

---

Bericht

**Governance von Desinformation in digitalisierten Öffentlichkeiten**

---

Stand: Ende Juni 2023

Autor:innen: Florent Thouvenin, Mark Eisenegger, Stephanie Volz, Daniel Vogler, Mariela Jaffé

---

Center for Information Technology,  
Society, and Law (ITSL)  
Universität Zürich

CENTER FOR  
INFORMATION  
TECHNOLOGY  
SOCIETY AND  
LAW — ITSL

Forschungszentrum Öffentlichkeit  
und Gesellschaft (fög)  
Universität Zürich

**fög**  
*Forschungszentrum  
Öffentlichkeit und Gesellschaft*

---

Bitte zitieren Sie diesen Bericht wie folgt:

Thouvenin, F.; Eisenegger, M.; Volz, S.; Vogler, D.; Jaffé, M., (2023). Governance von Desinformation in digitalisierten Öffentlichkeiten. Bericht für das Bundesamt für Kommunikation (BAKOM).

Die Autoren danken Prof. Dr. Franziska Oehmer-Pedrazzi ([franziska.oehmer@fhgr.ch](mailto:franziska.oehmer@fhgr.ch)) für die Mitarbeit am Bericht; sie hat die Konzeption der Darstellung des sozialwissenschaftlichen Forschungsstands (S. 10-26) erarbeitet und wesentlich zu den Texten dieses Teils beigetragen. Dank gebührt auch MLaw Alisa Zehner, Rechtsanwältin, für die umfangreichen Recherchen und Vorarbeiten zu den Ausführungen zum rechtlichen Rahmen in der EU.

<b>A.</b>	<b>Einleitung und Zielstellung</b>	<b>8</b>
<b>B.</b>	<b>Sozialwissenschaftliche Befunde zur Desinformation</b>	<b>10</b>
I.	Befunde zu Verfasser:innen und Verbreiter:innen	10
1.	Verfasser:innen	11
2.	Verbreiter:innen	13
3.	Verbreitung über (teil)automatisierte Accounts	15
II.	Befunde zu Kanälen	16
III.	Befunde zu Themen, Inhalten und Form	18
IV.	Befunde zur Nutzung	20
V.	Befunde zur Wirkung	20
1.	Wirkungen auf Individualebene	22
2.	Gesamtgesellschaftliche Wirkungen	23
VI.	Zwischenfazit: Sozialwissenschaftliche Befunde zur Desinformation (in der Schweiz)	24
<b>C.</b>	<b>Governance von Desinformation</b>	<b>27</b>
I.	Vorbemerkungen	27
1.	Allgemeines	27
2.	Besonderheiten der digitalen Desinformation	27
3.	Unterscheidung zwischen rechtmässiger und unrechtmässiger Desinformation	28
4.	Unterscheidung zwischen Verfasser- und Verbreiter:innen	28
II.	Rechtlicher Rahmen in der Europäischen Union	28
1.	Einleitung	28
2.	Rechtsrahmen für nicht rechtmässige Desinformationsinhalte: Der Digital Services Act («DSA»)	28
3.	Rechtsrahmen für rechtmässige Desinformationsinhalte	31
4.	Weitere Entwicklungen in der EU	33
5.	Auswirkungen des Digital Services Act auf die Schweiz	33
III.	Grundrechtlicher Rahmen	33
1.	Vorbemerkungen	34
2.	Einschränkung der Grundrechte	34
3.	Desinformation aus Sicht der Verbreiter:innen	35
4.	Durch Desinformation betroffene Grundrechte Dritter	38
5.	Die Grundsatzperspektive besonderer Verbreiter:innen	41
6.	Die Grundrechtsperspektive der identifizierten Verbreitungskanäle von Desinformation	41
7.	Zusammenfassung der grundrechtlichen Situation	42
IV.	Rechtsrahmen für rechtswidrige Desinformation in der Schweiz	43
1.	Vorbemerkungen	43

2.	Materielles Recht: Normen gegen Desinformation _____	43
3.	Einhaltung und Durchsetzung: Besondere Akteur:innen _____	45
4.	Einhaltung und Durchsetzung: Kanäle _____	46
5.	Zwischenfazit _____	52
V.	Rechtsrahmen für rechtmässige Desinformation in der Schweiz _____	52
1.	Vorbemerkung _____	52
2.	Materielles Recht: Normen gegen rechtmässige Desinformation _____	52
3.	Materielles Recht: Besondere Bestimmungen für gewisse Adressaten _____	54
4.	Einhaltung und Durchsetzung: Kanäle _____	56
5.	Massnahmen von zivilgesellschaftlichen Organisationen und Individuen _____	58
<b>D.</b>	<b>Empfehlungen _____</b>	<b>59</b>
I.	Vorbemerkungen _____	59
II.	Direkte Massnahmen _____	61
1.	Staatliche Regulierung _____	61
2.	Ko-Regulierung _____	65
3.	Selbstregulierung (insb. Branchenregulierung) _____	69
4.	Selbstorganisation _____	71
5.	Zivilgesellschaftliches und individuelles Engagement _____	77
III.	Indirekte Massnahmen _____	78
1.	Teilnahme an Europäischen Gremien und Initiativen _____	78
2.	Steigerung der Medien- und Digitalkompetenz _____	78
3.	Stärkung des Informationsjournalismus _____	80
4.	Gründung einer (kostenlosen) Medienplattform _____	80
<b>E.</b>	<b>Glossar _____</b>	<b>81</b>
<b>F.</b>	<b>Literatur _____</b>	<b>84</b>

## Management Summary

Mit der Digitalisierung hat Desinformation an Bedeutung gewonnen. Diese Entwicklung kann in demokratischen Gesellschaften eine Gefahr für den öffentlichen Diskurs darstellen. Auch in der Schweiz herrscht aktuell eine erhöhte Sorge um problematische Auswirkungen von Desinformation. Die Frage nach der Governance von Desinformation ist daher von wachsender Bedeutung.

Dieser Bericht lotet den Spielraum für die Governance von Desinformation in der Schweiz interdisziplinär, d.h. aus einer rechts- und sozialwissenschaftlichen Perspektive, aus. Diese gesamtheitliche Betrachtung des Instrumentariums zur Governance von Desinformation erlaubt es, auf neue gesellschaftliche Entwicklungen zu reagieren und bei dringlichen Forderungen aus der Politik oder der Gesellschaft empirisch fundierte Aussagen zu wirksamen und rechtlich zulässigen sowie sinnvollen Governance-Massnahmen zu machen.

Um dem Problem der Desinformation in der Schweiz zu begegnen, empfehlen wir die folgenden Massnahmen zur näheren Prüfung und allfälligen Umsetzung. Die Massnahmen werden im Bericht aus der rechts- und sozialwissenschaftlichen Forschung hergeleitet und in Kapitel D. ausgeführt und begründet:

**Fokus auf Social Media und Messengerdienste.** Bei der Governance von Desinformation sollten Social-Media-Plattformen und Messengerdienste im Zentrum stehen. Diese Kanäle nehmen bei der Verbreitung von Desinformation eine Schlüsselrolle ein.

**Schaffung eines:r Desinformations(schutz)beauftragten oder einer Desinformations-Monitoringstelle** als unabhängige Stelle oder als Stelle innerhalb einer Bundesbehörde. Die Person bzw. die Stelle könnte Beobachtungs-, Anlauf- und Koordinationsaufgaben wahrnehmen, wobei die Beobachtung der Entwicklung von Desinformation in der Schweiz im Vordergrund stünde.

**Gesetzliche Verankerung des Trennungsgebots bzw. einer Kennzeichnungspflicht für Werbung.** Wie in traditionellen Medien könnte ein Gebot zur Trennung von (kommerzieller und politischer) Werbung und redaktioneller Texte für alle Kanäle verankert werden.

**Kennzeichnungs- und Identifikationspflicht für politische Werbung.** Ergänzend zu einer allgemeinen Kennzeichnungspflicht für Werbung könnte eine Pflicht zur Kennzeichnung von Anzeigen als «politische Werbung» und zur Identifikation der hinter der Anzeige stehenden Person oder Organisation eingeführt werden.

**Schaffung eines (staatlichen) Streitschlichtungsverfahrens.** Die heutige Ko-Regulierung und die Selbstorganisation führen zu einer Verlagerung von der staatlichen zu einer privaten Rechtsdurchsetzung. Die Entscheidung über Recht und Unrecht sollte aus rechtsstaatlicher Sicht bei staatlichen Gerichten liegen. Aufgrund der immensen Anzahl potenzieller Fälle wäre

denkbar, dass der Staat ein unabhängiges und schnelles – auf erster Instanz allenfalls auch automatisiertes – Verfahren anbieten würde.

**Kennzeichnungspflicht für automatisierte Accounts (Social Bots).** Die Kennzeichnung würde Nutzer:innen darüber informieren, dass hinter einem Account ein Computerprogramm steht. Die Kennzeichnung von automatisierten Accounts und anderen algorithmischen Systemen wird auch in anderem Zusammenhang ganz allgemein gefordert.

**Prozedurale Vorgaben für (gewisse) Plattformen,** analog der Regelung im DSA. Dazu gehören u.a. die Einrichtung eines internen Beschwerdemanagements und die Verpflichtung, sich bei Streitfällen einem externen Streitbeilegungsverfahren zu unterwerfen, in welchem der interne Entscheid überprüft werden kann sowie die Verpflichtung von Angaben zu den entschiedenen Fällen.

**Transparenzvorgaben für (gewisse) Plattformen.** Auf Online-Plattformen entscheiden algorithmische Empfehlungssysteme über die Sichtbarkeit und damit über die Wirkung von Inhalten. Anbieter:innen von Online-Plattformen sollten deshalb bei der Verwendung von solchen Empfehlungssystemen die wichtigsten Parameter bekannt geben, welche das Ranking von Inhalten bestimmen.

**Recht auf Zugang für Nutzer:innen.** Alle Nutzer:innen, die sich an einem Diskurs beteiligen möchten, sollten sich auch an diesem beteiligen können. Zu diesem Zweck könnte ein «Recht auf Zugang» von Nutzer:innen zu (grossen) Plattformen oder zumindest eine Begrenzung der Möglichkeiten des Ausschlusses von Nutzer:innen eingeführt werden.

**Zugangsrecht zu (anonymisierten) Daten,** um empirische Forschung zu Desinformation zu ermöglichen. Initiativen der Selbstregulierung waren nicht erfolgreich. Deshalb sollte geprüft werden, wie ein Recht auf Zugang durch (Ko-)Regulierung oder gesetzliche Vorgaben umgesetzt werden kann.

**Schaffung eines Digitalrates** aus Plattformen und weiteren Unternehmen als Selbstregulierungsorgan, das analog zum Presserat gewisse gemeinsame Richtlinien und Standards definiert und unter Umständen auch durchsetzt. Der Digitalrat könnte auch als Beschwerdeinstanz dienen, die Beschwerden gegen das Sperren oder Löschen von Inhalten oder Nutzer:innen behandelt.

**Selbstorganisation der Plattformen** zur Regulierung der dort verbreiteten Inhalte ist sinnvoll. Zu den heute angewendeten Massnahmen zählen das Sperren von Nutzer:innen sowie das Löschen, Downranking und Prüfen von gewissen Inhalten. Die Wirkung dieser Massnahmen ist derzeit schwierig zu beurteilen, weil kein genügender Zugang zu Daten besteht und weil plattformspezifische Effekte von Bedeutung sind.

**Zivilgesellschaftliches und individuelles Engagement** gegen Desinformation, bspw. das Schaffen von Fact-Checking-Angeboten oder das Melden von Desinformation durch Bürger:innen, ist zu begrüßen. Der Gesetzgeber kann diese Massnahmen aber nicht initiieren, sondern nur angemessene Rahmenbedingungen schaffen.

**Teilnahme an Europäischen Gremien und Initiativen.** In Europa gibt es bereits zahlreiche Gremien und Initiativen, die sich mit Desinformation im weitesten Sinn befassen. In einigen dieser Gremien ist die Schweiz als Beobachterin dabei, in anderen wäre eine Mitarbeit oder Teilnahme hilfreich.

**Generelle Steigerung der Medien- und Digitalkompetenz.** Dieser Massnahme kommt aus unserer Sicht eine zentrale Rolle für die Erhöhung der Resilienz gegenüber Desinformation zu. Solche Massnahmen sollten umfassend und für verschiedene Zielgruppen angeboten werden, besonders aber für Jugendliche, ältere Personen und Menschen mit Migrationshintergrund.

**Stärkung des Informationsjournalismus.** Massnahmen, die den Informationsjournalismus in der Schweiz stärken, sind zu begrüßen, um den Zugang zu qualitativ hochwertiger Information für die Bevölkerung sicherzustellen. Der öffentliche Rundfunk spielt dabei eine zentrale Rolle, weil er eine hohe Qualität liefert und viele Menschen erreicht. Dieser Service public sollte aufrechterhalten werden und mit Blick in die Zukunft auch online über Webseiten und Social-Media-Angebote gewährleistet werden. Zudem ist der Zugang zu Qualitätsmedien für ein jüngeres Zielpublikum sicherzustellen, etwa durch (teilweise schon bestehende) vergünstigte Abonnements für Schüler:innen und Studierende.

**Gründung einer (kostenlosen) Medienplattform.** Der Staat oder ein Zusammenschluss von bestehenden Medien könnte eine Infrastruktur zur Verfügung stellen, die von Medien, Journalist:innen und anderen Kommunikator:innen mit (bestehenden) Inhalten befüllt wird. Auf diese Weise könnte man der Bevölkerung einen besseren Zugang zu qualitativ hochwertiger Information ermöglichen. Mit Blick auf die Sicherstellung der Qualität könnte der Zugang auf diejenigen Medien und Journalist:innen und andere Kommunikator:innen beschränkt werden, welche die journalistischen Grundsätze einhalten und sich den Richtlinien des Presserates unterwerfen.

## A. Einleitung und Zielstellung

Der Bericht fokussiert auf die Verbreitung von Desinformation. In der EU wird Desinformation verstanden als «nachweislich falsche oder irreführende Informationen, die *mit dem Ziel* des wirtschaftlichen Gewinns oder der vorsätzlichen Täuschung der Öffentlichkeit konzipiert, vorgelegt und verbreitet werden und öffentlichen Schaden anrichten können». Wir fassen in unserem Bericht den Begriff der Desinformation weiter und verstehen darunter jegliche in die Welt gesetzte oder verbreitete Falschinformation, die ein gesellschaftlich relevantes Thema oder Ereignis betrifft und geeignet ist, schädigende Wirkungen zu entfalten (vgl. Humprecht, 2019). Ob die Erstellung und Verbreitung dabei absichtsvoll geschehen, ist oftmals nicht nachvollziehbar, weshalb wir beide Fälle im Begriff Desinformation integrieren.

In modernen vernetzten Gesellschaften wie der Schweiz werden digital vermittelte Inhalte zunehmend zur wichtigsten Grundlage für Meinungs- und Willensbildungsprozesse (Udris et al., 2022). Vor allem für Jugendliche und junge Erwachsene sind die digitalen Kanäle zur zentralen Informationsquelle geworden. Webseiten, Social Media, Blogs oder Messengerdienste bieten die Möglichkeit, sich aus zahlreichen und diversen Quellen über eine Vielzahl an Themen, Ereignissen und Personen zu informieren. Diese Medien erlauben eine Interaktion mit den Inhalten – Personen können darauf reagieren, Kommentare hinterlassen, und Informationen mit anderen bekannten oder unbekanntenen Personen teilen. Dadurch erfüllen die digital vermittelten Inhalte nicht nur Informationsbedürfnisse, sondern ermöglichen auch sozialen Kontakt, die Zugehörigkeit zu Gruppen sowie das Signalisieren der eigenen sozialen Identität (Van Bavel & Pereira, 2018; Pereira et al., 2023). Digitale Medien sind wegen ihrer potenziellen Reichweite, des unbegrenzten Platzangebots, der weitgehend hürdenfreien Zugänglichkeit und der Anonymität gleichzeitig auch geeignete Kanäle und Katalysatoren für die Verbreitung von Desinformation (Steinebach et al., 2020; Valenzuela et al., 2019; Vosoughi et al., 2018).

Ob und gegebenenfalls in welchem Umfang die Menge an Desinformation in der Schweiz in den letzten Jahren zugenommen hat oder ob sie nur sichtbarer geworden ist, ist eine offene Frage. Fakt ist, dass rund die Hälfte der Schweizer:innen Desinformation aktuell als Problem wahrnimmt (Vogler et al., 2021). Allerdings zeigt die Forschung auch, dass die Schweiz vergleichsweise resilient gegen Desinformation ist (Humprecht et al., 2020; Oehmer et al., 2020). Begründet wird dies mit der Kleinräumigkeit und der entsprechend vergleichsweise hohen sozialen Kontrolle, der wenig polarisierten Gesellschaft, dem starken medialen Service public und dem nach wie vor qualitativ hochwertigen Informationsjournalismus.

Desinformation stellt eine Gefahr für demokratische Gesellschaften dar (McKay & Tenove, 2021). Denn eine funktionierende Demokratie setzt gut informierte Bürger:innen und Vertrauen in die staatlichen und medialen Organisationen voraus. Gerade die Schweiz mit ihrem direkt-

demokratischen System ist auf einen aufgeklärten öffentlichen Diskurs angewiesen. Faktentreue, Transparenz und Fairness sind für die Aushandlung von politischen Entscheidungen unabdingbar. Gestreute Desinformation, um politische oder wirtschaftliche Ziele zu erreichen, könnten in der Schweiz daher potenziell hohen Schaden anrichten. Es stellt sich deshalb die Frage, welche Governance-Massnahmen aus rechtlicher und sozialwissenschaftlicher Perspektive notwendig, umsetzbar und sinnvoll sind, um dem Problem der Desinformation zu begegnen.

Für die Governance von Desinformation lassen sich drei konzeptionell verschiedene Strategien unterscheiden: Erstens kann der Staat durch die Anwendung bestehender und den Erlass von neuen Rechtsnormen versuchen, das Verfassen und Verbreiten von Desinformationen zu unterbinden. Zweitens kann Desinformation auf Stufe der Verbreitenden mit den Mitteln der Ko- und Selbst-Regulierung sowie der Selbstorganisation begegnet werden. Die dritte Möglichkeit setzt auf der Stufe der Nutzer:innen bzw. der Media Literacy an: die Bevölkerung muss in der Lage sein, Desinformation zu erkennen und richtig darauf zu reagieren. Eine umfassende Governance von Desinformation sollte sozial-, kommunikations- und rechtswissenschaftliche Aspekte miteinander verbinden. Bislang fehlt allerdings eine solche, die Disziplinen verbindende Perspektive auf die Governance von Desinformation.

Dieser Bericht zielt auf Wirkungen in zwei Dimensionen: Inhaltlich soll erstmals eine interdisziplinäre Bestandesaufnahme der sozial- und rechtswissenschaftlichen Forschung und der rechtlichen Regelung von Desinformation (in der Schweiz) für die Mikro-, Meso- und Makroebene systematisch aufbereitet werden. Diese Erkenntnisse bilden die Grundlage für die Formulierung von Governance-Optionen, die staatliche, organisationale und individuelle Massnahmen berücksichtigen und gegebenenfalls den Anstoss für politische und gesetzliche Massnahmen geben können. Organisatorisch soll mit dem Forschungsprojekt, das zu diesem Bericht geführt hat, ein Netzwerk von Forschenden geschaffen und ein verstärkter Austausch zwischen Rechts- und Kommunikationswissenschaften etabliert werden, der sich vor allem für die wissenschaftliche Analyse und Begleitung von Media Governance-Massnahmen als tragfähig erweisen dürfte.

Der vorliegende Bericht bereitet zunächst die kommunikationswissenschaftlichen Befunde zur Desinformation mit besonderem Blick auf die Schweiz auf und wird ergänzt durch sozialpsychologische Forschungsergebnisse (**B**). Besonderes Augenmerk wird dabei auf die empirischen Befunde zu den Verfasser:innen und Verbreiter:innen, Kanälen, Inhalten, Nutzer:innen und Wirkungen von Desinformation gelegt. Auf Grundlage dieser Erkenntnisse werden die bestehenden schweizerischen Governance-Mechanismen dargelegt und durch rechtsvergleichende Betrachtungen zum europäischen Recht ergänzt (**C**). Auf dieser Grundlage werden Empfehlungen für Governance-Mechanismen formuliert (**D**). Dem interdisziplinären Charakter dieses Berichts wird auch bei den Governance-Empfehlungen Rechnung getragen. Die

vorgeschlagenen Massnahmen umfassen staatliche Regulierung, Ko-Regulierung und Selbstregulierung sowie Selbstorganisation zur Verhinderung der Verbreitung von Desinformation. Neben diesen direkten Massnahmen zur Verhinderung der Verbreitung werden auch indirekte Massnahmen zur Reduktion der Wirkung von Desinformation empfohlen, etwa die Stärkung der Medienkompetenz.

## **B. Sozialwissenschaftliche Befunde zur Desinformation**

Die Grundlage für die Ausführungen und Überlegungen zur Governance von Desinformation bilden die in der Kommunikationswissenschaft sowie in der Sozialpsychologie diskutierten Befunde zum Phänomen. Der Forschungsstand wird nach dem Kommunikationsmodell von Lasswell (1948) systematisiert und diskutiert:

- Verfasser:innen/Verbreiter:innen – Wer verbreitet Desinformation? (I.)
- Kanäle – Über welche Kanäle wird Desinformation verbreitet? (II.)
- Inhalte – Welche Inhalte sind von Desinformation betroffen (III.)
- Nutzung – Wer ist Desinformation ausgesetzt? (IV.)
- Wirkung – Welche (potenziellen) Effekte erzielt Desinformation? (V.)

Vorab sei darauf hingewiesen, dass ein Grossteil der kommunikationswissenschaftlichen Forschung das Phänomen der Desinformation aus einer konzeptionellen Perspektive behandelt (Bennett & Livingstone, 2018; Egelhofer & Lecheler, 2018; Zimmermann & Kohring, 2018). Empirische Studien, die für ein Land umfassende Erkenntnisse über das Ausmass, die Eigenschaften und Wirkungen von Desinformation liefern, gibt es kaum. Dies liegt erstens an der Vielzahl an möglichen Kanälen und Plattformen, die Desinformation verbreiten können. Diese eignen sich zweitens aufgrund ihres öffentlichen, teilöffentlichen oder auch privaten Settings und der eingeschränkten Datenzugänglichkeit nicht alle in gleichem Masse für empirische Analysen. Zudem sind drittens auf den Plattformen die Verfasser:innen, Verbreiter:innen und Inhalte von Desinformation so heterogen, dass nur ein Bruchteil identifiziert und analysiert werden kann. Die meisten Studien untersuchen deshalb einzelne oder mehrere Beispiele von Falschnachrichten, die von Fact-Checking-Seiten oder bereits bestehenden Listen von Falschnachrichten ausgewählt wurden (Sängerlaub, 2020; Humprecht, 2020; Allcott & Gentzkow, 2017; Steinebach et al., 2020). Die so gewonnenen Erkenntnisse bergen das Risiko, nicht das gesamte Spektrum der verbreiteten Desinformation abzubilden. Die Befunde können zudem in eine politische Richtung oder für einen Themenkomplex verzerrt sein (Allcott & Gentzkow, 2017).

### **I. Befunde zu Verfasser:innen und Verbreiter:innen**

Bei der Betrachtung von Desinformation ist es wichtig, zwischen den Verfasser:innen und Verbreiter:innen zu unterscheiden. Verfasser:innen (wenige) erstellen und veröffentlichen Desinformation erstmals, Verbreiter:innen (fallweise viele), hingegen tragen dazu bei, diese Desinformation durch Teilen, Kommentieren und Liken weiterzuverbreiten. Das Verfassen einer Falschnachricht steht am Anfang des Kommunikationsprozesses; dieses Verfassen kann durch eine bestimmte politische oder ökonomische Absicht geleitet sein. Die Verbreitung ist jedoch ein wesentlicher Aspekt. Sie trägt dazu bei, dass Falschnachrichten in bestimmten Fällen ein (sehr) grosses Publikum erreichen können.

Die Verbreitung wird einerseits durch das Verfassen der Information und durch die Interaktion der Nutzer:innen mit den Inhalten gesteuert; durch die individuelle Reaktion kann eine Nachricht weitergesendet oder anderen Personen mit einer höheren Wahrscheinlichkeit angezeigt werden. Die Verfasser:innen und Verbreiter:innen von Desinformation können dabei Kollektiv- oder Individualakteur:innen sein, die manchmal mit (teil)automatisierten Accounts (bspw. Social Bots) operieren. Andererseits wird die Verbreitung durch die Algorithmen der Social-Media-Plattformen gesteuert, welche oft Nachrichten mit hohem Interaktionslevel bevorzugen und damit wiederum für eine erhöhte Sichtbarkeit sorgen (Humprecht, et al., 2021). Insgesamt konnte gezeigt werden, dass Desinformation sich in den sozialen Medien durch die Verwendung von stark emotionalen und moralbesetzten (Brady et al., 2020, Brady et al., 2017) sowie negativen Begriffen (Robertson et al., 2023) stark verbreiten kann (Vosoughi et al., 2018). Wenn mit Desinformation mehr interagiert wird und Algorithmen Informationen mit hohem Engagement bevorzugt anzeigen, kann dies dazu führen, dass die von Plattformen verwendeten Algorithmen im Ergebnis zur Verbreitung von Desinformation beitragen (siehe Rathje, Van Bavel, & van der Linden, 2021, für eine gleiche Argumentation bei polarisierten Inhalten).

