdem, dass **journalistische Inhalte** besonders geschützt sein werden, zum Beispiel durch Ausnahmen für die eigenen Websites der Nachrichtenmedien und zusätzlich „robuste Schutzmaßnahmen für journalistische Inhalte, die auf Diensten innerhalb des Geltungsbereichs geteilt werden“. Als Grund dafür wird angegeben, dass Medienunternehmen „Bedenken geäußert haben, die Regulierung könne zu vermehrten Takedowns von journalistischen Inhalten führen“ (1.11). Es wurde noch nicht im Detail dargelegt, wie journalistische Inhalte definiert werden und wie dieser Schutz für Inhalte, die auf Diensten im Geltungsbereich geteilt werden, in der Praxis funktionieren soll.

Während alle Unternehmen, die in den Anwendungsbereich der Vorschläge fallen, eine gesetzliche Sorgfaltspflicht haben werden, fallen nur wenige Unternehmen mit hoher Reichweite und hohem Risiko (2.16) in die „**Kategorie 1**“ mit zusätzlichen Verpflichtungen. Für die Einstufung ist ein dreigliedriges Verfahren vorgesehen:

1. Die primäre Gesetzgebung wird Faktoren auf hoher Ebene festlegen (2.16), einschließlich der Größe der Zielgruppe und der angebotenen Funktionalität.

2. Die Regierung wird auf Anraten von Ofcom Schwellenwerte für diese Faktoren festlegen.

3. Die Bewertung von Ofcom erfolgt anhand dieser Faktoren und Schwellenwerte (2.18).

In der Konsultationsantwort wurden zwar keine spezifischen Unternehmen genannt, aber Medienbriefings deuten darauf hin, dass zu den Unternehmen der Kategorie 1 wahrscheinlich Facebook, TikTok, Instagram und Twitter gehören, neben anderen Plattformen wie YouTube. Die große Mehrheit der Dienste werden Unternehmen der „Kategorie 2“ sein, die keine zusätzlichen Verpflichtungen haben.

Sorgfaltspflicht

Ziel der Sorgfaltspflicht ist es, Nutzern von Online-Diensten mehr Sicherheit zu bieten, indem in erster Linie gegen Inhalte oder Aktivitäten vorgegangen wird, die „**erhebliche physische oder psychische Schäden [harms] bei Einzelpersonen verursachen**“ (2.7). Die Unternehmen werden „eine Bewertung der mit ihren Diensten verbundenen Risiken vornehmen und **angemessene Maßnahmen ergreifen, um die von ihnen festgestellten Risiken für das Auftreten von Schäden [‚harms‘] zu verringern**“. Zu diesen Maßnahmen gehören „Benutzerwerkzeuge, Inhaltsmoderation und Empfehlungsverfahren“ (2.9). Der Gesetzesrahmen wird **sowohl für öffentliche als auch für private Kommunikationskanäle und -dienste gelten**, beispielsweise für Messaging-Apps (29).

Ofcom wird nach Konsultation mit den Stakeholdern (2.50) **gesetzliche Codes of Practice herausgeben, wie Unternehmen die Sorgfaltspflicht erfüllen können** (2.48). Unternehmen können alternative Maßnahmen ergreifen, solange sie nachweisen können, dass diese gleichwertig sind oder die in den Codes dargelegten Standards übertreffen (2.48).

Unternehmen werden gesetzlich zu effektiven **Berichts- und Abhilfemechanismen** verpflichtet (2.12), aber die Regierung wird keine spezifischen Formen der Abhilfe vorschreiben (2.13), und es wird keine neuen Möglichkeiten für Einzelpersonen geben, Unternehmen zu verklagen (2.12). Ofcom wird zusätzlich Codes of Practice zu Rechtsbehelfsmechanismen veröffentlichen (2.12).

„Online Harms"

Der **Gesetzgeber wird schädliche Inhalte und Aktivitäten definieren**, die „erhebliche physische oder psychische Schäden [‚harms'] bei Personen verursachen" (2.7) und in den Anwendungsbereich der Regelung fallen, um Rechtssicherheit zu schaffen (2.1). Zu den ausdrücklich **ausgeschlossenen** Schäden gehören Verstöße in den Bereichen geistiges Eigentum, Datenschutz, Betrug, Verbraucherschutz und Cybersicherheit (2.4) mit der Begründung, dass sie von anderen rechtlichen Regelungen abgedeckt werden.

Zu den **vorrangigen Kategorien** schädlicher Inhalte und Aktivitäten, die in der Sekundärgesetzgebung festgelegt werden, gehören: i) vorrangige Straftaten, etwa sexuelle Ausbeutung und Missbrauch von Kindern (Child Sexual Exploitation and Abuse, CSEA), Terrorismus, Hasskriminalität; ii) für Kinder schädliche Inhalte, etwa Pornografie, gewalttätige Inhalte; iii) „vorrangige Kategorien schädlicher Inhalte und Aktivitäten, die legal sind, wenn sie von Erwachsenen aufgerufen werden, die für sie aber trotzdem schädlich sein können", wie z.B. Missbrauch, Essstörungen, Selbstverletzung (2.3 & 2.19). Rechtliche Verantwortlichkeiten in Bezug auf illegale Inhalte und Aktivitäten bleiben bestehen (2.23). Diese vorrangig zu behandelnden Kategorien umfassen sowohl illegale als auch legale, aber schädliche Inhalte und Aktivitäten.

Freiwillige vorläufige Verhaltenskodizes[11], die sich mit zwei vorrangigen Kategorien befassen, nämlich CSEA und Terrorismus, wurden zusammen mit der Konsultationsantwort veröffentlicht (2.51). Der CSEA-Kodex baut auf den Freiwilligen Grundsätzen zur Bekämpfung der sexuellen Ausbeutung und des Missbrauchs von Kindern im Internet auf (2.53). Ofcom wird die Befugnis haben, den Einsatz von „**automatisierter Technologie, die hochpräzise ist**" (2.59) zu verlangen, um CSEA-Inhalte und -Aktivitäten auf Plattformen zu identifizieren, wenn es „keine alternativen, weniger einschneidenden Ansätze [...] gibt" und dies verhältnismäßig ist (2.62). Ofcom wird verpflichtet sein, dem Innenminister jährlich über die Nutzung der Befugnis zu berichten, einschließlich über deren Effektivität und Genauigkeit (2.61). Ofcom wird die Befugnis haben, den Einsatz automatisierter Technologie zu verlangen, um illegale terroristische Inhalte und Aktivitäten zu identifizieren, zu kennzeichnen, zu blockieren oder zu entfernen (2.42), wenn dies effektiv, verhältnismäßig und notwendig ist (2.70).

Alle Unternehmen sind verpflichtet, die **Wahrscheinlichkeit zu bewerten, dass Kinder auf ihre Dienste zugreifen**. Wenn sie feststellen, dass Kinder wahrscheinlich auf ihre Dienste zugreifen, sind sie verpflichtet, zusätzliche Schutzmaßnahmen für Kinder, die diese Dienste nutzen, bereitzustellen. Dies entspricht dem Ansatz des Information Commissioner's Age Appropriate Design Code, der Standards für den Schutz personenbezogener Daten von Kindern vorgibt und so für Einheitlichkeit sorgt.

Wenn **Desinformation und Fehlinformation** „erhebliche physische oder psychische Schäden [‚harms'] bei Einzelpersonen" verursachen, fallen sie in den Anwendungsbereich des Rechtsrahmens. Ist es unwahrscheinlich, dass Desinformation oder Fehlinformation „diese Art von Schaden verursacht, fällt sie nicht in den Anwendungsbereich der Verordnung" (2.81). „[P]olitische Meinungen oder Kampagnen, die von inländischen Akteuren innerhalb des Gesetzes geteilt werden", sollen außerhalb des Anwendungsbereichs liegen (2.81). Ofcom wird verpflichtet, „eine Expertenarbeitsgruppe für Desinformation und Fehlinformation [...] einzurichten, um einen Konsens und technisches Wissen darüber aufzubauen, wie Desinformation und Fehlinformation zu bekämpfen sind" (2.85).

Internet-Ökosystem: Online Harms Framework

<table>
<tr><th>Außerhalb des Geltungsbereiches</th><th colspan="2">Innerhalb des Geltungsbereiches</th></tr>
<tr><td rowspan="3">• Internet Service Providers (ISPs)
• B2B-Dienstleistungen
• Hosting-Anbieter
• App-Stores</td><td colspan="2">Der Anwendungsbereich umfasst sowohl Suchmaschinen als auch Unternehmen, die:
• „nutzergenerierte Inhalte hosten, die von Nutzern im Vereinigten Königreich abgerufen werden können“ (und/oder)
• „öffentliche oder private Online-Interaktion zwischen Service-Benutzern erleichtern“.

Verpflichtungen:
• Übergreifende Sorgfaltspflicht, „zu verhindern, dass nutzergenerierte Inhalte oder Aktivitäten auf ihren Diensten zu erheblichen physischen oder psychischen Schäden [‚harms‘] bei Einzelpersonen führen“
• Ergreifung von Maßnahmen gegen relevante illegale Inhalte
• Beurteilung, ob Kinder den Dienst wahrscheinlich nutzen werden und ggf. Bereitstellung von Schutzmaßnahmen

Ausnahme:
• Solide Schutzmaßnahmen für journalistische Inhalte, die auf Diensten im Geltungsbereich geteilt werden</td></tr>
<tr><td>Kategorie 1</td><td>Kategorie 2</td></tr>
<tr><td>Zusätzliche Verpflichtungen zur Einhaltung von prioritären Schäden (‚priority harms‘) wie

1. prioritäre Straftaten wie CSEA, Terrorismus, Hassverbrechen;
2. Inhalte die für Kinder schädlich ist (z.B. Pornografie);
3. Inhalte die für Erwachsene legal aber schädlich ist (z.B. Missbrauch, Essstörungen, Selbstverletzung)

• Transparenzverpflichtungen</td><td>Keine zusätzlichen Verpflichtungen</td></tr>
</table>

Endnoten

Regulatorische Befugnisse und Durchsetzung

Ofcom wird eine Reihe von Funktionen haben, darunter die Festlegung, was Unternehmen tun müssen, um die Sorgfaltspflicht zu erfüllen, die Forderung an Unternehmen, effektive Rechtsbehelfsmechanismen zu haben, die Einrichtung von Mechanismen zur **Interessenvertretung der Nutzer** (4.39), die Bereitstellung einer ‚Superbeschwerdefunktion' für Interessenverbände (4.37) und die Förderung von Online-Sicherheit und -Innovation (Box 16).

Transparenzberichte (4.15) können von Unternehmen der Kategorie 1 verlangt werden. Darin werden diese darlegen, was sie tun, um die Sorgfaltspflicht zu erfüllen. Sie können aber auch von Unternehmen der Kategorie 2 verlangt werden. Dann werden sie wahrscheinlich Informationen über die interne Durchsetzung der Geschäftsbedingungen, Prozesse und Verfahren, den Einsatz automatisierter Tools, Risikobewertungen und Bemühungen zur Aufklärung der Nutzer enthalten (Box 17). Diese Transparenzberichte werden öffentlich zugänglich sein müssen (4.20).

Ofcom wird **Befugnisse zur Informationsbeschaffung** haben, etwa die Befugnis, Mitarbeiter zu befragen, sowie die Befugnis, „die Räumlichkeiten von Unternehmen zu betreten und Zugang zu Unterlagen, Daten und Geräten zu erhalten" (4.26). Zusätzlich wird Ofcom die Befugnis haben, von einem Unternehmen zu verlangen, einen externen „Sachverständigenbericht" einzureichen, der besonders nützlich sein wird, wenn externes technisches Fachwissen benötigt wird, etwa bei der Validierung „der Effektivität von automatisierten Moderationssystemen"(4.28).

ISD | Powering solutions to extremism and polarisation

Digital Policy Lab
Provokationspapier

Das freiheitlich-demokratische Internet – Fünf Modelle für eine digitale Zukunft

Alex Krasodomski-Jones

Über dieses Briefing

Dieses Provokationspapier untersucht Lösungsvorschläge für einen Ausgleich der politischen Kräfte, die für die Zukunft des Internets entscheidend sind. Es beleuchtet die Art und Weise, wie die Macht von Staaten, Unternehmen, Individuen und Maschinen das demokratischen Projekt fördert oder behindert. Desweiteren untersucht es das Gleichgewicht der Kräfte im Internet, von dem Regierungen unterschiedliche, teils gegensätzliche Vorstellungen haben. Die in diesem Papier geäußerten Ansichten sind die des Autors und spiegeln nicht unbedingt die Ansichten der Teilnehmer am Digital Policy Lab oder der vertretenden Regierungen wider.

Zur besseren Lesbarkeit wurde auf die gleichzeitige Verwendung männlicher und weiblicher Sprachformen verzichtet. Es wurde das generische Maskulinum verwendet, wobei alle Geschlechter gleichermaßen gemeint sind.

ISD | Powering solutions to extremism and polarisation

DEMOS

Beirut | Berlin | London | Paris | Washington DC

www.isdglobal.org

Über den Autor

Alex Krasodomski-Jones ist Direktor des von Demos und der University of Sussex geleiteten Centre for the Analysis of Social Media (CASM). Demos ist eine unabhängige, pädagogische Wohltätigkeitsorganisation, die in England und Wales registriert ist (Charity Registration no. 1042046).

Alex Krasodomski-Jones hat mehr als ein Dutzend Berichte über digitale Wahlintegrität, Praktiken der Inhaltsmoderation, digitale Regulierung und die Überschneidung von Technik und Politik verfasst. Dazu schreibt er regelmäßig Gastbeitrage für unter anderem die BBC, CNN, Wired und The Spectator.

Außerdem leitet er die Arbeit von Demos im Good Web Project, einer Koalition zivilgesellschaftlicher Organisationen, die daran arbeitet, eine liberal-demokratische Vision des Internets zu artikulieren. Sein Fokus liegt auf dem Aufbau einer Evidenzbasis, sowohl dafür, wie gute Praxis in der Architektur und den Praktiken zukünftiger Online-Räume aussieht, als auch für den öffentlichen und staatlichen Konsens über diese Zukunft.

Zusammenfassung

Ohne eine wertebasierte Vision für das Internet laufen unsere demokratischen Traditionen und Prinzipien, unsere Regierungen und Gesellschaften Gefahr, im Wettlauf um die Neugestaltung des wichtigsten internationalen politischen, kulturellen und sozialen Raums hinter autoritären Staaten, technopolistischen Industriegiganten und autonomen Technologien zurückzubleiben.

Das Internet wurde unter bestimmten Umständen geschaffen und entwickelte sich zunächst von einem nationalen US-Projekt zu einem westlichen Projekt bis es ein globales Phänomen wurde. Ein akademisches Projekt mit militärischen Ursprüngen, das zunächst von digitalen Visionären aufgebaut wurde, bevor die Macht an die Internet-Giganten überging, mit denen unser Leben nun völlig verwoben ist. Diese Entstehungsgeschichte spiegelt sich in den Prinzipien, die dem Internet, wie wir es kennen, zugrunde liegen.

Diese Prinzipien werden nun in Frage gestellt. Zuerst waren es autoritäre Regime, die misstrauisch gegenüber der emanzipatorischen Wirkung des Internets waren. Jetzt wetteifern Regierungen auf der ganzen Welt um eine neue Digitalpolitik. Klar ist: Um die Zukunft des Internets gibt es Richtungskämpfe, und von einem Interessenausgleich oder gar einer einheitlichen Lösung bestehender Probleme ist man weit entfernt.

Die Machtverhältnisse zwischen Staaten und Unternehmen, Unternehmen und Bürgern sowie der Gesellschaftsvertrag zwischen Staaten und ihren Bürgern sind online ständig in Bewegung. Mächtige Technologien – künstliche Intelligenz und ‚vertrauenslose' Technologie – stellen einen vierte Kraft dar in einer Zeit, in der unser Leben zunehmend von Maschinen und nicht mehr von Menschen bestimmt wird.

Dieses Provokationspapier untersucht Lösungsvorschläge für ein Gleichgewicht der politischen Kräfte und welche Rolle sie für die Zukunft des Internets spielen. Es beleuchtet die Art und Weise, wie die Macht von Staaten, Unternehmen, Individuen und Maschinen das demokratische Projekt fördert oder behindert. Außerdem untersucht es das Kräftegleichgewicht im Internet, von deren Ausgestaltung Regierungen unterschiedliche, teils gegensätzliche Vorstellungen haben. Das Papier fordert, dass wir unsere Vision einer liberalen Demokratie im digitalen Zeitalter neu ausrichten, die Öffentlichkeit dafür gewinnen und diese Vision in Zusammenarbeit mit unseren Freunden und Partnern umsetzten. Wir brauchen eine Vision von einem Internet, das für Demokratien funktioniert und nicht ihre Grundfesten aushöhlt.

Wo wir sind und wie wir hier herkamen

Technologie ist politisch. Keine Technologie könnte eine größere politische Bedeutung haben als die, auf der das Internet basiert. Technologie wird nicht im luftleeren Raum geschaffen oder eingesetzt. Sie ist von Natur aus politisch, wobei der politische, soziale und kulturelle Kontext bestimmt, welche Technologie entwickelt und zu welchen Zwecken sie eingesetzt wird.

Um zu verstehen, wie wir hierher gekommen sind, müssen wir einzelne Informationsschichten abtragen und uns die Technologie, die sie alle transportiert, genauer anschauen: die Internetprotokolle. Dieser beeindruckend klingende Begriff bezeichnet die Regeln, nach denen Informationen im Internet ausgetauscht werden.

Wir haben Offline-Regeln der Kommunikation. Wir geben uns zum Beispiel die Hand oder küssen uns zur Begrüßung auf die Wange (einmal in Peru, zweimal in Kroatien, dreimal in Belgien). Eine genauere Untersuchung dieser Gewohnheiten kann uns etwas über eine Gesellschaft verraten, genauso wie ein Blick auf die Internetprotokolle uns etwas über die Prinzipien des Internets in seiner jetzigen Form erzählen kann.