#### 1. Verfasser:innen

Mit dem Erstellen von Desinformation werden oft politische, gesellschaftliche oder ökonomische Intentionen verbunden wie bspw. die Destabilisierung von Gesellschaften durch das Untergraben der Glaubwürdigkeit in politische Institutionen oder durch den Versand von polarisierenden Informationen (Bennett & Livingston, 2018; Mourão & Robertson, 2019) oder das Generieren von Werbegeldern mittels stark beachteten Falschnachrichten (Hughes & Waismel-Manor, 2021).

In der Schweiz werden journalistische Medien kaum mit der Verbreitung von Desinformation assoziiert (Vogler et al., 2021). Auch Behörden wurden in der Befragung eine erhöhte Glaubwürdigkeit zugesprochen. Dies zeugt grundsätzlich von einem hohen Vertrauen in die Institutionen in der Schweiz.

Wie die Studie von Vogler et al. (2021) auf der Basis einer repräsentativen Befragung der Schweizer Bevölkerung zeigte, wird das Verfassen von Desinformation eher Individual- als Kollektivakteure:innen zugeschrieben. Aktivist:innen, Politiker:innen und gewöhnliche Bürger:innen werden neben Social Bots am ehesten als Verfasser:innen von Desinformation wahrgenommen. Zudem sind in der Regel wenige Nutzer:innen für den grossen Teil von Desinformation auf Social Media verantwortlich. So zeigte eine Analyse für die USA, dass im Frühjahr 2021 nur 12 User:innen für 65% der erfassten Falschnachrichten zur Covid-19-Impfung auf Facebook verantwortlich waren (Center for Countering Digital Hate, 2021). In diesem Kontext wird auch die Bedeutung von sogenannten Internet-Trollen beim Erstellen von Desinformation diskutiert (Linvill & Warren, 2018). Institutionalisierte Gruppen dieser Trolle (sog. Troll-Armeen oder Troll-Farmen) werden nachweislich eingesetzt, um die politische Meinungsbildung zu beeinflussen (Linvill & Warren, 2018). Genauso wie bei Social Bots wird jedoch die Wirkung von Trollen zu diesem Zweck infrage gestellt. In der Schweiz sind bisher keine grösseren Fälle der Einflussnahme auf den politischen Prozess durch Troll-Farmen bekannt.

Eine Manipulation des Online-Diskurses durch ausländische Regierungen als Verfasser:innen von Desinformation wird häufig diskutiert. Im Fokus stehen dabei aktuell Einflussversuche russischer Akteur:innen. Diese stehen aufgrund des Kriegs in der Ukraine wieder stärker im Fokus, allerdings eher im Zusammenhang mit geopolitisch sehr wichtigen Ländern wie die USA oder Deutschland und weniger mit der Schweiz (Bennett & Livingston, 2018; Borucki & Meisberger, 2019; Linvill & Warren, 2020). Doch auch in der Schweiz werden Vorfälle mit Desinformation aus pro-russischen Quellen diskutiert. Im Herbst 2022 tauchten im Netz Bilder eines angeblichen Plakates des Bundes auf, mit dem zum Denunzieren von Nachbar:innen aufgerufen wird, die zu viel Strom verbrauchten. Obwohl die genaue Urheberschaft nicht geklärt wurde, führten die Spuren nach Russland (McEvily, 2022; Mäder, 2022). Es stellt sich aber die Frage, ob die suggerierten Einflussversuche mit der tatsächlichen Bedeutung übereinstimmen. Eine Studie von Sänglerlaub et al. (2018) konnte zeigen, dass bei der Verbreitung von Desinformationen auf Online-Plattformen und Webseiten in Deutschland im Vorfeld der Bundestagswahl 2017 ausländische Akteur:innen eine vergleichsweise untergeordnete Rolle gespielt haben.

Gemäss der Studie von Vogler et al. (2021) wird alternativen Nachrichtenmedien eine besondere Relevanz als Verfasser:innen und Verbreiter:innen von Desinformationen zugeschrieben. Als alternative Medien werden Nachrichtenangebote verstanden, die sich selbst als Vertreter:innen einer Gegenposition zu den sogenannten Mainstream-Medien bezeichnen (Holt, Figenschou, & Frischlich, 2019; Schwaiger, 2021). Dazu zählen bspw. in der Schweiz die Portale uncut-news.ch, legitim.ch oder auch lesobservateurs.ch und internationale Portale wie Breitbart, RT oder Sputnik News (Schwaiger, 2021). Allerdings haben alternative Medien

in der Schweiz eine geringe Reichweite und werden nur von sehr wenigen Menschen regelmässig genutzt (Vogler et al., 2021).

Immer wieder werden auch Unternehmen oder Branchen der Verbreitung von Desinformation zur Durchsetzung ihrer Interessen bezichtigt, bspw. die Ölindustrie im Zusammenhang mit dem Klimawandel (Franta, 2021) oder die Tabakindustrie zu schädlichen Auswirkungen des Rauchens (McKee, 2017). Auch in der Schweiz sieht rund ein Drittel der Bevölkerung (30%) Unternehmen und Konzerne als Verfasser:innen von Desinformation.

## 2. Verbreiter:innen

Der Verbreitung von Desinformation kommt eine hohe Bedeutung zu. Denn erst mit der hohen Verbreitung kann Desinformation zu einem gesamtgesellschaftlichen Problem werden. Während das Verfassen von Desinformation meist auf wenige Individual- und Kollektivakteur:innen zurückzuführen ist, nehmen an der Verbreitung oft eine Vielzahl von Akteur:innen teil. Nicht alle teilen dabei die Falschinformationen absichtlich oder bewusst.

Ein hoher Social-Media-Konsum erhöht die Bereitschaft, Desinformationen zu verbreiten und zu liken (Vraga & Tully, 2021; Rossini et al., 2020; Humprecht et al., 2021). Der Konsum von traditionellen Medien mindert hingegen die Bereitschaft zur Verbreitung von Desinformation. Humprecht et al. (2021) haben festgestellt, dass geringes Nachrichtenvertrauen (vor allem in Service-Public-Medien) und ein homogenes Medienrepertoire positiv mit dem Verbreiten von Desinformation korrelieren (vgl. auch Zimmermann & Kohring, 2020). Dies gilt jedoch nicht für ihre Daten zur Schweiz. Zudem bleibt hier unklar, was als Ursache und was als Folge betrachtet werden kann: So wäre auch denkbar, dass Personen, die eine hohe Affinität zu Desinformationen aufweisen, sich aus genereller Skepsis gegenüber Nachrichtenmedien nur gezielt einigen wenigen Nachrichtenangeboten zuwenden.

Auf Individualebene haben sich die Rezipient:innen von alternativen Nachrichtenmedien als besondere Katalysatoren von Desinformation erwiesen: Sie interagierten (liken, teilen, kommentieren) deutlich häufiger mit Desinformation, die ihnen präsentiert wurde und erhöhten so deren Reichweite und Sichtbarkeit (Humprecht, 2020; Humprecht, 2021; Ständer et al., 2021).

Studien zeigen auch einen Zusammenhang zwischen polarisierender politischer Einstellung (meist rechtsextrem) oder der Präferenz für populistische Parteien und der Verbreitung von Desinformationen (Steinebach et al. 2020). Humprecht et al. (2021) konnten diese Annahme für Frankreich, Deutschland, Grossbritannien und die USA, jedoch nicht für die Schweiz bestätigen. Auch Funktionäre und Mitglieder populistischer Parteien sind häufig verantwortlich für die Verbreitung von Desinformation (Sängerlaub et al., 2018; Steinebach et al., 2020). Steinebach et al. (2020) resümierten auf der Basis ihrer Analyse, dass deutschsprachige Fake News im Untersuchungszeitraum zwar nicht ausschliesslich, aber doch überwiegend von

rechtsgerichteten Medienakteur:innen produziert wurden. Guess et al. (2019) zeigten zudem, dass es vornehmlich Personen über 65 Jahre sind, die Desinformation wahrscheinlicher teilen.

Aus den vorhandenen Studien lässt sich für die Schweiz ableiten, dass Personen, die regelmässig Desinformation verbreiten, eine Minderheit sind. Die beschriebenen Merkmale treffen nur auf wenige Personen zu.

Neben der Mediennutzung und der politischen Einstellung zeigen Forschungsergebnisse aus dem englischsprachigen Raum, dass eine Reihe von veränderbaren Faktoren das Erkennen und Teilen von Desinformation beeinflussen können. Vorwissen darüber und erste Erfahrungen damit wie Desinformation erstellt und besonders attraktiv gemacht werden kann, beeinflusst späteres Erkennen von Desinformation sowie die korrekte Unterscheidung von Information und resultiert in einer besseren Qualität der Entscheidungen in Bezug auf das Teilen von Informationen (vgl. Forschung um sogenannte «Impfung», vgl. Roozenbeek et al., 2022). In Medienkompetenz-Trainings vermittelbare Kompetenzen wie «kritisches Ignorieren» (z.B. laterales Lesen/Prüfen von Informationen sowie die Faustregel «Trolle» nicht zu füttern; Kozyreva et al., 2022) sowie aktives und offenes Denken (Newton et al., 2023) wirken sich ebenfalls positiv auf die Informationsnutzung aus.

Aktuelle Forschung zeigt ausserdem, dass die Nutzer:innen auch durch kurze Interventionen positiv beeinflusst werden können und bessere Entscheidungen in Bezug auf das Teilen von Desinformation und Information treffen. Generell besteht bei Personen eine hohe Zustimmung, dass es wichtig ist, nur richtige (orig. «accurate») Informationen weiterzuverbreiten (Pennycook et al., 2021). Bei einer Nutzung von digitalen Inhalten bzw. *sozialen* Medien steht das Bedürfnis nach Richtigkeit gegebenenfalls weniger stark im Vordergrund. Das Bedürfnis nach Zugehörigkeit und das Sicherstellen einer positiven sozialen Identität können Motivationen sein, gezielt Informationen auszuwählen und zu teilen, bei denen erwartet wird, dass das eigene Umfeld positiv darauf reagiert (Zugehörigkeit) und die eigene soziale Identität (z.B. die eigene Partei oder Interessensgruppe) gestärkt wird (van Bavel et al., 2023/in prep). Das Teilen von Desinformation kann auch in Fällen auftreten, in denen den Verbreiter:innen durchaus bewusst ist, dass es sich um Desinformation handelt – sie teilen die Information dennoch, um soziale Kosten (z.B. Ausschluss) gering zu halten (Lawson et al., 2023). Wird durch kurze Erinnerungsaufgaben die Aufmerksamkeit auf die Richtigkeit von Informationen gelenkt, führt dies wiederum dazu, dass Personen deutlich stärker zwischen Desinformation und Information unterscheiden und dies beim Teilen auch berücksichtigen können (vgl. Pennycook et al., 2021, Rathje et al., 2023, für eine Replikation über 16 Länder siehe Arechar et al., preprint). Massnahmen, die die Medienkompetenz verbessern und einen stärkeren Fokus auf die Richtigkeit (im Vergleich zu sozialen Motiven) lenken, erscheinen daher vielversprechend.

### 3. Verbreitung über (teil)automatisierte Accounts

Für das Verbreiten von Desinformation werden auch (teil)automatisierte Accounts wie bspw. Social Bots oder hyperaktive Accounts verantwortlich gemacht (Bennett & Livingstone, 2018; Bessi & Ferrara, 2016; Ferrara, 2017; Howard & Kollanyi, 2016, Hegelich & Thielges, 2019; Shao et al., 2018). Social Bots haben im Kontext der Verbreitung von Desinformation eine hohe öffentliche Aufmerksamkeit erhalten. Obwohl Social Bots auch Inhalte erstellen können, wenn sie entsprechend programmiert sind, liegt ihre Rolle vor allem in der Verbreitung von Inhalten. Sie sollen in Debatten auf Social Media gewisse Frames oder Positionen verstärken oder abschwächen, so die Hoffnung ihrer Verfasser:innen.

Im Kontext von Desinformation werden unter dem Begriff Social Bots Computerprogramme verstanden, die in sozialen Netzwerken oder Onlineforen automatisiert Inhalte produzieren und verbreiten. Die meisten Bots sind so programmiert, dass sie Inhalte (z.B. Tweets) mit bestimmten Stichworten (z.B. Hashtags) oder Inhalte von spezifischen Accounts weiterverbreiten. Damit können sie die Bedeutung von Hashtags oder Accounts künstlich erhöhen. Andere Bots werden so programmiert, dass sie online verfügbare Inhalte nach definierten Regeln zusammenstellen und als Post veröffentlichen. Beispiele dafür sind Accounts, die regelmässig Wetterdaten, Sendezeit oder Sportresultate veröffentlichen. Manchmal werden sie so programmiert, dass sie versuchen, unter Vortäuschung menschlicher Identität bzw. Nachahmung menschlichen Verhaltens, mit menschlichen Nutzer:innen zu interagieren. Wenn Interaktionen stattfinden, dann aber meist mit anderen Bots (sogenannten Bot-Netzwerken). Aktivitäten von Social Bots können (möglicherweise auch) Meinungen oder Verhalten im Zusammenhang mit politischen Prozessen, meist Wahlen und Abstimmungen, beeinflussen (Ferrara et al., 2016; Stieglitz, Brachten, Ross, et al., 2017; Oehmer & Pedrazzi, 2020).

Bot-Aktivitäten konnten für verschiedene Länder nachgewiesen werden (z.B. Deutschland: Neudert et al., 2017; Keller & Klinger, 2019; Großbritannien: Howard & Kollanyi, 2016), insbesondere in Konfliktsituationen z.B. in der Ukraine (Hegelich & Janetzko, 2016) oder Syrien (Abokhodair et al., 2015). Dem stehen Erkenntnisse entgegen, die zeigen, dass die Reichweiten von Social Bots vorsichtig zu bewerten sind (Gallwitz & Kreil, 2021; Rauchfleisch & Kaiser, 2021; Reuter, 2019; Stier & Jungherr, 2019), unter anderem weil Bots wenig Follower haben und ihre Kommunikationsaktivitäten in empirischen Messungen zwar erscheinen, die Informationen aber von Nutzer:innen nicht gesehen bzw. nicht geglaubt werden. Viele Bot-Aktivitäten verstossen zudem gegen die Nutzungsbedingungen von Plattformen, weshalb sie gelöscht werden. Es ist also schwer, Bot-Netzwerke zur wirkungsvollen Verbreitung von Desinformation zu betreiben.

Meist wurden die Inhalte, die durch Social Bots verbreitet wurden, nicht analysiert. Aussagen über das Ausmass von durch Social Bots verbreiteter Desinformation sind damit nur bedingt

möglich. Die wenigen Studien für die Schweiz zeigen, dass die prominente Thematisierung von Social Bots in der Öffentlichkeit vermutlich nicht mit der tatsächlichen Problematik einhergeht. Viele Diskurse fokussieren auf Twitter, das in der Schweiz eher wenig genutzt wird. Studien für die Schweiz fanden bislang keine bis sehr wenig Hinweise für problematische Bot-Aktivität (Rauchfleisch & Vogler, 2018) auch nicht mit Bezug zu Desinformation während der Covid-19-Pandemie (Rauchfleisch et al., 2020). Aus dem gegenwärtigen Erkenntnisstand lässt sich damit keine hohe Priorisierung von Social Bots für die Governance von Desinformation ableiten.

## **II. Befunde zu Kanälen**

Im Fokus dieses Kapitels liegen die Kanäle, die zur Verbreitung von Desinformation verwendet werden. Kernargument ist dabei, dass Verfasser:innen oder Verbreiter:innen Desinformation veröffentlichen und streuen, während Kanäle die Infrastruktur für die Verbreitung zur Verfügung stellen und diese ggf. durch ihre Algorithmen sogar begünstigen. Diese Unterschiede im Wissen und Prozedere sind mit jeweils unterschiedlichen Governance-Fragen verbunden: Bei Verfasser:innen und Verbreiter:innen wird vor allem diskutiert, wie man sie davon abhalten kann Desinformation zu veröffentlichen und wie man sie befähigen kann, Desinformation zu erkennen und nicht weiter zu verbreiten. Bei Kanälen ist hingegen zu klären, inwiefern sie für die auf ihren Kanälen verbreitete Desinformation verantwortlich gemacht bzw. zur Beteiligung an Massnahmen gegen die Verbreitung von Desinformation, bspw. durch eine Anpassung der Infrastruktur (Algorithmen), verpflichtet werden können und sollen.

Wie Studien zeigen, werden vor allem Social Media wie Facebook, Instagram, TikTok, Twitter und Youtube mit der Verbreitung von Desinformation assoziiert (Allcott et al, 2019; Vogler et al. 2021). Zunehmend rücken aber – bedingt durch das weitgehende Fehlen von (Selbst)Regulierungsmassnahmen und auch aufgrund ihrer Reputation als sicherer Ort für Personen, die von anderen Plattformen gesperrt wurden (Gill & Rojas, 2020; Rogers, 2020; Urman & Katz, 2020; Walther & McCoy, 2021) – Messengerdienste wie Telegram und WhatsApp als Verbreitungskanäle von Desinformation in den Vordergrund. So brachten 28% der Teilnehmenden einer repräsentativen Umfrage in der Schweiz Messengerdienste mit Desinformation in Verbindung (Vogler et al., 2021).

Bei Kommunikation über Messengerdienste handelt es sich weitestgehend um private Kommunikation. Gruppenmitglieder oder Organisatoren bestimmen selbst, wie öffentlich die Debatten und Inhalte sind. Sie können als öffentliche Gruppe oder komplett abgeschlossene Gruppe aufgesetzt werden. Dies steht im Gegensatz zur Kommunikation in journalistischen Medien oder auf grossen Plattformen wie Facebook, Instagram und Twitter oder der Videoplattform YouTube, die weitestgehend öffentlich ist. Diese Privatisierung des öffentlichen Diskurses (Eisenegger, 2021) ist ein Problem für den demokratischen Diskurs. Radikalität und

Verbreitung von Desinformation verschiebt sich in den privaten Bereich (Rogers, 2020; Tantuco, 2021). Allfällige problematische Entwicklungen sind weniger visibel, bleiben aber bestehen. Das macht Messengerdienste zu einer grossen Herausforderung für die Governance von Desinformation. Forschung zu öffentlichen Gruppen in Messengerdiensten zeigt nur die «Spitze des Eisbergs». Problematische Entwicklungen wie die Verbreitung von Desinformation bleiben hingegen oftmals im Verborgenen, das heisst, sie finden in geschlossenen Gruppen statt.

Angebote von Bedeutung sind WhatsApp, Facebook Messenger und Telegram. In der Schweiz hat Telegram im Zuge der Coronapandemie an Bedeutung gewonnen, vor allem zur Mobilisierung und Artikulation von Protest gegen Coronamassnahmen (Nachtwey et al., 2020). Die Nutzung ist noch eher gering, seit 2020 aber mit Bedeutungszuwachs (Udris et al., 2022). Vor allem die Visibilität und Bekanntheit der Messenger-App ist gestiegen, mitunter auch als Folge der prominenten Thematisierung in etablierten journalistischen Medien. Der wachsende Erfolg von Telegram wird unter anderem durch die fehlenden Einschränkungen wie unbegrenzte Gruppengrössen und die Möglichkeit zur Ein-Weg-Kommunikation, erklärt (Mossburger, 2021). «Letztlich erinnert Telegram (...) viel mehr an ganz klassische soziale Medien wie Twitter, Instagram oder Facebook, wo User einem anderem folgen und dessen Posts, seien es Links, Videos oder Fotos, eher passiv konsumieren» (Mossburger, 2021).

Telegram wird oft in Verbindung mit der Verbreitung von Desinformation und Verschwörungstheorien gebracht. Befunde zur Schweiz sind spärlich und teilweise anekdotischer Natur. Eine Ausnahme sind die Befunde einer Befragungsstudie von Vogler et al. (2021). Die Untersuchung zeigt unter anderem, dass die Nutzung von Messengerdiensten mit höherer Affinität gegenüber Verschwörungstheorien und der Nutzung von alternativen Medien korreliert. Weitere wichtige Erkenntnisse zur Schweiz liefern die ethnografischen Studien von Nachtwey und Frei (Nachtwey et al., 2020). Sie identifizieren Telegram als zentralen Kanal für Kommunikation und Mobilisierung für Proteste gegen die Coronamassnahmen. Ergänzend kann die Analyse der zentralen Schweizer Telegram-Gruppen in der Neuen Zürcher Zeitung genannt werden (Schoop et al., 2021).

Der Fokus in der internationalen Forschung liegt auf WhatsApp. Telegram ist hingegen (noch) wenig erforscht. Die Forschung fokussiert auf den Einfluss auf politische Einstellungen (Gill und Rojas, 2020; Rossini et al., 2020; Rossini et al., 2021) und Prozesse, unter anderem im Vorfeld von Wahlen, bspw. prominent in Brasilien (Abdin, 2019). In jüngerer Zeit ist auch ein starker Fokus auf die Erforschung von Desinformation und von Protesten im Zusammenhang mit der Coronapandemie zu sehen (Bowles et al., 2020; Elias & Catalan-Matamoros, 2020; mit Bezug zur Schweiz: Frei & Nachtwey, 2020; Schoop et al., 2021; Vogler et al., 2021). Eine wichtige Rolle spielt in diesem Zusammenhang die mobile Mediennutzung. Eine neuere Studie zeigt, dass die Kombination von mobilen Geräten und Messenger-Apps aufgrund des privaten

Settings (privates Gerät, private Kommunikation) das Teilen von Desinformation begünstigt (Herrero-Diz et al., 2021).

Neben den Messengerdiensten fokussiert Forschung aus dem englischsprachigen Raum besonders auf Social-Media-Plattformen und deren Infrastruktur. Algorithmen auf diesen Plattformen bevorzugen Inhalte mit vielen Interaktionen, die dann wiederum mehr Nutzer:innen angezeigt werden. Mit Desinformation bzw. Informationen, die eine verzerrte Wahrnehmung von sozialen Normen kreieren (wie Unhöflichkeit oder öffentliche Empörung; González-Bailón & Lelkes, 2022), wird oftmals stark interagiert, sodass diese im Vergleich zu neutralen Informationen häufiger und breiter verbreitet werden (Vosoughi et al., 2018; Brady et al., 2020, Brady et al., 2017). Durch die Infrastruktur von sozialen Medien (besonders durch die Möglichkeit, Inhalte zu teilen) werden weiterhin eher soziale Motive in den Vordergrund gerückt, während das Bedürfnis nach Richtigkeit weniger salient zu sein scheint (vgl. Epstein et al., 2023). Soziale Motive können erfüllt werden, indem Informationen geteilt werden, die die eigene Identität positiv herausstellen oder von als wichtig empfundenen anderen Personen positiv bewertet werden (unabhängig davon, ob die Information wahr oder falsch ist).

Auch durch das Bereitstellen von Interaktionsmöglichkeiten wie «Gefällt mir»-Schaltflächen an Stelle von «Vertrauenswürdig» / «Nicht vertrauenswürdig» (vgl. Globig et al., 2023) und die kaum vorhandene Betonung der Wichtigkeit von Richtigkeit (z.B. Rathje et al., 2023) können Nutzer:innen auf Social-Media-Plattformen dazu verleiten, dass mehr Desinformation verbreitet wird. Durch eine Anpassung des Algorithmus sowie der Plattform-internen Infrastruktur könnte die Bedeutsamkeit von Richtigkeit in den Vordergrund gestellt werden und somit aktiv das Verhalten von Nutzer:innen positiv beeinflusst werden.

Darüber hinaus könnte durch eine Einflussnahme auf den Feed von Nutzer:innen die Diversität innerhalb angezeigter Nachrichten erhöht werden. Aktuell enthalten Feeds vor allem Informationen von Personen, zu denen eine Zugehörigkeit bei den Nutzer:innen besteht. Meistens umgeben sich Menschen eher mit Personen, die einen ähnlichen Hintergrund und ähnliche Einstellungen haben, sodass viele der angezeigten Informationen zu den eigenen Einstellungen passen (vgl. González-Bailón & Lelkes, 2022 bzw. allgemeiner das Similarity-Attraction Paradigma; vgl. Byrne, 1961, 1971). Forschung zeigt allerdings auch, dass etwa 24 – 35% der Informationen aus anderen Bereichen kommen können (González-Bailón & Lelkes, 2022), sodass ein gewisses Mass an Informationsdiversität gegeben ist. Diese Zahl könnte, wenn gewollt, durch gezielte Vorschläge, wem eine Person folgen könnte, erhöht werden. Ein Feed mit mehr Informationsdiversität könnte dann Nutzer:innen ermöglichen, sich eine differenziertere Meinung zu Themen zu bilden.

### **III. Befunde zu Themen, Inhalten und Form**

Wie das Beispiel des Ukraine Konflikts zeigt, ist auch Desinformation einem zeitlichen Wandel ausgesetzt, zumindest was ihre Sichtbarkeit und die wahrgenommene, gesamtgesellschaftliche Bedeutung betrifft. Während vor einigen Monaten im Kontext von Covid vornehmlich innenpolitische Themen und Gesundheitsthemen im Fokus standen, liegt die Aufmerksamkeit im Zusammenhang mit dem Krieg in der Ukraine nun die internationale Komponente von Desinformation (im Sinne klassischer Propaganda). Trotz der aktuellen Entwicklung lassen sich in der Forschung einige Schwerpunkte im Zusammenhang mit Desinformation erkennen.

Inhaltlich lässt sich Desinformation häufig den Themenschwerpunkten «Innere Sicherheit (Kriminalität und Justiz)» und – meist auch in Kombination – «Migration und Flucht» zuordnen (Bennett & Livingstone, 2018; Humprecht, 2019; Sänglerlaub, 2018; Steinebach et al., 2020). Damit spiegeln die Themen auch klassische Schwerpunkte populistischer und rechtskonservativer Akteur:innen (Steinebach et al., 2020), die auch als massgebliche Quellen von Desinformation identifiziert worden sind. Für die USA haben Allcott & Gentzkow (2017) im Rahmen einer Untersuchung der auf Fact-Checking-Seiten veröffentlichten Desinformation bspw. feststellen können, dass diese sich dem rechtskonservativen Spektrum zuordnen lässt und sich in klarer Unterstützung zum damaligen Präsidenten Donald Trump positioniert. Daneben kann Desinformation auch im Zusammenhang mit Gesundheitsthemen festgestellt werden (Brennan et al., 2021; Humprecht, 2019). Welches Themenspektrum jeweils durch Desinformation adressiert wird, hängt – wie eine ländervergleichende Analyse von Humprecht (2019) zeigt – massgeblich auch von den jeweils national diskutierten Themenagenden ab.

Desinformation, die sich dem Themenfeld Wirtschaft widmet, wurde in der Forschung hingegen kaum thematisiert (Franta, 2021; Johann & Wagner, 2020). Dies muss jedoch nicht zwingend ein Beleg dafür sein, dass in diesem Bereich wenig Desinformation existiert. Vielmehr kann dies auch eine Folge der methodischen Herangehensweise vieler Studien (Allcott & Gentzkow, 2017; Brennan et al., 2021; Humprecht, 2019; Sänglerlaub et al., 2018) sein, die meist nur Inhalte analysieren, die auf Fact-Checking-Seiten zusammengetragen wurden und nicht repräsentativ für den Themenhaushalt einer Gesellschaft sind.

Desinformation zu unterschiedlichen Themen kann in verschiedenen Formen aufbereitet bzw. verbreitet werden. Klassischerweise als Text, der sich ohne grossen Aufwand produzieren und über verschiedene Kanäle einfach verbreiten lässt. Viele emotionsbezogene Worte und moralische Inhalte erhöhen dabei die Aufmerksamkeit in Bezug auf die Inhalte sowie die Interaktion mit diesen (Brady et al., 2020, Brady et al., 2017). Mit der technischen Entwicklung rückt immer stärker auch Desinformation in Form von Audio- und Videoinhalten in den Fokus (Brennan et al., 2020; Dan et al., 2021). Manipulierte oder aus dem Kontext gerissene Bilder und Videos werden verwendet, um Falschnachrichten besonders echt und glaubwürdig wirken zu lassen. Aus der sozialpsychologischen Forschung um den Truthiness-Effekt ist bekannt, dass Aussagen, die um nicht-informative Bilder ergänzt werden (also Bilder ohne weiteren

Informationsgehalt), eher geglaubt werden als Aussagen ohne die Bilder (Newman & Zhang, 2021) – die Anreicherung von Desinformationen mit Bildern ist daher besonders besorgniserregend. Jüngst wurden auch sogenannte Deep Fakes prominent diskutiert (Vaccari & Chadwick, 2020). Deep Fakes sind realistisch wirkende Bilder oder Videos, bspw. mit prominenten Personen, die oft durch Techniken der künstlichen Intelligenz (KI) erstellt, abgeändert oder verfälscht werden.

#### **IV. Befunde zur Nutzung**

Dass Desinformation eine potenziell schädigende Wirkung entfalten kann, setzt voraus, dass sie überhaupt wahrgenommen wird (wobei jedoch deutlich gemacht werden muss, dass die bloße Wahrnehmung nicht mit der Nutzung oder dem Glauben von Desinformation gleichzusetzen ist).

Von (bekannten) Beispielen abgesehen, ist es oftmals schwer zu bewerten, wie Desinformation tatsächlich genutzt wird und ob Desinformation mit Verhalten in kausalem Zusammenhang steht. Je nach Methoden und analysiertem Land lassen sich zu diesen Fragen unterschiedliche Erkenntnisse gewinnen: In der Einschätzung der Teilnehmenden einer Schweizer Befragungsstudie (Vogler et al., 2021) stiessen knapp ein Viertel (25%) der Befragten oft oder sehr oft auf Desinformation und rund ein Fünftel (21%) schätzte das Ausmass an Desinformation in der Schweiz als gross oder sehr gross ein. Für Deutschland konnten Sänglerlaub et al. (2018) nur eine moderate Reichweite von Desinformation feststellen (vgl. auch Marchal et al., 2019; Neudert et al., 2017). Für die Wahlkampfphase in den USA kalkulierten Allcott und Gentzkow (2017) auf der Basis von Befragungsdaten und Daten zu Interaktionen auf Facebook, dass jeder Erwachsene mit mindesten einer Desinformation in Kontakt geraten ist.

Diese Studien, die die Reichweite von Desinformation messen, haben auf Befragungsdaten (Vogler et al. 2021; Allcott & Gentzkow, 2017) oder die Analyse der Interaktionslevel (Likes, Shares, Comments; vgl. Sänglerlaub, 2020; Allcott et al., 2019; Allcott & Gentzkow, 2017) der Beiträge, die Desinformation enthalten, zurückgegriffen. Beide methodischen Herangehensweisen bieten nur Näherungswerte zur tatsächlichen Reichweite von Desinformation, die auf Basis dieser Ansätze eher überschätzt wird.

#### **V. Befunde zur Wirkung**

Welche Wirkungen Desinformation haben kann, ist nicht hinreichend geklärt. Bisher werden vor allem maligne Effekte im Bereich von Politik, Wirtschaft, Wissenschaft und Gesundheit auf der Makro-, Meso- und Mikroebene diskutiert: So zählt bspw. der sicherheitspolitische Bericht des Bundesrates Desinformation ausländischer Akteur:innen zu den zentralen Bedrohungslagen der Schweiz. Sie würde genutzt, «um politische Prozesse zu beeinflussen

oder zu sabotieren, die Glaubwürdigkeit von Institutionen und Medien anzugreifen und Zweifel an Informationen schlechthin zu säen» (S. 8); Szenarien, die seit dem Krieg in der Ukraine wieder an Bedeutung gewonnen haben. Aber auch innenpolitisch wird Desinformation als Bedrohung wahrgenommen. In Zeitungsartikeln wird mit Überschriften wie «Bedrohen Fake News die direkte Demokratie?» (20 Minuten, 16.11.21) die Furcht vor einem Einfluss von Desinformation auf Wahlen und Abstimmungen zum Ausdruck gebracht. Zudem wird ein Zusammenhang zwischen der Rezeption von Desinformation und dem Vertrauensverlust in politische und gesellschaftliche Institutionen sowie dem nachlassenden gesellschaftlichen Zusammenhalt konstatiert.

Im Gesundheitsbereich wird das Verweigern notwendiger und sinnvoller medizinischer Massnahmen wie bspw. das Impfen gegen COVID-19 aufgrund von Desinformation besonders kritisch gesehen (vgl. Bund, 04.06.2021). Traurige Berühmtheit hat in diesem Kontext bspw. die in der hoch angesehenen medizinischen Fachzeitschrift «The Lancet» publizierte und mittlerweile widerrufen Studie zum Zusammenhang der Impfung gegen Mumps, Masern und Röteln und psychischen bzw. Entwicklungserkrankungen erreicht. Die in der Studie vorgestellten Zusammenhänge wurden als unwissenschaftlich bewertet, dennoch wurden die Inhalte in den Medien weit gestreut und sie stehen in zeitlichem Zusammenhang mit einem Rückgang der Impfquote bei Kindern. Dies hat wiederum zu einer Zunahme an Erkrankungen geführt und kann eine Gefahr besonders für diejenigen bedeuten, die auf eine hohe Impfquote in der Gesellschaft angewiesen sind (vgl., van der Linden, 2023).

Aus ökonomischer Perspektive werden Reputationsverluste und ein Einbrechen der Aktienkurse von Unternehmen aufgrund von Desinformation als mögliche Bedrohungsszenarien diskutiert («Baukonzern-Aktie stürzt nach Fake-News», handelszeitung.ch, 23.11.2016).

Insgesamt zeigt die Persuasionsforschung geringe und zielgruppenspezifische Effekte von Mediennutzung auf individuelle Einstellungen. Es gibt wenig Gründe anzunehmen, dass die im Bereich der Desinformation anders ist (Allcott & Gentzkow, 2017; Bail et al., 2020). Allerdings gibt es einzelne Beispiele von desinformativen Inhalten, die hohe Reichweite erhalten, weil sie von Medien aufgegriffen und verteilt werden (siehe Veränderung von Impfquoten aufgrund von unwissenschaftlichen Studien; van der Linden, 2023). Kommt ausserdem noch die auf den sozialen Medien mögliche Zielgruppen-spezifische Ausrichtung und Formulierung von Informationen (sog. Micro-Targeting), könnte es sein, dass sich Effekte von Desinformation potenzieren (vgl. van der Linden, 2023, bzw. Haenschen, 2022, und Walker et al., 2020, zu experimentellen Studien in diesem Bereich).

Die Wirkungen von Desinformation sollten somit eher vorsichtig eingeschätzt werden (Adams et al., 2023). Zudem können Wirkungsgrade sehr stark in Abhängigkeit der konsumierenden

Person variieren – Personen, die vielleicht erste Zweifel am politischen Vorgehen haben, können bspw. offener für Desinformation in diesem Kontext sein und dadurch in ihrem Denken beeinflusst werden.

Im folgenden Kapitel werden (potenzielle) Wirkungen von Desinformation beleuchtet. Dabei wird zwischen Wirkungen auf Individual- (1.) und gesellschaftlicher Ebene unterschieden (2.). So können Faktoren identifiziert werden, die eine Wirkung von Desinformation allenfalls begünstigen könnten.

### 1. Wirkungen auf Individualebene

Generell geht die Forschung von eher geringen Effekten von Desinformation auf der Individual-ebene aus (z.B. Bail et al., 2020). Auf individueller Ebene haben sich Studien meist im Rahmen von Befragungsstudien im Experimentaldesign mit der Frage auseinandergesetzt, ob Desinformation Glauben geschenkt wird (Allcott & Gentzkow, 2017; Holzer & Sengl 2020; Hohlfeld 2020). Studien zeigen, dass Desinformation von Personen wahrscheinlicher für wahr erachtet wird, die folgende nach a) Mediennutzung, b) politischer Einstellung & Engagement und c) soziodemographischen Merkmale systematisierte Charakteristiken aufweisen:

Desinformation wirkt eher bei Menschen, die Social Media (Allcott & Gentzkow, 2017) oder alternative Medien (Holzer & Sengl 2020; Ständer et al. 2021) als Informationsquelle nutzen und etablierten Nachrichtenmedien nicht vertrauen (Zimmermann und Kohring, 2020). Weiter zeigt sich, dass Desinformation bei wiederholter Konfrontation eher geglaubt wird (vgl. Hohlfeld 2020, Pennycook et al., 2018).

Desinformation erzeugt eher eine Wirkung bei Menschen, deren Einstellung und Weltanschauung mit der verbreiteten Desinformation korrespondiert (Allcott & Gentzkow, 2017; Sänglerlaub et al. 2018; Holzer & Sengl 2020; Hohlfeld 2020), sowie bei Menschen, die gesellschaftlich ausgegrenzt sind (Graeupner & Coman, 2017), sich politisch nicht engagieren (Hohlfeld 2020) und kein Vertrauen in Politik haben (Zimmermann & Kohring, 2020).

Desinformation zeigt auch eher Effekt bei Menschen, die eine geringere Bildung aufweisen (Allcott & Gentzkow, 2017, Hohlfeld 2020; aber: kein Effekt von Bildung bei Holzer & Sengl 2020). Bezüglich Alter zeigen einige Studien, dass vor allem jüngere Menschen darauf ansprechen (Allcott & Gentzkow, 2017; Hohlfeld 2020). Allerdings zeigen andere Studien, dass Desinformation gerade auch bei älteren Menschen eine Wirkung erzeugen kann.

Neben diesen individuellen Faktoren scheinen auch situative Merkmale einen Einfluss auf die Glaubwürdigkeit von Desinformation zu haben. So konnten Ständer et al. (2021) in einer Befragungsstudie im Experimentaldesign während der Coronapandemie feststellen, dass Personen, die besonders unsicher waren, wie hoch ein Infektionsrisiko einzuschätzen ist, eher geneigt waren, Desinformation zuzustimmen.

Bei der Wirkung von Desinformation ist weiterhin zu berücksichtigen, dass eine Korrektur von Informationen mit Aufwand verbunden ist. Psychologischen Forschung zeigt, dass je häufiger eine Information gelesen wird, sie eher als wahr beurteilt wird. Dieser Zusammenhang zwischen wiederholter Darbietung und wahrgenommener Glaubwürdigkeit tritt auch dann auf, wenn die Information selbst mit einer gewissen Wahrscheinlichkeit als falsch bzw. als Desinformation gekennzeichnet ist (vgl. Pennycook et al., 2018). Auch wenn der einzelne Kontakt mit einer Desinformation wenig Auswirkungen haben mag, können wiederholte Kontakte dazu führen, dass Informationen vertrauter erscheinen und daher eher geglaubt werden.

Wenn einer Information erst einmal Glauben geschenkt wird, ist es nachträglich schwierig, ein Urteil zu revidieren. Personen sind oftmals motiviert, die eigenen Ansichten zu bestätigen und suchen daher nach entsprechender Information. Zudem lässt sich eine Abneigung feststellen, bestehende Informationen zu prüfen, wenn die Überprüfung gegebenenfalls zu einer Änderung der eigenen Meinung führen müsste. Fact-Checking wird daher eher verwendet, wenn davon ausgegangen werden kann, dass die eigenen Positionen bestätigt werden (Chia et al., 2022).

Wie bei der Verbreitung von Desinformation lässt sich für die Schweiz sagen, dass Personen, bei denen gemäss Studien eine Prädisposition für Desinformation besteht, eine Minderheit sind. Mit Blick auf die Befunde von Hohlfeld (2020, S. 188) kann aber generell von einer «gewisse[n] Anfälligkeit für Desinformation» gesprochen werden. In seiner Befragungsstudie im Experimentaldesign konnten nur die Hälfte der nach dem Zufallsprinzip präsentierten, Desinformation enthaltenden Nachrichten zweifelsfrei als solche identifiziert werden. Die Befunde von Pennycook et al. (2018) weisen darauf hin, dass besonders der wiederholte Kontakt mit Desinformation problematisch sein kann. Weiterhin erweist sich eine Korrektur von auf Desinformation basierten Ansichten oftmals als schwierig, da Personen neue Informationen nicht neutral suchen und ihr Wissen integrieren, sondern hier gewissen motivationalen Verzerrungen in der Suche und Verarbeitung unterliegen.

## 2. Gesamtgesellschaftliche Wirkungen

Erkenntnisse zur Bedeutung von Desinformation für gesamtgesellschaftliche Prozesse sind empirisch aufgrund der meist multikausalen Zusammenhänge kaum überprüfbar (siehe auch Adams et al., 2023). Aussagen über potenzielle Folgen für die Gesamtbevölkerung basieren daher meist auf eher theoretisch-konzeptionellen Überlegungen: Häufig werden dabei die potenziellen Folgen von Desinformationen auf Meinungs- und Willensbildungsprozesse in Demokratien, bspw. während Abstimmungen und Wahlen, thematisiert (u.a. Bennett & Livingstone, 2018; Saurwein & Spencer-Smith, 2020; Vogler et al., 2021). Befürchtet wird, dass durch die Verbreitung von Desinformation Wahl- und Abstimmungsergebnisse

manipuliert werden können. Allcott und Gentzkow (2017) sprechen von einem existierenden, aber eher marginalen Einfluss von Desinformation auf den Wahlausgang.

## **VI. Zwischenfazit: Sozialwissenschaftliche Befunde zur Desinformation (in der Schweiz)**

Kommunikationswissenschaftliche und sozialpsychologische Studien zum Phänomen der Desinformation bieten Erkenntnisse zu den Verfasser:innen, Verbreiter:innen, Kanälen und Inhalten sowie zur Nutzung und Wirkung.

Übergeordnet lässt sich sagen, dass Desinformation aufgrund der hohen Reichweite und der schnellen, teilautomatisierten und gegebenenfalls auch teilpersonalisierten Verbreitung von Inhalten vor allem ein Problem des digitalen Medien- und Informationskonsums ist. Dabei spielen Social Media und Messengerdienste wie WhatsApp und Telegram eine grosse Rolle, da in sozialen Medien oftmals neben dem Bedürfnis nach Richtigkeit andere Bedürfnisse (z.B. Zugehörigkeit, positive soziale Identität) aktiviert sind und dazu führen, dass mit Informationen anders interagiert wird. Das Verfassen von Desinformation ist in der Regel auf wenige Individual- und Kollektivakteur:innen zurückzuführen, die damit politische oder ökonomische Ziele verfolgen können. Bei der Verbreitung von Desinformation können hingegen, zumindest fallweise, sehr viele Akteur:innen involviert sein, die problematische Inhalte auf Social Media und über Messengerdiensten teilen.

Die Themen spiegeln dabei jeweils die gesellschaftliche Entwicklung, eine spezifische Ereignislage sowie die politische Agenda wider (Humprecht 2018). So zirkulierten bspw. 2020 und 2021 Desinformationen zum Thema Corona. Aktuell ist mit einer starken Verbreitung von Desinformationen zum Krieg in der Ukraine auszugehen. Desinformationen werden von Personen mit hohem Social-Media-Konsum, geringen Vertrauen in Politik und etablierte Nachrichtenmedien wahr- und ernstgenommen. Gerade in Zeiten grösserer Unsicherheit ist mit einer zunehmenden Verbreitung und Überzeugungskraft von Desinformation zu rechnen (vgl. Ständer et al., 2021).

Deutlich wurde, dass die Verbreitungsmuster und die potenzielle Wirkmacht von Desinformation nach Länderkontexten und damit nach politischen, medialen und ökonomischen Rahmenbedingungen variiert: So sind bspw. Länder mit starker politischer Polarisierung oder ausgeprägtem Populismus besonders anfällig für Desinformation. Hingegen stärken ein heterogenes Mediennutzungsrepertoire und eine starke Zuwendung zu Service-public-Inhalten die Resilienz gegenüber Desinformation (Humprecht et al., 2021). Die wenige Forschung, die auch die Schweizer Kommunikation berücksichtigt, zeigt, dass die Schweiz vergleichsweise resilient gegen Desinformation ist (Humprecht et al., 2020). Als ursächlich

können der geringe Grad an Polarisierung sowie die Stärke des Service public angeführt werden.

Einige Problembereiche, die aktuell stark im Zusammenhang mit Desinformation diskutiert werden, sind vermutlich mit Blick auf die Schweiz überschätzt. Es gibt wenig Anzeichen, dass Social Bots eine grosse Rolle spielen. Auch alternative Medien werden in der Schweiz selten genutzt und ihre Reichweite ist beschränkt.

Einschränkend muss jedoch auf die Grenzen der bestehenden empirischen Forschung und damit deren Aussagekraft hingewiesen werden: Mit Blick auf die Analyse der Kanäle und Verfasser:innen ist zu konstatieren, dass die Forschung meist den (technologischen) Entwicklungen hinterherhinkt. Untersucht werden zudem vor allem solche Plattformen, die den Forschenden Zugang über sogenannte APIs gewähren. Andere Kanäle bleiben aufgrund der Zugangsschwierigkeiten meist unberücksichtigt. Zudem sind neue Plattformen, bspw. TikTok noch wenig erforscht. Schliesslich lassen sich Social Bots nicht immer zuverlässig identifizieren, gerade wenn Tools (bspw. Botometer) verwendet werden, die in einem englischsprachigen Kontext entwickelt wurden (Gallwitz & Kreil, 2021; Rauchfleisch & Kaiser, 2021).

Erkenntnisse zu den Inhalten und Themenbereichen von verbreiteter Desinformation müssen insofern als wenig repräsentativ angesehen werden, da sie meist nur auf Analysen basieren, die Desinformationen berücksichtigen, die auf Fact-Checking-Seiten oder auf bereits bestehenden Listen anderen Forschenden zugänglich gemacht wurden. Im Bereich der sozialpsychologischen Forschung wird dies in experimentellen Studien oftmals so gehandhabt, dass korrekte und falsche Informationen verwendet oder erstellt werden und ihre Überzeugungskraft überprüft wird – dies findet dann aber ausserhalb der entsprechenden Plattformen und Kanäle statt. Der Fokus liegt hier und in der kommunikationswissenschaftlichen Forschung zumeist auf Themen und -bereichen, die von gesellschaftlicher Relevanz sind. Desinformationen aus dem Wirtschaftsbereich, die im Zusammenhang mit einzelnen Unternehmen oder Branchen verbreitet werden, werden meist nicht erfasst. Zudem steht auch die Forschung vor der generellen Herausforderung, dass sie einschätzen muss, welche Informationen als klar falsch und als klar wahr gelten können. Zudem werden auch Gruppendynamiken noch zu wenig berücksichtigt. Die sozialen Dynamiken spielen gerade bei der Verbreitung von Desinformation auf Social Media und in Messengerdiensten eine zentrale Rolle (siehe z.B. Lawson et al., 2023).

Auch zur Reichweite von Desinformation und damit zur Frage, wer überhaupt in Kontakt mit Desinformation gerät, können bislang auf der Basis des bestehenden Forschungsstandes lediglich eher unbefriedigende Aussagen gemacht werden: Studien, die auf Befragungsdaten basieren, müssen vor dem Hintergrund des Phänomens der sozialen Erwünschtheit sowie

möglicher Erinnerungslücken interpretiert werden. Dient die Anzahl Likes, Shares und Comments eines Posts mit Desinformation als Indikator für die Reichweite, werden diejenigen übersehen, die den Post gelesen, aber nicht mit ihm interagiert haben.

Es fehlen auch gesicherte empirische Erkenntnisse zur Wirkung von Desinformation. Dies ist aus Governance Perspektive ein besonders gravierendes Problem. Denn Massnahmen lassen sich nur rechtfertigen und machen aus einer Kosten-Nutzen-Perspektive nur Sinn, wenn tatsächlich ein relevantes Problem besteht. Gerade Studien, welche die Quantität von Desinformation untersuchen, können keine Aussagen zu deren Wirkung machen. Ein Beispiel dafür sind Social Bots. Diese erzeugen zwar viel Kommunikation (z.B. Tweets) und können Inhalte multiplizieren. Allerdings ist damit noch nicht gesagt, dass diese Inhalte auch wahrgenommen werden, wovon ihre Wirkung letztendlich abhängt. Einige neuere Studien gehen deshalb von einer geringen Wirkung von Desinformation auf die Meinungsbildung aus (u.a. Bail et al., 2021, siehe aber van der Linden, 2023).

## C. Governance von Desinformation

### I. Vorbemerkungen

#### 1. Allgemeines

Nachstehend wird die bestehende Governance von Desinformation in der Schweiz aufgezeigt. Als Governance wird mit Mayntz (2004, 66) die «Gesamtheit aller nebeneinander bestehenden Formen der kollektiven Regelung gesellschaftlicher Sachverhalte: von der institutionalisierten zivilgesellschaftlichen Selbstregelung über verschiedene Formen des Zusammenwirkens staatlicher und privater Akteur:innen bis hin zu hoheitlichem Handeln staatlicher Akteur:innen» verstanden (vgl. auch Puppis, 2010).

Wir werden nachfolgend zunächst den grundrechtlichen Rahmen aufzeigen, an dem sich Governance Massnahmen zu orientieren haben (II.); danach werden wir Regulierung von Desinformation in der Europäischen Union und einigen Mitgliedstaaten der EU aufzeigen, deren Auswirkungen auf die Schweiz nicht zu unterschätzen sein dürfte (nachfolgend III.). Schliesslich folgt eine Darstellung der geltenden Rechtslage in der Schweiz, zuerst für rechtmässige (IV.), hernach für nicht rechtmässige Desinformation (V.). Die Darstellung ist bei Social Media grundsätzlich auf YouTube, Facebook und Instagram beschränkt; zudem verweisen wir punktuell auf weitere Netzwerke wie TikTok. Bei den Messengerdiensten fokussieren wir auf Whatsapp, Threema und Telegram.

#### 2. Besonderheiten der digitalen Desinformation

Das geltende Recht kennt zahlreiche Normen, die sich gegen Desinformation richten und deren Verbreitung Schranken setzen. Weil sich Desinformation in digitalisierten Öffentlichkeiten jedoch viel schneller und weiterverbreitet als im analogen Kontext, ergeben sich gewisse Probleme, insb. im Bereich der Rechtsdurchsetzung.