Die **Internet Protocol Suite (IPS)**, die nach ihren beiden bekanntesten Protokollen oft als **TCP/IP (Transmission Control Protocol/Internet Protocol)** bezeichnet wird, trägt die politischen und kulturellen Narben ihrer Entstehungsgeschichte. Die Geschichte des Internets ist eine Geschichte von Streitigkeiten und Konflikten über seine Ausgestaltung: Über den Grad der möglichen zentralen Kontrolle, über das Ausmaß der Beteiligung von Militär und Regierung und darüber, wer davon profitieren sollte.[1, 2] Paul Barans frühe Projekte zum Aufbau einer belastbaren globalen Kommunikation bei der RAND Corporation fanden vor dem Hintergrund des Kalten Krieges statt, und das spätere ARPANET, das in den 1960er-Jahren entstand, wurde von Akademikern aufgebaut und mit Militärgeldern finanziert. Staaten, Unternehmen und Institutionen wetteiferten um die Macht, während das Netzwerk immer schneller immer größer wurde. Anleger witterten eine Goldgrube. Kurz nach dem Platzen der Dotcom-Blase entwickelten nach Google auch Facebook und andere das optimale Geschäftsmodell, um mit dieser Technologie Geld zu verdienen. Es dauerte nicht einmal zwanzig Jahre, bis ihre Anwendungen eine neue Basis für die globale Wirtschaft, Kultur, Gesellschaft und Politik kreiert hatten.

Je größer das Internet wurde, desto tiefer wurden die Wurzeln der Protokolle, Prinzipien und Normen auf denen es fußt. Sie formten die Welt neu.

Die Protokolle dieses Netzwerks unterstützen eine „stumme, vertrauenswürdige Mitte“ mit „intelligenten, anonymen Hosts an den Rändern“.[3] Das Internetprotokoll, unter Fachleuten schlicht IP genannt, kümmert sich weder darum, welche Informationen über das Netzwerk gesendet werden, noch kann es diesen Prozess kontrollieren. Es kümmert sich auch nicht um die Absender oder Empfänger. Dies ist von Grundgedanken her eine Barriere gegen zentralisierte Kontrolle oder Überwachung. **Das System basiert auf Vertrauen. Es setzt gutes Verhalten voraus und bietet begrenzte Sanktionsmöglichkeiten, wenn sich Menschen sittenwidrig verhalten, gegen Regeln verstoßen oder die Technologie missbrauchen.** Die Endnutzer sind bevollmächtigt, bleiben für die meisten Teilnehmer anonym und haben nach Belieben Zugang zum Netz: **Skalierbarkeit hat Vorrang vor jeglicher zentralisierter Kontrolle.** In ihrer Geschichte des TCP/IP erzählt Rebekah Larsen die Geschichte von Vint Cerfs Schalter: Einer der Väter des Internets hatte einen An/Aus-Schalter für das gesamte Netzwerk, über den Updates des ursprünglichen ARPANET durchgesetzt werden konnten. Von dieser Art der Kontrolle sind wir heute weit entfernt.

Anonym, frei, offen, vertrauensvoll, dezentralisiert und resistent gegen zentrale Kontrolle.
Das sind die Grundprinzipien des Internets, wie sie in der Technologie, die es zusammenhält, festgeschrieben sind.

Wie gut diese Prinzipien bei Staaten und Bürgern ankommen, verändert sich kontinuierlich. Manchen Menschen sind sie durchaus willkommen. Ein Unternehmer aus dem Silicon Valley zu Zeiten der Dotcom-Blase oder ein mitfühlender Beobachter der

Ereignisse im arabischen Frühling 2011 zeigte sich vermutlich begeistert angesichts der Möglichkeiten des Internets - ein Betrüger oder ein Rekrutierer für Terrororganisationen vielleicht auch. Einem Diktator, der die Rede-, Presse- oder Versammlungsfreiheit seiner Bürger fürchtet, gefällt das Internet vielleicht weniger. Gleiches gilt für Eltern, die sich Sorgen um die Surfgewohnheiten ihrer Kinder machen, oder für Sicherheitsbehörden, die sich mit neuen Formen der Informationskriegsführung und digital unterstützter Kriminalität auseinandersetzen.

Diese Bedenken prägen den Kampf um die Zukunft des Internets. Er wird in Aktionärsversammlungen und auf Titelseiten ausgetragen, in vertraulichen Tech-‚Roundtables' sowie im Heim jedes einzelnen Internetnutzers sowie an jedem Schauplatz staatlichen Handelns: Investitionen, Krieg, Handel, Regulierung, Sicherheitsvorkehrungen und vieles mehr.

Ein spannender Bericht findet sich in Wendy Halls und Kieron O'Haras bahnbrechendem Aufsatz Four Internets, der die Geschichten dieser Perspektiven erzählt.[4] Er beschreibt das offene Internet des Silicon Valley, den Enkel des frühen Internets, in dem Technologie und Profit die Innovation vorantreiben und die Prinzipien Freiheit und ungehinderter Zugang bestehen bleiben, wenn auch vorbehaltlich kommerzieller Zwänge. In diesem Modell hält sich der Staat bei der Festlegung der Spielregeln hinter den Unternehmen und der Technologie zurück.

Im Gegensatz dazu positioniert sich „Pekings autoritäres Internet" ideologisch als ein Werkzeug der Überwachung und Kontrolle. Autoritäre Staaten wie China zeigen wenig Begeisterung für die "dumme, vertrauensvolle Mitte", die als Bollwerk gegen staatliche Überwachung - wohlwollend oder nicht - fungiert. Private Unternehmen sind in diesem politischen Raum eine Erweiterung des Staates. Chinesische Internet-Giganten unterstehen der Regierung, nicht umgekehrt.

Neben diesen Monolithen decken bestimmte Visionen die ganze Bandbreite politischer Ideologien ab. Da gibt es beispielsweise Freiheitsabsolutisten, die sogar die Abschaffung bestehender Regulierungen oder Protokolle fordern, insofern sie als wettbewerbsfeindlich gelten. Diese Vision für das Internet wird in Four Internets von zahlreichen Republikanern in Washington, D.C. verkörpert, findet sich aber vielleicht auch in der eher visionären Welt der Krypto-Anarchisten und Alt-Techs, die wild entschlossen sind, die individuelle Freiheit gegenüber der Verantwortung zu maximieren.

Vor dem Hintergrund dieser gegensätzlichen Visionen erhebt sich eine neue Macht, die selbstbestimmende Technologie. Das ist eine Technologie, die Entscheidungen automatisiert oder bisweilen sogar ihre eigenen Regeln aufstellt. Wir sind noch nicht ganz am Zeitpunkt der technologischen Singularität angekommen.[5] Trotzdem treffen Maschinen für uns bereits seit Jahren viele kleine und große Entscheidungen – in der Regierung und in der Wirtschaft, von Krebsvorsorgeuntersuchungen über Prüfungsergebnisse und Nachrichten die wir lesen, bis hin zu Dingen, die wir kaufen sollen. Oft beruhen diese algorithmischen Empfehlungen auf Entscheidungen die wenig transparent sind und dadurch die Arbeit von Regulierungs- und Justizbehörden erschweren. Die Mathematik, die hinter der Ende-zu-Ende-Verschlüsselung (end-to-end encryption) steht, minimiert das Vertrauen, das ein Benutzer in zentrale Behörden setzen muss, ähnlich wie die Prinzipien, die der öffentlichen Blockchain-Technologie zugrunde liegen. **Der Nutzer vertraut auf die Technologie, nicht auf die Regierung, die Gesellschaft oder ein Unternehmen.** Wo immer wir sehen, dass Regierungen, die Gesellschaft oder Unternehmen mit der Technologie zu kämpfen haben, ist es vielleicht an der Zeit, innezuhalten und zu hinterfragen, ob es die Technologie selbst ist, die das eingangs beschriebene Machtverhältnis in Frage stellt.

Die Prinzipien des sich ständig verändernden Internets stellen liberale Demokratien vor Herausforderungen. Manche sind zu begrüßen. Dass ein gewisser Anstand erwartet wird (manche würden das naiv nennen), gehört zu den gängigen Werten in liberalen Demokratien als auch unter den Architekten des ursprünglichen Internets. Rechte auf Privatsphäre, Rede- und Versammlungsfreiheit sind Gründungsprinzipien liberaler Demokratien. In diesem Sinne haben wir 2011 den Anblick von Smartphones auf dem ägyptischen Tahrir-Platz gefeiert.

Andere Prinzipien haben Anlass zur Sorge gegeben. Demokratien beruhen auf einem Gesellschaftsvertrag, dem Vertrauen in die Regierung. Das Internet hat

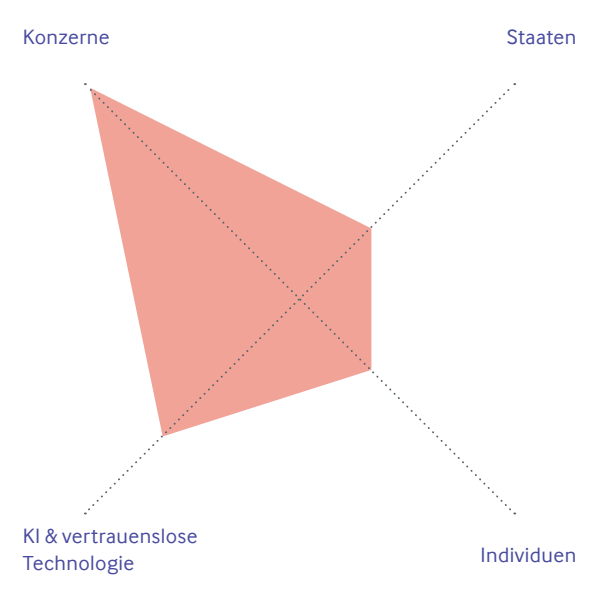
wiederholt die Grenzen der Staatsmacht, ihren Willen durchzusetzen, ausgereizt. Die „dumme Mitte“ und nachfolgende Technologien zur Verbesserung der Privatsphäre, etwa durch Verschlüsselung, stellen die Fähigkeit des Staates in Frage, eine seiner grundlegenden Pflichten zu erfüllen, nämlich die Sicherheit seiner Bürger zu gewährleisten. Sie machen sich gleichzeitig Sorgen, dass die Privatsphäre von Bürgern durch Technologieunternehmen verletzt wird, deren Unternehmenssitz oft außerhalb ihrer Gerichtsbarkeit liegt.[6]

Insgesamt identifizieren wir vier Kräfte, die das Internet in Zukunft gestalten können, vier Mächte, denen wir die Verantwortung für das digitale Leben geben müssen. Diese Mächte sind Staaten, Konzerne, Individuen und Maschinen.

Staaten, Konzerne, Individuen und Maschinen

Vier Mächte werden für die Form und Qualität des Internets verantwortlich sein: Staaten, Konzerne, Individuen und Maschinen: KI und ‚vertrauenslose' Technologie.

Silicon Valley



Eine Skizze der Machtverhältnisse im technozentrisch-unternehmerischen Silicon-Valley-Modell

Staatliche Macht des Internets kann viele Formen annehmen, umfasst aber in diesem Modell Bemühungen zur Regulierung und Kontrolle der Gestaltung des Internets durch nationale Regierungen und durch internationale Zusammenarbeit. In einer Demokratie geht es vor allem um Rechtsstaatlichkeit und deren Durchsetzung.[7] In China sieht das wohl sehr anders aus, als es in Europa oder nach einem UN- oder anderen internationalen Vertrag aussehen könnte. Es geht aber immer darum, das Internet und die ihm zugrundeliegende Technologie Regeln zu unterwerfen, die von Regierungen aufgestellt werden.

Unternehmensmacht ist anders. Hier machen die Unternehmen die Regeln. Ob es um die Geschwindigkeit oder den Umfang der Inhaltsmoderation geht oder um die Art der Inhalte, die zu den einzelnen Nutzern durchgelassen werden, die Regeln und Mechanismen, die diese Prozesse steuern, werden in den Vorstandsetagen festgelegt, bevor die Gesetzgeber ins Spiel kommen. Das Verhältnis zwischen staatlicher und unternehmerischer Kontrolle ist kompliziert. In einigen Fällen können Staaten die Regelsetzung an Unternehmen übertragen, mit der Begründung, dass es sich um private Unternehmen handelt, die das Recht haben, ihre eigenen Standards zu setzen. In anderen Fällen kann es sein, dass Regierungen einfach nicht die Macht oder Zuständigkeit haben, eine Plattform zu Änderungen zu zwingen, entweder weil letztere nicht willens oder einfach nicht in der Lage ist, das umzusetzen, was von ihr verlangt wird. Der anhaltende Kampf um urheberrechtlich geschützte Inhalte ist ein gutes Beispiel dafür. Einige Plattformen sind nicht bereit, urheberrechtlich geschützte Inhalte zu entfernen, während anderen die Technologie fehlt, um sie schnell genug zu erkennen und zu entfernen. In beiden Fällen ist die Macht des Staates zweitrangig.

Die Macht des Einzelnen stellt die Verantwortung und Fähigkeit der Bürger in den Vordergrund, die Online-Welt zu verstehen, zu beeinflussen und zu kontrollieren. Die informellen Vereinbarungen, die in die Protokolle des frühen Internets geschrieben wurden, sind Hinweise darauf, dass seine frühen Architekten großen Wert auf die Freiheit und Macht seiner Nutzer legten. Den Nutzern die Macht zu geben, ihre persönlichen Daten besser zu verwalten und die Räume, die sie online nutzen, entsprechend zu pflegen, ist dabei ein Kernpunkt. Die Sicherstellung der Macht

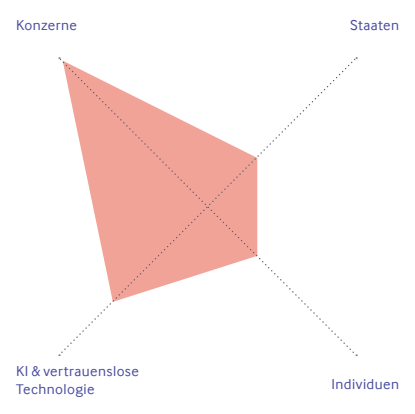
und Autonomie der Menschen im Internet war keine Priorität des kommerzialisierten Internets, wie wir es kennen, mit seinem Wirtschaftsmodell der gezielten Werbung, das stark von der Datenextraktion und einer willfährigen Nutzerbasis abhängig ist. Schutz der Nutzer vor Schäden sowie Chancengleichheit tauchen bestenfalls als Randnotiz in den Firmenbilanzen auf.

Und schließlich gibt es die Macht der Maschinen, genauer gesagt, der künstlichen Intelligenz und der ‚vertrauenslosen' Technologien wie Verschlüsselung, Blockchain und Kryptowährungen. Obwohl diese Technologien mitunter sehr unterschiedlich sind, haben sie eines gemeinsam: Es ist nicht ein Mensch, der die Regeln festlegt oder durchsetzt, sondern eine Technologie. Eine KI kann Krebs diagnostizieren, ein Bitcoin kann gekauft oder verkauft oder gehandelt werden, und Verträge können über die Blockchain abgeschlossen werden – alles ohne die Aufsicht einer zentralen Behörde. In den letzten Jahren haben wir gesehen, wie die wachsende Macht dieser Technologien unser Leben gestaltet. Ihre Existenz impliziert Entscheidungen in die weder Unternehmen noch Staaten Einblick haben. Die Mathematik hinter der Ende-zu-Ende-Verschlüsselung ist allgemein bekannt. Theoretisch kann sie jeder nutzen und es ist nahezu unmöglich, sie zu zensieren oder zu verbieten. Aber ihre Verwendung schafft Kanäle, die per Definition für Staaten, Unternehmen oder andere Personen nicht zugänglich sind. Die Blockchain-Technologie wird häufig als Übung zur Entfernung staatlicher und unternehmerischer Kontrolle aus einem System angepriesen: Bitcoin braucht keine Zentralbank. Künstliche Intelligenzen, selbst solche, die sich nominell in den Händen von Staaten oder Unternehmen befinden, sind häufig so kompliziert, dass ihre Entscheidungen nicht einfach erklärt, berechnet oder rückentwickelt werden können. Bürger geben ihre Entscheidungsgewalt bereits täglich an Algorithmen ab, wenn sie einkaufen oder navigieren, und auch Regierungen auf der ganzen Welt greifen bei ihrer Entscheidungsfindung zunehmend auf Algorithmen zurück. Staaten und Unternehmen mögen zwar glauben, dass KI kaum mehr als eine Erweiterung ihrer eigenen Entscheidungsbefugnis ist, aber diese Schlussfolgerung ist kurzsichtig. Als Erstes kommt die KI, die Beamte und Marketingverantwortliche zwar nicht verstehen, aber trotzdem einsetzen. Darauf folgt die KI, die so kompliziert ist, dass selbst ihre Schöpfer nicht genau wissen, wie sie funktioniert. Schließlich kommt eine zukünftige KI, die mächtig genug ist, um ihre eigenen Regeln zu schreiben und Regierungsaktivitäten weitaus effektiver durchzuführen als jede menschliche Organisation und die sich als resistent gegen jede Aufsicht, Rechenschaftspflicht oder Erklärung erweist.

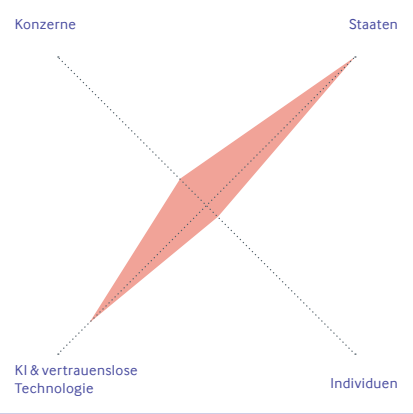
Maschinenmacht ist kein Schreckgespenst. Auch im Jahr 2021 sind Milliarden von Menschen in allen Bereichen ihres Lebens maschinellen Entscheidungen unterworfen. Diese Macht zu ignorieren und ihre Nutzung nicht zu regulieren, wäre ein Fehler. Zu den ersten Bemühungen in diese Richtung gehören die laufende Arbeit an einer KI-Ethik und der ethischen Nutzung von KI sowie Proteste gegen den Einsatz von KI in Gerichten, in der Polizeiarbeit und im Bildungswesen.

Die Grenzen sind fließend. Die Kontrolle durch Unternehmen kann die staatliche Aufsicht durch Einführung von Verschlüsselung aushebeln, wie es Facebook zu tun droht, während Staaten vielleicht willfährige Unternehmen heranzüchten, wie Sina Weibo oder WeChat in China. Dennoch bieten uns diese vier Arten von Macht nützliche Anhaltspunkte für die Vision eines zukünftigen Internets.