Massnahmen, die gegen Desinformation ergriffen werden, setzen deshalb oft auf der Ebene der Durchsetzung an. Weil aufgrund der Menge der auf Social Media und Messengerdiensten verbreiteten Desinformation eine Verfolgung durch staatliche Gerichte in den meisten Fällen unrealistisch ist, zielt ein Grossteil der gegen Desinformation ergriffenen Massnahmen auf die Inpflichtnahme von Diensten, welche ihren Nutzer:innen die Verbreitung von Informationen ermöglichen. Damit kommt es zu einer Verlagerung von der staatlichen zur privaten Rechtsdurchsetzung. Dies ist zwar effizient, aus rechtsstaatlicher Sicht aber nicht unproblematisch. Immerhin bestehen für die staatliche Gerichtsbarkeit in den Verfassungen der Staaten und in den Prozessgesetzen klare, transparente und in einem rechtsstaatlichen Verfahren erlassene Regeln, die bei Privaten fehlen. Aus diesem Grund werden die Dienste zumindest in der EU dazu verpflichtet, gewisse prozessuale Grundsätze einzuhalten (dazu nachfolgend II.3.b.).

### 3. Unterscheidung zwischen rechtmässiger und unrechtmässiger Desinformation

Aus rechtlicher Sicht ist die Unterscheidung zwischen rechtswidrigen und rechtmässigen Desinformationsinhalten zentral. Rechtswidrige Inhalte sind solche, die gegen Rechtsnormen verstossen, rechtmässige Desinformationen sind nach geltendem Recht zulässig. Dieser Unterscheidung folgt auch der von der EU erlassene Digital Services Act (dazu nachfolgend II.2 und II.3).

So wäre die – nichtzutreffende – Behauptung, ein:e Politiker:in hätte sich einer Straftat schuldig gemacht, als Verleumdung (Art. 173 StGB) und entsprechend rechtswidrige Desinformation zu qualifizieren, die Behauptung, eine Abstimmung fände am Montag statt, obwohl sie am Sonntag stattfindet, ist eine rechtmässige Desinformation.

### 4. Unterscheidung zwischen Verfasser- und Verbreiter:innen

Die Befunde zu den Verfasser- und Verbreiter:innen haben ergeben, dass Desinformation in der Regel von Einzelpersonen verfasst und verbreitet wird, dass aber auch Kollektivakteur:innen als Verfasser- oder Verbreiter:innen infrage kommen.

Aus rechtlicher Sicht kann die Unterscheidung relevant sein. So bestehen für gewisse Akteur:innen bei der Durchsetzung Sonderregeln, bspw. für Medien. Auch gibt es im Bereich der rechtmässigen Desinformation für gewisse Akteur:innen besondere Bestimmungen, insbesondere was deren Verfolgbarkeit angeht.

## II. Rechtlicher Rahmen in der Europäischen Union

### 1. Einleitung

In der EU ist die Rechtslage davon abhängig, ob es sich um rechtmässige oder unrechtmässige Desinformation handelt. Der nicht rechtmässige, «illegale Charakter» von Inhalten ergibt sich in der EU aus dem Unionsrecht oder aus dem nationalen Recht. Der Begriff der Illegalität ist dabei weit auszulegen. Er soll nach Ansicht des europäischen Gesetzgebers alle Inhalte, Produkte, Tätigkeiten oder Dienstleistungen erfassen, die entweder an sich rechtswidrig sind oder deshalb, weil sie mit rechtswidrigen Handlungen zusammenhängen (Erw. 12 DSA).

### 2. Rechtsrahmen für nicht rechtmässige Desinformationsinhalte: Der Digital Services Act («DSA»)

#### a) Grundlagen

Bei nicht rechtmässiger Desinformation steht die Durchsetzung des anwendbaren Rechts im Vordergrund. Die einschlägigen Normen finden sich in der unmittelbar anwendbaren, bereichsübergreifenden Verordnung des Europäischen Parlaments und des Rates über einen Binnenmarkt für digitale Dienste (Gesetz über digitale Dienste, Digital Services Act und zur Änderung RL 2000/31/EG, COM (2020) 825 final). Der DSA soll einen horizontalen Rahmen für alle Kategorien von Inhalten, Produkten, Dienstleistungen und Tätigkeiten im Zusammenhang mit

Vermittlungsdiensten setzen. Neben der Fortentwicklung geltender Vorschriften enthält der DSA neue Durchsetzungsvorschriften und neue Anforderungen für Online-Plattformen.

Der Anwendungsbereich des DSA lässt sich wie folgt zusammenfassen:

- Sachlich anwendbar ist der DSA grundsätzlich nur auf nicht rechtmässige, «illegale Inhalte» (Art. 2 Bst. g DSA). Eine Ausnahme besteht insoweit, als er zusätzliche Compliance-Vorschriften für sehr grosse Online-Plattformen betreffend rechtmässige Inhalte statuiert (dazu II.2.e.)
- Der persönliche Anwendungsbereich umfasst 1) Anbieter:innen von Vermittlungsdiensten, 2) Hosting-Diensteanbieter, 3) Online-Plattformen, 4) sehr grosse Online-Plattformen («VLOP»), d.h. solche mit mehr als 45 Mio. aktiven Nutzer:innen in der EU pro Monat. Die EU hat bisher 17 VLOP identifiziert. Die hier untersuchten Social-Media-Kanäle und Messengerdienste zählen zu den Hosting Diensteanbietern und zur Unterkategorie der Online-Plattformen, Facebook, Instagram und Youtube gehören zu den VLOP.
- Regelungsgegenstände sind: 1) Allgemeines Haftungsregime (entspricht weitgehend der sog. E-Commerce RL); 2) Besondere Sorgfaltspflichten; 3) Vorschriften zur (besseren) Durchsetzung.

Die Festlegung der Sanktionen für Verstösse der Anbieter:innen von Vermittlungsdiensten gegen den DSA obliegt den Mitgliedstaaten (Art. 42 DSA). Dies bedeutet, dass es im Wesentlichen von den einzelnen Mitgliedstaaten abhängen wird, mit welcher Durchsetzungskraft die Regulierung versehen wird.

#### b) Qualifikation von Messengerdiensten

Messengerdienste gelten als Vermittlungsdienste im Sinne von Art. 2 Bst. f (a) DSA und unterstehen damit gewissen (Grund-)Pflichten. Umstritten ist, ob und inwieweit Messengerdienste unter den Begriff der Online-Plattform fallen. Nach Ansicht des Unionsgesetzgebers soll dies zumindest für Anbieter:innen von Diensten der Fall sein, die ihren Nutzer:innen die Möglichkeit geben, Informationen für eine potenziell unbegrenzte Zahl von Nutzer:innen, die nicht von der Nutzer:in selbst bestimmt werden, bereitzustellen. Dies ist etwa bei offenen Benutzergruppen und offenen Kanälen der Fall (DSA, Erw. 14; Kuhlmann & Trute, 115). Entscheidend ist, ob die Personen der Gruppe ohne menschliche Entscheidung oder Auswahl, d.h. automatisch, beitreten können.

Nicht als Online-Plattformen sollen Messengerdienste gelten, deren Funktion sich auf die Zurverfügungstellung eines Services für interpersonelle Kommunikation zwischen einer begrenzten Anzahl von Nutzer:innen beschränkt (Kuhlmann & Trute, 115). Ausgenommen sind zudem auch diejenigen Fälle, in denen das Angebot offener Nutzungsgruppen eine untergeordnete Funktion des Hauptdienstes darstellt.

### c) Instrumente zur Bekämpfung von rechtswidrigen Inhalten

Der DSA sieht vor, dass alle Anbieter:innen von Vermittlungsdiensten in ihren allgemeinen Geschäftsbedingungen (AGB) Beschränkungen anzugeben haben, die sie möglicherweise bei der Nutzung ihres Dienstes auferlegen. Darin enthalten sein müssen Angaben zu allen Leitlinien, Verfahren, Massnahmen und Werkzeugen, die vom Vermittlungsdienst zur Inhaltsmoderation eingesetzt werden (Art. 14 DSA).

Hostingdiensteanbieter:innen, einschliesslich Online-Plattformen sind verpflichtet ein Verfahren einzurichten, mit dem mutmasslich rechtswidrige Inhalte gemeldet werden können (Art. 16 DSA). Das Verfahren muss leicht zugänglich und benutzerfreundlich sein und eine Übermittlung von Meldungen ausschliesslich auf elektronischem Weg ermöglichen. Kommt die Meldung von einem «vertrauenswürdigen Hinweisgeber» (Trusted Flagger) ist er gemäss Art. 22 DSA vorrangig und unverzüglich zu bearbeiten.

Wird als Reaktion auf eine Meldung ein Inhalt gesperrt, entfernt oder herabgestuft, ist die betreffende Nutzer:in darüber in Kenntnis zu setzen. Zudem ist die Person darüber zu informieren, ob die Massnahme gestützt auf eine Meldung von Dritten oder aus Eigeninitiative der Anbieter:in erfolgt, ob automatisierte Mittel zum Einsatz kamen und auf welcher Grundlage (Rechtsgrundlage oder AGB des Anbieters) die Massnahme ergriffen wurde. Auch die Rechtsbehelfe, die der Nutzer:in zur Verfügung stehen, sind zu nennen (Art. 17 DSA).

Für die Behandlung von Beschwerden gegen Entscheide über Inhalte oder Personen müssen Online-Plattformen *ein internes Beschwerdemanagementsystem* einrichten (Art. 20 DSA). Ergänzend müssen die Plattformen zur Beilegung von Streitigkeiten mit Nutzer:innen mit zugelassenen *aussergerichtlichen Streitbeilegungsstellen* zusammenarbeiten (Art. 21 DSA).

Nutzer:innen, die häufig und offensichtlich rechtswidrige Inhalte veröffentlichen, sollen nach vorgängiger Verwarnung für eine gewisse Zeit vom Dienst ausgeschlossen werden (Art. 23 Abs. 1 DSA). Dasselbe gilt für Nutzer:innen, die häufig und unbegründet Meldungen erstatten und Beschwerden einreichen (Art. 23 Abs. 2 DSA). Die Sperre kann im Rahmen des internen Beschwerdeverfahrens angefochten werden (Kuhlmann & Trute, 117).

Zur Förderung der Transparenz muss die Online-Plattform Berichte über ihre Tätigkeiten zur Entfernung und Sperrung von Informationen, die als illegale Inhalte betrachtet werden oder gegen ihre AGB verstossen, veröffentlichen (Art. 24 DSA).

### d) Durchsetzungsmechanismen im Besonderen

Der DSA enthält staatliche und private Durchsetzungsmechanismen. Zu Ersteren gehört die Bestimmung der rechtlichen Zuständigkeit zur Durchsetzung des DSA. Diese liegt beim EU-Mitgliedstaat, in dem sich die Hauptniederlassung der Anbieter:in befindet (Art. 56 DSA). Die Mitgliedsstaaten sind zur Ernennung einer zuständigen nationalen Behörde, einschliesslich eines Koordinators für digitale Dienste, verpflichtet (Art. 49 f. DSA). Dem Koordinator werden besondere Untersuchungs- und Durchsetzungsbefugnisse übertragen; so kann er etwa bestimmte Informationen verlangen, die Einstellungen von Zuwiderhandlungen anordnen und

Massnahmen bis hin zu Geldbussen, verhängen (Art. 50 DSA). Zudem kann er sich an gemeinsamen Untersuchungen beteiligen (Art. 53 DSA).

Ergänzt werden die staatlichen Durchsetzungsmechanismen durch *private Durchsetzungsmechanismen*. Alle Anbieter:innen werden dazu verpflichtet, eine zentrale *Kontaktstelle* einzurichten (Art. 11 DSA). Soweit sie keinen *Sitz in einem Mitgliedsstaat* haben, müssen sie einen *Rechtsvertreter* in der EU bezeichnen (Art. 13 DSA).

#### e) Transparenzvorschriften

Anbieter:innen von Online-Plattformen haben gewisse Transparenzvorschriften bezüglich Werbung und Empfehlungssysteme zu beachten. Die bereits für Werbetreibende bestehende Verpflichtung zur Kennzeichnung von Werbung wird auf die Anbieter:innen von Online-Plattformen ausgedehnt (Art. 26 DSA; Gerdemann & Spindler, 119). Online-Plattformen, die Empfehlungssysteme nutzen, müssen in ihren AGB in klarer und verständlicher Sprache die wichtigsten Parameter nennen, die diese Systeme verwenden und die Nutzer darüber informieren, dass (und wie) sie diese Parameter ändern oder beeinflussen können (Art. 27 DSA). Weitergehende Vorschriften gelten für VLOP und sehr grosse Suchmaschinen. Diese müssen ihren Nutzer:innen bei den Empfehlungssystemen zumindest eine Option zur Verfügung stellen, die nicht auf Profiling beruht. Zudem müssen sie in der Vergangenheit angezeigte Werbung über eine gewisse Zeit hinaus öffentlich zugänglich machen.

#### f) Pflichten für VLOP

Weitergehende Pflichten bestehen für VLOP. Diese sind verpflichtet, unter gewissen Bedingungen dem Koordinator für digitale Dienste oder der Kommission und zugelassenen Forschern *Zugang zu Daten gewähren* (Art. 40 DSA). Zudem müssen sie einen *Compliance-Beauftragten* benennen (Art. 41 DSA) und *zusätzliche Transparenzvorschriften* einhalten (Art. 39, 42 DSA).

### 3. Rechtsrahmen für rechtmässige Desinformationsinhalte

#### a) Entwicklungen

Erste Massnahmen für ein koordiniertes Vorgehen der EU-Mitgliedsstaaten gegen rechtmässige Desinformation wurden im Jahr 2018 im *Aktionsplan gegen Desinformation* festgehalten. Dieser sah verschiedene Massnahmen vor, unter anderem den Ausbau der Fähigkeiten, Desinformation mittels regionaler Taskforces und Analyse-Einheiten zu erkennen, zu untersuchen und zu enthüllen. Von Bedeutung ist die Europäische Beobachtungsstelle für digitale Medien (EDMO), eine unabhängige Drehscheibe, die Faktenprüfer und akademische Forschende mit Fachwissen auf dem Gebiet der Online-Desinformation sowie von Social-Media-Plattformen, journalistischen Medien und Medienkompetenz zusammenbringen soll (Näheres dazu: European Digital Media Observatory, EDMO, abrufbar unter <https://digital-strategy.ec.europa.eu/en/policies/european-digital-media-observatory>). Seit März 2019 besteht zudem ein Frühwarnsystem auf EU-Ebene.

Weiter gab es Bestrebungen zur Mobilisierung des Privatsektors zur Selbstregulierung bei der Bekämpfung von Desinformation. Dazu wurde bereits vor Erlass des Aktionsplans der Verhaltenskodex für Desinformation von 2018 aufgestellt. Mit diesem haben sich die Plattformen zum ersten Mal weltweit freiwillig auf Selbstregulierungsstandards zur Bekämpfung von Desinformation geeinigt. Der Verhaltenskodex wurde am 16. Juni 2022 in überarbeiteter Form vorgelegt («Strengthened Code of Practice on Desinformation», Verhaltenskodex für Desinformation 2022; siehe dazu nachfolgend II.3.c.).

#### b) Digital Services Act

Auch der DSA enthält Bestimmungen zu rechtmässiger Desinformation. Für Anbieter von VLOP sehen Art. 34 Abs. 1 Bst. b und Bst. c eine Pflicht zur Risikobewertung vor, sowohl für illegale Inhalte aber auch für systematische Risiken wie die nachteiligen Auswirkungen auf Grundrechte, das Diskriminierungsverbot und Kinderrechte oder die vorsätzliche Manipulation der Plattform durch unauthentische oder automatisierte Nutzung, wenn sich daraus nachteilige Auswirkungen auf gesellschaftliche Rechtsgüter wie öffentliche Gesundheit oder die Demokratie ergeben können.

Dies bedeutet, dass die VLOP jährlich darlegen müssen, ob ihre Moderations- und Empfehlungssysteme und der Umgang mit Werbeanzeigen ein systemisches Risiko bergen, welche Folgen die Verbreitung der Desinformation auf die Informations- und Meinungsfreiheit sowie auf das Privatleben von Menschen hat und inwiefern sich der Gebrauch von Fake-Accounts negativ auf die Gesundheit, auf Minderjährige und auf den zivilen Diskurs sowie Wahlen auswirkt. Die Bewertung soll gegebenenfalls zur Einführung von Massnahmen führen, wie bspw. zum Stopp der Auszahlung von Werbegeldern für einschlägigen Content oder zu einer erweiterten Sichtbarkeit zuverlässiger Informationsquellen. Zudem müssen sich die Plattformen *externen und unabhängigen Prüfungen unterziehen* (Art. 37 DSA).

#### c) Verstärkter Verhaltenskodex für Desinformation

Das Selbstregulierungsinstrument soll insbesondere finanzielle Anreize für Anbieter:innen von Desinformation reduzieren, indem verhindert wird, dass sie von Werbeeinnahmen profitieren. Die Platzierung von Werbung neben Desinformation soll möglichst vermieden werden. Die Plattformen sollen des Weiteren:

- Transparenzmassnahmen ergreifen, die es den Nutzer:innen ermöglichen, politische Anzeigen leicht zu erkennen;
- Initiativen zur Verbesserung der Medienkompetenz der Nutzer:innen ergreifen;
- Forscher:innen einfacher Zugang zu Forschungsdaten gewähren.

Ein wichtiger Schritt ist zudem die Zusammenarbeit der Plattformen zur Verhinderung von manipulativem Verhalten. Dazu erarbeiten die Plattformen ein dienstübergreifendes Verständnis von unzulässigen manipulativen Taktiken, Techniken und Verfahren (TTP), die von Akteur:innen eingesetzt werden. Dieses wird regelmässig auf Aktualität überprüft. Damit sollen

klare Strategien implementiert werden, die das Spektrum der ermittelten Verhaltensweisen und Praktiken abdecken. (<https://digital-strategy.ec.europa.eu/de/policies/code-practice-disinformation>, zuletzt besucht 25. Mai 2023).

Die fakultative Umsetzung dieser Massnahmen wird durch die Kommission mit der Unterstützung der Gruppe Europäischer Regulierungsstellen für audiovisuelle Medien überwacht (Aktionsplan gegen Desinformation, S. 10).

#### 4. Weitere Entwicklungen in der EU

In der EU gibt es weitere Bestrebungen, um Desinformation einzuschränken. Einige Massnahmen sind Teil des sog. «Aktionsplans Demokratie für Europa». In deren Rahmen ist eine Verordnung in Planung, welche die Transparenz für politische Werbung und deren Finanzierungsquellen erhöhen sollen (Verordnung über die Transparenz und das Targeting politischer Werbung vom 25. November 2021, COM/2021/731 final). Zudem besteht das europäische Kooperationsnetzwerk für Wahlen, welche Manipulationen im Zusammenhang mit Wahlen verhindern soll.

Weitere Regulierungsbestrebungen der EU betreffen die Medien. Das geplante Medienfreiheitsgesetz soll den Medienpluralismus und die Unabhängigkeit der Medien fördern. Geplant sind zudem Bestimmungen zur Moderierung von Inhalten; namentlich sollen für die Entfernung von Inhalten von Medien auf Plattformen strengere Anforderungen gelten.

#### 5. Auswirkungen des Digital Services Act auf die Schweiz

Der DSA gilt nur für die EU und findet in der Schweiz keine Anwendung. Einzuhalten wäre er jedoch von in der Schweiz ansässigen digitalen Dienstleister:innen, die in den persönlichen Anwendungsbereich des DSA fallen und die ihre Dienste in der EU anbieten. Solche Dienste gibt es aber derzeit nicht. Von grösserer Bedeutung ist die Frage nach der Auswirkung des DSA auf die Nutzer:innen in der Schweiz.

Insbesondere stellt sich die Frage, ob international tätige Online-Plattformen künftig spezifische Regelungen für die Schweiz vorsehen werden. Aufgrund der beschränkten Grösse des schweizerischen Marktes ist davon eher nicht auszugehen; wahrscheinlicher ist, dass die Plattformen für Nutzer:innen in der Schweiz (weiterhin) dieselben Regeln anwenden werden wie für die EU. Darauf deuten namentlich die Nutzungsbedingungen hin, die im Wesentlichen den Anforderungen des DSA entsprechen.

### III. Grundrechtlicher Rahmen

Grundlage und Rahmen aller Empfehlungen zur Governance von Desinformation sind die in der schweizerischen Bundesverfassung und in der Europäischen Menschenrechtskonvention (Konvention vom 4. November 1950 zum Schutz der Menschenrechte und Grundfreiheiten, SR 0.101, «EMRK») verankerten Grundrechte. Dabei sind die Grundrechte der Meinungs- und

Informationsfreiheit, der Medienfreiheit und der Wirtschaftsfreiheit von zentraler Bedeutung. Dasselbe gilt für die Bestimmungen zu den politischen Rechten und für das Fernmeldegeheimnis. Diese Normen werden nachfolgend kurz dargestellt. Die grundrechtliche Perspektive orientiert sich an den betroffenen Personen, beginnend mit der Perspektive der Verbreiter:innen von Desinformation über die Rezipient:innen bis hin zu den Kanälen, die als Mittler zwischen Verbreiter:innen und Rezipient:innen wirken und sich auf eine eigene grundrechtlich geschützte Position stützen können.

## 1. Vorbemerkungen

Die bestehende oder zu schaffende Regulierung von Desinformation ist in den verfassungsrechtlichen Rahmen eingebettet. Der öffentliche Meinungsbildungsprozess soll nach weit verbreiteter Auffassung (dazu ausführlich Cueni, S. 4 ff.) grundsätzlich ohne staatliche Einmischung stattfinden. Es gibt jedoch gewisse Fälle, in denen eine staatliche Einmischung geboten sein kann, um den Meinungsbildungsprozess zu schützen (Biaggini, Art. 35 N 4; Besson, S. 49 ff.; Dröge, S. 13 ff., 90 ff. u. 123 ff.). Der Staat hat damit gewisse Schutz- und Gewährleistungspflichten, welche die Ausübung und Verwirklichung der Grundrechte durch Individuen erst möglich machen (Schefer & Cueni, S. 26; Müller, 2018, N 15). Ein regulierender Eingriff des Staates ist insb. in den Fällen angezeigt, in denen zwischen Privaten ein Machtgefälle besteht, weil hier ein Schutz der schwächeren Partei erforderlich ist (Kiener, N 33).

Der Bedarf nach (gesetzlichen) Massnahmen ergibt sich nicht zuletzt deshalb, weil Privatpersonen in der Regel nicht an die Grundrechte gebunden sind (keine sog. direkte Drittwirkung), selbst wenn sie im Einzelfall über eine ähnlich starke Stellung gegenüber den Individuen verfügen wie der Staat (Häfelin et al., N 281 ff.).

Die Verfassung selbst sieht allerdings in Art. 35 Abs. 3 BV eine indirekte Drittwirkung vor, indem sie dem Staat die Pflicht auferlegt, dafür zu sorgen, dass die Grundrechte, soweit sie sich dazu eignen, auch unter Privaten wirksam werden. In der Praxis wirkt sich die Drittwirkung dadurch aus, dass sie bei der Auslegung unbestimmter Rechtsbegriffe des Privatrechts heranzuziehen sind, wodurch eine verfassungskonforme Auslegung gewährleistet wird (Häfelin et al., N 282). In bestimmten Bereichen ist der Staat zudem gesetzgeberisch tätig geworden und hat durch den Erlass von – privatrechtlichen oder strafrechtlichen – Normen die Grundrechte zwischen Privaten wirksam gemacht. So schützt beispielsweise die strafrechtliche Norm von Art. 179<sup>bis</sup> StGB das Fernmeldegeheimnis zwischen Privaten.

## 2. Einschränkung der Grundrechte

Grundrechte gelten nicht unbeschränkt. Vielmehr können sie beim Vorliegen einer gesetzlichen Grundlage, eines öffentlichen Interesses und unter Wahrung der Verhältnismässigkeit eingeschränkt werden, wenn der Kerngehalt des Grundrechtes nicht betroffen ist (Art. 36 BV).

Die Ausübung der vorliegend betroffenen Kommunikationsgrundrechte dient selbst einem öffentlichen Interesse, weshalb Einschränkungen nur beim Vorliegen von besonders wichtigen öffentlichen Interessen zulässig sind (Müller & Schefer, S. 378 f.; Weber & Nobel, S. 187).

### 3. Desinformation aus Sicht der Verbreiter:innen

#### a) Meinungs- und Informationsfreiheit

Aus Sicht der Verbreitenden steht die Meinungs- und Informationsfreiheit im Vordergrund. Sie ist die Grundnorm der Kommunikationsgrundrechte (Weber/Nobel, S. 176). Als wesentliche Grundlage für einen öffentlichen Meinungsdiskurs in der Gesellschaft wurde sie auch in internationalen Übereinkommen verankert, so in Art. 10 EMRK und in Art. 19 des internationalen Pakts über bürgerliche und politische Rechte (Internationaler Pakt vom 16. Dezember 1996 über bürgerliche und politische Rechte, SR 0.103.2).

Geschützt ist die «Meinung», ein Begriff, der weit verstanden wird und die «Gesamtheit der Mitteilungen menschlichen Denkens» (BGE 127 I 145, E. 4.b.) erfasst. Der grundrechtliche Schutz umfasst subjektive (eigenen oder fremde) Meinungen und Werturteile, aber auch Tatsachenbehauptungen. Inhaltlich sind auch provozierende oder schockierende Äusserungen (BGE 138 I 274, E. 2.2.1) und selbst unwahre Tatsachenbehauptungen geschützt (Schefer, Art. 16 N 8; Cueni, S. 4 ff.; Kiener, Kälin, Wytenbach, § 18 N 17; BK BV-Hertig, Art. 16 N 9; Peduzzi, S. 186–188; Müller & Schefer, S. 393).

Der Schutz der Meinungsfreiheit umfasst nicht nur den Inhalt einer Äusserung, sondern auch die Mittel und Form der Äusserung (BGE 127 I 145, E. 4.b. m.w.H.; BK BV-Hertig, Art. 16 N 12). Der Kanal der Äusserung ist unerheblich, womit auch Äusserungen auf Social Media (Gunjic, S. 189; Kiener, Kälin, Wytenbach, § 18 N 12) oder in Messengerdiensten durch die Meinungsfreiheit geschützt sind.

Die Meinungs- und Informationsfreiheit erfasst eine innere Dimension, welche das Recht umfasst, eine Meinung zu haben, eine Meinung zu bilden und Informationen zu empfangen sowie eine äussere Dimension, verstanden als Recht, die eigene Meinung und Informationen zu äussern und zu verbreiten. Eingriffe in die Meinungs- und Informationsfreiheit müssen den Anforderungen von Art. 36 BV genügen (dazu vorne I.2.a.). Besonders hohe Anforderungen sind dabei an Massnahmen zu stellen, die an den Inhalt einer Äusserung anknüpfen (BV-Hertig, Art. 16 N 38; Schefer, N 83). Präventive Eingriffe wie Publikationsverbote sind deshalb nur im Einzelfall bei Vorliegen einer unmittelbaren Bedrohung für elementare Rechtsgüter zulässig (BK BV-Hertig, Art. 16 N 41).

#### b) Medienfreiheit

Relevant sein kann ferner die Medienfreiheit, welche die öffentliche Verbreitung von (zumindest vorwiegend) ideellen Inhalten unter Zuhilfenahme technischer Hilfsmittel (BK BV-Zeller & Kiener, Art. 17 N 9) schützt. Der Wortlaut von Art. 17 BV schützt grundsätzlich Presse, Radio

und Fernsehen sowie andere Formen fernmeldetechnisch verbreiteter Kommunikation, was auch die Kommunikation über das Internet erfasst. Der Geltungsbereich der Medienfreiheit ist eng auszulegen und auf journalistische Äusserungen zu beschränken. Nicht-journalistische Äusserungen sind durch die allgemeine Meinungsfreiheit geschützt. Die Medienfreiheit spielt deshalb im Zusammenhang mit Desinformation eine untergeordnete Rolle, zumal es sich bei den Verbreitenden von Desinformation in der Regel nicht um Journalist:innen handelt, die sich auf die Medienfreiheit berufen könnten.

#### c) Versammlungsfreiheit

Betroffen ist mitunter auch die Versammlungsfreiheit, welche der kollektiven gesellschaftlichen und politischen Meinungsbildung und Kundgabe dient (BK BV-Hertig, Art. 22 N 1; Wyss, S. 395). Von besonderer Bedeutung ist sie für Minderheiten, «die ihre politische Meinung innerhalb der bestehenden demokratischen Einrichtungen nicht oder nicht genügend zur Geltung bringen können» (BK BV-Hertig, Art. 22 N 1). Eine Mehrheit der Lehre befürwortet unserer Ansicht nach zu Recht die Erstreckung des grundrechtlichen Schutzes auf virtuelle Versammlungen (Befürwortend BK BV-Hertig, Art. 22 N 6, m.w.H.). Charakteristika der Versammlungsfreiheit sind die kollektive Meinungsbildung und -kundgabe (BK BV-Hertig, Art. 22 N 6). Unter die Versammlungsfreiheit fällt neben live stattfindenden Veranstaltungen auch die Teilnahme an Chatforen, soweit diese Foren der Willensbildung dienen.

Die Versammlungsfreiheit schützt das Recht, Versammlungen zu organisieren, an diesen teilzunehmen oder ihnen fernzubleiben. Geschützt ist unter anderem der freie Zugang zu einer Veranstaltung (BK BV-Hertig, Art. 22 N 11; Wyss, S. 407). Verfassungsmässig geschützte Versammlungen können auch online in Chatrooms oder in sozialen Netzwerken stattfinden. Wenn die Verbreitung von Desinformation zur Folge hat, dass eine Person an einer virtuellen Veranstaltung nicht teilnehmen kann, ist ihre Versammlungsfreiheit berührt.

#### d) Wirtschaftsfreiheit

Zu beachten ist zudem die Wirtschaftsfreiheit, welche jegliche privatwirtschaftliche Tätigkeit schützt, die auf die Erzielung eines Einkommens oder Gewinns ausgerichtet ist (BGE 140 I 218, E. 6.3). Art. 27 Abs. 2 BV nennt exemplarisch „die freie Wahl des Berufes sowie den freien Zugang zu einer privatwirtschaftlichen Erwerbstätigkeit und deren freie Ausübung“. Davon erfasst sind auch alle mit der Erwerbstätigkeit zusammenhängenden Aspekte (Biaggini, Art. 27 N 9). Vorliegend ist insb. der Schutz von kommerziellen Äusserungen relevant.

Diese können neben der Meinungs- und Informationsfreiheit im Sinn von Art. 16 BV (siehe dazu D.II.2.a.) auch der in Art. 27 BV statuierten Wirtschaftsfreiheit zugeordnet werden (BGE 139 II 173, E. 5.1; Biaggini, Art. 16 N 6 m.w.H.). Einschränkungen der Wirtschaftsfreiheit, bspw. von kommerziellen Äusserungen, müssen entsprechend den allgemeinen Regeln (Art. 36 BV), auf einer gesetzlichen Grundlage beruhen – etwa dem UWG – im öffentlichen Interesse liegen

und verhältnismässig sein. Staatliche Massnahmen, die in den freien Wettbewerb eingreifen, sind nur zulässig, wenn sie in der Bundesverfassung vorgesehen sind oder auf kantonalen Hoheitsrechten beruhen (Art. 94 Abs. 4 BV; vgl. auch BGE 131 I 223, E. 4.2; BGE 128 I 3, E. 3.a; BGE 143 I 403, E. 5.2). Unzulässig sind folglich staatliche Massnahmen, die den freien Wettbewerb behindern, um bestimmte Wirtschaftszweige oder Formen der wirtschaftlichen Tätigkeit zu sichern oder zu begünstigen oder das Wirtschaftsleben nach einem bestimmten Plan zu lenken (BGE 143 I 403, E. 5.2). Abzugrenzen sind staatliche Massnahmen, die ordnungs- und sozialpolitische Ziele verfolgen oder nicht primär wirtschaftlichen Interessen dienen. Letztere fallen von vornherein nicht in den Schutzbereich der Wirtschaftsfreiheit (Art. 94 BV; BGE 143 I 403, E. 5.2).

#### e) Fernmeldegeheimnis

Das Fernmeldegeheimnis ist ein Teilgehalt der persönlichen Freiheit (Art. 13 Abs. 1 BV). Das Brief-, Post- und Fernmeldegeheimnis schützt die Vertraulichkeit der Kommunikation. Es erfasst nicht nur traditionelle Kommunikationsmittel wie Post und Telefon, sondern auch neuere wie E-Mail (BGE 126 I 50, E. 6.a.; BK BV-Diggelmann, Art. 13 N 29). Auch die Kommunikation über Messengerdienste fällt unter das Fernmeldegeheimnis (SK StPO-Hansjakob & Pajarola, Art. 269 N 12 f.; Equey, S. 685 ; sinngemäss BGE 140 IV 181, E. 2.6). Der Schutz des Fernmeldegeheimnisses beschränkt sich nicht auf Inhalte, sondern umfasst auch sog. Kommunikationsranddaten; im Bereich des Internets sind dies insb. Informationen über die Verwendung von E-Mail- und IP-Adressen sowie über Adressaten und Empfangszeitpunkte von Mitteilungen (BK BV-Diggelmann, Art. 13 N 29).

Kommunikation über Messengerdienste, wie z.B. Telegram, ist längst nicht mehr immer privat, sondern oft an eine grössere Gruppe gerichtet. Wenn Verbreiter:innen von Desinformation ihre Mitteilungen an einen offenen Kreis von Personen richten, zu denen kein Vertrauensverhältnis besteht oder wenn sie die Kontrolle über den Wirkungsbereich der verbreiteten Informationen aufgeben (BK StGB-Fiolka, Vor-Art. 258 N 15), entfällt die Privatheit der Kommunikation und damit auch der Schutz des Fernmeldegeheimnisses.

Art. 13 Abs. 1 BV bindet bei der Beförderung von Briefen und elektronischen Mitteilungen vorzugsweise den Staat; Beschränkungen des privaten Zugriffs auf die Kommunikation müssen auf dem Weg der Gesetzgebung erfolgen (Botschaft VE 96, 153; SG BV-Breitenmoser & Schweizer, Art. 13 N 66). Entsprechend enthält das Strafrecht verschiedene Bestimmungen (Art. 179 ff. StGB), welche die private Kommunikation schützen.

#### f) Verfahrensgrundrechte

Die Verfahrensgrundrechte der Bundesverfassung (Art. 29, 29a, 30 und 32 BV) können im Zusammenhang mit Desinformation relevant werden. Staatliche Verfahren haben gewisse verfahrensrechtliche Vorgaben zu beachten, die in den Art. 29 ff. BV aufgeführt sind. Die zentrale

Bestimmung ist Art. 29 BV, die *Verfahrensgerechtigkeit* verspricht und damit einen fundamentalen rechtsstaatlichen Leitgedanken festhält (Biaggini, Art. 29 N 2). Die Bestimmung lässt sich in verschiedene, nicht abschliessende Teilgehalte gliedern, wie das Verbot der Rechtsverweigerung und der Rechtsverzögerung; (ii) das Verbot des überspitzten Formalismus; (iii) der Anspruch auf rechtliches Gehör, bestehend aus zahlreichen Unterkategorien; und der (iv) Anspruch auf unentgeltliche Rechtspflege (Biaggini, Art. 29 N 7). Weitere Verfahrensgrundrechte betreffen die allgemeine Rechtsweggarantie bei Rechtsstreitigkeiten (Art. 29a BV) und Rechte vor gerichtlichen Behörden (Art. 31 BV), namentlich den Anspruch auf einen unabhängigen und unparteiischen Richter. Art. 32 BV enthält spezifische Rechte im Strafprozess, die hier von geringerer Bedeutung sind.

Die Garantien von Art. 29 ff. BV sind in allen (zivil-, straf- und verwaltungsrechtlichen) staatlichen Verfahren zu beachten, in denen über individuelle Rechte und Pflichten entschieden wird (vgl. BGE 129 I 232; BGE 119 Ia 141). Geschützt sind sowohl natürliche als auch juristische Personen, wenn auch bei letzteren gewisse Einschränkungen bestehen.

Die Verfahrensgrundrechte sind anders als die bisher erläuterten Grundrechte keine Freiheits- oder Abwehrrechte. Sie nehmen von ihrer Struktur her eine Sonderstellung ein (SG BV-Steinmann, Art. 29 N 6 m.w.H.). Bei der Bekämpfung von Desinformation besteht eine zentrale Frage darin, ob (gewisse) Plattformen (minimale) Verfahrensrechte gewähren müssen, die inhaltlich den klassischen Verfahrensgrundrechten entsprechen. Dies wird teilweise mit der Meinungsbildungsrelevanz der grossen Plattformen begründet (Schefer, N 56). Für einen solchen Ansatz mögen auch das Machtgefälle zwischen Plattformen und Nutzer:innen und die unter Umständen weitreichenden Auswirkungen von Entscheidungen sprechen, etwa beim Ausschluss einer professionellen Influencerin von einer wichtigen Plattform.

#### 4. Durch Desinformation betroffene Grundrechte Dritter

Betrachtet man Desinformation aus Sicht der Rezipient:innen, sind vor allem die Bestimmungen des verfassungsmässigen Persönlichkeitsschutzes relevant, dessen Gehalt sich aus verschiedenen Normen (Art. 7 BV, 10 Abs. 1 BV, 13 Abs. 2 BV) ergibt. Dabei stellt sich allerdings die Frage, ob die blossе Konfrontation mit Desinformation den verfassungsmässigen Persönlichkeitsschutz bereits zu verletzen vermag. Erfolgt eine solche Verletzung durch Private, ist das Erfordernis von staatlichen Schutzpflichten zu diskutieren. Ein besonderer Schutz besteht durch Art. 34 Abs. 2 BV im Bereich von Wahlen und Abstimmungen (siehe dazu D.III.3.e.).

##### a) Menschenwürde

Die Verbreitung von Desinformation kann unter Umständen die Menschenwürde verletzen. Nach Art. 7 BV ist die Würde des Menschen zu achten und zu schützen. Obwohl es sich bei dieser Bestimmung um ein subsidiäres Auffanggrundrecht handelt, kann die Menschenwürde

im Rahmen der persönlichen Freiheit als eigenständige Garantie angerufen werden (BBI 1997 I 140 f.).

Art. 7 BV enthält nach seinem Wortlaut eine Handlungsanweisung: Die Menschenwürde ist beim staatlichen Handeln ganz allgemein zu achten und zu schützen. Die Bestimmung hat insofern die Bedeutung eines Leitsatzes für jede staatliche Tätigkeit. Sie bildet als innerster Kern die Grundlage der Freiheitsrechte und dient daher zu deren Auslegung und Konkretisierung (BGE 127 I 6, E. 5.a). Inhaltlich hat Art. 7 BV einen offenen Gehalt und entzieht sich einer abschliessenden positiven Festlegung. Die Menschenwürde ist immer betroffen, wenn in irgendeiner Weise in den unantastbaren Innenraum einer Person eingedrungen wird, was bspw. dann der Fall ist, wenn der Glaube manipuliert wird (BK BV-Belser & Molinari, Art. 7 N 61; Müller, 2005, S. 42). Ein Eingriff in das Grundrecht der Menschenwürde kann nicht gerechtfertigt werden, weil die Menschenwürde unantastbar ist.

#### b) Recht auf persönliche Freiheit

Das Recht auf persönliche Freiheit (Art. 10 Abs. 2 BV) schützt neben den explizit genannten Teilgehalten der körperlichen und geistigen Unversehrtheit sowie der Bewegungsfreiheit alle elementaren Erscheinungen der Persönlichkeitsentfaltung (BGE 133 I 110, E. 5.2; BK BV-Tschentscher, Art. 10 N 32). Im Zusammenhang mit Desinformation ist insb. das Recht auf geistige Unversehrtheit von Bedeutung, die auch als geistige bzw. psychische Integrität bezeichnet wird (BGE 90 I 29, E. 3; BGE 118 Ia 427, E. 4.b). Denn das Recht auf persönliche Freiheit schützt unter anderem die *Entscheidungsfreiheit* der Individuen (Biaggini, Art. 10 N 21). Sie garantiert jeder Person die Freiheit, «eine bestimmte Situation nach eigener Einschätzung zu beurteilen und aufgrund dieser Einschätzung zu handeln» (BGE 133 I 110 E. 5.2; BGE 127 I 6 E. 5a, 11; Botschaft VE 96, 148). Das Recht auf geistige Unversehrtheit garantiert die Freiheit der Willensbildung, nicht aber der Willensbetätigung, insb. ergibt sich daraus keine allgemeine Handlungsfreiheit (BK BV-Tschentscher, Art. 10, N. 53; anders z.B. Auer, Malinvern, Hottelier, N 340 f.; BGE 124 I 85, E. 2.a; BGE 122 I 153 E. 6.b/bb; BGE 119 Ia 460, E. 5.a, m.w.H.).

Grundsätzlich wäre es denkbar, dass Desinformation die Freiheit der Willensbildung einer betroffenen Person beeinträchtigt, weil diese ihre Entscheidung aufgrund falscher Grundlagen fällt und ihren Willen damit nicht frei bilden kann. Dies kann bspw. bei systematischen und anhaltenden Desinformationskampagnen durch den Staat der Fall sein.

#### c) Schutz der Privatsphäre

Ergänzt wird der Schutz der Persönlichkeit durch die Bestimmung von Art. 13 Abs. 1 BV, die insb. den Schutz der Privatsphäre garantiert. Die Abgrenzung von Art. 10 Abs. 2 und Art. 13 Abs. 1 BV kann im Einzelfall schwierig sein, manchmal überschneiden sich die Schutzbereiche. Im Zusammenhang mit Desinformation dürfte Art. 10 Abs. 2 BV in der Regel vorgehen, doch

ist Art. 13 Abs. 1 BV der Vollständigkeit halber ebenfalls zu erwähnen. Geschützt ist insb. das Privatleben einer Person; als positiver Schutzbereich gehört zum Privatleben auch die psychische und physische Integrität (SG BV-Breitenmoser, Art. 13 N 19). Auch hier stellt sich die Frage, ob die Verbreitung von Desinformation die psychische Integrität der Rezipient:innen zu beeinträchtigen vermag.

#### d) Meinungsfreiheit

Wie dargelegt umfasst die Meinungs- und Informationsfreiheit auch das Recht auf freie Meinungsbildung. Dieses betrifft die passive Seite der Meinungsfreiheit und die Freiheit, eine innere Überzeugung zu haben (Weber & Nobel, S. 181). Ergänzt wird die Meinungsfreiheit durch die Informationsfreiheit, die neben der positiven Seite, d.h. dem Recht des Individuums, ungehindert auf allgemein zugängliche Informationen zugreifen zu können, auch die negative Informationsfreiheit umfasst, also den Schutz vor unerwünschten oder aufgedrängten Informationen. Für den Staat bedeutet dies, dass er den Bürger:innen keine Meinungen durch Indoktrinieren oder auf andere Weise aufdrängen darf (Weber & Nobel, S. 181). Aufgrund der objektiv-rechtlichen Dimension kann man sich fragen, ob der Staat dafür zu sorgen hat, dass den Grundrechtsträger:innen keine ungewünschten Informationen aufgedrängt werden.

#### e) Wahl- und Abstimmungsfreiheit

Wie erwähnt schützt Art. 34 Abs. 2 BV die freie Willensbildung der Stimmberechtigten und deren unverfälschte Willensabgabe. Gemäss bundesgerichtlicher Rechtsprechung soll kein Abstimmungs- oder Wahlergebnis anerkannt werden, das nicht den freien Willen der Stimmbürger:innen zuverlässig und unverfälscht zum Ausdruck bringt (BGE 124 I 55, E. 2.a; BGE 138 I 61, E. 6.2). Von der freien Willensbildung umfasst ist unter anderem der Schutz vor unzulässiger Einflussnahme (BK BV-Tschannen, Art. 34 N 32). Der Staat hat sich auf kurze und sachliche Erläuterungen zu beschränken. Private können unter dem Schutz der Meinungs- und Informationsfreiheit grundsätzlich ohne Einschränkung auch Desinformation verbreiten (siehe dazu: Besson & Boillet, S. 246). Davon gibt es allerdings zwei Ausnahmen: Einerseits besteht als Ausdruck staatlicher Schutzpflichten eine Pflicht zur reaktiven Richtigstellung von irreführender oder falscher Propaganda, auch in Wahl- und Abstimmungskämpfen (BK BV-Tschannen, Art. 34 N 34 f.). Eine weitere Ausnahme gilt, wenn die private Einwirkung auf eine Wahl oder Abstimmung in einem derart späten Zeitpunkt mit unwahren oder irreführenden Angaben erfolgt, dass es den Stimmberechtigten nicht mehr möglich ist, sich aus anderen Quellen ein zuverlässiges Bild von den tatsächlichen Verhältnissen zu machen (BGE 135 I 292, E. 4.1.; BGE 119 Ia 271, E. 3.c; BGer 1C\_472/2010 vom 20.01.2011, E. 4; Schefer, N 96). Dieses sog. «Überrumpelungsverbot» geht indessen nicht so weit, dass die Privaten in diesen Konstellationen an die Grundrechte gebunden wären; es auferlegt dem Staat aber die Pflicht, das Resultat einer solchen Abstimmung oder Wahl nicht anzuerkennen (BK BV-Tschannen, Art. 34 N 38).

## 5. Die Grundsatzperspektive besonderer Verbreiter:innen

### a) Ausländische Regierungen und Grundrechte

Die schweizerische Bundesverfassung regelt nicht, wer Träger eines Grundrechts ist. Grundrechte sind in der Regel auf den Schutz von Individuen, d.h. natürlichen und teilweise auch juristischen Personen vor staatlicher Macht und der damit verbundenen Ausübung von Willkür ausgerichtet. Nach bundesgerichtlicher Rechtsprechung stehen die Grundrechte deshalb nur Privatpersonen zu, Gemeinwesen als Inhaber staatlicher Gewalt sind vom Schutz ausgenommen (BGE 125 I 173, E. 1.b). Von diesem Grundsatz gibt es jedoch verschiedene Ausnahmen. So kommt die Meinungsfreiheit auch öffentlich-rechtlichen Angestellten und Amtsträgern in der Schweiz zu, wobei sie sich unter Umständen stärkere Einschränkungen gefallen lassen müssen (Biaggini, Art. 16 N 5). Zur grundrechtlichen Stellung von ausländische Regierungen hat sich das Bundesgericht noch nicht geäußert. Aufgrund des Gesagten ist jedoch davon auszugehen, dass auch sie sich auf die Meinungsäusserungsfreiheit berufen können, unter Vorbehalt gewisser gesetzlicher Einschränkungen.

### b) Social Bots und Grundrechte

Ob die Verbreitung von Meinungen durch Social Bots unter die Meinungsfreiheit fällt, ist umstritten. Die Äusserungen von Bots lassen sich auf diejenigen Menschen zurückführen, welche die Bots programmiert haben (Schefer, N 62); sie sind deshalb von der Meinungsfreiheit der hinter einem Bot stehenden Personen erfasst und entsprechend geschützt (Oehmer, 2019). Gegen den Schutz von Äusserungen von Bots spricht, dass durch deren massenhafte Verwendung ein falscher Eindruck des Meinungsbildes entstehen könnte (Löber & Rossnagel, 158). Dies würde dem Zweck der Meinungsfreiheit – der Gewährleistung des Meinungsdiskurses – zuwiderlaufen. Allerdings liegen zumindest heute in der Schweiz keine Hinweise vor, dass die Verbreitung von Desinformation über Bots ein relevantes Problem ist. Insofern besteht kein Anlass, die von Bots verbreiteten Aussagen vom Schutz der Meinungsfreiheit auszunehmen. Eine staatliche Einschränkung der Verwendung von Bots müsste, wie stets, den Voraussetzungen von Art. 36 BV genügen. Bei der Abwägung der verschiedenen Kriterien könnte durchaus zwischen der direkten eigenen Verbreitung von Meinungen und der Verbreitung mithilfe von Bots unterschieden und Einschränkungen bei einer Verbreitung mithilfe von Bots könnten an weniger strenge Voraussetzungen geknüpft werden.

6. Die Grundrechtsperspektive der identifizierten Verbreitungschanäle von Desinformation Neben der Perspektive der Verbreiter:innen und Rezipient:innen von Desinformation ist auch die grundrechtliche Stellung der identifizierten Verbreitungschanäle zu analysieren, denn Massnahmen, welche die Verbreitungschanäle in die Pflicht nehmen, haben auch deren Grundrechte zu beachten.

Social-Media-Plattformen und Messengerdienste werden in der Regel von Unternehmen und nicht von Einzelpersonen betrieben; diese sind als Gesellschaften organisiert und in aller Regel als juristische Personen zu qualifizieren. Juristische Personen sind vom Grundrechtsschutz nicht ausgeschlossen, soweit die Grundrechte dazu geeignet sind.

Bei Social-Media-Plattformen und Messengerdiensten steht die Wirtschaftsfreiheit im Vordergrund. Grundsätzlich steht es diesen frei, ihr Geschäftsmodell nach eigenem Gutdünken auszugestalten. Staatliche Massnahmen, die sie in dieser Freiheit einschränken, müssen sich an den Voraussetzungen von Art. 36 BV messen.

Die Medienfreiheit ist wegen ihres engen Schutzbereichs weder auf Social Media noch auf Messengerdienste anwendbar (siehe dazu oben D.II.4.). Die Betreiber von Social Media und Messengerdiensten können sich aber, soweit sie eigene Meinungen kundtun, auf die Meinungsäusserungsfreiheit berufen, die natürliche und juristische Personen gleichermassen schützt (CR CF-Cottier, Art. 16 N 16; Weber & Nobel, S. 177).