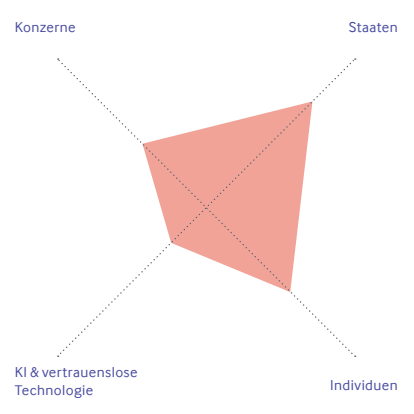
Wo soll die Macht liegen?
Sechs Modelle einer digitalen Zukunft



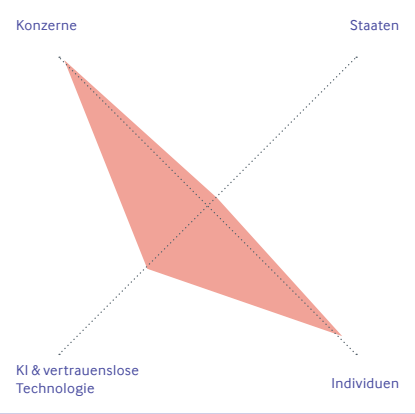
Ein Silicon Valley Internet



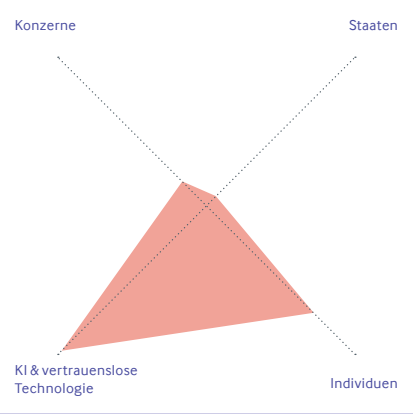
Ein autokratisches Internet



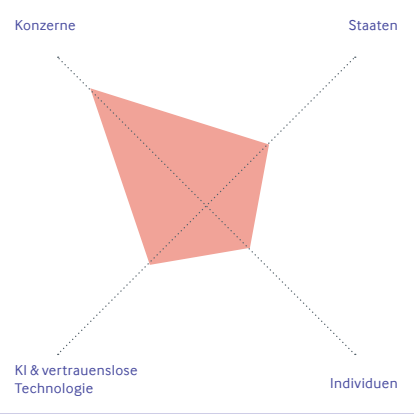
Ein EU Internet



Ein libertäres Internet



Ein maschinelles Internet



Großbritannien in 2021

Die Form dessen, was kommen wird

Die hier gezeigten Diagramme sind Karikaturen. Sie zeigen aber, wie unterschiedlich die konkurrierenden Visionen für das Internet aussehen könnten. Jede dieser Visionen birgt eine eigene Reihe von Bedrohungen und Chancen.

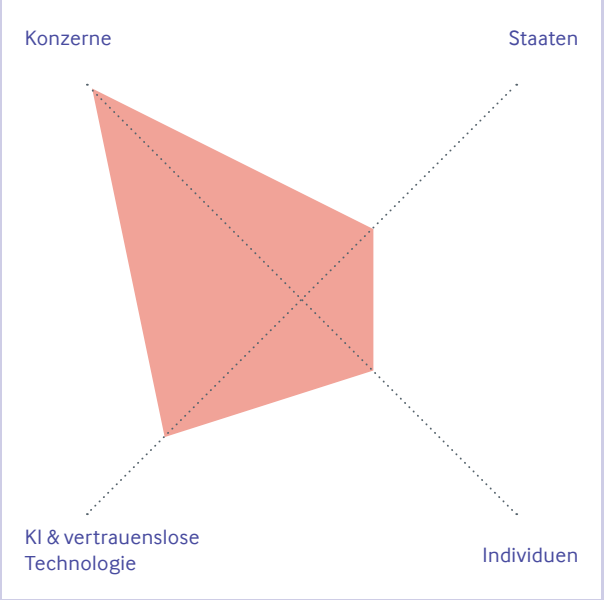
Ein Corporate Internet

Das **Corporate Internet** kommt dem Status quo der westlichen Welt am nächsten. Nach diesem Modell haben die großen internationalen Konzerne – Facebook, Amazon oder Google und seine Tochtergesellschaft YouTube – den größten Einfluss auf die Gestaltung, die Kultur und die Regeln im Internet. Jenseits von klar definierbaren illegalen Inhalten werden die Grenzen der freien Meinungsäußerung in den Vorstandsetagen des Silicon Valley festgelegt oder über die Konstellation kleinerer Plattformen entschieden, die mit Werbeeinnahmen finanziert und deren Mechanismen häufig von Google, Amazon oder Facebook bereitgestellt werden.

Es gibt zwar Spannungen zwischen den Infrastrukturanbietern und den Anwendungen, die auf ihnen laufen. In diesem Modell aber werden die Spannungen im privaten Sektor gelöst. Die Ausweitung des Starlink-Programms von Tesla zum unternehmenseigenen internationalen Anbieter von Internetzugängen ist ein guter Indikator für das, was auf uns zukommt: **Unternehmen umgehen die staatlich auferlegten infrastrukturellen Grenzen ihrer Aktivitäten.**

Die Beziehung zwischen Plattform und Staat ist eine einseitige Angelegenheit. Die Regulierung kommt nur langsam voran, wird ständig angefochten, und die Rechtsanwendung wird erschwert durch einen Mangel an Transparenz und sinnvollen Möglichkeiten, das zu messen oder zu überwachen, was zu einem bestimmten Zeitpunkt auf einer Plattform geschieht. Der Datenzugang und die Datenerfassung seitens des Staates sind im Vergleich zu den Möglichkeiten der Plattformen eher schwach. Für den einzelnen Nutzer sieht es noch schlechter aus. Die angebotenen Dienste

Ein Silicon Valley Internet



sind außergewöhnlich und nominell kostenlos, werden aber zu Bedingungen angeboten, die ihre Nutzer entmachten. Beschwerde- oder Kontrollsysteme, die den Nutzern auf den Plattformen angeboten werden, sind kaum mehr als ein Furnier, das die **Asymmetrie der Macht** verdeckt.

Die Technologie spielt hier eine entscheidende Rolle. Verschlüsselung erschwert den Überblick und wird eingesetzt, um Marktanteile zu schützen, aber auch, um eine gewisse Distanz zwischen der Plattform und den auf ihr zirkulierenden Inhalten zu schaffen. KI und algorithmische Kuratierung sind der einzig gangbare Weg, um so große Räume zu verwalten, die Datenerfassung zu optimieren und Werbeeinnahmen zu maximieren. Die Funktionsweise dieser Algorithmen ist undurchsichtig und ihre Entscheidungen sind weitgehend unanfechtbar.

Ein autokratisches Internet

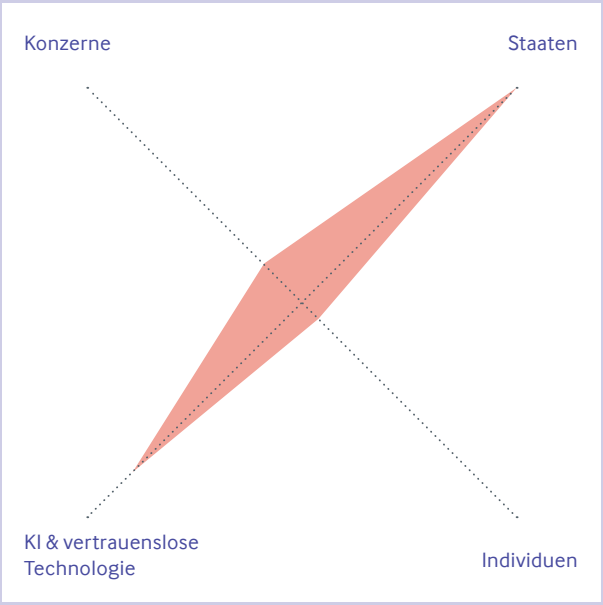
Die Online-Welt, die in China und unter in seinem zunehmenden Einfluss auch in den Entwicklungsländern entsteht, steht im Gegensatz zum „Technopol“ des korporativen Internets, wie wir es im Westen kennen. Hier hat der Staat das Sagen, und die Plattformtechnologie ist eher eine Erweiterung der Staatsmacht als ein Dorn in deren Auge. Die Macht der Individuen ist minimal.

Protokolle und Infrastruktur sind staatszentriert und stellen die staatliche Souveränität in den Mittelpunkt. In ihrer anspruchsvollsten Form schließen sie harte Begrenzungen für das Internet ein, wie etwa das Internet in Nordkorea. In ihrer einfachsten Form sind sie ein An/Aus-Schalter.

Die Rechte des Einzelnen bleiben eingeschränkt. In diesem Modell erleichtern die Plattformen der Regierung den Datenzugriff, während enormer Mengen verknüpfter Daten der Internetnutzer Überwachung, soziale Punktesysteme und experimentelle Technologien untermauern. Die Unterwerfung der Uiguren in China wird vor allem durch Technologien erleichtert.[8]

Dieselben Daten erschließen das volle Potenzial staatlich angepasster künstlicher Intelligenzen, die ihrerseits eine zusätzliche Erweiterung der staatlichen Macht darstellen, insofern ihre Ergebnisse und Operationen bekannt und verständlich bleiben. Die Möglichkeiten der Bürger, Rechtsmittel einzulegen, sind gering, egal ob eine Entscheidung von einer KI oder dem Staat getroffen wird, wobei der Unterschied zunehmend verschwimmt.

Maschinen in den Dienst des Überwachungspanoptikums zu stellen, bedeutet, einige Technologien zu unterbinden und andere zu fördern. Chinas Verschlüsselungsgesetz von 2020 führt einen abgestuften Ansatz zur Verschlüsselung ein, den Kritiker als gleichbedeutend mit dem Verbot einer Ende-zu-Ende-Verschlüsselung für alle außer der Regierungspartei betrachten.[9] Kryptographische Anwendungen wie Tor, Telegram, WhatsApp, Mastodon oder Virtuelle Private Netzwerke (VPNs) sind in dem Land verboten.



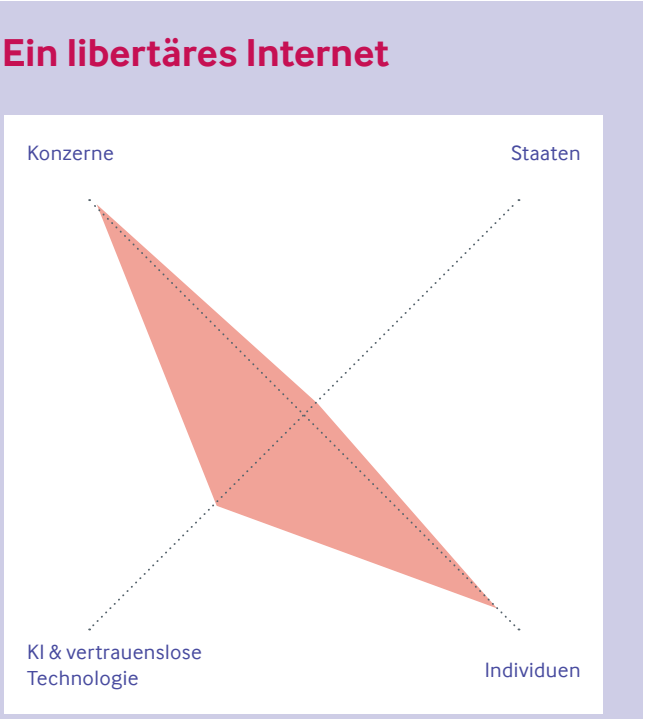
Ein libertäres Internet

Die Position der USA zur Zukunft des Internets wird oft mit den Ideen und Ambitionen der Internet-Giganten in Verbindung gebracht, die hier zu Hause sind – sprich das Corporate-Internet. Aber es gibt eine Spaltung im Land, und die Entscheidungsträger in der Politik haben eine konkurrierende Vision für die Seele des Internets entwickelt: Washington D.C.'s kommerzielles Internet, ein Internet, das privaten Akteuren von der Plattform bis zur Bereitstellung der Infrastruktur den Vorrang gibt, ein Markt, frei von jeglicher Regulierung. **Meinungsfreiheit wird in diesem Modell als Freiheit von staatlichen Eingriffen interpretiert und nicht als staatlich garantierte Chancengleichheit.**

Nach diesem Modell sollte es nichts geben, was eine Person daran hindert, online digitale Dienste zu erstellen, auf sie zuzugreifen oder anderweitig an ihnen teilzunehmen. Der Einzelne übernimmt die Verantwortung für sein Verhalten unter Bedingungen, die von anderen Einzelpersonen festgelegt und durchgesetzt werden. Der Schutz der Privatsphäre oder der eigenen Online-Rechte obliegt den Einzelnen: Ihren Fähigkeiten oder den Diensten, die sie nutzen. Einschränkungen der Freiheit, wie gesetzliche Kodizes der Rede oder der Meinungsäußerung, sind hier ebenso tabu wie Regeln, die Chancengleichheit fordern.

Nach diesem **eigentumsbasierten Modell** haben Unternehmen kein Interesse daran, etwas Anderes anzubieten als das, was ihre Kunden vielleicht wollen. Internet Service Provider (ISPs) sollen in der Lage sein, ihre Gewinne zu maximieren. Altbewährte Internet-Prinzipien wie die Netzneutralität stehen diesem Ziel im Weg. Nach diesem Modell gibt es keine Anforderungen in Bezug auf das öffentliche Wohl oder die Transparenz von Plattformen. Auch die Interoperabilität von Internetseiten und Diensten wird meist nicht befördert. Dies läuft den Hoffnungen der frühen Internetpioniere zuwider, nach deren Ansicht das Internet ein einziger, verbundener Informationsraum sein sollte. Vielmehr hat eine Fragmentierung des Internets in profitorientierte „Walled Gardens" stattgefunden.[10]

Hier teilen sich Unternehmen und ihre Kunden die Macht, frei von jeder staatlichen Aufsicht. Es ist ein Modell, auf das sich diejenigen berufen, die staatliche Eingriffe auf einer eher mikroskopisch kleinen Ebene



ablehnen. Der Druck auf die großen Plattformen, „Online Harms" zu reduzieren, hat vermehrt zu sogenannten „**Oasen der freien Meinungsäußerung**" geführt, alternativen Technologieplattformen wie Parler und Gab, die tendenziell extremistische politische Positionen bedienen, die durch die Nutzungsbedingungen der Silicon-Valley-Giganten verboten sind. Die infrastrukturellen Schwächen dieser Alternativen kamen Anfang 2021 ans Licht, als Parler nach den Anschlägen auf das US-Kapitol von Amazon Internet Services gesperrt wurde. Das libertäre Internetmodell fördert diese alternativen Plattformen. Es gibt einen Markt für sie, und so sollte es ihnen erlaubt sein, diesen Markt zu befriedigen. Druck auf Dienstanbieter, sie zu zensieren, wird ihr Wachstum und ihre Verbreitung vermutlich nur beschleunigen.

Vertrauenswürdige Technologie wird bei diesem Modell lediglich zu einem weiteren Produktmerkmal. Wenn Kunden Sicherheit verlangen, sollte es keine Hindernisse geben, eine leistungsstarke Verschlüsselung in den Dienst zu integrieren, wenn man damit den eigenen Kundenstamm vergrößern und die Mitbewerber hinter sich lassen kann.

Ein maschinelles Internet

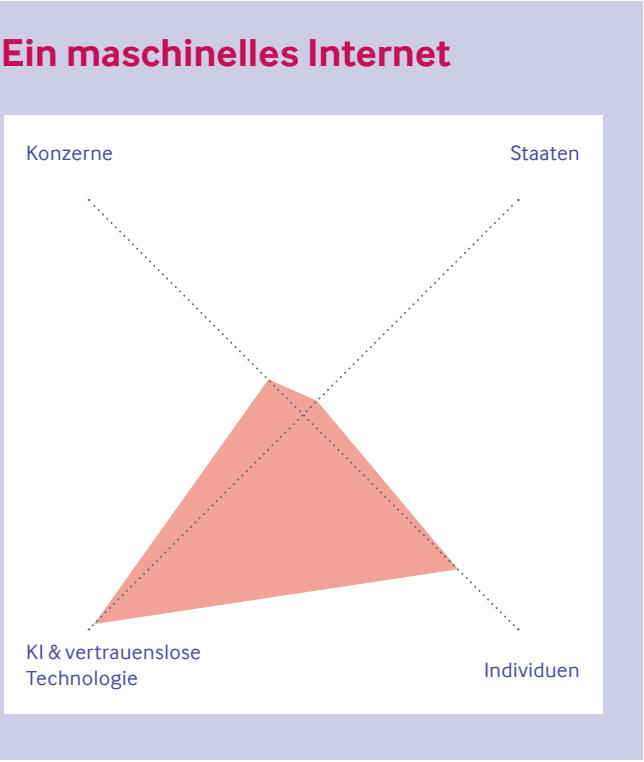
Schließlich skizzieren wir einen vierten Rahmen: Ein maschinelles Internet. Hier werden Politik, Gesellschaft und Kultur von Regeln bestimmt, die nicht Menschen, sondern künstliche Intelligenz festgelegt haben.

Nach dieser Vorstellung legt die Technologie selbst die Regeln fest und setzt sie durch – zunächst auf Geheiß eines Staates oder Unternehmens, aber schließlich jenseits aller Unternehmens- oder Staatsinteressen. Diese Vision des Internets ist am weitesten von unserer aktuellen Vorstellung und Realität entfernt, aber die Entwicklung der maschinengestützten Entscheidungsfindung und das anhaltende Wachstum von Krypto-Assets machen eine Welt erforschenswert, in der Code das Gesetz ist.

Governance durch KI ist auf dem Vormarsch. Ausreichend leistungsfähige KI wird zukünftig – und teils bereits – eingesetzt, um Entscheidungen über immer mehr Bereiche unseres Lebens zu treffen, angefangen bei den Wegen, die wir zur Arbeit nehmen, über unsere Möglichkeiten, einen Kredit zu erhalten oder ein Haus zu kaufen, bis hin zu KI-gestützter Strafverfolgung, nationaler Sicherheit und Bereitstellung öffentlicher Dienstleistungen.

Das Anfechten von Entscheidungen, die von Algorithmen getroffen werden, ist angesichts des dafür erforderlichen technischen Fachwissens schon jetzt schwierig. Computergestützte Entscheidungsfindung hat sich in einigen Bereichen bereits als effektiver erwiesen als menschliche Entscheidungen, etwa bei der Diagnose von Erkrankungen oder bei der Erkennung von Betrug. In einem maschinellen Internet wird erwartet, dass KI-Systeme die Funktionen der Regierung immer effektiver und effizienter nachbilden und schließlich Stück für Stück ersetzen. Dies hat ungeklärte Auswirkungen auf Fragen der demokratischen Wahl und der politischen Repräsentation. Voreingenommene, undurchsichtige Algorithmen erfüllen nicht die uns geläufigen demokratischen Standards.