#### 7. Zusammenfassung der grundrechtlichen Situation

Die grundrechtliche Situation ist vielschichtig, weil verschiedene Grundrechte der von Desinformation betroffenen Gruppen betroffen sind. Bei den Verfasser:innen und Verbreiter:innen von Desinformation steht die Meinungsäusserungsfreiheit (Art. 16 BV) im Vordergrund. Diese steht teilweise im Konflikt mit der Informationsfreiheit der Rezipient:innen und möglicherweise auch mit dem verfassungsmässigen Persönlichkeitsschutz (Art. 7, Art. 10 Abs. 1, Art. 13 Abs. 2 BV). Geht es um die politische Willensbildung, kann auch Art. 34 Abs. 2 BV einschlägig sein. Bei den Kanälen steht die Wirtschaftsfreiheit im Vordergrund. Zudem können sich Social-Media-Anbieter:innen, Messengerdienste und Medien auch selbst auf die Meinungsfreiheit berufen, wenn sie eigene Inhalte verbreiten.

## IV. Rechtsrahmen für rechtswidrige Desinformation in der Schweiz

### 1. Vorbemerkungen

Normen, die spezifisch dem Schutz vor Desinformation dienen, existieren in der Schweiz nicht. Es gibt aber eine Reihe von Normen, die Sachverhalte erfassen, die sich als Desinformation bezeichnen lassen. Diese dienen primär dem Schutz von Individuen. Einige Normen bezwecken aber auch den Schutz von Systemen, insb. den Schutz der demokratischen Willensbildung, des Funktionierens des Marktes und des öffentlichen Friedens oder der Gesundheit.

### 2. Materielles Recht: Normen gegen Desinformation

#### a) Individuelle Ebene

Das Schweizer Recht enthält eine Reihe von Normen, die Individuen davor schützen, Gegenstand von Desinformation zu werden. Gewisse Desinformationsinhalte können rechtlich eine Persönlichkeits- oder eine Ehrverletzung i.S.v. Art. 28 ff. des Schweizerischen Zivilgesetzbuchs («ZGB») oder der Art. 173 ff. des Schweizerischen Strafgesetzbuchs («StGB») sein. Der zivilrechtliche Ehrbegriff schützt sowohl das berufliche als auch das gesellschaftliche Ansehen einer Person (BK ZGB I-Meili, Art. 28 N 17; BGE 129 III 715, E. 4.1). Der strafrechtliche Begriff der Ehre ist enger; geschützt wird hier nach der Rechtsprechung des Bundesgerichts nur der «Ruf, ein ehrbarer Mensch zu sein, d.h. sich so zu benehmen wie nach allgemeinen Anschauungen ein charakterlich anständiger Mensch sich zu verhalten pflegt.» (BGE 93 IV 21; BGE 103 IV 158; BGE 105 IV 112; BGE 105 IV 195; BGE 117 IV 28). Damit erfasst der strafrechtliche Schutz die berufliche Ehre nicht (BGE 119 IV 44, E. 2.a, m.w.H.; dazu grundlegend: PK StGB-Trechsel & Lehmkuhl, Vor Art. 173 N 3).

Ein Verstoss gegen diese Normen kann unter gewissen Voraussetzungen gerechtfertigt werden, so etwa wenn überwiegende Interessen vorliegen. Ein überwiegendes Interesse an der Verbreitung von Unwahrheiten dürfte jedoch in den wenigsten Fällen bestehen (BGE 138 III 641, E. 4.1.2).

Aus zivilrechtlicher Sicht stehen dem verletzten Individuum negatorische Ansprüche auf Beseitigung, Unterlassung und Feststellung der Persönlichkeitsverletzung sowie Berichtigung und Urteilspublikation zu sowie auch reparatorische Ansprüche auf Schadensersatz, Genugtuung und Herausgabe des Gewinns (Art. 28b ZGB). Die Ansprüche können gegenüber jeder Person geltend gemacht werden, die an der Verletzung mitwirkt (Art. 28 Abs. 1 ZGB; siehe dazu BGer 5A\_792/2011 vom 14. Januar 2013, [kein Haftungsprivileg für Hostingprovider]). Die Rechtsdurchsetzung ist im Online-Kontext – auch bei Persönlichkeitsverletzungen – allerdings schwierig (siehe dazu Rosenthal, 2013).

Strafrechtlich verfolgt werden können nicht nur die Verfasser:in, sondern jede Person, welche die Aussage weiterleitet. Darunter fällt gemäss bundesgerichtlicher Rechtsprechung auch das Liken und Teilen von Inhalten (BGer 6B\_1114/2018 vom 29.01.2018, dazu die Entscheidung Simmler M., AJP 2020, S. 658; Riedo & Beglinger, 1249).

## b) Gesamtgesellschaftliche Ebene

### i. *Wettbewerb*

Marktbezogene Desinformation kann das Funktionieren des Wettbewerbs beeinträchtigen. Die Verbreitung von marktrelevanten Informationen muss deshalb den Anforderungen des Bundesgesetzes gegen den unlauteren Wettbewerb («UWG») genügen.

Dazu muss das Verbreiten von Desinformation als Wettbewerbshandlung zu qualifizieren sein, d.h. die Information muss objektiv auf die Beeinflussung des Wettbewerbs gerichtet sein und nicht in einem völlig anderen Zusammenhang erfolgen. Erfasst werden damit nur marktrelevante, marktgeneigte und wettbewerbsgerichtete Informationen (BGE 120 II 76, E. 3.a). Gegen das UWG können nicht nur Anbieter von Waren oder Dienstleistungen, sondern alle Personen verstossen, auch Aussenstehende, deren Verhalten den Wettbewerb zu beeinflussen vermag. Ein Wettbewerbsverhältnis zwischen den Beteiligten ist nicht vorausgesetzt (BGE 120 II 76, E. 3.a; BK UWG-Rüetschi, Art. 11 N 3 m.w.H.). Auch Äusserungen, die in sozialen Medien oder über einen anderen Kanal getätigt werden, unterstehen dem UWG.

In materieller Hinsicht stehen bei Desinformation die Tatbestände der Herabsetzung (Art. 3 Abs. 1 lit. a UWG), der Irreführung (Art. 3 Abs. 1 lit. b UWG) und der vergleichenden Werbung (Art. 3 Abs. 1 lit. e UWG) im Vordergrund. Sind die Voraussetzungen dieser Tatbestände nicht erfüllt, kann allenfalls die Generalklausel (Art. 2 UWG) Anwendung finden. Durchgesetzt werden kann das UWG auf zivilrechtlichem Weg (Art. 9 ff. UWG); für die hier infrage stehenden Tatbestände steht auch eine strafrechtliche Durchsetzung zur Verfügung (Art. 23 UWG).

### ii. *Öffentliche Gesundheit*

Normen, welche die öffentliche Gesundheit vor Beeinflussung durch Desinformation schützen sollen, finden sich im Bundesgesetz über Arzneimittel und Medizinprodukte («HMG») und im Bundesgesetz über Lebensmittel und Gebrauchsgegenstände («LMG»). Für verschreibungspflichtige Medikamente darf keine Publikumswerbung gemacht werden, für die übrigen Arzneimittel gilt ein Verbot von irreführender oder der öffentlichen Ordnung und den guten Sitten widersprechender Werbung (Art. 32 HMG). Ähnliche Restriktionen gelten gemäss Art. 69 Abs. 2 MepV mit dem Verbot irreführender Angaben, insb. über Zweckbestimmung, Sicherheit und Leistungsfähigkeit eines Produktes. Auch für Lebensmittel (inkl. Bedarfsgegenstände und Kosmetika) besteht ein Verbot irreführender Angaben (Art. 18 LMG). Diese Bestimmungen schützen umfangreich vor Desinformation im Bereich von Heil- und Lebensmitteln und dienen dadurch unter anderem dem Schutz der öffentlichen Gesundheit.

Die Durchsetzung dieser Bestimmungen erfolgt durch die zuständige öffentliche Behörde, im Arzneimittelbereich etwa durch Swissmedic, im Bereich der Lebensmittel durch das Bundesamt für Lebensmittelsicherheit und Veterinärwesen und die zuständigen kantonalen Vollzugsstellen (Kantonsärzt:in, -chemiker:in sowie Lebensmittelinspekteur:innen und -kontrolleur:innen).

### iii. *Öffentlichen Frieden*

Zum Schutz des öffentlichen Friedens stehen verschiedene strafrechtliche Normen zur Verfügung. Das Strafrecht schützt nicht nur das Individuum, sondern teilweise auch kollektive Rechtsgüter, wie den öffentlichen Frieden, also «das Vertrauen der Gesamtheit oder von Teilen der Bevölkerung in den Bestand der und in den Schutz durch die Rechtsordnung» (Müller, 1996, S. 664; BK StGB-Trechsel & Vest, Art. 261<sup>bis</sup> N 6, m.w.H. auf abweichende Ansichten). Nach den Delikten gegen den öffentlichen Frieden macht sich strafbar, wer der Bevölkerung wider besseres Wissen eine Gefahr für Leib, Leben oder Eigentum vorspiegelt (Art. 258 StGB, «Schrecken der Bevölkerung»). Auch macht sich strafbar, wer öffentlich gegen eine Person oder eine Gruppe von Personen wegen eines geschützten Merkmals wie Rasse oder Religion (nicht nationale Zugehörigkeit allein; siehe dazu: BK StGB-Schleiminger Mettler, Art. 261<sup>bis</sup> N 16; PK StGB-Trechsel & Vest, Art. 261<sup>bis</sup> N 12, je m.w.H.) zu Hass oder Diskriminierung aufruft (Art. 261<sup>bis</sup> Abs. 1 StGB), wer gewisse herabsetzende oder verleumdende Ideologien öffentlich verbreitet (Art. 261<sup>bis</sup> Abs. 2 StGB) oder Verbrechen gegen die Menschlichkeit leugnet, verharmlost oder rechtfertigt (Art. 261<sup>bis</sup> Abs. 4 StGB; grundlegend zu Art. 261<sup>bis</sup> StGB, «Diskriminierung und Aufruf zu Hass»: Müller, 1996, S. 665 ff.).

### iv. *Verfassungsmässige Ordnung und Demokratie*

Auch Desinformation, welche auf die Störung der verfassungsmässigen Ordnung abzielt (Art. 275 StGB, «Angriffe auf die verfassungsmässige Ordnung») ist strafbar. Weiter kann mit Strafe belegt werden, wer durch die Verbreitung von Desinformation die Beziehungen zum Ausland stört, etwa durch die Beleidigung eines fremden Staates (Art. 296 StGB) oder einer zwischenstaatlichen Organisation (Art. 297 StGB).

## 3. Einhaltung und Durchsetzung: Besondere Akteur:innen

### a) Journalistische Medien

In materieller Hinsicht gelten für journalistische Medien dieselben gesetzlichen Bestimmungen für die Verbreitung von Desinformation wie für Individualpersonen. Insb. können sie für Persönlichkeitsverletzungen zivil- und strafrechtlich belangt werden. Auch Medien können gegen die Vorgaben des UWG verstossen, wenn sich die verbreiteten Informationen auf den Wettbewerb auswirken können. Ein besonderer Rechtsbehelf besteht gegenüber periodisch erscheinenden Medien: die betroffene Person kann hier von den Medien bei Vorliegen einer Tatsachenbehauptung eine Gegendarstellung verlangen (Art. 28g ZGB).

Das Strafrecht enthält für die Verfolgung zwei Sonderbestimmungen für journalistische Medien: Art. 28 StGB sieht bei einer strafbaren Handlung durch die Veröffentlichung in einem Medium, die sich in der Veröffentlichung erschöpft (sog. Mediendelikte) eine Kaskadenordnung vor; strafbar ist allein der Autor. Nur wenn dieser nicht ermittelt oder in der Schweiz nicht vor ein Gericht gestellt werden kann, ist der zuständige Redaktor oder jene Person, die für die Ver-

öffentlichung verantwortlich ist, strafbar. Durchgesetzt wird die Bestimmung durch Art. 322 StGB, welcher den Medien eine Auskunftspflicht über die verantwortlichen Personen auferlegt. Ergänzend enthält das Strafgesetzbuch eine Bestimmung, welche die vorsätzliche oder fahrlässige Nichtverhinderung einer strafbaren Veröffentlichung sanktioniert (Art. 322<sup>bis</sup> StGB).

#### b) Ausländische Regierungen

Für ausländische Regierungen, die als Verfasser- oder Verbreiter:innen von Desinformation infrage kommen, gelten Besonderheiten bei der Verfolgbarkeit. Das Völkergewohnheitsrecht sieht für Staats- und Regierungschefs sowie Aussenminister während ihrer Amtsdauer eine unbegrenzte Immunität in zivil- und strafrechtlichen Angelegenheiten vor (BK StGB-Fiolka, Art. 264n N 6). Diese diplomatische Immunität ist in Art. 29 und Art. 38 des Wiener Übereinkommens über diplomatische Beziehungen (SR 0.191.01) ausdrücklich verankert. Eine Verfolgung von Handlungen, die von ausländischen Regierungsvertretern begangen wurden, ist somit nicht möglich. Desinformationskampagnen werden sich allerdings ohnehin nur in den wenigsten Fällen direkt auf ausländische Regierungen zurückführen lassen.

Bei Online-Aktivitäten von fremden Nachrichtendiensten kann auf Grundlage des Nachrichtendienstgesetzes («NDG») vorgegangen werden (Heimgartner, S. 1218). Wenn Gegenmassnahmen ergriffen werden sollen, z.B. durch Eindringen in ein Computersystem oder dessen Zerstörung, ist eine Genehmigung des Bundesverwaltungsgerichts erforderlich (Art. 26 lit. d i.V.m. Art. 29 NDG) für Massnahmen, die in der Schweiz durchgeführt werden sollen oder des Bundesrates bzw. der VBS-Vorsteherin (Art. 37 NDG) für Massnahmen, die im Ausland durchgeführt werden sollen (Heimgartner, S. 1220).

#### c) Social Bots

Das schweizerische Recht kennt keine Normen, die sich spezifisch an Social Bots richten. Diese sind weder als natürliche noch juristische Person zu qualifizieren; Social Bots sind damit nicht rechtsfähig und können weder zivil- noch strafrechtlich belangt werden. Verantwortlich für die Inhalte ist die hinter dem Bot stehende Person, auch wenn es oft schwierig oder gar unmöglich sein dürfte, diese Person zu identifizieren.

### 4. Einhaltung und Durchsetzung: Kanäle

#### a) Verantwortung für Inhalte Dritter

Seit dem Aufkommen von Hosting Providern und Online-Plattformen wird die Frage diskutiert, ob und inwieweit diese Akteur:innen für rechtswidrige Inhalte, die von ihren Nutzer:innen verbreitet werden, verantwortlich gemacht werden können.

Aus zivilrechtlicher Sicht geht man bei Social Media davon aus, dass zumindest negatorische Ansprüche, bspw. der Anspruch auf Sperrung oder Löschung eines Beitrages, gegen den Anbieter eines Social-Media-Dienstes geltend gemacht werden können (siehe dazu: Thouvenin, S. 155 ff.; zur Problematik der Rechtsdurchsetzung siehe auch: Heinemann & Althaus,

N 15 ff.). Weil Messengerdienste keine Möglichkeit haben, von den über sie verbreiteten Inhalte Kenntnis zu nehmen, können keine Ansprüche gegen sie geltend gemacht werden.

Aus strafrechtlicher Sicht steht eine mögliche Gehilfenschaft (Art. 25 StGB) von Social Media im Vordergrund. In einem älteren Entscheid hat das Bundesgericht entschieden, dass sich der Betreiber einer Fernmeldeeinrichtung der Gehilfenschaft (Art. 197 StGB, Pornographie) strafbar machte, weil er wusste, dass er die Haupttat aktiv förderte (BGE 120 IV 109, E. 3.b). Eine aktive Gehilfenschaft einer Social-Media-Plattform wird in den meisten Fällen aber abgelehnt, weil es am erforderlichen Vorsatz fehlt. Ebenso besteht keine strafrechtliche Pflicht, Informationen vor Veröffentlichung auf ihre Rechtmässigkeit zu überprüfen. Neben der aktiven Gehilfenschaft kann die Förderung der Haupttat auch durch Unterlassung erfolgen, wenn eine Pflicht zum Eingreifen oder Handeln besteht, wenn also eine sog. Garantenstellung vorliegt. Ob und unter welchen Voraussetzungen eine solche anzunehmen ist, ist derzeit noch ungeklärt (Riedo & Beglinger, N 64).

#### b) Selbstorganisation: Eigene Massnahmen der Durchsetzung

Bereits vor Erlass der einschlägigen Regelungen durch die EU haben Plattformen in ihren Nutzungsbedingungen gewisse Regeln zur Moderation von Inhalten festgelegt, die den Umgang mit Desinformation betreffen. Erlassen wurden sowohl inhaltsbezogene Regeln, die gewisse Leitplanken für zulässige und unzulässige Inhalte setzen, als auch prozessuale Regeln, die festlegen, was passiert, wenn ein:e Nutzer:in gegen die Inhaltsregeln verstösst. Die Rechtsnatur dieser Nutzungsregeln, auch als «Community Guidelines» oder «Community Standards» bezeichnet, ist allerdings unklar.

Häufig dürfte es sich dabei um AGB handeln, also um vorformulierte, nicht einzeln ausgehandelte vertragliche Bestimmungen, die meist beim Abschluss einer Vielzahl von Verträgen eingesetzt werden (zum Begriff der AGB jüngst: BGer 4A\_330/2021 vom 5. Januar 2022, E. 2; grundlegend: Schnyder, S. 40). Voraussetzung für diese Qualifikation ist allerdings, dass zwischen den Social-Media-Anbieter:innen bzw. den Messengerdiensten und den Nutzer:innen ein Vertragsverhältnis besteht, was insb. bei Anbietern, bei welchen die Nutzung nicht gegen Bezahlung erfolgt, nicht ohne weiteres angenommen werden kann. Insb. kann es in solchen Fällen vonseiten der Nutzer:innen an einem Rechtsbindungswillen fehlen, also am Willen, einen Vertrag einzugehen und sich dadurch rechtlich zu binden. Da nun sogar der Digital Services Act von AGB spricht, scheint sich jedoch das Verständnis durchgesetzt zu haben, dass zwischen der Plattform und den Nutzer:innen ein Vertrag zustande kommt. Die in den Community Standards festgelegten Regeln sind in der Schweiz oft das Resultat (ko-)regulatorischer Vorgaben auf EU-Ebene (siehe zu diesen vorstehend: II.). Dies ändert aber nichts daran, dass es in der Schweiz keine besonderen staatlichen Anforderungen an die Community Standards der Plattformen gibt, die neben den allgemeinen vertraglichen Grundsätzen gelten.

Damit AGB Geltung erlangen, müssen sie in den Vertrag übernommen werden. Dazu müssen die Nutzer:innen auf die AGB hingewiesen werden und von deren Inhalt in zumutbarer Weise Kenntnis genommen haben. Die Schweiz kennt keine offene Inhaltskontrolle von AGB-Klauseln. Nach der Gerichtspraxis gelten aber die sog. Ungewöhnlichkeits- und die Unklarheitenregel sowie die Vorgaben von Art. 8 UWG. Nach der Ungewöhnlichkeitsregel werden ungewöhnliche Klauseln, d.h. solche, die einen geschäftsfremden Inhalt aufweisen, nicht Vertragsinhalt, wenn die übernehmende Partei nicht gesondert auf die Klausel hingewiesen wurde. Nach der Unklarheitenregel wird eine unklar formulierte Bestimmung im Zweifelsfall zuungunsten derjenigen Partei ausgelegt, welche diese verfasst hat («in dubio contra stipulatorem») (BGE 115 II 264; BGE 124 III 155). Schliesslich sind gemäss Art. 8 UWG AGB nichtig, die in Treu und Glauben verletzender Weise zum Nachteil der Konsument:innen ein erhebliches und ungerechtfertigtes Missverhältnis zwischen den vertraglichen Rechten und den vertraglichen Pflichten vorsehen. Soweit ersichtlich, genügen die heutigen AGB der Plattformen und Messengerdiensten diesen Anforderungen.

i. *Sperren von Nutzer:innen*

Der DSA verpflichtet die ihm unterstehenden Plattformen, Nutzer:innen, die häufig und offensichtlich rechtswidrige Inhalte verbreiten, zumindest temporär zu sperren. Social Media und Messengerdienste behalten sich das Sperren von Nutzer:innen denn auch schon länger vor. YouTube sieht nach mehreren Verwarnungen («strikes»), welche ggf. mit temporären Sperrungen verknüpft sind, auch die Möglichkeit einer definitiven Sperrung von Nutzer:innen vor. Dies setzt einen Verstoss gegen die Community Guidelines und die darin enthaltenen «Misinformation Policies» voraus. Facebook und Instagram geben pauschal an, Nutzerkonten zu sperren, die mehrfach gegen Gemeinschaftsstandards oder Werberichtlinien verstossen.

Auch Messengerdienste behalten sich das Recht vor, Accounts im Falle von Verstössen gegen die Nutzungsbedingungen zu sperren. Nach den Nutzungsbedingungen von WhatsApp kann «das Veröffentlichen von Unwahrheiten, Falschdarstellungen oder irreführenden Aussagen» zu einer «Deaktivierung» oder «Suspendierung» des Accounts führen. Diese erfolgen nach den Bestimmungen über die Kündigung, wobei ein Verstoss gegen die Nutzungsbedingungen ausdrücklich als Kündigungsgrund genannt wird. Eine Kündigung darf allerdings grundsätzlich erst nach einer Verwarnung bzw. einer Abhilfefrist erfolgen; vorbehalten bleibt die Möglichkeit zur fristlosen Kündigung bei besonders schwerwiegenden Verstössen. Die Nutzungsbedingungen von Threema sehen die Möglichkeit der temporären oder dauerhaften Sperrung der Nutzer:innen vor. Die Verbreitung von Desinformation wird nicht explizit als Grund für eine Sperrung genannt, wohl aber illegale Handlungen oder ein Verstoss gegen die Nutzungsbedingungen. Bei Telegram finden sich keine Ausführungen zu den Voraussetzungen für die Sperrung eines Nutzeraccounts.

## ii. *Löschung und Sperrung von Inhalten*

Die konsequenteste Massnahme, die gegen (illegale) Desinformation ergriffen werden kann, besteht darin, die Information zu sperren bzw. zu löschen. In sozialen Netzwerken geschieht dies in der Regel über ein Notice-and-Take-Down-Verfahren, bei welchem die Nutzer:innen den Anbietern die rechtswidrigen Inhalte durch ein Meldeverfahren (dazu unten IV.4.b.viii) melden, worauf diese die Inhalte löschen.

Youtube löscht oder sperrt rechtswidrige und rechtmässige Inhalte, die gegen die Community Guidelines und die darin enthaltenen «Misinformation Policies» verstossen. Dasselbe gilt für TikTok. Ähnlich agiert Facebook.

Auch Telegram sieht die Möglichkeit vor, gegen Inhalte vorzugehen. Der Anbieter differenziert dabei zwischen illegalen Inhalten in «Telegram- und Gruppenchats» und solchen in öffentlichen «Sticker-Paketen, Kanälen und Bots.» Erstere sind gemäss Nutzungsbedingungen «Privatsache der jeweiligen Nutzer», weshalb Telegram auch keine diesbezüglichen Anfragen bearbeitet. Bei den öffentlich zugänglichen Sticker-Paketen, Kanälen und Bots sind Anfragen zu als illegal erachteten Inhalten an Telegram via E-Mail oder Meldung über das Profil zulässig. Telegram behält es sich vor, «Sticker, die Rechte an geistigem Eigentum verletzen oder Porno-Bots offline [zu] nehmen». Dies bedeute indes gemäss Angaben von Telegram «nicht, dass dies durch die lokalen Beschränkungen der Redefreiheit geschieht.» Diese etwas unklare Formulierung ist wohl dahingehend zu verstehen, dass die Löschung und Sperrung von Inhalten sehr zurückhaltend vorgenommen wird.

Bestimmungen betreffend die Löschung und Sperrung von Inhalten fehlen in den Nutzungsbedingungen von Threema und Whatsapp; beide Messengerdienste sehen aber eine Meldemöglichkeit vor (dazu unten IV.4.b.viii).

## iii. *Downranking*

Die Aufmerksamkeitsökonomie von sozialen Netzwerken führt dazu, dass vor allem diejenigen Inhalte sichtbar sind, welche häufig und zuerst erscheinen. Indem gewisse Inhalte weniger oder in einem Feed weiter unten angezeigt werden, kann faktisch eine ähnliche Wirkung erreicht werden wie durch eine Löschung. Downranking wird als Massnahme gegen die Verbreitung von Desinformation eingesetzt, so etwa auf TikTok, Facebook und Instagram. So gibt Meta (für Facebook und Instagram) an, dass «Inhalte, bei denen es sich um Fehlinformationen handeln könnte, [...] bis zu ihrer Überprüfung möglicherweise vorübergehend weiter unten im Feed angezeigt [werden]».

Die Nachrichten in Messengerdiensten werden in chronologischer Reihenfolge angezeigt, Downranking ist folglich nicht möglich.