Dies ist keine Science-Fiction. Tagtäglich sind Milliarden von Menschen weltweit Entscheidungen unterworfen, die von Maschinen getroffen werden – Maschinen, die sie nicht verstehen, über die sie



keine Macht haben und von denen sie keinerlei Schadensersatz erwarten können. Tagtäglich werden Regierungen mit Technologien konfrontiert, die ihre Macht einschränken.

Der Wert von Kryptowährungen wie Bitcoin wurde bisher eher durch Spekulation angetrieben, jedoch wird ihr Einsatz als Möglichkeit gesehen, Finanzdienstleistungen für Menschen anzubieten, die kein Vertrauen in Unternehmen oder zentrale Behörden haben. Neue, sogenannte erlaubnisfreie Systeme, auch **digitale autonome Organisationen (DAOs) genannt** mit intelligenten Verträgen sind alle darauf ausgelegt, den Transfer von Geld und die kommerzielle Zusammenarbeit zwischen Nutzern zu ermöglichen, und zwar ganz ohne die Beteiligung oder Aufsicht Dritter, seien es Unternehmen oder Staaten.[11] Die Architekten dieser Systeme stellen sich eine Welt vor, in der die digitale Technologie die Nationalstaaten vollständig ersetzt, indem sie es Individuen ermöglicht, allein durch Technologie zu kooperieren.

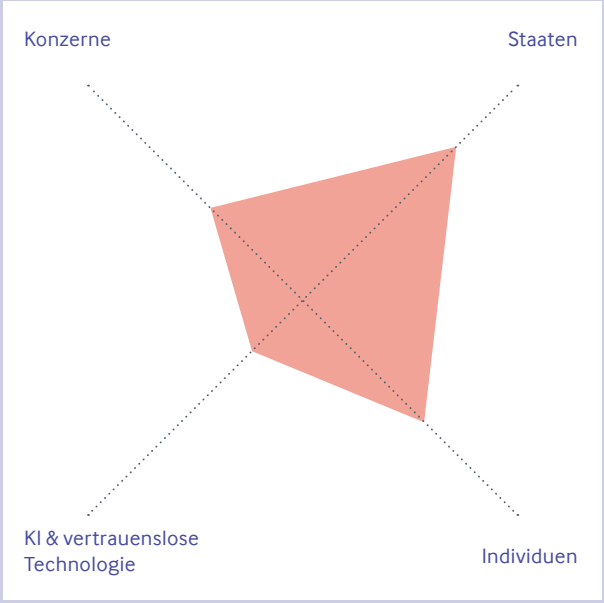
Das EU-Internet: Ein freiheitlich-demokratisches Internet?

Dieses Modell – Hall und O'Haras „Bourgeois Internet" – lässt sich am besten mit „der Staat schlägt zurück" beschreiben. Es könnte auch das sein, was aus Sicht von Regierungen auf der ganzen Welt einem liberal-demokratisches Netz am nächsten kommt. Die von der EU und ihren Mitgliedsstaaten ausgehende digitale Regulierung ist reaktiv und versucht, wahrgenommene Schäden und Bedrohungen abzustellen, die durch das unternehmerische Internet entstehen. Obwohl die EU-Regulierung zunehmend in proaktiver Sprache formuliert wird, ist sie in erster Linie eine Abhilfemaßnahme. Die Allgemeine Datenschutzverordnung (GDPR), der Fall Google Spain vs. AEPD (die spanische Datenschutzbehörde), das NetzDG in Deutschland und zuletzt der Digital Services Act (DSA) der EU sind gute Beispiele für staatliche Maßnahmen um das Verhalten oder die Pläne von (hauptsächlich US-) Technologieplattformen im Zaum zu halten.

Es zeichnet sich nun ab, wie dieses bürgerliche Internet aussehen könnte. In dieser Vision ist die Staatsmacht der wichtigste Verteidiger von Recht und Freiheit der Bürger. Von den Bürgern wird erwartet, dass sie ihr Vertrauen in nationale und internationale Institutionen setzen. Die Rolle der Bürger wird gestärkt, mit der Erwartung, dass sie sich im Gegenzug anständig und tolerant verhalten. Das Recht auf Schutz persönlicher Daten wird dahingehend gestärkt, dass die Bürger mehr Kontrolle über die Verwendung und den Wert der von ihnen produzierten Daten bekommen.

Ungezügelte Maschinenmacht stellt eine Bedrohung für die Vormachtstellung des Staates dar. Das zeigt sich im bürgerlichen Internet ebenso wie anderswo. Die EU hat sich an die Spitze derer gestellt, die **ethische Standards für künstliche Intelligenz** fordern, und damit den zunehmenden Einsatz dieser Form der Entscheidungsfindung anerkannt. Zivilgesellschaftliche Organisationen rufen lautstark nach **algorithmischer Transparenz, Regressansprüchen und Vorsicht bei der Implementierung von KI-gestützten Technologien wie Gesichtserkennung**. Während die Einführung vertrauenswürdiger Technologien toleriert wird, werden Gesetze rund um Werbung für und Bereitstellung von Kryptowährungsdiensten

Ein EU Internet



umgesetzt. In der Debatte über datenschutzfreundliche Technologien wie die Ende-zu-Ende-Verschlüsselung ist weiterhin keine Lösung in Sicht. **Es existiert ein Dilemma zwischen Sicherheit und Schutz auf der einen und dem Recht auf Privatsphäre und freie Meinungsäußerung sowie kommerzielle Fragen auf der anderen Seite.**

Das freiheitlich-demokratische Internet

Staaten, Unternehmen, Individuen und Maschinen Macht zu geben, stellt sowohl eine Bedrohung als auch eine Chance für die freiheitlich-demokratische Entwicklung dar. Durch diese Ambivalenz und dieses Dilemma zu navigieren ist kein einfaches Unterfangen, und die Zeit ist knapp. Der Moment, in dem man das Internet als mächtiges Werkzeug zur Projektion liberaler Werte hätte feiern können, ist vorbei. Es war nie unvermeidlich, nie das Ende der Geschichte. Der Umgang mit Sprache und Informationen in einer liberalen demokratischen Gesellschaft ist eine mühsame Übung in langsamer Regulierung, Sorgfalt und Vorsicht. Langsamkeit und Geduld werden in der schnelllebigen Welt der Technologie leicht ausgenutzt.

Die Aufgabe liberaler Demokratien wird es sein, die Technologien, die Gestaltungsprinzipien und die Governance zu identifizieren, die ein Kräftegleichgewicht gewährleisten, das freiheitlich-demokratischen Werten entspricht. Ausmaß und Tiefe der Herausforderung sind gewaltig.

Auf jeder Ebene der Technologie, aus dem das Internet besteht, und auf der Ebene der Rechte und Freiheiten des Einzelnen bis hin zu den großen politischen Fragen wie internationale Sicherheit und nationale Souveränität, gibt es viel zu tun.

In den folgenden Abschnitten wird dieses Dilemma beschrieben und Gefahren und Chancen in drei großen Bereichen identifiziert: Der digitale Bürger, das digitale Gemeinwesen sowie Sicherheit und Souveränität. Für jeden Bereich zeigen wir auf, wo liberale Demokratien ihre Verteidigung und Unterstützung verstärken sollten und wo Bedrohungen durch die Macht von Unternehmen, Staaten, Individuen oder Maschinen besondere Wachsamkeit erfordern.

Digitale Staatsbürgerschaft

Die Verteidigung und Förderung von Rechten und Freiheiten der Bürger im Internet und ihre aktive Teilnahme am Online-Leben ist die grundlegende Herausforderung, vor der liberale Demokratien stehen, wenn sie die Online-Welt neugestalten wollen.

Es ist kaum eine Übertreibung, die demokratische Entmachtung des durchschnittlichen Online-Nutzers als die größte Tragödie des Internets zu bezeichnen. In den meisten westlichen Ländern wurde die aktive Teilnahme der Bürger am politischen und zivilen Leben vollständig unter die Vorrechte von Monopolplattformen und dem Wirtschaftsmodell, das ihrem Design zugrunde liegt, untergeordnet. Der durchschnittliche Internetnutzer hat keine Macht, den digitalen Raum umzugestalten oder zu kultivieren. Er ist durch willkürliche, verwirrende oder inkonsistente Nutzungsbedingungen und deren Durchsetzung von Plattformen eingeschränkt. Diejenigen wählen zu können, die uns regieren, ist ein zentraler Grundsatz der freiheitlichen Demokratie, aber Online-Nutzer haben weder das Recht noch die Möglichkeit, Entscheidungen anzufechten, die nach dem Standard-Plattformmodell von nicht demokratisch legitimierten „höheren" Mächten getroffen werden. „Innerhalb dieses Rahmens", schreibt Giovanni De Gregorio, „führt das Fehlen jeglicher Rechte oder Rechtsmittel der Nutzer dazu, dass Online-Plattformen die gleiche Ermessensfreiheit über ihre Gemeinschaft ausüben wie eine absoluten Macht."[12] Shoshana Zuboff nennt dies „die sozialen Beziehungen einer vormodernen absolutistischen Autorität".[13] Andere haben das Plattformmodell als feudalistisch oder hobbesianisch bezeichnet, als ein System, in dem man seine Rechte im Austausch für Produkte und Dienstleistungen aufgibt.[14] Was auch immer es ist, die aktuelle Situation passt nicht zu unserer Vorstellung von Bürgern in einer Demokratie.

Die Verteidigung der Rechte und Pflichten der Bürger durch die Staaten im Sinne eines traditionellen Gesellschaftsvertrags wurde in liberalen Demokratien durch die Macht der Unternehmen vereitelt und war unter autoritären Regimen nie eine Perspektive. Die COVID-19-Pandemie hat ein Schlaglicht darauf geworfen, welche Gefahren in der digitalen Welt drohen, wenn die staatliche Macht ungehindert ausgeübt wird. KI-gesteuerte Kameras, Datenerfassung und -analyse sowie Gesichtserkennungssoftware sorgen dafür, dass die Bürger sorgfältig überwacht werden und Verstöße gegen ein Gesetz oder eine Richtlinie mit deutlich höherer Wahrscheinlichkeit entdeckt werden.[15]

Das Rechtsstaatsprinzip selbst wurde geschwächt. Seine Durchsetzung wird durch uneinheitliche Leistungsfähigkeit, veraltete Gesetze, eingeschränkten Zugang zu Beweisen und schwache internationale Koordination beeinträchtigt. Darüber hinaus werden die Online-Grenzen akzeptablen Verhaltens durch Geschäftsbedingungen festgelegt, und zwar lange bevor ein Gesetz verabschiedet und durchgesetzt werden kann.

Schließlich stellt der Aufstieg der maschinellen Entscheidungsfindung eine neue Bedrohung für traditionelle Vorstellungen von der Macht der Bürger dar. Schon jetzt hat eine ‚vertrauenslose' Technologie wie die Ende-zu-Ende-Verschlüsselung die Regeln für Menschenrechte neu geschrieben. Sie ist ein Segen für das Recht auf Privatsphäre und bedeutet gleichzeitig ein Risiko für das Recht auf Sicherheit. Hinzu kommt, dass die kryptografische Technologie einer ausgewählten Gruppe von technologisch versierten Personen neue Befugnisse verliehen hat.

Code wird Gesetz. Unsere neuen Gesetzgeber sind erst Ingenieure, dann künstliche Intelligenzen. Die Möglichkeiten der politischen und gesellschaftlichen Teilhabe und die Rechte und Freiheiten der Teilhabenden werden nicht etwa durch menschliche Aufsicht definiert, sondern durch die Technologie selbst. Das wirft die Frage auf, wie Menschen in einer Welt der Maschinen Macht ausüben können.

Technologie sollte stattdessen für die Verteidigung der Rechte der Bürger eingesetzt und in den Dienst ihrer Ermächtigung gestellt werden. Darauf zu achten, dass die Macht der Unternehmen durch Gesetze und Regierungsgewalt kontrolliert wird, ist ein wesentlicher erster Schritt, um sicherzustellen, dass selbige Rechte verteidigt werden und dass die Bürger die Räume, in denen sie online leben, verstehen, beeinflussen und herausfordern können. Gegenwärtig mögen Internet-Giganten das Ziel dieser Bestrebungen sein, aber die

gleichen Fragen müssen in Zukunft auch zur Macht von Maschinen gestellt werden, um sicherzustellen, dass Algorithmen und künstliche Intelligenzen nach Richtlinien arbeiten, die mit freiheitlich-demokratischen Prinzipien vereinbar sind. Die Macht von Online-Gemeinschaften zu stärken ist ein wesentlicher Schritt, und das Internet liefert tausende von Beispielen, wie man dies effektiv tun kann.

Staatliche Macht und das Rechtsstaatsprinzip sollten die Bürger vor unternehmerischem Fehlverhalten schützen und sicherstellen, dass die digitale Gestaltung ihre Rechte, Pflichten und Freiheiten berücksichtigt. Unternehmen, die entsprechend befugt sind, werden konkurrierende Modelle für das Online-Leben entwickeln, die den Bürgern echte Wahlmöglichkeiten bieten, wo ihr Online-Leben stattfinden soll, und liberal-demokratische Technologien zu neuen Zielgruppen auf der ganzen Welt bringen. Bürger, die entsprechend befähigt sind, Verantwortung für ihr Online-Leben zu übernehmen, werden Wege zu einer sinnvollen digitalen Bürgergesellschaft finden und neue Gemeinschaften und Beziehungen zu den Bedingungen ihrer Wahl bilden und kultivieren. ‚Vertrauenslose' Technologie kann eingesetzt werden, um Rechte und Freiheiten dort zu schützen, wo autokratische Staaten und Konzerne ihre Macht missbrauchen. Die Blockchain-Technologie wurde bereits eingesetzt, um das Recht auf Eigentum an Orten zu schützen, an denen dieses Recht nicht garantiert ist.[16, 17] Sorgfältig entworfene künstliche Intelligenzen können die Fähigkeit der Bürger, ein erfülltes und freies Leben zu führen, durchaus erhöhen, indem sie die Entscheidungsfindung, den Zugang zu Informationen und neue Modelle der Arbeit und sozialen Unterstützung verbessern. Der ethische Einsatz von KI wurde häufig als etwas angepriesen, das für liberale Demokratien durchaus von Vorteil sein könnte.[18]

Fallstudie

Gesichtserkennungen

Wenn Bürger nicht über ausreichend Macht verfügen, kann Technologie nur allzu leicht von Staaten und Unternehmen als Waffe gegen Einzelpersonen und Gemeinschaften eingesetzt werden. Es ist bereits bekannt, dass die chinesische Regierung digitale Überwachung einsetzt, um Gräueltaten gegen das uigurische Volk in Xinjiang zu begehen.[19] Kürzlich wurde auch bekannt, dass das Unternehmen Huawei an der Erprobung von KI-Gesichtserkennungstechnologien zur Identifizierung der ethnischen Zugehörigkeit von Menschen beteiligt war, die einen „Uiguren-Alarm" an die Polizei senden konnten, wenn ein Mitglied der Minderheitengruppe identifiziert wurde.[20]

Die Machtausübung von Konzernen, die staatliche Unterdrückung unterstützen, ist ein globales Problem. Einige Unternehmen haben versucht, sich davon zu distanzieren. Die Tatsache, dass Strafverfolgungsbehörden in den USA Gesichtserkennungssysteme eingesetzt haben, von denen bekannt ist, dass sie schwere geschlechts- und rassenspezifische Verzerrungseffekte haben, hat dazu geführt, dass Amazon, Microsoft und IBM den Verkauf ihrer Gesichtserkennungstechnologien an Strafverfolgungsbehörden eingestellt haben.[21, 22]

Der Einsatz verschiedener Gesichtserkennungssysteme durch Strafverfolgungsbehörden und private Unternehmen war Gegenstand von Gerichtsverfahren von South Wales[23] bis Illinois.[24]

Wir befinden uns in einer Situation, in der staatliche Macht – und staatliche Unterdrückung – durch den Einsatz von nicht rechenschaftspflichtigen Technologien in großem Umfang verstärkt werden kann und in der wir uns zur Einschränkung von Missbrauch auf den guten Willen der Unternehmen, oder schwerfällige rechtliche Prozess, verlassen. Eine freiheitlich-demokratische Ordnung kann sich nicht damit begnügen, ständig hinter der rasanten Entwicklung und Einführung von Technologien hinterherzuhinken, und auch nicht damit, dass die demokratische Kontrolle immer nur ein nachträglicher Gedanke ist, wenn überhaupt.

Ein digitales Gemeingut

Wir sind in das 21. Jahrhundert mit einer Reihe von Annahmen darüber eingetreten, was einen „demokratischen" Informations- und Medienraum ausmacht und was eine vollständige, freie und faire öffentliche Debatte unterstützt: Meinungsfreiheit, Pluralismus, die Metapher vom Marktplatz der Ideen. Aber Online-Räume haben diese Ideale häufig nicht erfüllt und stellen eine neue Herausforderung für die Kohärenz dieser Prinzipien dar.

Die gängige Analogie ist die des Einkaufszentrums: Räume, die sich in der Offline-Welt wie öffentliche Räume anfühlen, in denen aber eigene Regeln gelten, die im Privaten aufgestellt und von privaten Sicherheitsdiensten durchgesetzt werden. Die Zentralisierung und Homogenisierung des digitalen öffentlichen Raums durch eine Handvoll US-Firmen hat dazu geführt, dass die Ausgestaltung, die kulturellen Normen und die Form des öffentlichen Diskurses, der durch diese Medien ermöglicht wird, in den Händen der Unternehmensmacht geblieben sind. Der Gestaltungsimperativ hinter diesen Räumen ist klar. Die Maximierung der Shareholder-Values erfordert ein Panoptikum gesammelter Daten und die Priorisierung aufmerksamkeitsstarker Inhalte. Diese Räume werden streng überwacht, um diejenigen zu schützen, die mächtig, wohlhabend oder gewieft genug sind, um ihrer Stimme Gehör zu verleihen. Die öffentlich-rechtlichen Medien werden zunehmend zu einem Modell, das auf Wohlwollen angewiesen ist.

Maschinen spielen eine wesentliche Doppelrolle bei der Aufrechterhaltung dieser Kontrolle. Sie dienen der Aufrechterhaltung der riesigen „privaten Allmende", die von Social-Media-Plattformen repräsentiert wird. Komplizierte Algorithmen priorisieren Inhalte aus Profitgründen, schwerfällige Algorithmen zensieren Sprache und Informationen, Personalisierungsalgorithmen segmentieren und liefern maßgeschneiderte Informationen bis zu dem Punkt, an dem zwei Bürger in völlig unterschiedlichen Realitäten leben könnten.

Die demokratischen Staaten bleiben in der Defensive. Als Geldgeber für Plattformwerbung haben die Staaten einen Weg gefunden, das Beste aus dieser neuen Weltordnung zu machen, ohne ihre Prinzipien ernsthaft in Frage zu stellen. Und die Versuche, die Prinzipien des öffentlichen Raums zu wahren, beschränken sich auf reaktive Regelungen, die auf „Online Harms" abzielen. Lösungen sind nicht einfach, und gut gemeinte Versuche der digitalen Regulierung bedienen sich häufig dem autoritären Vokabular von Take-Downs, Sperren, Verboten und Zensur.

Autoritäre Staaten hingegen haben sich die Digitalisierung des öffentlichen Raums zunutze gemacht, indem sie entweder auf den Überwachungszug aufgesprungen sind oder dessen Schwächen innerhalb und außerhalb ihrer Grenzen ausgenutzt haben.

Bei der Gestaltung des öffentlichen Raumes ist die Macht des Einzelnen stark eingeschränkt, wenn die Welt ein einziges großes Einkaufszentrum ist. Die Schaffung oder Aufrechterhaltung von öffentlichem Raum im Internet ist eine undankbare Aufgabe für diejenigen, die nicht in der Lage sind, ihn zu Geld zu machen. Die wenigen Menschen, die in der Lage sind, als Kommentator, Journalist, öffentliche Figur oder Talking Head dauerhaft online präsent zu sein, tun dies sozusagen als Pächter durch Premium Snapchats, auf Amazons Twitch oder Google's YouTube.

Die Neugestaltung der öffentlichen Sphäre muss in Demokratien Priorität haben, und ein liberales demokratisches Internet erfordert Veränderungen in allen vier Ecken der obigen Skizzen.

Die unternehmerische Bereitstellung von öffentlichem Raum, die üblicherweise unter dem Motto der Vernetzung der Welt ausgeweitet wird, könnte als mächtige Arena für die Projektion demokratischer Werte fungieren. Aber neue Modelle zum Erhalt des öffentlichen Raums und der Stimmen darin sind unerlässlich. Eine Regulierung zugunsten alternativer Modelle öffentlicher Medien und die Wiederherstellung und Erhaltung von Finanzierungsmodellen außerhalb von Werbeeinnahmen sind wichtige Wege, um sicherzustellen, dass die Medien pluralistisch, verantwortungsvoll und nachhaltig sind.

Unkontrollierbare und undurchsichtige Maschinenmacht kann nicht mit der Governance der digitalen Allmende betraut werden. Wo der Raum notwendigerweise von Algorithmen aufrechterhalten wird, sollte durch die Änderung ihres Designs und die Erhöhung ihrer Transparenz eine durch Technologie ermöglichte öffentliche Sphäre geschaffen werden, in der Maschinen zur Verteidigung der Stimmen von Minderheiten und zur Bewahrung freier und offener Medien eingesetzt werden. Richtig eingesetzt und verwaltet, können Technologien wie Verschlüsselung und dezentrale Netzwerke als weitere Verteidigungslinie der öffentlichen Sphäre gegenüber denjenigen dienen, die sie überwacht, kontrolliert oder abgeschaltet sehen wollen.

Ermächtigte Bürger sind Hüter und Teilhaber der digitalen Allmende. Mit den richtigen Anreizen wird ein freiheitlich-demokratisches Internet die Transformation seiner Nutzer von digitalen Leibeigenen zu digitalen Bürgern erleben, die in der Lage sind, zu einer gesunden und lebendigen öffentliche Sphäre beizutragen und sie zu gestalten.

Fallstudie

Das Section 230-Rätsel

Eine der offensichtlichen Merkwürdigkeiten der US-Wahl 2020 war, wie Präsident Trump und Präsidentschaftskandidat Biden trotz ihrer unterschiedlichen Positionen zur Online-Welt zum gleichen Schluss kommen konnten, nämlich dass der Paragraph Section 230 reformiert werden müsse. Dieser entbindet Internetunternehmen von der Haftung für Inhalte, die auf ihren Plattformen gehostet werden.[25] Trumps langjähriger (unbewiesener) Vorwurf an die großen Tech-Unternehmen war, sie würden einseitig konservative und rechtsgerichtete Stimmen „zensieren“, wenn sie gegen Hassrede und Extremismus auf ihren Plattformen vorgehen.[26] Biden wiederum sagt, die Verbreitung von Fehlinformationen und Desinformationen im Internet seien ein Grund, die Schutzmaßnahmen zu überdenken.[27]

Klar ist: Ohne eine gemeinsame Vorstellung davon, was ein guter öffentlicher Online-Raum sein soll (keine Fehlinformationen? keine Zensur?), werden die Ansätze zur Beseitigung des Machtungleichgewichts in öffentlichen Online-Räumen Stückwerk bleiben. Und die Tatsache, dass diese Räume im Besitz privater Unternehmen sind, die ohne Aufsicht oder nennenswerte Transparenz agieren, bedeutet, dass diese Art von widersprüchlichen Schlussfolgerungen wahrscheinlich ist. Regierungen kämpfen mit all ihren zur Verfügung stehenden Mitteln darum, die Macht von den Unternehmen zurückzuerlangen. Ob dies letztendlich die Bürger ermächtigt ist ungewiss.

Sicherheit und Geopolitik

In den letzten fünf Jahren ging der kalte Krieg im Internet in die heiße Phase. Dem Kampf um digitale Souveränität gingen innenpolitische Entwicklungen voraus. Nationalstaaten wie Russland und China drängten auf mehr Kontrolle über das Internet innerhalb ihrer Grenzen. Das Internet wurde auch zu einem Vektor für internationale geopolitische Ziele – sowohl durch die Bewaffnung offener Online-Räume und die Verbreitung von Desinformation als auch durch das Wettrüsten digitaler Infrastruktur auf der ganzen Welt. Parallel zu diesen groß angelegten Plänen stellt Cyberkriminalität die am schnellsten wachsende Bedrohung für Bürger dar. Sie reicht von Betrug und Identitätsdiebstahl bis hin zur Rekrutierung von Extremisten und der sexuellen Ausbeutung von Kindern im Internet. Die liberalen Demokratien haben zu langsam auf diese Bedrohungen reagiert.

Der Einfluss der Unternehmen, verkörpert durch die Policy- und Resilience-Teams der großen Plattformen, wurde aufgedeckt. Die Plattformen hatten geschlafen. Entweder waren sie sich nicht bewusst, wie ihre Plattformen ausgenutzt wurden, konnten dem nicht entgegenwirken, oder sie hatten sich entschieden, es zu ignorieren.

Einzelpersonen wurden zu Kanonenfutter degradiert. Von Plattformen, die auf Vernetzung und Wachstum aus sind, in die vorderste Reihe gedrängt und ohne digitale Kompetenz, waren und sind sie eine leichte Beute für Gruppen und Einzelpersonen, die sie ausnutzen wollten. Bildungsinitiativen und Faktenchecker bleiben ein schwaches Mittel zur Selbstverteidigung. Es wird angenommen, dass jeder dritte Amerikaner von Betrug und Cyberkriminalität betroffen ist.[28]

Die Fähigkeit des Staates, seine Bürger zu schützen, wurde immer wieder dort in Frage gestellt, wo sich unser Leben online abspielt, mit verschlüsselten Geräten und Kommunikationsplattformen sowie weiteren Hürden für die Strafverfolgung, die mit der Bekämpfung von digital bedingten Schäden beauftragt ist. Wie bereits erwähnt, schränkt eine Technologie, die sich einer zentralen Kontrolle und Aufsicht entzieht, zwangsläufig die Macht einer zentralen staatlichen oder unternehmerischen Autorität ein.

Nationalstaaten, die sich im Kampf gegen den Einfluss von Plattformen auf reaktive Regulierung verlassen haben, waren ebenfalls langsam und machtlos, wenn es darum ging, die neue Informationslandschaft und ihre willkürlichen Hüter aus dem Silicon Valley gegen ausländische Akteure zu verteidigen. Eine lasche Herangehensweise an die infrastrukturelle Entwicklung hat dazu geführt, dass Länder auf Infrastrukturimporte von autoritären Regimen in ihrem eigenen Hinterhof angewiesen sind. Es gibt einen Mangel an wettbewerbsfähigen Infrastrukturangeboten auf dem internationalen Markt, beispielsweise im Vergleich zu dem Ausmaß und den Ambitionen der chinesischen Belt and Road Initiative. Wir stehen zwar auf der Schwelle zum Zeitalter des Internets der Dinge, aber Fragen zu den Sicherheitsimplikationen der Geräte, die an Millionen verkauft werden, bleiben weiterhin offen.

Die internationale Cyber-Vorherrschaft wird zu einem großen Teil von der Macht der Maschinen bestimmt werden. In den Händen von Staaten und Unternehmen bedeutet dies die Entwicklung künstlicher Intelligenz. Wie in Four Internets erwähnt, könnte die Fähigkeit autoritärer Regime, datenschutzrechtliche Bedenken auszuräumen und riesige, vernetzte Datensätze anzuhäufen, mit denen KI trainiert werden kann, ihnen einen Vorteil bei der Entwicklung überlegener Produkte verschaffen. In chinesischem Besitz befindliche Apps wie TikTok finden bereits ein westliches Publikum, während Anwendungen aus dem Silicon Valley innerhalb der chinesischen Grenzen verboten oder limitiert werden.

Demokratien müssen eine **liberale Doktrin der Sicherheit und Souveränität** definieren, die Bedrohungen durch Informationsoperationen und Cyberangriffe, sowohl aus dem Ausland als auch aus dem Inland, sowie Online-Kriminalität als solche

erkennt, aber auch Freizügigkeit und den freien Fluss von Informationen über Grenzen hinweg garantiert.

Staaten und multilaterale Institutionen zu befähigen, ein offenes Internet zu sichern und zu verteidigen, ist ein wichtiger Schritt, um ihre Souveränität in der Online-Welt wieder zu behaupten. Dies erfordert eine Veränderung und Verbesserung der Netzwerkarchitektur des Internets, sowohl um das offene Internet gegenüber Protokollen zu stärken, die darauf abzielen, es zu fragmentieren, als auch um sicherzustellen, dass sich freiheitlich-demokratische Prinzipien weiterhin in der zugrundeliegenden Technologie widerspiegeln. Die Transparenz der Gremien für digitale Standards und die Beteiligung multilateraler Institutionen müssen verbessert werden. Wo Unternehmensmonopole als Schwachstelle für die nationale und internationale Sicherheit identifiziert werden, müssen diese Schwachstellen angegangen werden, um sicherzustellen, dass globale Konzerne eine Vorhut der freiheitlich-demokratischen Werte sind, anstatt sie zu untergraben.

Außerdem müssen die Staaten autoritäre Begriffe wie Take-Downs, Sperren, Verbote und Zensur hinter sich lassen und aufhören, eifersüchtig über den Zaun auf die scheinbaren Erfolge zu blicken, die autoritäre Regime bei der Unterdrückung von missliebiger Online-Sprache und entsprechendem Verhalten haben. Stattdessen muss ein freiheitlich-demokratischer Ansatz für die Polizeiarbeit und die Online-Sicherheit eingeführt werden, der dafür sorgt, dass die Sicherheitsdienste in der Lage sind, die Bürger zu schützen, und dies auf eine Weise tun, die verhältnismäßig ist und Augenmaß zeigt. Die Stärkung der nationalen Sicherheit hängt von einer mündigen, gebildeten Zivilgesellschaft ab, die von Sicherheitsdiensten geschützt wird, die im Rahmen der Rechtsstaatlichkeit arbeiten, aber dennoch handlungsfähig sind.

Demokratisches, von Staaten und Unternehmen gefördertes Infrastrukturwachstum ist von entscheidender Bedeutung. Dazu gehören auch der Export und die Förderung von Infrastruktur, die das offene Internet rund um den Globus stärkt. Es ist inakzeptabel, die Standards und den Ausbau der digitalen Infrastruktur an kompromittierte Anbieter und autoritäre Regime zu übergeben.

Case Study

Internet-Abschaltungen

Regierungen auf der ganzen Welt betrachten es als ihr Hoheitsrecht, gegen das offene Internet vorzugehen: im Extremfall durch Internetabschaltungen, mehr oder weniger ernsthaft zum Schutz der nationalen Sicherheit, von Recht und Ordnung oder zur Verhinderung von „Online Harms“.[29, 30] Abschaltungen werden von den Betroffenen jedoch als „Kollektivstrafe“[31] aufgefasst und beeinträchtigen darüber hinaus nicht nur die grundlegende Informations- und Meinungsfreiheit, sondern haben auch erhebliche negative wirtschaftliche und gesundheitliche Auswirkungen.[32, 33] Die monatelangen Internetbeschränkungen in Myanmar wurden vor allem im Jahr 2020 kritisiert, weil sie den Zugang zu wichtigen Informationen über die Covid-19-Pandemie blockierten.[34]

Solange es keine klaren globalen Standards und Verbindlichkeiten bezüglich dessen gibt, was ein Internetzugang sein sollte und wann Beschränkungen legitim oder illegitim sind, führt der plattformübergreifende Umgang der Staaten mit Problemen dazu, dass die Rechte der Bürger eher ausgehöhlt als geschützt und dass ihre Möglichkeiten, sich darüber zu äußern, beschnitten werden.

Schlussfolgerungen

Die liberale Demokratie als Balanceakt ist ein Klischee, aber hier sind wir wieder. Vier Kräfte sind für die Gestaltung des zukünftigen Internets verantwortlich: Die Macht des Staates, die Macht der Konzerne, die Macht des Einzelnen und die Macht der Maschinen. Sie alle müssen im Namen der liberalen Demokratie im Zaum gehalten werden. Diese Kräfte im Gleichgewicht zu halten, ist die Herausforderung für liberale Demokratien. Die hier vorgestellten Beispiele zeigen, dass eine zu starke Ausrichtung in eine Richtung das Projekt als Ganzes untergräbt und dass eine Politik, die die Bedeutung einer dieser Mächte ignoriert, unzureichend ist.

Wohin man auch schaut, sieht man Beweise für dieses Versagen. Die durch das Internet in seinen aktuellen Ausprägungen verursachten Schäden und Misserfolge sind gut dokumentiert. Wir sprechen weniger oft über Internet-Erfolgsgeschichten, aber auch das ist eine Frage von Beweisen. Auf dem Weg zur proaktiven Vision eines freiheitlich-demokratischen Internets müssen wir die vielen brillianten Designer und Architekten im Internet unterstützen und sicherstellen, dass sich alle Teile des Internets an die Standards seiner Erfolgsgeschichten halten. Man kann Lektionen von Wikipedia, der offenen Coding Community StackOverflow und von den Legionen virtueller Gemeinschaften lernen, die Erfolg haben, obwohl sie kaum Schlagzeilen machen.

Andere Lektionen kann man von den Internet-Giganten lernen. Sie sind in den letzten Jahren zu Recht wegen ihrer Versäumnisse in die Kritik geraten, haben aber mehr als alle anderen dazu beigetragen, das Internet für die Welt zu öffnen. Wenn sie sich für die liberale Demokratie einsetzen, werden sie vielleicht wieder als Vorreiter freiheitlich-demokratischer Werte in der Welt wahrgenommen.

Die dringlichsten Fragen müssen auf dem Gebiet der staatlichen und individuellen Macht beantwortet werden. **Das Internet wird der Ort sein, an dem die Demokratie im 21. Jahrhundert neu definiert wird, aber dies erfordert eine radikale Verbesserung der staatlichen und multilateralen Kontrolle der Online-Welt und der ihr zugrundeliegenden Technologie. Sicherzustellen, dass Individuen auch online von ihren Rechten Gebrauch machen können, ist ein wichtiger Schutz vor staatlicher und unternehmerischer Übervorteilung.**

In der Zukunft wird es nicht mehr nur das Trilemma von Staaten, Individuen und Privatwirtschaft geben. Die sich beschleunigende Entwicklung der Maschinenmacht, von der künstlichen Intelligenz bis zur erlaubnisfreien Technologie, wird dazu führen, dass die Maschinenmacht eine Herausforderung für alle darstellt. Angesichts des Tempos, in dem sich alles, was mit Global Governance zu tun hat, weiterentwickelt, ist es von entscheidender Bedeutung, dass die getroffenen Maßnahmen den wachsenden Einfluss von Maschinen in unserem sozialen, wirtschaftlichen und politischen Leben widerspiegeln.

Mehr denn je brauchen wir jetzt eine Möglichkeit, die liberalen Demokratien hinsichtlich der Förderung und Befürwortung ihrer eigenen Vision des Internets zu vereinen. Während autoritäre Mächte ihre Vision immer einheitlicher vertreten, sind die Demokratien derzeit uneins. Es bestehen grundlegende Unterschiede in der Herangehensweise zwischen Nordamerika, Europa und Asien. Es gibt aber auch Grundwerte und Interessen, die uns vereinen und die artikuliert werden müssen.

Ohne Beweise dafür, was online funktioniert, und ohne eine prinzipientreue Vision für das Internet laufen unsere demokratischen Traditionen und Werte, unsere Regierungen und Gesellschaften Gefahr, im Wettlauf um die Neugestaltung des wichtigsten internationalen politischen, kulturellen und sozialen Raums hinter autoritären Staaten, Industriegiganten und mächtigen Technologien zurückzubleiben. Wir dürfen nicht den Fehler machen, uns ausschließlich auf das zu konzentrieren, was wir nicht wollen, und darüber das zu vergessen, was wir wollen.