#### iv. *Meldung/Flagging durch andere Nutzer:innen (User Flagging)*

Nutzer:innen von Social Media und Messengerdiensten kann die Möglichkeit eingeräumt werden, falsche oder irreführende Inhalte bei einer internen Stelle zu melden. Für Hostinganbieter:innen im Sinne des DSA ist die Einrichtung eines Meldeverfahrens zwingend (Art. 16 DSA, dazu vorne II.2.b). Anknüpfend an die Meldung können die sozialen Netzwerke oder Messengerdienste sodann weitere Massnahmen ergreifen wie Tagging (Kennzeichnung) oder auch Downranking oder Löschung. Solche Meldemöglichkeiten bestehen bei Youtube, Facebook und Instagram.

Auch Messengerdienste sehen die Möglichkeit vor, gewisse Inhalte zu melden. Bei WhatsApp können «Urheberrechts-, Marken- oder sonstige Verletzung bezüglich des geistigen Eigentums Dritter» via E-Mail oder auf dem Postweg gemeldet werden, auch bei können unzulässige Inhalte via E-Mail gemeldet werden. Explizite Vorschriften betreffend die Meldung von Desinformationsinhalten – die zwar zu einer Sperrung des Nutzeraccounts führen können (siehe dazu C.V.2.) – bestehen in den Nutzungsbedingungen nicht.

Auf Telegram können öffentliche «Sticker-Pakete, Kanälen und Bots» via E-Mail oder Meldung über das Profil gemeldet werden. Bei Threema können Nutzer:innen, die feststellen, dass ein anderer Nutzer unzulässige Inhalte über den Dienst kommuniziert, dies via E-Mail melden.

#### v. *Flagging durch besonders qualifizierte Personen (Trusted Flagger)*

Eine weitere Massnahme kann darin liegen, dass (nur) besonders qualifizierte Dritte zweifelhafte Inhalte melden können (Trusted Flagging) oder solche Meldungen schneller behandelt werden. Der DSA verlangt von Online-Plattformen, dass Meldungen von besonders vertrauenswürdigen Hinweisgebern vorrangig behandelt werden (Art. 22 DSA). Die von uns untersuchten Social-Media-Anbieter haben entsprechend Trusted Flagger. Bei Youtube gibt es bspw. ein spezifisches Trusted Flagger Programm.

Die Nutzungsbedingungen der von uns betrachteten Messengerdienste enthalten keine Bestimmungen zu Trusted Flaggers.

#### vi. *Automatisierte Entfernung von Desinformation*

Viele Social-Media-Anbieter:innen verwenden technische Hilfsmittel, um Inhalte, die gegen die Community Guideline verstossen, zu identifizieren, so etwa YouTube oder TikTok. Die automatisierte Prüfung kann präventiv, d.h. vor Erscheinen auf der Plattform erfolgen (Upload Filter) oder laufend, wenn die Inhalte bereits auf der Plattform aufgeschaltet sind (z.B. Copyright Match Tool oder Content ID von Youtube, die bei der Identifikation von urheberrechtsverletzenden Inhalten helfen).

Die Inhalte von Messengerdiensten werden nicht automatisiert geprüft.

vii. *Kennzeichnung und Einschränkung der Möglichkeit, Inhalte und Nachrichten weiterzuleiten*

Die Frage der Kennzeichnung und Einschränkung der Möglichkeit, Nachrichten weiterzuleiten, stellt sich primär bei Messengerdiensten. Bei Social Media geht es um das Teilen und Liken von Inhalten, die zu deren Weiterverbreitung führen. Die Kennzeichnung und/oder Einschränkung weitergeleiteter Nachrichten kann sinnvoll sein, um auf die massenhafte Verbreitung von Desinformation aufmerksam zu machen und dieser entgegenzuwirken.

Auf WhatsApp werden weitergeleitete Nachrichten als solche gekennzeichnet. Wenn eine Nachricht über mehr als fünf Personen weitergeleitet wurde, wird die Nachricht als «häufig weitergeleitet» gekennzeichnet. Ein Versand an mehrere Personen ist bei Whatsapp nur über die Broadcast-Funktion möglich und beschränkt sich auf 256 Empfänger:innen.

Auf Telegram können nur Nachrichten aus «geheimen Chats» nicht weitergeleitet werden. Darüber hinaus bestehen, soweit ersichtlich, keine Einschränkungen. Die Nutzungsbedingungen von *Threema* enthalten keine Bestimmungen zur Weiterleitung von Nachrichten.

viii. *Kennzeichnungspflicht oder Verbot von Social Bots*

Social Bots lassen sich nicht immer leicht von menschlichen Nutzer:innen unterscheiden. Durch ein Verbot oder eine Kennzeichnungspflicht von Social Bots soll verhindert werden, dass Personen über den/die Verfasser:in eines Inhaltes getäuscht werden.

Youtube sieht keine Vorschriften zu Social Bots vor; die Video-Plattform eignet sich allerdings auch wenig zur Verwendung von Bots. Die Nutzungsbedingungen von Meta untersagen es den Nutzer:innen, sich als eine andere Person auszugeben, zudem ist unauthentisches Verhalten unzulässig.

Die Nutzungsbedingungen von WhatsApp statuieren implizit ein Verbot von Social Bots, indem es untersagt ist, «Accounts für unsere Dienste über nicht autorisierte oder automatisierte Mittel erstellen». Anders verhält es sich auf Telegram; hier findet man auf der Webseite sogar eine Anleitung zur Erstellung von Bots. Eine Kennzeichnungspflicht für Social Bots wird nicht statuiert. *Threema* verbietet den Zugriff mittels Bots, wenn dies nicht über die offiziellen APIs erfolgt.

ix. *Implementierung von Beschwerdeverfahren*

Bereits vor dem Inkrafttreten des DSA boten viele Plattformen von sich aus ein Beschwerdeverfahren an. Massnahmen der Plattformen wie Löschungen oder Sperrungen sind für die betroffenen Personen mit einer Rechtsfolge verbunden; entsprechend sind, wenn die Massnahmen automatisiert erfolgen, die datenschutzrechtlichen Vorgaben zu den automatisierten Entscheidungen anwendbar (Art. 21 DSGVO; Art. 22 DSGVO), die den betroffenen Personen unter Umständen die Möglichkeit zur Stellungnahme einräumen. Darüber hinaus besteht ein

Anspruch auf Information über die Logik, auf der die Entscheidung beruht (Art. 25 Abs. 2 lit. f DSGVO; Art. 13 Abs. 2 Bst. f; Art. 14 Abs. 2 Bst. g; Art. 15 Abs. 1 Bst. h DSGVO).

## 5. Zwischenfazit

Der bestehende Rechtsrahmen in der Schweiz zum Schutz vor Desinformation scheint aus materiell-rechtlicher Sicht genügend. Probleme bestehen auf Ebene der Einhaltung und Durchsetzung. Hier kann die Schweiz teilweise von der neuen EU-Regulierung profitieren. Es bestehen aber einige Gesetzeslücken, weshalb man sich überlegen kann, ergänzende Vorschriften zu erlassen, bspw. eine Pflicht zur Benennung eines Ansprechpartners.

## V. Rechtsrahmen für rechtmässige Desinformation in der Schweiz

### 1. Vorbemerkung

Nicht nur unrechtmässige, sondern auch an sich zulässige «legale» Desinformationsinhalte können gefährlich sein, bspw. weil sie die öffentliche Meinungsbildung manipulieren. Dadurch kann das Vertrauen in staatliche Institutionen und traditionelle Medien untergraben werden, was im schlimmsten Fall zu einer Spaltung der Gesellschaft führt. Aus diesem Grund haben verschiedene Akteur:innen auch Massnahmen gegen legale Desinformation ergriffen.

### 2. Materielles Recht: Normen gegen rechtmässige Desinformation

#### a) Schutz der politischen Willensbildung

##### i. *Politische Rechte, Art. 34 Abs. 2 BV*

Art. 34 Abs. 2 BV schützt bei eidgenössischen Wahlen und Abstimmungen unter anderem «die freie Willensbildung und die unverfälschte Stimmabgabe» der Individuen. Zweck der Bestimmung ist jedoch auch der Schutz des demokratischen Systems als solches (vgl. BK BV-Tschannen, Art. 34 N 2); nach der Rechtsprechung des Bundesgerichts soll die Norm zudem sicherstellen, dass «kein Abstimmungs- oder Wahlergebnis anerkannt wird, das nicht den freien Willen der Stimmbürger zuverlässig und unverfälscht zum Ausdruck bringt» (BGE 141 II 297, E. 5.2, m.w.H; erstmals: BGE 75 I 244). Ein Teilgehalt des Schutzes der freien Willensbildung und der unverfälschten Stimmabgabe ist der Schutz vor unzulässiger Einflussnahme, etwa durch Desinformation vonseiten des Staates und der Behörden (vgl. BK BV-Tschannen, Art. 34 N 3; BGE 140 I 338, E. 5; Besson & Boillet, S. 238 ff.; vgl. auch die gesetzlichen Anforderungen an die Information der Stimmberechtigten in eidgenössischen Stimmrechtssachen, wie etwa in Art. 10a des Bundesgesetzes über die politischen Rechte («BPR») ausgeführt).

Soweit der Desinformationsinhalt nachweisbar entscheidende Auswirkung auf das Ergebnis einer Wahl oder Abstimmung hatte, muss diese wiederholt werden (BGE 112 Ia 129, E. 3.a; grundlegend m.w.H. Biaggini, Art. 34 N 8). Da Desinformation in der Schweiz primär durch private Personen verfasst und verbreitet wird (siehe dazu B.I.1.), verschaffen Art. 34 Abs. 2 BV und seine Ausführungsbestimmungen in den eidgenössischen, kantonalen oder

kommunalen Erlassen wenig Abhilfe zur Bekämpfung von Desinformation, weil sie vor allem vor Einflussnahme vonseiten des Staats schützen sollen.

Ein staatliches Eingreifen und eine Wiederholung einer Wahl oder Abstimmung sind bei privater Einflussnahme nur in seltenen Fällen zulässig bzw. erforderlich. Voraussetzung ist, dass objektiv falsche Tatsachen verbreitet wurden (keine blossen Werturteile), die erheblich für die Abstimmung sind, aufgrund des späten Vorbringens nicht mehr richtiggestellt werden können und deren Auswirkung auf die Wahl oder Abstimmung «ausser Zweifel steht oder zumindest als sehr wahrscheinlich erscheint» (BGE 135 I 292, E. 4.1, E. 4.3 f.; BGE 119 Ia 271, E. 3.c; siehe dazu grundlegend: Besson & Boillet, S. 239; Buser, S. 8 ff.). Das Bundesgericht hat die Wiederholung einer Wahl aufgrund privater Einflussnahme etwa in einem Fall bejaht, in dem die Stimmberechtigten keine Möglichkeit hatten, einen für die Abstimmung relevanten Vertrag einzusehen und zu prüfen, welcher unter Nennung objektiv unzutreffender Tatsachenbehauptungen in einem äusserst späten Zeitpunkt (in der Gemeindeversammlung am Tag der Wahl) präsentiert wurden (BGE 135 I 292, E. 4.1, E. 4.3). Soweit ersichtlich, gibt es bis heute keine Entscheide des Bundesgerichts zu privater Einflussnahme auf Wahlen oder Abstimmungen über digitale Kanäle. Da im digitalen Kontext eine Richtigstellung – insb. durch Kommentare anderer Nutzer:innen – rasch und jederzeit möglich ist, werden die Anforderungen an eine Wiederholung von Wahlen oder Abstimmungen aufgrund von über digitale Kanäle verbreiteter Desinformation hoch sein.

#### ii. *Strafrechtlicher Schutz der politischen Willensbildung*

Der Schutz der politischen Willensbildung ist teilweise strafrechtlich abgesichert durch die Bestimmungen über die Vergehen gegen den Volkswillen (Art. 279-283 StGB). Diese Tatbestände schützen in erster Linie die Ausübung des Stimm- und Wahlrechts, nicht aber die Meinungsbildung im Vorfeld von Abstimmungen (vgl. Art. 279-286 StGB sowie Egli & Rechsteiner, S. 253). Auch sanktioniert keine dieser Bestimmungen die Beeinflussung des Willens durch reine Desinformation. Vorausgesetzt wird vielmehr, dass ein Desinformationsinhalt zugleich eine Androhung ernstlicher Nachteile darstellt, welche z.B. dazu geeignet ist, dass eine Versammlung, Wahl oder Abstimmung behindert oder gestört wird (Art. 279 StGB). Eine Androhung ernstlicher Nachteile liegt vor, wenn nach der Darstellung des Täters der Eintritt des Nachteils als von seinem Willen abhängig erscheint und wenn das Androhen geeignet ist, den Betroffenen in seiner Entscheidungsfreiheit einzuschränken (BGE 120 IV 19). Die wenigsten Desinformationsinhalte stellen eine solche Androhung ernstlicher Nachteile dar, weshalb die Bestimmungen zum Schutz des Volkswillens keinen Schutz gegen Desinformation vermitteln.

#### b) Selbstregulierung durch Lauterkeitskommission

Im Bereich des Wettbewerbsrechts besteht mit der schweizerischen Lauterkeitskommission (SLK) ein Organ der (branchenübergreifenden) Selbstregulierung. Die SLK ist das Selbstregu-

lierungsorgan der Werbewirtschaft. Sie wendet ihre eigenen Grundsätze und das UWG an. Ihr gehören alle bedeutenden Organisationen der Werbewirtschaft an. Sie beurteilt unlautere Sachverhalte im Bereich der kommerziellen Kommunikation; darunter kann auch Desinformation fallen (Über uns, abrufbar unter [www.faire-werbung.ch/de/ueberuns/](http://www.faire-werbung.ch/de/ueberuns/), zuletzt besucht am 29. Juni 2023).

Jede Privatperson kann vermeintlich unlautere kommerzielle Desinformation bei der SLK anzeigen (Art. 8 des Geschäftsreglements Lauterkeit in der kommerziellen Kommunikation, «Geschäftsreglement SLK»). Die Beschwerde kann sich auch gegen eine natürliche oder juristische Person richten, die nicht Mitglied der Lauterkeitskommission ist.

Unabhängig von einem Entscheid der SLK kann der gleiche Streitgegenstand oder die gleiche Tat vor einem ordentlichen Gericht zivil- oder strafrechtlich (parallel) geltend gemacht werden.

### 3. Materielles Recht: Besondere Bestimmungen für gewisse Adressaten

#### a) Journalistische Medien

##### i. *Medienrechtliche Vorgaben*

Das Bundesgesetz über Radio und Fernsehen («RTVG») statuiert in Art. 4 RTVG Mindestanforderungen an Programminhalte in Radio und Fernsehen. Insb. müssen redaktionelle Sendungen mit Informationsgehalt «Tatsachen und Ereignisse sachgerecht darstellen, sodass sich das Publikum eine eigene Meinung bilden kann» (Art. 4 Abs. 2 RTVG).

Redaktionelle Sendungen mit Informationsgehalt sind solche, die Bedeutung für die Meinungsbildung haben (LRTV-Masmejan, Art. 4 N 33; RTVG-Weber, Art. 4 N 12 ff.). Als nicht redaktionelle Inhalte gelten etwa Werbung oder blosser Unterhaltungssendungen (LRTV-Masmejan, Art. 4 N 33; RTVG-Weber, Art. 4 N 16). Das sog. Sachgerechtigkeitsgebot ist verletzt, wenn ein Beitrag in Missachtung der journalistischen Sorgfaltspflichten die Zuschauer in dem Sinn manipuliert, dass sie sich kein persönliches Bild mehr machen können (BGE 138 I 107, E. 2 m.w.H.; jüngst: BGer 2C\_483/2020 vom 28. Oktober 2020, E. 4.2). Alle Sender sind verpflichtet, das Sachgerechtigkeitsgebot einzuhalten, auch solche mit offenkundig politischer Ausrichtung (LRTV-Masmejan, Art. 4 N 32). Das Sachgerechtigkeitsgebot verlangt allerdings nicht, dass alle Standpunkte qualitativ und quantitativ genau gleichwertig dargestellt werden; entscheidend ist, dass die Zuschauer erkennen können, dass und inwiefern eine Aussage umstritten ist, und sie in ihrer Meinungsbildung nicht manipuliert werden (BGE 138 I 107, E. 2 m.w.H.). Vom Sachgerechtigkeitsgebot ist das Vielfaltsgebot zu unterscheiden. Dieses verpflichtet einzig die konzessionierten Programme, nicht die bloss meldepflichtigen Veranstalter (BGE 138 I 107, E. 2). Die konzessionierten Programme müssen nach dem Vielfaltsgebot in der Gesamtheit ihrer redaktionellen Sendungen die Vielfalt der Ereignisse und Ansichten programmatisch angemessen zum Ausdruck bringen (Art. 4 Abs. 4 RTVG; BGE 138 I 107, E. 2 m.w.H.). Desinformationsinhalte, die über Radio oder Fernsehen verbreitet werden, sind als Verstoß gegen das Sachgerechtigkeitsgebot zu qualifizieren.

Bei einem Verstoß gegen das Sachgerechtigkeitsgebot kann die Ombudsstelle angerufen werden. Danach kann Beschwerde an die unabhängige Beschwerdeinstanz für Radio und Fernsehen (UBI) erhoben werden.

## ii. *Presserat*

Der Presserat erlässt den sog. Journalistenkodex, bestehend aus der Erklärung der Pflichten und Rechte der Journalist:innen sowie den dazugehörigen Richtlinien. Mit Blick auf Desinformation ist insb. die «Richtlinie 1.1 Wahrheitssuche» relevant. Demnach stellt «die Wahrheitssuche (...) den Ausgangspunkt der Informationstätigkeit dar. Sie setzt die Beachtung verfügbarer und zugänglicher Daten, die Achtung der Integrität von Dokumenten (Text, Ton und Bild), die Überprüfung und die allfällige Berichtigung voraus» (Richtlinien 1.1).

Die Zuständigkeit des Presserats ist umfassend: Sie «erstreckt sich auf den redaktionellen Teil oder damit zusammenhängende berufsethische Fragen sämtlicher öffentlicher, periodischer und/oder auf die Aktualität bezogener Medien» (Art. 2 Geschäftsreglement des Presserates («Geschäftsreglement PR»). Social Media fallen nicht in die Zuständigkeit des Presserates, weil sie keine eigenen Inhalte erstellen, sondern als Kommunikationskanal zur Verbreitung von Inhalten Dritter dienen (vgl. dazu: Stellungnahme des Presserates vom 11.5.2=11, E. 15). Der Presserat spricht nur Empfehlungen aus – er hat keine Sanktionsmöglichkeiten gegenüber Journalist:innen (Art. 17 Geschäftsreglement PR). Nach der Präambel der Erklärung der Journalistinnen und Journalisten sind die Medien aber verpflichtet, eine sie betreffende Stellungnahme zu veröffentlichen. Social Media sind dem Presserat nur insoweit unterworfen, als der Presserat auch Empfehlungen zu Beiträgen von Journalist:innen in den sozialen Medien ausspricht (Entscheid Nr. 2/2019). Social-Media-Anbieter:innen sind selbst nicht Adressat der Empfehlungen, sondern nur die darin publizierenden Journalist:innen.

### b) Anbieter:innen von Fernmeldediensten: Messengerdienste?

Anbieter:innen von Fernmeldediensten sind verpflichtet, unlautere Massenwerbung per E-Mail («Spam») zu bekämpfen (Art. 45a FMG i.V.m. Art. 3 Abs. 1 lit. o UWG). Der Gesetzgeber sieht in Art. 83 der Verordnung über Fernmeldedienste (FDV) vor, dass Fernmeldediensteanbieter:innen eine Meldestelle für Spam betreiben und Massnahmen gegen Spam ergreifen müssen.

Die Meldestelle soll dazu dienen die Verbreiter:innen von Spam zu identifizieren. Wenn eine Fernmeldediensteanbieter:in Kenntnis vom Versand oder von der Weiterleitung unlauterer Werbung hat, muss sie den Versand dieser Nachrichten umgehend sperren und den Aufbau der entsprechenden Verbindungen verhindern (Art. 83 Abs. 4 FDV).

Messengerdienste sind als sog. Over-the-Top-(OTT)-Anbieter:innen grundsätzlich als Fernmeldediensteanbieter:innen zu qualifizieren und unterliegen den entsprechenden Regelungen.

Für die Verbreitung von Desinformation sind die Bestimmungen allerdings nur relevant, wenn die Desinformation als Spam zu qualifizieren ist.

#### 4. Einhaltung und Durchsetzung: Kanäle

##### a) Sperren von Nutzer:innen

Die Nutzungsbedingungen von Facebook und Instagram unterscheiden bezüglich der zu ergreifenden Massnahmen nicht nach rechtmässigen und unrechtmässigen Inhalten; ein Verstoss gegen geltendes Recht wird vielmehr einem Verstoss gegen die Nutzungsbedingungen bzw. Community Rules gleichgestellt. Wenn eine Person mehrfach und absichtlich gegen die Nutzungsbedingungen verstösst, kann sie nach entsprechender Warnung temporär oder dauerhaft gesperrt werden. Dieselben Regeln gelten bei Whatsapp und Threema (dazu oben IV.4.b.i.). Telegram sieht keine Sperrung vor.

##### b) Löschung und Sperrung von Inhalten

Facebook entfernt verschiedene Inhalte, die zwar nicht illegal sind, aber gegen die Community Guidelines verstossen, bspw. gefährliche Gesundheitsfehlinformationen oder Informationen, die auf die Beeinflussung von Wählern und/oder Volkszählungen ausgerichtet sind, wie Fehlinformationen zu Terminen, Orten, Zeiten und Methoden für die Stimmabgabe.

Messengerdienste verzichten mit Ausnahme von Telegram bei öffentlichen Kanälen auf die Löschung und Sperrung von illegalen Inhalten. Entsprechend sind auch keine Massnahmen für rechtmässige Inhalte vorgesehen.

##### c) Downranking

Facebook behält sich ausdrücklich vor, die Verbreitung problematischer Inhalte zu reduzieren. Auch Instagram zeigt gewisse Inhalte weiter unten im Feed und in den Stories an, wenn diese als problematisch qualifiziert werden.

Bei Messengerdiensten ist kein Downranking möglich.

##### d) Meldung/Flagging durch andere Nutzer:innen (User Flagging)

Bei den untersuchten Social-Media-Anbietern und bei den Messengerdiensten besteht (ausser bei Telegram) eine Möglichkeit zur Meldung nicht nur bei potenziell illegalen Inhalten, sondern auch bei einem möglichen Verstoss gegen die Nutzungsbedingungen. Damit wird zumindest bei den betrachteten Social-Media-Anbietern auch rechtmässige Desinformation erfasst.

##### e) Flagging durch besonders qualifizierte Personen (Trusted Flagger)

Trusted Flagger können neben unrechtmässige auch rechtmässige Desinformation flaggen.

##### f) Automatisierte Entfernung von Desinformation

In den Bestimmungen der von uns untersuchten Social-Media-Anbieter finden sich keine spezifischen Angaben über die automatisierte Entfernung von Inhalten. Es ist jedoch davon auszugehen, dass bei der automatisierten Entfernung von grundsätzlich rechtmässigen Inhalten mehr Zurückhaltung ausgeübt wird.

#### g) FactChecking

Ein besonderes Mittel gegen rechtmässige Desinformation ist das sog. Fact-Checking. Der Begriff beschreibt die systematische Prüfung von Inhalten auf ihren Wahrheitsgehalt (Dreyer et al., S. 55). Einige soziale Netzwerke haben ein internes Fact-Checking etabliert. Die meisten sozialen Netzwerke arbeiten aber mit externen Faktenprüfern zusammen. So geben Facebook und Instagram an, externes Fact-Checking zu nutzen, um die Richtigkeit der am häufigsten verbreiteten Inhalte zu überprüfen und zu bewerten. Dazu wurde ein eigenes Prüfprogramm etabliert.

Die analysierten Messengerdienste sehen kein internes Fact-Checking vor. Auf WhatsApp besteht die Möglichkeit, Desinformation direkt an Faktenprüfer zu melden, die auf WhatsApp vertreten sind.

#### h) Tagging

Auch Tagging (Kennzeichnen) ist im Gegensatz zum Löschen von Inhalten ein Mittel, das vor allem bei rechtmässiger Desinformation angewendet wird. Das Tagging von Desinformationen ist eine Massnahme, welche aufgrund einer Meldung (Flagging) von Nutzer:innen oder von Trusted-Flaggern erfolgen kann. Dadurch soll erreicht werden, dass sich die Nutzer:innen mit einem kritischeren Blick mit den Inhalten auseinandersetzen (Dreyer et al., S. 54). YouTube kennzeichnet, soweit ersichtlich, keine gemeldeten Videos. Anderes gilt für die sozialen Netzwerke von Meta. So werden auf Facebook und Instagram Beiträge, die von Faktenprüfungsorganisationen als gefälscht identifiziert wurden, mit einem Warnhinweis versehen werden, der sie als unglaubwürdig einstuft. Dieser enthält einen Link zum entsprechenden Artikel und eine Begründung der Entscheidung.

Die analysierten Messengerdienste sehen kein Tagging vor.

#### i) Debunking

Beim Debunking wird Desinformation durch korrigierende oder den Inhalt widerlegende Informationen ergänzt. Die zur Widerlegung notwendige Information wird in der Regel durch Fact-Checking-Organisationen zur Verfügung gestellt (Dreyer et al., 57). Dadurch sollen die Rezipient:innen auf den strittigen Inhalt einer Information aufmerksam gemacht und es soll eine kritischere Auseinandersetzung mit den Inhalten gefördert werden.