Endnoten

1. R. Larsen, The Political Nature of TCP/IP, Science, Technology & Society Program, University of Pennsylvania; https://repository.upenn.edu/cgi/viewcontent.cgi?article=1004&context=momentum (p.27)
2. Zum Beispiel der Kampf zwischen TCP und dem Open Systems Interconnection-Protokoll.
3. Larsen (p.47)
4. K O'Hara & W Hall, Four Internets: The Geopolitics of Digital Governance, CIGI Paper No. 206, Dezember 2018; https://www.cigionline.org/publications/four-internets-geopolitics-digital-governance
5. Technologische Singularität beschreibt einen hypothetischen Zeitpunkt. „an dem künstliche Intelligenz (KI) die menschliche Intelligenz übertrifft und sich dadurch rasant selbst verbessern und neue Erfindungen machen würde, wodurch der technische Fortschritt irreversibel und derart beschleunigt würde, dass die Zukunft der Menschheit nach diesem Ereignis nicht mehr vorhersehbar wäre." Siehe https://de.wikipedia.org/wiki/Technologische_Singularit%C3%A4t
6. Vint Cerf selbst gab zu, dass das Einzige, was eine frühere Einführung von Verschlüsselungsstandards im Internet verhinderte, die militärische Sicherheitseinstufung der erforderlichen Technologie war; https://www.youtube.com/watch?v=17GtmwyvmWE&feature=share&t=23m1s
7. In Ländern, in denen die Rechtsstaatlichkeit selbst kompromittiert oder fremd ist, ist es einfach die Fähigkeit des Staates, die ultimative Macht auszuüben.
8. The New York Times, How China Uses High-Tech Surveillance to Subdue Minorities, 22nd Mai 2019; https://www.nytimes.com/2019/05/22/world/asia/china-surveillance-xinjiang.html
9. S Dickson, China's New Cryptography Law: Still No Place to Hide, Harris Bricken, 7th November 2019; https://www.chinalawblog.com/2019/11/chinas-new-cryptography-law-still-no-place-to-hide.html
10. J L Zittrain, The Future of the Internet -- And How to Stop It, Yale University Press & Penguin UK, 2008; https://dash.harvard.edu/bitstream/handle/1/4455262/Zittrain_Future+of+the+Internet.pdf?sequence=1
11. Eine nützliche Liste von Beispiel-DAOs finden Sie unter https://hackernoon.com/what-is-a-dao-c7e84aa1bd69
12. G. De Gregorio, Democratising Online Content Moderation: A Constitutional Framework, Computer Law and Security Review, April 2020;
13. S. Zuboff, Big Other: Surveillance Capitalism and the Prospects of an Information Civilization (2015) p.83
14. B. Schneier, Data and Goliath (2015) p.58
15. Reuters, Coronavirus brings China's surveillance state out of the shadows, 7th Februar 2020; https://www.reuters.com/article/us-china-health-surveillance-idUSKBN2011HO
16. D Daniel & C I Speranza, The Role of Blockchain in Documenting Land Users' Rights: The Canonical Case of Farmers in the Vernacular Land Market, Frontiers in Blockchain, Mai 2020; https://www.frontiersin.org/articles/10.3389/fbloc.2020.00019/full
17. L Tombs, Could blockchain be the future of the property market?, HM Land Registry Blog, Mai 2019; https://hmlandregistry.blog.gov.uk/2019/05/24/could-blockchain-be-the-future-of-the-property-market/
18. European Commission, Ethics guidelines for trustworthy AI, April 2019; https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai
19. The Guardian, China's hi-tech war on its Muslim minority, 11th April 2019; https://www.theguardian.com/news/2019/apr/11/china-hi-tech-war-on-muslim-minority-xinjiang-uighurs-surveillance-face-recognition
20. The Washington Post, Huawei worked on several surveillance systems promoted to identify ethnicity, documents show, 12th Dezember 2020; https://www.washingtonpost.com/technology/2020/12/12/huawei-uighurs-identify/
21. The Atlantic, Defund Facial Recognition: I'm a second-generation Black activist, and I'm tired of being spied on by the police, 5th Juli 2020; https://www.theatlantic.com/technology/archive/2020/07/defund-facial-recognition/613771/
22. The Verge, Amazon verbietet der Polizei die Nutzung seiner Gesichtserkennungstechnologie für das nächste Jahr, 10th July 2020; https://www.theverge.com/2020/6/10/21287101/amazon-rekognition-facial-recognition-police-ban-one-year-ai-racial-bias
23. BBC News, Facial recognition use by South Wales Police ruled unlawful, 11th August 2020; https://www.bbc.co.uk/news/uk-wales-53734716
24. The Verge, ACLU sues facial recognition firm Clearview AI, calling it a 'nightmare scenario' for privacy, 28th Mai 2020; https://www.theverge.com/2020/5/28/21273388/aclu-clearview-ai-lawsuit-facial-recognition-database-illinois-biometric-laws

Endnoten

25. Council on Foreign Relations, Trump and Section 230: What to Know, 2nd Dezember 2020; https://www.cfr.org/in-brief/trump-and-section-230-what-know

26. CNN, Trump says right-wing voices are being censored. The data says something else, 28th Mai 2020; https://edition.cnn.com/2020/05/28/media/trump-social-media-conservative-censorship/index.html

27. The Verge, Joe Biden wants to revoke Section 230, 17th January 2020; https://www.theverge.com/2020/1/17/21070403/joe-biden-president-election-section-230-communications-decency-act-revoke

28. Demos, The Great Cyber Surrender: How police and governments abandon cybercrime victims, November 2020; https://demos.co.uk/project/the-great-cyber-surrender-how-police-and-governments-abandon-cybercrime-victims/

29. Chatham House, Asia's Internet Shutdowns Threaten the Right to Digital Access, Februar 2020; https://www.chathamhouse.org/2020/02/asias-internet-shutdowns-threaten-right-digital-access

30. DW, India's internet shutdowns function like 'invisibility cloaks', November 2020; https://www.dw.com/en/indias-internet-shutdowns-function-like-invisibility-cloaks/a-55572554

31. Ibid; https://www.dw.com/en/indias-internet-shutdowns-function-like-invisibility-cloaks/a-55572554

32. Human Rights Watch, Shutting Down the Internet to Shut Up Critics, 2020; https://www.hrw.org/world-report/2020/country-chapters/global-5

33. Ibid; https://www.dw.com/en/indias-internet-shutdowns-function-like-invisibility-cloaks/a-55572554

34. The Wire, Human Rights Groups Criticise 'World's Longest Internet Shutdown' in Myanmar, Juni 2020; https://thewire.in/south-asia/myanmar-worlds-longest-internet-shutdown

ISD | Powering solutions to extremism and polarisation

Digital Policy Lab
Diskussionspapier

Überlegungen zur Zukunft der Online-Regulierung

About This Briefing

Dieses Diskussionspapier skizziert mehrere Schlüsselaspekte einer künftigen Digitalpolitik, die in den kommenden Jahren weiter erforscht, diskutiert und konsensfähig gemacht werden müssen. Es wird gezeigt, in welchen Bereichen es noch an Klarheit mangelt, selbst nachdem die EU und Großbritannien Ende 2020 wesentliche Vorschläge von politischer Tragweite gemacht haben. Das Papier untersucht auch die Spannungen und Kompromisse, die sich aus neuen Regulierungen ergeben könnten, da die Internet-Governance zunehmend in die Zuständigkeit von Staaten und deren Regulierungsbehörden fällt.

ISD | Powering solutions to extremism and polarisation

Beirut | Berlin | London | Paris | Washington DC

www.isdglobal.org

Einführung

Wie bereits in dem begleitenden Diskussionspapier *Nationale und internationale Modelle zur Online-Regulierung* dargelegt, haben sich die Haftungsregelungen für Online-Plattformen („Intermediäre") seit den 1990er- und frühen 2000er-Jahren in vielen Rechtsräumen, einschließlich der USA und der EU, kaum verändert. Die bestehenden Regelungen basierten auf der Annahme, die Redefreiheit und das Wachstum der digitalen Wirtschaft seien gefährdet, wenn Plattformen direkt für nutzergenerierte Inhalte verantwortlich gemacht würden. Jedoch haben viele Unternehmen und insbesondere soziale Medien Geschäftsmodelle geschaffen, die in der Öffentlichkeit für Verrohung, Verzerrung oder Spaltung sorgen.

Diese Geschäftsmodelle diktieren immer stärker den Erfolg oder Misserfolg bestimmter Inhalte, und bevorzugen sensationslüsterne oder kontroverse Beiträge, um Nutzer-Engagement und im Gegenzug Werbeeinnahmen zu maximieren. Im Laufe der Zeit hat diese Dynamik zu einer Zunahme von Hass, Extremismus, Terrorismus und Desinformation im Internet beigetragen, die der Gesellschaft und der Demokratie beträchtlichen Schaden zufügt. Regierungen haben Technologieunternehmen dahingehend unter Druck gesetzt, dass sie mehr Verantwortung für schädliche Aktivitäten auf ihren Plattformen übernehmen sollen und zur Rechenschaft gezogen werden können. Die informellen, freiwilligen oder branchengeführten Reaktionen darauf waren jedoch häufig reaktiv oder themenspezifisch und in vielen Fällen unzureichend. Trotz Verbesserungen sind die übergreifenden Strukturen und Prozesse, die schädliche Inhalte fördern, verstärken oder empfehlen, beziehungsweise schädliche Aktivitäten ermöglichen, in der Regel undurchsichtig geblieben.

Um diese durch Online-Plattformen ermöglichten negativen Effekte zu bekämpfen, nehmen Regierungen immer mehr Abstand von informellen, freiwilligen oder branchengeführten Selbstregulierungsbemühungen. Stattdessen wenden sie sich der Gesetzgebung zu. In unserem vorherigen Papier haben wir die neuen Vorschriften grob in zwei Kategorien eingeteilt:

- **Inhaltsbezogene Ansätze** zielen oft auf bestimmte „Online Harms"[1] wie Hassrede oder Desinformation bei Wahlen ab und konzentrieren sich auf die wirksame und rechtzeitige Entfernung dieser Inhalte, wo es angemessen scheint (oft als „Notice-and-Takedown-Modell" bezeichnet).

- **Bei systemischen Ansätzen** müssen Online-Plattformen nachweisen, dass ihre Richtlinien, Prozesse und Systeme so konzipiert und umgesetzt werden, sodass sie potenziell negative Folgen möglicher „Online Harms" berücksichtigen.

Die Tatsache, dass bisher wenige Fortschritte bei der Eindämmung von „Online Harms" gemacht wurden, hat zu einer wachsenden Unterstützung für systemische Ansätze geführt. Diese zielen auf die Entwicklung eines einheitlichen Aufsichtsregimes ab, mit dem illegale und in einigen Fällen legale, aber schädliche Aktivitäten angegangen werden können. Das Ziel ist ein Einwirken auf die internen Prozesse der marktbeherrschenden Technologieunternehmen, die eine wichtige Rolle bei der Verschärfung der Probleme spielen können.

Im Dezember 2020 präsentierte die Europäische Kommission ihren Entwurf zum Digital Services Act (DSA). Hierbei handelt es sich um ein Regulierungssystem für soziale Medien, Online-Marktplätze und andere in der EU tätige Online-Plattformen. Der DSA regt neue Verpflichtungen für digitale Dienste an, die als Vermittler agieren, mit dem Ziel, „dass die Menschen weniger mit illegalen Aktivitäten und gefährlichen Gütern zu tun bekommen, und [...] den Schutz der Grundrechte [zu] gewährleisten, sodass ein sichereres Online-Umfeld entsteht, damit die Bürgerinnen und Bürger frei ihre Ideen äußern, miteinander kommunizieren und online einkaufen können".[2] Als solcher aktualisiert der DSA die seit 2000 geltende E-Commerce-Richtlinie und signalisiert eine **Verlagerung hin zu einem präventiven (statt reaktiven) Modell**. Darüber hinaus sind diese neuen Regeln laut Europäischer Kommission „ein wichtiger Schritt zur Verteidigung europäischen Werte im Online-Raum" und sollen „neue Maßstäbe, an denen sich die Regulierungsansätze für Online-Vermittler auch auf weltweiter Ebene messen lassen" setzen.[3]

Ebenfalls im Dezember 2020 veröffentlichte die britische Regierung in einem umfassenden

Konsultationsbericht ihre neuesten Vorschläge zur Bekämpfung von „Online Harms“. Unter diesen Begriff fallen nach Auffassung der britischen Regierung Hassrede, Extremismus und Terrorismus sowie Belästigung, Desinformation und Kindesgefährdung. Die Vorschläge sollen eine Grundlage für die kommende „Online Safety Bill“ bilden. Das Gesetz soll Großbritannien zum „sichersten Ort der Welt [..] machen, um online zu gehen, und zum besten Ort, um ein digitales Unternehmen aufzubauen und zu gründen“. Dies soll durch eine neue gesetzliche „Sorgfaltspflicht“ erreicht werden, die von einer unabhängigen Regulierungsbehörde durchgesetzt wird (gegenwärtig ist dafür die bestehende Regulierungsbehörde für Rundfunk und Telekommunikation, Ofcom, vorgesehen).[4] Diese Vorschläge sind in der begleitenden **Zusammenfassung des EU Digital Services Act & UK Online Safety Bill** genauer beschrieben.

Weder die Vorschläge der EU noch die Großbritanniens sind bisher endgültig und werden in den kommenden Jahren von Interessengruppen aus Politik, Wirtschaft und Zivilgesellschaft noch heftig diskutiert. Daher werden beide Gesetzentwürfe vor ihrer endgültigen Verabschiedung Änderungen erfahren, die aufgrund der rasanten Entwicklung des Online-Ökosystems wahrscheinlich fortlaufend vorgenommen werden müssen.

Globale Ambitionen

Die Europäische Kommission und die britische Regierung bekunden ausdrücklich einen globalen Anspruch, die Richtung der digitalen Politik über ihre Grenzen hinaus mitzubestimmen und im Namen ihrer Bürger künftige Präzedenzfälle für die Regulierung von Schlüsselbereichen des öffentlichen Internets zu schaffen. Im Kern versuchen beide, das derzeitige (Un-)Gleichgewicht der Macht demokratisch gewählter Regierungen und Institutionen auf der einen und privater Unternehmen auf der anderen Seite zugunsten demokratisch legitimierter Macht zu verändern und Unternehmen gleichzeitig zu veranlassen, bei der Gestaltung ihrer Produkte und Dienstleistungen mehr Sorgfalt walten zu lassen.

Diese Ambitionen deuten auf einen anhaltenden geopolitischen Wettbewerb hin, bei dem es darum geht, die Zukunft der Internet-Governance zu definieren, einschließlich der Machtbalance zwischen Regierungen (sowohl demokratischen als auch autoritären), privaten Unternehmen, Bürgern und der Technologie selbst. Wir untersuchen diese Themen in dem begleitenden Papier „**Das freiheitlich-demokratische Internet – Fünf Modelle für eine digitale Zukunft**“, und weisen dort darauf hin, wie dringend eine gemeinsame Definition für eine prinzipienfeste Vision des Internets ist. Diese Vision muss wichtige demokratische Werte in der Infrastruktur und den öffentlichen Räumen der Online-Welt verankern, um mit aufkommenden autoritären Modellen der Internet-Governance effektiv konkurrieren zu können.

In Anbetracht der außerordentlichen Bandbreite politischer Systeme und kultureller Traditionen sollte man nicht erwarten, dass Gesetzgebung und Regulierung einheitlich sind. Es besteht jedoch das Bedürfnis, die gemeinsamen Werte und Interessen freiheitlich-demokratischer Länder kollektiv zu bekräftigen und sie besser auf die Online-Welt anzuwenden. Es ist daher von entscheidender Bedeutung, über eine US- und eurozentrische Sichtweise auf „Online Harms“ und aus ihnen folgende Reaktionen hinauszugehen und zu einem wirklich globalen, aber anpassungsfähigen, **freiheitlich-demokratischen Modell für die Internet-Governance** zu gelangen.

Blick in die Zukunft: Wichtige Fragen, die sich aus den Gesetzesvorschlägen ergeben

Dieses Papier untersucht einige der wichtigsten Streitpunkte, die sich aus den aktuellen Bemühungen um eine kontextübergreifende Regulierung ergeben. Zu den aktuellen Diskussionspunkten gehören:

- Vom Inhalt zum System: Haftung & „Safety by Design“
- Rechtliche Zuständigkeit: Wer legt die Regeln fest und setzt sie durch?
- Der Umgang mit „Legal Harms“
- Regulierungsumfang: Wer wird reguliert?

Ohne die derzeitigen Bemühungen zu kritisieren oder zu implizieren, dass es eine Reihe von einfachen Lösungen gibt, untersuchen die folgenden Abschnitte die der Debatte inhärenten Spannungen und Kompromisse. Sie zeigen auch potenzielle Probleme auf, die in Zukunft an Bedeutung gewinnen könnten, wenn Regierungen anfangen die Online-Plattformen stärker zu regulieren. Wir hoffen, dass dies die Grundlage für zukünftige Diskussionen innerhalb des DPL-Netzwerks bildet und damit zur laufenden digitalpolitischen Debatte beitragen kann.

Vom Inhalt zum System: Haftung & „Safety by Design“

Vorschläge wie die von Großbritannien und der EU bewegen sich zunehmend weg von inhaltsbasierten „Notice-and-Takedown“-Ansätzen hin zu systemischen Modellen der Regulierung. Ersteres hat zu einigen Verbesserungen bei der Entfernung illegaler (z.B. terroristischer) Inhalte geführt und Unternehmen veranlasst, mehr Ressourcen für die Bekämpfung solcher Probleme bereitzustellen. Bestehende politische Initiativen in diesem Bereich werden in dem begleitenden *Diskussionspapier: Nationale und internationale Modelle zur Online-Regulierung* besprochen. Sie haben sich jedoch bisher als weitgehend unwirksam erwiesen, wenn es darum geht, die gesamte Bandbreite von „Online Harms“ einzudämmen. Sie bergen auch die Gefahr, die Plattformbetreiber zum übermäßigen Entfernen von Inhalten zu verleiten, um potenzielle Geldstrafen zu vermeiden (sogenanntes „Overblocking“).

Als „systemische Ansätze“ bezeichnen wir Regelungen, die versuchen ein derart restriktives Verhalten von Plattformen zu vermeiden. Anstatt sich auf einzelne illegale oder schädliche Inhalte zu fixieren, zielt ein systemischer Ansatz darauf ab, die legitime freie Meinungsäußerung zu schützen. So sollen Anreize für proaktive Maßnahmen zur Risikoprävention bei der Gestaltung von Plattformprodukten, Richtlinien und Prozessen geschaffen werden. In der EU und in Großbritannien fügen die Vorschläge den bestehenden Haftungsbestimmungen eine weitere Ebene hinzu. Hierbei wird das zugrundeliegende Prinzip beibehalten, nach dem Plattformen nicht direkt für nutzergenerierte Inhalte haftbar gemacht werden sollten. Die in den Vorschlägen skizzierten Maßnahmen sollen Plattformen zwingen, zu „**Safety by Design**“-Ansätzen überzugehen. Dies fordert eine proaktive Berücksichtigung potenzieller Risiken für ihre Nutzer oder negativer Auswirkungen auf die Gesellschaft, die durch die Nutzung ihrer Produkte oder Dienste entstehen könnten. Die Vorschläge enthalten auch verschiedene Bestimmungen, die sicherstellen sollen, dass Unternehmen über einheitliche Kennzeichnungs-, Berichts- und Moderationssysteme verfügen, einschließlich Einspruchsverfahren und verbesserter Transparenzanforderungen. Diese würden von den Unternehmen verlangen, dass sie den Betroffenen ihre Entscheidungen besser erklären. Nach beiden Regelungen würden die Systeme großer Plattformen einer Prüfung durch unabhängige Regulierungsbehörden unterzogen, die auch die Funktion ihrer Kernalgorithmen einschließt. Solche Prüfungen würden ergriffene Maßnahmen, sämtliche Daten in Bezug auf ihre Wirksamkeit wie auch künftige Schritte zu ihrer Verbesserung detailliert aufzeigen.

Diese systemischen Ansätze lassen eine wesentliche Verschiebung der Machtdynamik zwischen Regierungen, Technologieunternehmen und den Nutzern ihrer Dienste erkennen. Ziel ist es, die bestehenden kommerziellen Anreize an der Notwendigkeit zu messen, illegale Inhalte und Aktivitäten besser zu bekämpfen. In einigen Fällen gilt es auch die „Harms“ zu reduzieren, die durch legale Inhalte entstehen können. Diese systemischen Ansätze sollten auch die strukturelle Natur von „Online Harms“ angehen, und zwar durch verstärkte Prüfung und Überwachung der zugrundeliegenden Verbreitungsprozesse. Hierzu gehören Empfehlungs- und Newsfeed-Algorithmen, die eine wichtige Rolle für die Reichweite und den Einfluss von Online-Inhalten spielen.

Längerfristig sollten diese neuen Regelungen dazu beitragen, Wettbewerb, Innovation und Investitionen in die Online-Sicherheit zu fördern. Unternehmen werden einen Wettbewerbsvorteil haben, wenn sie diese Anforderungen effektiver und effizienter erfüllen. Auch Nutzer suchen möglicherweise verstärkt nach Produkten und Dienstleistungen, die ein sichereres Online-Erlebnis bieten. Die vorliegenden Vorschläge verfolgen einen abgestuften Ansatz für die Regulierung in Abhängigkeit von der Unternehmensgröße mit mehr Verpflichtungen und einer verstärkten Aufsicht für größere Unternehmen. Es muss sichergestellt werden, dass solche Maßnahmen die gegenwärtige Marktposition der Internetgiganten nicht dauerhaft festschreiben. Diese Unternehmen stehen aufgrund des Umfangs ihrer Nutzerbasis zweifellos vor einer größeren Herausforderung, verfügen aber auch über deutlich mehr Ressourcen als viele kleinere Unternehmen, die ähnliche Produkte und Dienstleistungen anbieten. Um Kartell- und Monopolmacht kümmern sich andere Regulierungsinstrumente, etwa das EU-Gesetz für digitale Dienste (Digital Markets Act – DMA). Nichtsdestotrotz werden die Regulierungsbehörden eine wichtige Rolle spielen, indem sie bewährte Verfahren (z.B. Techniken zur Inhaltsmoderation)

austauschen. Dadurch können die Behörden kleinere Unternehmen bei der Erfüllung ihrer neuen Verpflichtungen unterstützen. Auch der private Sektor sollte ermutigt werden, branchenübergreifend zu kooperieren und Ressourcen und Fachwissen zu teilen. Die Einführung neuer Regulierungen sollte eine signifikante Wirkung auf die „Angebotsseite" von „Online Harms" haben und illegale oder schädliche Online-Inhalte deutlich reduzieren. Systemische Ansätze müssen jedoch auch Vorkehrungen für die „Nachfrageseite" beinhalten, etwa größere gesamtgesellschaftliche Investitionen in die digitale und generelle Medienkompetenzerziehung. Zu den bestehenden Initiativen, die dies berücksichtigen, gehören der Europäische Aktionsplan für Demokratie der EU und die entstehende „Media Literacy Strategy" in Großbritannien.[5] Neue gesetzliche Regelungen sind notwendig, werden aber nicht unbedingt ausreichen, um das gesamte Spektrum der „Online Harms" zu bekämpfen, auf das diese Vorschläge abzielen.

Sinnvolle Transparenz gestalten

Sowohl die freiwilligen Initiativen der Industrie als auch gesetzgeberische Initiativen haben deutlich gemacht, dass Transparenz in Bezug auf Richtlinien, Prozesse und Ergebnisse der Plattformaktionen eine immer wichtigere Rolle spielt. Entsprechend der Argumentation im begleitenden *Diskussionspapier: Transparenz, Datenzugriff und „Online Harms“*, sollte Transparenz das öffentliche Bewusstsein für „Online Harms“ verbessern und die wirksamsten politischen Gegenmaßnahmen sichtbar machen. Es muss klare Regelungen bezüglich des Datenzugangs für Behörden geben, damit der Erfolg oder Misserfolg jeglicher Regulierungsanstrengungen sinnvoll bewertet werden kann.

Im Sinne der sich abzeichnenden Gesetzesvorschläge würden erweiterte Transparenzanforderungen den Regulierungsbehörden und Forschenden wichtige Daten zu den Auswirkungen von Unternehmensmaßnahmen zur Bekämpfung illegaler Inhalte liefern. Weiterhin beinhalten sie die Schaffung zusätzlicher Klarheit für diejenigen Nutzer, deren Inhalte (zu Recht oder Unrecht) entfernt wurden. Es bleibt jedoch abzuwarten, wie effektiv diese Forderungen nach Transparenz in Bezug auf die Plattform-Richtlinien und die Maßnahmen der Unternehmen wirklich sind. Die vorgeschlagenen Anforderungen legen den Fokus erneut eher auf das Vertrauen in die Transparenzberichtsverfahren des privaten Sektors, anstatt auf externe demokratische oder gerichtliche Kontrolle. Betroffene Nutzer müssen in der Regel erst den Rechtsweg durch die Systeme der Unternehmen gehen, bevor sie sich um eine wirklich unabhängige Schlichtung und Wiedergutmachung bemühen können. Es bleibt das Spannungsverhältnis zwischen bestehenden Transparenzansätzen (die meist von den Unternehmen selbst ausgestaltet werden) und dem Ansatz, neue externe Prüfmechanismen für Aktionen, Prozesse und Ergebnisse von Unternehmen zu erfinden. Der Gesetzgeber steht vor der Wahl zwischen

a) praktischen, aber potenziell begrenzten Regeln, die darauf abzielen, bestehende Transparenzansätze des privaten Sektors zu verbessern oder zu ändern, und

b) ehrgeizigen Bemühungen, völlig neue Anforderungen an den Informationszugang und die Bewertung dieser Art von Herausforderungen zu stellen.

Rechtliche Zuständigkeit: Wer legt die Regeln fest und wer setzt sie durch?

Die EU und Großbritannien teilen den Ehrgeiz, richtungsweisend für die weltweite Internetregulierung zu sein. Ihre jeweiligen Vorschläge thematisieren die Herausforderung, nationale oder regionale Rechtsprechung auf inhärent globale Internetdienste anzuwenden. Beide Ansätze werfen Fragen hinsichtlich des potenziell enormen Umfangs, der Komplexität und der Durchsetzbarkeit durch die jeweiligen Regulierungsbehörden auf. Die wesentliche Frage bleibt, wie man eine nationale oder regionale Regulierung für Unternehmen aufbaut, deren Nutzerbasis oft keine identifizierbaren Standorte hat, die transnationale Gesetzesverstöße ermöglichen können und die ihren Firmensitz nur einem oder zwei Ländern haben, aber global agieren.

Terroristische und extremistische Inhalte, Hassrede und Desinformation sind allesamt grundsätzlich globale Phänomene. Staatliche und nichtstaatliche Desinformationsnetzwerke sind meist grenzüberschreitend. Das bedeutet, dass viele Inhalte, die von einem nationalen Publikum konsumiert werden, ihren Ursprung außerhalb der nationalen Rechtsordnung haben. Solche Formen krimineller Aktivitäten sind auch durch eine internationale Uneinheitlichkeit nationaler Rechtsprechungen gekennzeichnet. Beispielhaft dafür sind nationale Verbote bestimmter politischer Organisationen, die in anderen Ländern erlaubt sind.

Was den **Anwendungsbereich betrifft**, so sind große Teile der Inhalte in sozialen Medien für jeden Nutzer des Dienstes verfügbar (und in vielen Fällen für jeden Internetnutzer, unabhängig davon, ob er ein Konto bei diesem Dienst hat, z. B. YouTube). Dies gilt unabhängig von der Herkunft der Inhalte oder dem Standort des Konsumenten. Dieser Umstand wirft Fragen hinsichtlich der Durchführbarkeit einer Regulierung in Bezug auf national definierte Bedrohungen auf. Angesichts der großen Vielfalt innerhalb der EU[6] und Großbritannien[7] konsumiert das jeweilige Publikum Inhalte in einer großen Anzahl von Sprachen, die unweigerlich auch illegale Inhalte wie Hassrede, aber auch legale und dennoch schädliche Inhalte wie Desinformation enthalten. **Es gibt zwar Präzedenzfälle, in denen Regulierungsbehörden ein breites Spektrum an Inhalten und Sprachen abdecken (beispielsweise reguliert Ofcom die Anbieter von Satellitenfernsehen in Großbritannien). Das enorme Ausmaß an Online-Inhalten stellt jedoch eine erhebliche Veränderung in der Größenordnung der Datenmengen dar.**

Fallstudie

Die Bemühungen des DSA zur Vereinheitlichung der Ansätze in der EU

- Die DSA-Gesetzesentwürfe gelten für alle Vermittler, die ihre Dienstleistungen in der EU erbringen, unabhängig davon, wo sie rechtlich ihre Hauptniederlassung haben.
- Dies umfasst Unternehmen mit Sitz in der EU, Plattformen mit einer signifikanten Anzahl von EU-Nutzern und Dienste, die auf EU-Mitgliedstaaten ausgerichtet sind (z. B. wegen der Verwendung einer Landessprache, einer Währung oder einer Domain wie .fr, .de usw.).
- Im Rahmen der DSA muss jedes Unternehmen einen Rechtsvertreter mit Sitz in der Region benennen, der bei Nichteinhaltung oder anderen Problemen sofort zur Rechenschaft gezogen werden kann.
- Die Durchsetzung und Überwachung obliegt dem Koordinator für digitale Dienste (DSC) im jeweiligen Land, der bei Bedarf auch das EU-Board einschalten kann. Wenn Unternehmen versäumen, die notwendigen Schritte zu unternehmen (etwa nach einer unabhängigen Prüfung durch den DSC), kann Druck auf EU-Ebene ausgeübt werden, einschließlich der Verhängung erheblicher Bußgelder auf Basis des Jahresumsatzes.

Fallstudie

Die Antwort der britischen Regierung auf ihre „Online Harms“-Konsultation

- Die Antwort der britischen Regierung vom Dezember 2020 deutet darauf hin, dass Unternehmen in den Anwendungsbereich der kommenden „Online Safety Bill“ fallen, wenn sie nutzergenerierte Inhalte hosten, die britischen Nutzern zur Verfügung stehen, oder Online-Interaktionen (entweder öffentlich oder privat) ermöglichen, bei denen einer oder mehrere Teilnehmer in Großbritannien ansässig sind.

Die Online Harms-Vorschläge sehen vor, dass britische Gesetze auf alle Online-Inhalte anzuwenden sind, wenn diese Inhalte für britische Nutzer verfügbar sind. Der DSA wird als horizontale EU-Gesetzgebung auf einem **komplexen Flickenteppich aus nationalen Gesetzen** von EU-Mitgliedstaaten mit unterschiedlichen Rechtssystemen und Traditionen beruhen. Zwar besteht in der Regel weitgehende Einigkeit darüber, dass das, was offline illegal ist, auch online illegal sein sollte, jedoch könnten nationale Unterschiede innerhalb der EU - angesichts des transnationalen Charakters des Internets - zu neuen Spannungen führen. Beispielsweise haben einige EU-Mitgliedstaaten bestimmte terroristische oder gewalttätige extremistische Gruppen und ihre Symbole verboten, die aber in anderen nationalen Kontexten erlaubt sind. Inhalte, die den Holocaust leugnen oder das Naziregime loben, sind in Deutschland illegal, nicht aber in anderen EU-Ländern. In der Praxis könnte dies dazu führen, dass Unternehmen alle Kategorien von Inhalten, die irgendwo illegal sind, in ihre überall geltenden Nutzungsbedingungen aufnehmen. Dies wäre eine einfache Lösung für große Plattformen um die verschiedenen Rechtsbestimmungen der verschiedenen nationalen Märkte umzusetzen.

In bestimmten Bereichen ist dies bereits der Fall. Beispielsweise hat Facebook 2020 erstmals explizit weltweit die Leugnung des Holocausts auf ihrer Plattform durch eine neue AGB verboten.[8] Auch wenn Facebook’s neue Nutzungsbedingungen in diesem Fall zu begrüßen sind, kann dieser Ansatz dazu führen, dass nationale Gesetze de facto über die Landesgrenzen hinaus extrapoliert werden. Dies könnte sich innerhalb der EU als problematisch erweisen, da beispielsweise die Regierungen bestimmter Mitgliedstaaten unterschiedliche Ansichten zu LGBTQ+ oder Geschlechtergleichstellung haben. Außerdem wird zusätzlicher Druck auf Unternehmen ausgeübt, sich über ihre AGBs an inhaltsbezogene Gesetze zu halten, selbst wenn die Gesetze den Grundwerten der EU widersprechen (z.B. Blasphemiegesetze).

Eine Auswirkung könnte sein, dass Plattformen „geofencing“ auf ihre Dienste anwenden, um Länder oder Regionen auszuschließen, die eine Einhaltung der Regulierung verlangen. Damit ließe sich eine entsprechende Regelkonformität vermeiden, wie bei dem sich abzeichnenden Patt im australischen Fall.[9] In einem anderen Szenario könnten die Plattformen übervorsichtig werden und sämtliche Inhalte über ihre AGB auf Grundlage der strengsten nationalen Gesetze beschränken, und zwar weltweit. Dies könnte zwar zu einer gewissen Verbesserung der Sicherheit führen, hätte aber auch einen Verlust an Zugang und Auswahl für die Nutzer zur Folge und würde den Wettbewerb und die Meinungsfreiheit beeinträchtigen.

Was die **Durchsetzung** betrifft, so enthalten die neuen Gesetzesentwürfe auch Bestimmungen, die eine Haftung leitender Mitarbeiter von Diensten beinhaltet. Hierzu gehört beispielsweise die Anforderung, eine Kontaktstelle und einen gesetzlichen Vertreter zu benennen, der bei Nichteinhaltung haftbar gemacht werden könnte. Dieser Ansatz wirft jedoch die Frage auf, ob dies wirksam angewendet und durchgesetzt werden kann, wenn die betreffenden Personen nicht dauerhaft in der EU oder in Großbritannien ansässig sind. Es könnte auch dazu führen, dass sich die Dienste aus einem der beiden Märkte zurückziehen, um diese Regulierung zu umgehen.

Die Vorteile und Herausforderungen einer Politikgestaltung auf transnationaler Ebene werden an dem schwierigen Umfeld, mit dem sich der DSA in der EU konfrontiert sieht, deutlich. Der gewählte Ansatz legt eine beträchtliche Verantwortung für Aufsicht und Durchsetzung in die Hände bestimmter nationaler Regulierungsbehörden, wenn Plattformen innerhalb der EU ansässig sind (beispielsweise sind viele der größten Unternehmen in der Republik Irland ansässig). Dies könnte zu einem unverhältnismäßigen Machtungleichgewicht zwischen den verschiedenen nationalen Regulierungsbehörden führen. „Harms", die in einem Land entstanden sind, könnten in die Verantwortung der Regulierungsbehörde in einem anderen Land fallen, die aber gegenüber der nationalen Regierung des Verursachers nicht rechenschaftspflichtig wäre. Denkt man über Europa hinaus, werden solche Spannungen zwischen Rechtsräumen auch eine Herausforderung für den Aufbau supranationaler Regulierungsstrukturen darstellen.

Der Umgang mit „Legal Harms“

Regierungen bedenken zurecht, dass es ein Risiko ist, spezifische Maßnahmen gegen so genannte „Legal Harms“ gesetzlich vorzuschreiben, da so Innovationen erstickt, marginale Stimmen ausgeschlossen oder das Grundrecht auf Meinungsfreiheit negativ beeinflusst werden können. Um ein Gleichgewicht zwischen dem Schutz der Meinungsfreiheit und der Gewährleistung eines fairen und sicheren Umfelds für Meinungsfreiheit zu erreichen, ohne dabei die Risiken für physische oder psychische „Harms“ zu erhöhen, muss klar definiert werden, was „Harm“ bedeutet. Der Schutz der Grundrechte, einschließlich derer, die sich auf die Sicherheit, die Unversehrtheit und die Meinungsfreiheit beziehen, muss durch ein verhältnismäßiges und evidenzbasiertes Verständnis von den Grenzen der Verantwortlichkeit einer Plattform begleitet werden. Regulierungsansätze, die legale aber „schädliche“ Inhalte oder Aktivitäten einbeziehen und somit definieren möchten, müssen eine Reihe von Schlüsselfragen beantworten. Dazu gehören:

- **Schadensmessung**: Wie wird festgestellt, ob es einen Kausalzusammenhang zwischen Online-Inhalten oder -Aktivitäten und dem physischen oder psychischen „Harms“ für eine oder mehrere Personen gibt? Welche Offline- und Online-Nachweise könnten bei dieser Art der Bewertung hilfreich sein? Was können wir von anderen Bereichen der Regulierung lernen, die sich mit den externen Effekten von Unternehmen oder Konzernen befassen? Wie können Regierungen Schwellenwerte festlegen, um nachteilige psychische Auswirkungen von Online-Inhalten oder -Aktivitäten zu messen?

- **Schadensumfang**: Sollten gesellschaftliche Gefahren – also solche, die beispielsweise Institutionen, Abläufe oder Prozesse der Demokratie bedrohen – neben individuellen „Harms“ in den Anwendungsbereich der Regulierung einbezogen werden? Was sind die zusätzlichen Herausforderungen bei der Identifizierung, Messung und somit Verhinderung oder Sanktionierung dieser gesellschaftlichen „Harms“?

Je breiter gefasst die angewandte Definition ist, desto höher ist zwangsläufig das Risiko einer unverhältnismäßigen Auswirkung auf die Redefreiheit. Wenn eine weit gefasste Definition verwendet wird, hätte die Regulierungsbehörde mehr Spielraum zu bestimmen, was als schädlich anzusehen ist. Dies würde einen größeren Anreiz für Plattformen und Dienste schaffen, ein weites Netz auszuwerfen, wenn sie Inhalte oder Aktivitäten in Betracht ziehen, die möglicherweise in den Anwendungsbereich fallen, und könnte ihnen helfen, eine weitere behördliche Prüfung oder Durchsetzung von Regularien zu vermeiden.[10] **Wird hingegen nicht versucht, legale von illegalen Äußerungen abzugrenzen, entsteht ein unangemessener Druck auf die Nutzungsbedingungen der Plattformen. Die AGBs privater Unternehmen sollten aber nicht darüber entscheiden, wo die Grenzen der öffentlichen Meinungsäußerung liegen.** Fast zwangsläufig führt der Versuch, die Grenzen legaler, aber schädlicher Sprache weiter zu definieren, zu erheblichen regulatorischen Grauzonen.

Ein möglicher Ausweg sind Regulierungsbemühungen, die legale aber schädliche Inhalte von ihrem Zuständigkeitsbereich ausschließen. Sie könnten dennoch das Vorhandensein, beziehungsweise das Ausmaß dieser „Harms“ im Internet beeinflussen. Systemische Ansätze zur Regulierung illegaler Inhalte auf digitalen Plattformen können Anforderungen zur Verbesserung der Transparenz, zur Umsetzung des „Safety-by-Design“-Prinzips und das Ermöglichen unabhängiger Prüfungsbefugnisse zur Überwachung der Richtlinien, Prozesse und Ergebnisse von Unternehmensmaßnahmen beinhalten. In diesen Fällen müssen die Plattformen möglicherweise alle Prozesse im Hinblick auf die Transparenz und Sicherheit der Moderation und Kuratierung von Inhalten auf ihren Diensten anpassen und verbessern. **Während diese Schritte von den Plattformen unternommen würden, um ihrer Verantwortung bezüglich illegaler Inhalte gerecht zu werden, könnten die Verbesserungen der Transparenz auch das Vorhandensein von „legal harms“ einschränken – oder zumindest das öffentliche Verständnis dafür verbessern, wie und wo diese Art von „harms“ entstehen.**

Fallstudie

Die Antwort der britischen Regierung auf ihre „Online Harms“-Konsultation sieht vor, Plattformen der Kategorie 1 zu verpflichten, legale aber schädliche Inhalte zu bekämpfen:

- Alle Plattformen werden verpflichtet, gegen illegale Inhalte (z. B. Terrorismus oder Kindesmissbrauch, kurz CSEA) vorzugehen und die Risiken für Kinder durch legale aber schädliche Inhalte zu berücksichtigen. Unternehmen der Kategorie 1 müssen auch legale aber schädliche Inhalte für alle Nutzer in ihren Nutzungsbedingungen (ABGs) ansprechen und durch Transparenzberichte nachweisen, dass sie ihre AGBs konsequent durchsetzen.

- Während die britische Regierung Entwürfe zu „Codes of Practice“ für terroristische und CSEA-Inhalte veröffentlicht hat, die weitgehend unverbindliche Ansätze zur Bekämpfung dieser Kategorie von illegalen Inhalten enthalten, obliegt es der Regulierungsbehörde, Kodizes für die Kategorie „legal aber schädlich“ zu erstellen.

- Die Vorschläge der britischen Regierung basieren im Wesentlichen auf dem „Harms“-Konzept, das dadurch definiert wird, dass Online-Inhalte oder -Aktivitäten „ein begründetes vorhersehbares Risiko einer erheblichen nachteiligen physischen oder psychischen Auswirkung auf Einzelpersonen bergen“. Um zusätzliche Klarheit zu schaffen, wird die kommende Gesetzgebung die Definition schädlicher Inhalte und Aktivitäten, welche in den Anwendungsbereich fallen ergänzen, und zwar durch vorrangige „Harms“-Kategorien, die in der Sekundärgesetzgebung enthalten sind, sowie durch spezifische Ausnahmen, für die es bereits Regelungen gibt (z. B. Urheberrecht, Daten- und Verbraucherschutz, Cyber-Betrug, Hacking).

Fallstudie

Der DSA versucht nicht, legale aber schädliche Inhalte explizit zu regulieren, und schlägt eine Reihe bestehender bzw. zusätzlicher Gesetze vor, um illegale Inhalte abzudecken, die ansonsten nicht definiert sind. Hierzu gehören etwa bestehende Gesetze der Union oder der Mitgliedsstaaten und die vorgeschlagene Verordnung zu terroristischen Online-Inhalten.

- Der DSA behält die bestehenden Haftungsregeln für Anbieter von Vermittlungsdiensten bei, nach denen ein Hosting-Dienst nur dann zum Handeln verpflichtet ist, wenn er tatsächlich Kenntnis von einer Rechtswidrigkeit erlangt. Dann muss er allerdings zeitnah handeln und den Zugang zu diesem Inhalt sperren oder den Inhalt entfernen.[11]

- Die Kategorie „schädliche Informationen und Aktivitäten“ wird als „ein heikler Bereich mit schwerwiegenden Auswirkungen auf den Schutz der Meinungsfreiheit“ beschrieben.[12] Die DSA-Vorschläge erkennen jedoch an, dass der Umfang und die Allgegenwärtigkeit bestimmter Plattformen „ihre Rolle bei der Vermittlung und Verbreitung rechtswidriger oder anderweitig schädlicher Informationen und Aktivitäten gestärkt hat“.[13]

- Die Gesetzesentwürfe würden somit Sorgfaltspflichten schaffen, die auch die Inhaltsmoderation von Plattformen umfassen. Dies betrifft sowohl illegale als auch legale aber potenziell schädliche Inhalte, die gegen ihre Nutzungsbedingungen verstoßen. Dabei wird jedoch betont, dass die beiden Kategorien unterschiedlich behandelt werden und Regeln, die ein Entfernen legaler Inhalte vorschreiben, nicht enthalten sein sollten.[14]

- Die DSA-Vorschläge heben die breiteren Risiken hervor, die von den größten Plattformen ausgehen, die „im Allgemeinen auf eine Optimierung ihres oft werbegestützten Geschäftsmodells ausgerichtet [sind] und kann Anlass zu gesellschaftlichen Bedenken geben [...], ohne die Risiken und den gesellschaftlichen und wirtschaftlichen Schaden, den sie verursachen können, zu erkennen und wirksam zu mindern“.[15] Infolgedessen werden diese Plattformen verpflichtet, Bewertungen durchzuführen, die alle systemischen Risiken im Zusammenhang mit ihren Diensten abdecken, einschließlich des potenziellen Missbrauchs durch Nutzer, und dann entsprechende Maßnahmen zur Abmilderung zu ergreifen.

- Dies betrifft illegale Inhalte und Aktivitäten, aber auch andere negative Auswirkungen auf Grundrechte, etwa die Achtung des Privat- und Familienlebens, die Meinungs- und Informationsfreiheit, das Diskriminierungsverbot und die Rechte von Kindern. Es umfasst auch „die absichtliche und oft koordinierte Manipulation des Dienstes der Plattform mit vorhersehbaren Auswirkungen auf die Gesundheit, den zivilen Diskurs, Wahlprozesse, die öffentliche Sicherheit und den Jugendschutz, unter Berücksichtigung der Notwendigkeit, die öffentliche Ordnung zu schützen, die Privatsphäre zu wahren und betrügerische und irreführende Geschäftspraktiken zu bekämpfen“.[16]

Die Corona-Pandemie gezeigt, dass außergewöhnliche Umstände eintreten können, unter denen auf Plattformen neue Arten von schädlichen oder bedrohlichen Inhalten und Aktivitäten auftauchen. Einige Vorschläge für die digitale Regulierung verweisen auf „**Krisenprotokolle**“, die unter solch außergewöhnlichen Umständen aktiviert werden könnten. Zu diesen außergewöhnlichen Umständen gehören auch Ereignisse, die die öffentliche Sicherheit oder Gesundheit betreffen. Beispiele hierfür sind Naturkatastrophen, Pandemien, Terroranschläge, aber auch wichtige Wahlen oder, mit den Worten der DSA, Fälle in denen „Online-Plattformen für die schnelle Verbreitung illegaler Inhalte oder Desinformation missbraucht werden können oder in denen die Notwendigkeit einer schnellen Verbreitung zuverlässiger Informationen besteht“.[17] Die genauen Schwellenwerte für „Krise“ und die zu erwartende Reaktion bleiben unklar, könnten aber in Zusammenarbeit mit der Zivilgesellschaft und anderen Expertengremien definiert werden. Angesichts der Tatsache, dass sich die Vorhersage solcher Notfälle schwierig gestaltet, werden den Plattformen unter solchen Umständen wahrscheinlich „Light-Touch“-Verpflichtungen auferlegt und sie werden sich bei deren Auslegung, Durchsetzung und Transparenz auf ihre eigenen Nutzungsbedingungen stützen.

Regulierungsumfang: Wer wird reguliert?

Ein zentrales Spannungsfeld bei der Gestaltung einer effektiven, aber auch effizienten und verhältnismäßigen Regulierung für Online-Plattformen ist die Frage, welche Dienste und/oder Unternehmen unter bestimmte Verpflichtungen fallen. **Ein zentraler Streitpunkt bei den sich abzeichnenden politischen Vorschlägen für diesen Bereich ist die Frage, ob „Harm" über die Reichweite (d. h. gefährliche Inhalte, die ein möglichst breites Publikum erreichen) oder über die Extremität (d. h. wie gefährlich diese Inhalte sind) definiert werden sollte.**

Wenn das Ziel darin besteht, für die Online-Sicherheit zu sorgen, ist die Anzahl der Plattformnutzer kein genauer Gradmesser für „Online Harms". In den letzten Jahren gab es Abwanderungen von größeren Plattformen - die unter dem Druck der Regierungen (zumindest in bestimmten Bereichen, z. B. bei terroristischen Inhalten) schrittweise die Mäßigung und Durchsetzung verbessert haben - hin zu kleineren Plattformen. Letztere sind entweder schlecht ausgerüstet oder nicht bereit, Maßnahmen zu ergreifen. Da weitere Regulierungen eingeführt werden, die sich vor allem auf die größten Plattformen konzentrieren, wird sich dieser Abwanderungstrend wahrscheinlich fortsetzen. Die jüngst gewachsenen Nutzerzahlen von Parler und Telegram machen das deutlich.

Solche Plattformen stellen je nach Art des „Harms" unterschiedlich große Herausforderungen für die Sicherheit dar. Extremistische Gruppen fangen teilweise auf größeren Plattformen an, jedoch ist die eigentliche Radikalisierung und Rekrutierung in kleineren, ideologisch homogeneren und nicht moderierten Räumen effizienter und findet somit auch eher dort statt. Im Gegensatz dazu können Desinformationskampagnen nichtstaatlicher Akteure zwar auf kleineren Plattformen organisiert werden, brauchen aber, um maximale Wirkung zu erzielen, ein breiteres Publikum auf größeren Plattformen.

Die Größe einer Plattform ist kein perfektes Analogon für digitale Dienstleistungen. Nichtsdestotrotz ist es ermutigend, dass neue Vorschläge in liberalen Demokratien zunehmend die große Vielfalt verschiedener Dienstleistungen, die von Plattformen und auf verschiedenen Ebenen des technischen „Stacks" des Internets angeboten werden, berücksichtigen. Ein Beispiel hierfür ist, dass einige Gesetzesvorschläge zwischen Infrastrukturanbietern (z.B. Cloudflare oder AWS) und Plattformen (z.B. Facebook oder Twitter) unterscheiden. Wenngleich noch im Anfangsstadium, ist diese Art von Regulierung dem Ansatz einer nur auf die großen soziale Medien abzielenden Regulierung vorzuziehen.

Die Vorschläge der EU und Großbritanniens gehen darauf ein, dass die durch Internetdienste ausgelösten Bedrohungen für die Online-Sicherheit stark variieren können. Auch verfügen nicht alle Internetdienste über die gleichen internen Ressourcen und das Fachwissen, um Herausforderungen an die Online-Sicherheit zu bewältigen. Darüber hinaus betonen die EU und Großbritannien, dass Wachstum und Innovation nicht behindert und kleinere Unternehmen und Start-ups nicht unverhältnismäßig belastet werden dürfen. Daher sehen beide Vorschläge einen abgestuften Ansatz für die Regulierung vor, bei dem größere Plattformen im Vergleich zu ihren kleineren Pendants zusätzliche Verpflichtungen haben und einer stärkeren Aufsicht unterliegen. Diese werden in der begleitenden **Zusammenfassung des DSA und des „Online Harms White Paper"** näher erläutert.

Im Laufe der Zeit werden die verbesserten Transparenzbestimmungen beider Entwürfe Regulierungsbehörden und Forschern einen besseren Zugang zu Daten ermöglichen und zu einem besseren Verständnis für die Rolle von Online-Diensten bei der Entstehung und Verbreitung von „Online Harms" beitragen. **Das Ziel ist eine nuancierte Zuteilung von Verantwortlichkeiten und Verpflichtungen an Plattformbetreiber und Regulierungsbehörden.**

Blick in die Zukunft

Dieses Papier hat nur an der Oberfläche der Herausforderungen gekratzt, die dann entstehen, wenn Regierungen nach möglichst effektiven, verhältnismäßigen und realisierbaren Wegen zur Internetregulierung suchen. Auf keine dieser Herausforderungen gibt es eine einfache Antwort. Stets muss der Kontext, in dem eine Regulierung in Betracht gezogen wird, berücksichtigt werden. Komplizierte Regulierungsgeflechte werden die Art und Weise, wie Plattformen ihren Nutzern dienen können und müssen, im kommenden Jahrzehnt dramatisch verändern. Abgesehen von den oben genannten Aspekten stehen die Regierungen vor der zusätzlichen Frage, wie die Regulierung mit parallel stattfindenden Bemühungen in den Bereichen Wettbewerb und Datenschutz in Einklang gebracht werden kann. Beim Wettbewerb geht es unter anderem um kartellrechtliche Maßnahmen und um potenzielle Probleme in Zusammenhang mit der bestehenden Gesetzgebung zu Geschäftsgeheimnissen, welche die Transparenz interner Unternehmensprozesse einschränken kann. Auch der Datenschutz ist relevant für die laufenden politischen Debatten, um die Regulierung von verschlüsselten Messaging-Plattformen zu organisieren, während gleichzeitig die Nutzerrechte im Hinblick auf private Kommunikation gewahrt bleiben.

Der jeweilige rechtliche, politische und kulturelle Kontext wird die Lösungen prägen müssen, die jede Regierung erarbeitet, um sicherzustellen, dass die Regulierung der Öffentlichkeit so effektiv wie möglich dient. Die regulatorischen Herausforderungen sind nicht gänzlich neu. In vielen Bereichen der unternehmerischen und politischen Tätigkeit wurden komplexe Aufsichtsmechanismen entwickelt, die scheinbar unlösbare Spannungen zwischen Rechten, Verantwortlichkeiten und Risiken mit sich bringen. Multidisziplinäre Gespräche zwischen freiheitlich-demokratischen Regierungen können bei der Entscheidungsfindung in diesen Fragen hilfreich sein, wenn sie Lektionen aus anderen Bereichen und über Grenzen hinweg berücksichtigen.

Endnoten

1. Der Begriff „Online Harms“ geht auf das „Online Harms White Paper“ der britischen Regierung zurück, welches im April 2019 veröffentlicht wurde. „Online Harms“ steht sowohl für konkrete als auch potenzielle Schäden (sprich Gefahren), die z.B. durch das Verbreiten von problema¬tischen Inhalten in sozialen Medien entstehen. Da der englische Begriff die deutschen Begriffe Schäden, Gefahren und Risiken vereint, wird in diesem Papier der englische Originalbegriff verwendet.
2. https://ec.europa.eu/commission/presscorner/detail/de/QANDA_20_2348
3. https://ec.europa.eu/commission/presscorner/detail/de/QANDA_20_2348
4. Online Harms White Paper: Full government response to the consultation. https://www.gov.uk/government/consultations/online-harms-white-paper/outcome/online-harms-white-paper-full-government-response#part-1-who-will-the-new-regulatory-framework-apply-to
5. Initiativen zur Förderung digitaler Medienkompetenz sollten nicht allein auf Schüler abzielen, sondern auch Bildungsangebote für Erwachsene anbieten. Ein Beispiel dafür ist der Business Council for Democracy (#BC4D) - ein Schulungsangebot für Arbeitnehmerinnen und Arbeitnehmer. Das Pilotprojekt wird vom ISD gemeinsam mit der Hertie-Stiftung und der Robert Bosch-Stiftung durchgeführt.
6. 24 offizielle Sprachen und über 60 regionale Sprachen oder Dialekte, z. B. Katalanisch oder Baskisch: https://europa.eu/european-union/about-eu/eu-languages_en
7. Neben Englisch, Walisisch und Gälisch listet die britische Volkszählung 2011 Polnisch, Panjabi, Urdu, Bengali, Gujarati, Arabisch, Französisch, Chinesisch (alle Dialekte), Portugiesisch und Spanisch als die nächsthäufigsten Sprachen auf. Der British Council schätzt, dass allein in London über 300 Sprachen gesprochen werden: https://www.ons.gov.uk/peoplepopulationandcommunity/culturalidentity/language & https://study-uk.britishcouncil.org/moving-uk/student-life/language
8. https://www.bbc.co.uk/news/technology-54509975
9. https://www.bbc.com/news/world-australia-56107028
10. Dies ist im Wesentlichen die Gefahr des „Overblocking“, die im Mittelpunkt der Kritik an inhaltsbezogenen Regulierungsansätzen und insbesondere am deutschen Netzwerkdurchsetzungsgesetz steht.
11. DSA (Artikel 5.1, §22)
12. DSA (S.9)
13. DSA (§5)
14. DSA Artikel 2.p
15. DSA (§56)
16. DSA (Artikel 26, §57)
17. DSA (§71)

ISD | Powering solutions to extremism and polarisation

Beirut | Berlin | London | Paris | Washington DC

www.isdglobal.org