YouTube weist darauf hin, wenn ein «Publisher» einen «unabhängigen Faktencheck» durchgeführt hat (siehe dazu vorstehend, IV.4.b.). Facebook und Instagram führen bei Inhalten, die von Fact-Checking-Organisationen als unglaubwürdig eingestuft wurden, einen Link zum entsprechenden Artikel der Organisation sowie eine Begründung der Einstufung als unglaubwürdig auf.

Messengerdienste bieten keine Debunking Massnahmen im weiteren Sinne an.

#### j) Selbstbeschränkungen für Werbung

Plattformen werden oft mit dem Vorwurf konfrontiert, dass sie mit Desinformation Geld verdienen. Einerseits wird argumentiert, dass die hohen Werbeeinnahmen der Plattformen in den letzten Jahren unter anderem auch mit der Verbreitung von Desinformation erwirtschaftet wurden (Global Disinformation Index, 2019). Andererseits wird vor allem Google vorgeworfen, dass der Konzern Werbung auf Webseiten ausspielt, die Desinformation verbreiten (Campaign for Accountability, 2017).

Im Rahmen des Kodex gegen Desinformation von 2022 haben sich die teilnehmenden Plattformen verpflichtet, problematische Inhalte nicht mit Werbung zu bespielen, um die finanziellen Anreize für das Verbreiten von Desinformation zu verringern.

#### 5. Massnahmen von zivilgesellschaftlichen Organisationen und Individuen

Zivilgesellschaftliche und individuelle Initiativen umfassen Massnahmen, die von nicht-staatlichen und nicht-kommerziellen Individual- und Kollektivakteur:innen ergriffen werden. Dazu zählen bspw. stiftungsbasierte Fact-Checking-Redaktionen oder Evaluationshilfen für News, genauso wie das Melden von Desinformation durch die Nutzer:innen selbst. Mit Letzterem wird dem Umstand Rechnung getragen, dass auf digitalen Plattformen die Nutzer:innen selbst massgeblich zur Verbreitung von Desinformation beitragen (Helberger et al, 2018; Pedrazzi & Oehmer, 2020).

Im Vergleich zum europäischen Ausland gibt es in der Schweiz allerdings nur vereinzelte zivilgesellschaftliche und individuelle Initiativen, die von nicht-staatlichen und nicht-kommerziellen Individual- und Kollektivakteur:innen ergriffen werden (können), um Desinformation zu bekämpfen.

## **D. Empfehlungen**

### **I. Vorbemerkungen**

Auf Grundlage der vorliegenden Erkenntnisse kommen wir zum Schluss, dass Desinformation in der Schweiz bislang (noch) keine grösseren Probleme verursacht hat. Die jüngsten Entwicklungen zeigen aber, dass sich dies jederzeit ändern kann (siehe dazu auch vorn, B.I). Ein verstärktes Monitoring von Desinformation in der Schweiz und weitere empirische Forschung sind deshalb wichtig. Ein solches Monitoring könnte ein:e (staatliche:r) Desinformationsbeauftragte:r oder eine Meldestelle für Desinformation leisten (siehe dazu sogleich, D.X.5.i). Die Erkenntnisse könnten quasi als Seismograf für die Bedeutung von Desinformation in der Schweiz dienen und weitere Massnahmen anleiten, wenn sich ein Bedarf ergeben sollte.

Auch wenn kein akuter Handlungsbedarf bestehen mag, sollte die Schweiz mit dem Ergreifen geeigneter Massnahmen nicht zuwarten bis sich Desinformation in massgeblicher Weise auf die öffentliche Meinungsbildung auswirkt. In der Folge skizzieren wir deshalb eine Reihe von Massnahmen, mit denen die Verbreitung und die Auswirkungen von Desinformation in der Schweiz erfasst und verringert werden können. Wir unterscheiden dabei zwischen direkten und indirekten Massnahmen. Die direkten Massnahmen zielen darauf ab, das Problem der Desinformation unmittelbar und spezifisch zu erfassen, namentlich durch staatliche Regulierung, Ko-Regulierung, Selbstregulierung und Selbstorganisation sowie zivilrechtliches Engagement. Die indirekten Massnahmen sollen in erster Linie die Resilienz der Bevölkerung gegenüber Desinformation stärken, insb. durch die Erhöhung der Medien- und Digitalkompetenz sowie durch die Stärkung des Qualitätsjournalismus und einen verbesserten Zugang zu qualitativ hochwertigen Medieninhalten.

Die Empfehlungen basieren auf dem Grundsatz, dass sich der Staat nicht in die Inhalte der Kommunikation einmischen darf und sich neutral zu verhalten hat, soweit die Inhalte als solche nicht rechtswidrig sind. Massnahmen sind nur zu ergreifen, wenn sie erforderlich sind, um das Funktionieren der (öffentlichen) Kommunikationsprozesse zu gewährleisten. Ein Eingreifen des Rechts auf Ebene der Inhalte erachten wir grundsätzlich als problematisch und potenziell kontraproduktiv. Massnahmen auf dieser Ebene greifen vergleichsweise stark in Grundrechte ein, insb. ins Grundrecht auf freie Meinungsäusserung, eines der höchsten Güter in demokratischen Gesellschaften. Sobald sich eine Massnahme auf Inhalte bezieht, stellt sich zudem die Frage, wer über Richtigkeit bzw. Unrichtigkeit eines bestimmten Inhalts entscheiden kann und soll.

Desinformation tritt verstärkt innerhalb der digitalen öffentlichen Kommunikation auf, insb. über Social Media und Messengerdienste. Wir empfehlen deshalb, die Governance-Massnahmen auf diese Angebote auszurichten.

Zum Problem wird Desinformation erst durch ihre massenhafte Verbreitung. Das bloss Bestehen von falschen Informationen ist für die öffentliche Meinungsbildung nicht per se problematisch. Personen, die problematische Meinungen vertreten oder Desinformationen verfassen und verbreiten, wird es immer geben. Das muss eine demokratische Gesellschaft bis zu einem gewissen Grad aushalten können. Heikel wird es erst, wenn falsche Informationen ein grosses Publikum erreichen und schädliche Wirkungen entfalten, indem sie z.B. das Vertrauen in die demokratischen Institutionen untergraben. Governance-Massnahmen sollten deshalb in erster Linie bei der massenhaften Verbreitung von Desinformation und bei den Verbreiter:innen ansetzen. Dieser Ansatz greift zudem weniger tief in die Meinungsfreiheit ein, weil so sichergestellt wird, dass auch weiterhin (im Rahmen des geltenden Rechts; siehe dazu vorn, C.II.) beliebige Meinungen frei geäussert werden können und nur die (massenhafte) Verbreitung von falschen Informationen eingeschränkt wird.

Die Schweiz ist beim Ergreifen von Massnahmen gegen Desinformation nicht allein, ganz im Gegenteil. Wie vorstehend aufgezeigt (siehe dazu, C.IV.1.), hat die EU verschiedene Initiativen ergriffen und Vorgaben erlassen, um gegen Desinformation vorzugehen. Aus heutiger Sicht spricht viel für die Annahme, dass die meisten Anbieter:innen die Vorgaben des europäischen Rechts in gleicher oder ähnlicher Weise auch in der Schweiz umsetzen werden, weil es für sie wenig Sinn ergibt, für die Schweiz ein eigenes Regime zu entwickeln. Sicherheit besteht in diesem Punkt allerdings nicht. Vielmehr ist denkbar, dass die Anbieter:innen den Spielraum nutzen werden, der ihnen verbleibt, wenn die Schweiz keine eigenen Massnahmen ergreift.

Vor diesem Hintergrund stellt sich die grundsätzliche Frage, ob die Schweiz überhaupt eigene Massnahmen vorsehen soll oder ob man auf die Einhaltung der geltenden EU-Regeln auch in der Schweiz vertrauen kann. Grundsätzlich wäre auch eine inhaltliche Übernahme der EU-Regeln denkbar, um den Gleichlauf mit dem EU-Recht sicherzustellen. Obschon diese Optionen naheliegend sind, schlagen wir vor, weder auf eigene Massnahmen zu verzichten noch die geltenden EU-Regeln zu übernehmen, sondern eigene, auf die schweizerischen Verhältnisse zugeschnittene Massnahmen zu treffen. Für ein blosses Abwarten oder Übernehmen ist der Prozess der öffentlichen Meinungsbildung für eine direkte Demokratie schlicht zu wichtig. Zudem unterscheidet sich die öffentliche Meinungsbildung in der Schweiz recht stark von den entsprechenden Prozessen in den Mitgliedstaaten der EU. Ein spezifisch schweizerischer Ansatz erscheint deshalb angezeigt.

Die schweizerische Rechtsordnung enthält bereits viele Normen, die es erlauben, bestimmte, aus unterschiedlichen Gründen als problematisch qualifizierte Inhalte zu erfassen und deren Verbreitung zivil- und strafrechtlich zu sanktionieren (siehe dazu vorn, C.II.). Diese Normen beruhen jeweils auf einer Abwägung von Grundrechten. Verbote und Rechte der Betroffenen

bestehen nur dort, wo diese Abwägung dazu geführt hat, dass die Meinungsfreiheit aus überwiegenden Gründen eingeschränkt werden muss, bspw. zum Schutz der Persönlichkeit der betroffenen Personen. Unsere Analyse hat nicht nahegelegt, dass die bestehenden inhaltsbezogenen Normen um weitere Normen ergänzt werden sollten, um Desinformation wirksam erfassen und eindämmen zu können. Zur Unterbindung der Verbreitung von Desinformation sollten aber verschiedene direkte und indirekte Massnahmen ergriffen werden. In der Folge werden die Massnahmen, die wir zu Governance von Desinformation empfehlen, näher ausgeführt. Gleichzeitig diskutieren wir auch, ob sich die Übernahme einiger EU-Vorschriften in der Schweiz anbietet.

## **II. Direkte Massnahmen**

### **1. Staatliche Regulierung**

Die schweizerische Rechtsordnung enthält keine Normen, die spezifische Desinformation adressieren (siehe dazu vorn, C.II.). Sie kennt aber eine Vielzahl von Normen, die Sachverhalte erfassen, die unter den Begriff der Desinformation fallen können, wie z.B. der zivil- und strafrechtliche Schutz der Persönlichkeit. Auch auf Ebene des Systemschutzes gibt es verschiedene, namentlich strafrechtliche Vorschriften, die gesellschaftliche Rechtsgüter wie z.B. den öffentlichen Frieden schützen.

Denkbar wäre die Einführung eines gesetzlichen Verbots von Desinformation. Ein solches erscheint uns aber als zu weitgehend. Zum einen bestehen auf Ebene des Individualschutzes schon weitgehende Bestimmungen; zum andern scheint die öffentliche Meinungsbildung durch Desinformation nur im Ausnahmefall gefährdet, sodass weitere Bestimmungen auf Ebene des Systemschutzes in der Schweiz nicht erforderlich erscheinen.

#### **a) Schaffung eines:r Desinformations(schutz)beauftragten oder einer Desinformations-Monitoringstelle**

Eine mögliche Massnahme zur Bekämpfung von Desinformation ist die Schaffung eines:r «Desinformations(-schutz)beauftragten» («Beauftragte:r») nach dem Vorbild des Eidgenössischen Datenschutz- und Öffentlichkeitsbeauftragten (EDÖB). Die zentrale Aufgabe dieser Person würde darin bestehen, das Phänomen der Desinformation in der Schweiz (Umfang, Verfasser:innen, Verbreiter:innen, Quellen, Kanäle, Relevanz, Wirkung etc.) zu beobachten. Diese Tätigkeit sollte nicht auf das Verfassen und Verbreiten von Desinformation durch Private beschränkt sein, sondern auch die Informationstätigkeit der öffentlichen Organe mit umfassen, um (ungewollte) Desinformation durch Behörden zu erkennen (Bsp.: «Masken nützen nichts») und falsche Aussagen gegebenenfalls durch aktive Kommunikation richtiger Information zu berichtigen. Zudem müsste der:die Beauftragte auch allfällige Tätigkeiten des Geheimdienstes zur Abwehr von Desinformation durch ausländische Regierungen (Propaganda) beobachten und überwachen. Um diese Tätigkeit ausüben zu können, müsste der:die Beauftragte über ein

hohes Mass an Unabhängigkeit verfügen. Dies sollte, wie beim EDÖB, durch eine direkte Wahl durch das Parlament und weitere Bestimmungen zur Absicherung der Unabhängigkeit sichergestellt werden.

Der:die Beauftragte könnte verschiedene weitere Aufgaben übernehmen. Namentlich könnte er:sie: (i) die Einhaltung der Vorgaben der Ko- und Selbstregulierung bzw. Selbstorganisation durch die Anbieter:innen von Social Media und Messengerdiensten überwachen; (ii) als Beschwerdeinstanz eingesetzt werden, an die sich Nutzer:innen von (sehr grossen) Plattformen bei einem Ausschluss von einer Plattform oder beim Löschen oder Sperren (ev. auch beim Downgrading) von Inhalten wenden könnten; und (iii) als Anlaufstelle für Wissenschaftler:innen, investigative Journalist:innen und weiteren Interessierten wirken, die Zugang zu relevanten Daten der (sehr grossen) Plattformen verlangen. Diese Aufgaben könnte der:die Beauftragte als blasse:r Vermittler:in übernehmen, der:die nur Empfehlungen aussprechen und die Öffentlichkeit informieren kann. Denkbar wäre aber auch, dem:der Beauftragten die Kompetenz zum Erlass von Verfügungen einzuräumen.

Weiter könnte der:die Beauftragte (i) als Akkreditierungsstelle für Fact-Checking-Organisationen dienen, die gewisse Qualitätsstandards einhalten oder ein Label für akkreditierte Organisationen vergeben, das beim Flagging mit angezeigt werden könnte, damit Plattformen erkennen können, dass es sich um Trusted Flaggers handelt. Zudem könnte der:die Beauftragte (ii) als Meldestelle für Desinformation wirken, an die sich beliebige Personen wenden können, wenn sie problematische Inhalte entdecken; er:sie könnte diese Inhalte von Fact-Checking-Organisationen überprüfen lassen und/oder die betroffenen Individuen, die Plattformen oder den Presserat informieren.

Der:die Beauftragte könnte zudem lediglich als Beobachtungs-, Anlauf- und Koordinationsstelle wirken. An diese Stelle könnten sich Nutzer:innen, Wissenschaftler:innen, Journalist:innen und weitere interessierte Personen im Zusammenhang mit Desinformation wenden, um Auskunft über die Rechtslage und die Möglichkeiten für ein Vorgehen gegen Desinformation zu erhalten. Der:die Beauftragte könnte zudem den laufenden Austausch mit den Anbieter:innen sicherstellen und als Schnittstelle zu den Behörden wirken. Schon allein das Erfüllen dieser Aufgaben wäre wertvoll, weil sich so sicherstellen liesse, dass das Phänomen der Desinformation in der Schweiz laufend beobachtet und dokumentiert wird, sodass problematische Entwicklungen frühzeitig erkannt werden können. Sollte sich Bedarf nach weiteren (gesetzgeberischen) Massnahmen ergeben, liessen sich diese auf fundiertere Grundlagen stützen.

Alternativ zur Schaffung eines:einer unabhängigen «Desinformations(schutz)beauftragten» könnte auch lediglich eine Desinformations-Monitoringstelle innerhalb einer Bundesbehörde

geschaffen werden, naheliegenderweise beim Bakom. Diese Stelle könnte insb. als Beobachtungs-, Anlauf- und Koordinationsstelle tätig sein.

b) Gesetzliche Verankerung des Trennungsgebots bzw. Kennzeichnungspflicht für Werbung

In den traditionellen Medien gilt das Trennungsgebot, wonach Werbung und redaktioneller Text zu trennen sind. Verankert ist das Trennungsgebot in Art. 9 RTVG i.V.m. Art. 12 RTVV und im Presserecht (Punkt 10 der Erklärung der Rechte und Pflichten von Journalisten) sowie in den Grundsätzen der SLK (Grundsatz 3.12). Ein Verstoß gegen das Trennungsgebot wird nach herrschender Lehre auch von Art. 2 UWG erfasst (Volz, S. 67 f. m.w.H.). Bei der Anwendung des Trennungsgebots ist den jeweiligen Besonderheiten eines Inhaltes Rechnung zu tragen. Auf welche Kanäle und Inhalte das Trennungsgebot im online Bereich konkret Anwendung findet, ist bislang weitgehend ungeklärt. Eine mögliche gesetzgeberische Massnahme könnte deshalb darin bestehen, ein allgemeines Trennungsgebot bzw. eine Kennzeichnungspflicht für alle Kanäle zu verankern, wobei neben der kommerziellen insb. auch die politische «Werbung» erfasst werden sollte (dazu auch Besson & Boillet, S. 251).

Durch eine solche Massnahme könnte insb. im politischen Kontext sichergestellt werden, dass sich die Rezipient:innen bewusst sind, dass sie nicht mit objektiven Informationen konfrontiert werden. Dadurch liesse sich die Wirkung der Desinformation im Kontext der Werbung mindern.

c) Kennzeichnungspflicht / Identifikationspflicht für politische Werbung (Werbung von politischen Parteien)

In eine ähnliche Richtung geht die Einführung einer Kennzeichnungspflicht für politische Werbung, unter Umständen verbunden mit einer Identifikationspflicht, d.h. die Bezeichnung der verantwortlichen Person für politische Werbung. In gewissen Situationen kann die Gefahr der Beeinflussung der öffentlichen Meinungsäusserung besonders gross sein, bspw. weil einer Äusserung besonderes Vertrauen entgegengebracht wird. Dies kann bei Äusserungen von (bekannten) politischen Parteien und Akteur:innen der Fall sein. Viele Menschen messen politischen Äusserungen besonderes Gewicht zu. Desinformation zu politischen Fragen kann in der Schweiz besonders problematisch sein, weil hier politische Äusserungen oft direkt im Zusammenhang mit anstehenden Abstimmungen oder Wahlen erfolgen, wodurch Desinformation einen direkten Einfluss auf die politische Willensbildung haben kann.

Somit könnte ergänzend zu einer allgemeinen Kennzeichnungspflicht für Werbung eine Pflicht zur Kennzeichnung von Anzeigen als «politische Werbung» und zur Identifikation der hinter der Anzeige stehenden Person oder Organisation eingeführt werden. So können die Rezipient:innen politische Äusserungen besser einordnen.

d) Verpflichtung zur Einhaltung gewisser Regeln für Personen mit einer gewissen Reichweite

Die heutigen Gesetze wie der DSA knüpfen an der Reichweite von Plattformen an. Denkbar wäre es, anstatt der Plattformen direkt die Personen in die Pflicht zu nehmen, die Inhalte verbreiten, wenn diese eine gewisse Reichweite – d.h. Anzahl an Follower – haben. Inhaltlich wären Sorgfalts-, Klarheits- und Wahrheitspflichten in verschiedenen Abstufungen denkbar.

Eine solche Massnahme wäre allerdings ein grosser Eingriff in die Meinungsäusserungsfreiheit und möglicherweise in die Wirtschaftsfreiheit der betroffenen Personen und müsste damit den Voraussetzungen von Art. 36 BV genügen. Grössere Einschränkungen der Meinungsäusserungsfreiheit sind allerdings für gewisse Personen schon heute zulässig (dazu BSK BV-Hertig, Art. 16 N 42; Müller & Schefer, S. 403 ff.).

e) Werbeverbot im Kontext desinformativer Inhalte

Im aktuellen Kodex zur Bekämpfung von Desinformation verpflichten sich die teilnehmenden Plattformen, die finanziellen Anreize für die Verfasser:innen und Verbreiter:innen von Desinformation zu reduzieren, indem die Platzierung von Werbung neben Desinformationsinhalten vermieden wird. Denkbar wäre die Einführung eines entsprechenden gesetzlichen Verbots zur Verstärkung und Verstetigung der bis anhin auf Selbstregulierung basierenden Massnahme.

Unserer Ansicht nach genügen die Massnahmen auf Selbstregulierungsebene. Der neue Kodex wurde erst Mitte 2022 eingeführt; seine Wirksamkeit lässt sich heute noch nicht beurteilen. Die Einführung eines gesetzlichen Werbeverbots würde im heutigen Zeitpunkt wohl zu stark in die Wirtschaftsfreiheit der Plattformen eingreifen, ohne dass im Vergleich zur Selbstregulierung eine erhöhte Wirksamkeit erreicht werden könnte.

f) Schaffung eines (staatlichen) Streitschlichtungsverfahrens

Die heutige Ko-Regulierung und die Selbstorganisation führen zu einer Verschiebung von der staatlichen zu einer privaten Rechtsdurchsetzung. Damit ist ein (unfreiwilliger) Machtgewinn der Plattformen verbunden (dazu oben C.1.2.).

Die Entscheidung über Recht und Unrecht sollte aus rechtsstaatlicher Sicht bei den Gerichten liegen. Aufgrund der immensen Anzahl potenzieller Fälle sind Entscheidungen von staatlichen Gerichten allerdings kaum realistisch. Denkbar wäre aber, dass der Staat ein unabhängiges und schnelles – auf erster Instanz auch automatisiertes – Verfahren anbieten würde, indem streitige Fälle entschieden werden könnten. Ein solche Verfahren hätte den Vorteil, dass die heutige Personalunion zwischen Regelsetzer:innen und Regelanwender:innen aufgelöst werden und die Vorgaben des Schweizer Rechts im Rahmen eines solchen Verfahrens plattformübergreifend einheitlich angewendet werden könnte. Zudem würde die Möglichkeit geschaffen, wie in Gerichtsverfahren üblich, der unterliegenden Partei einen Teil der Verfahrens-

kosten aufzuerlegen, was die Hemmschwelle für die absichtliche Verbreitung von rechtswidrigen Inhalten erhöhen könnte.

#### g) Massnahmen hinsichtlich automatisierter Accounts

Das Verbot von automatisierten Accounts ist das effektivste Mittel, um die Verbreitung von Desinformation durch Bots zu verhindern (Löber & Rossnagel, S. 186). Ein staatliches Verbot würde allerdings in die Meinungsfreiheit der hinter den Accounts stehenden Personen eingreifen. Social Bots sind zudem in der Schweiz bisher kein relevantes Problem (siehe dazu vorn, B.I.3). Ein Verbot wäre daher wenig zielführend, schwierig umzusetzen und nicht ausreichend legitimiert. Anstelle eines Verbots von automatisierten Accounts könnte die Anzahl der Aktivitäten eines Accounts begrenzt werden, insb. die Anzahl der verfassten oder geteilten Inhalte innerhalb eines bestimmten Zeitraums. Damit liesse sich der Einfluss von Bots auf einer quantitativen Ebene begrenzen.

Eine mildere Massnahme wäre die Einführung einer Kennzeichnungspflicht für automatisierte Accounts. Die Kennzeichnung würde die Nutzer:innen darüber informieren, dass hinter einem Account ein Computerprogramm steht. Die Kennzeichnung von automatisierten Accounts und anderen algorithmischen Systemen wird auch in anderem Zusammenhang ganz allgemein gefordert (Thouvenin et al., S. 3; Schefer & Cueni, S. 71). Denn unabhängig von der Verbreitung von Desinformation erscheint es wichtig, dass betroffene Personen erkennen können, ob sie mit einem Menschen oder einer Maschine interagieren. Der Einführung einer solchen Kennzeichnungspflicht stehen keine relevanten grundrechtlichen Bedenken entgegen. Sie wäre einem Verbot als mildere Massnahme vorzuziehen. Eine Kennzeichnungspflicht könnte auch von den Anbietern selbst eingeführt werden, etwa im Rahmen von Ko- und Selbstregulierung oder im Rahmen der Selbstorganisation.

## 2. Ko-Regulierung

Ko-Regulierung gibt es in der Schweiz im Bereich der Desinformation bisher nicht. Denkbar wären die folgenden Massnahmen:

#### a) Prozedurale Vorgaben für Plattformen

Der DSA verpflichtet Online-Plattformen zur Einhaltung gewisser prozeduraler Vorgaben, etwa zur Einrichtung eines internen Beschwerdemanagements; zudem bestehen gewisse Begründungspflichten (dazu oben C.II.2). Die Plattformen sind zudem verpflichtet, sich bei Streitfällen einer aussergerichtlichen Streitbeilegungsstelle zu unterwerfen, welche den internen Entscheid überprüfen kann (Art. 21 DSA). Online-Plattformen müssen zudem gewisse Angaben zu den entschiedenen Fällen veröffentlichen (Art. 24 DSA).

Auch wenn prozedurale Vorgaben an sich sinnvoll erscheinen, bestehen gewisse Bedenken. Ein Grund für die Auferlegung der prozeduralen Vorgaben ist die Verlagerung der staatlichen auf die private Rechtsdurchsetzung. Diese Verschiebung ist nicht unproblematisch (siehe dazu

vorn II.C.1.). Deshalb sollen Private, wenn sie quasi staatliche Aufgaben wahrnehmen, auch dieselben prozessrechtlichen Vorgaben beachten müssen wie der Staat. Die Verpflichtung zur Gewährleistung eines Beschwerdeverfahrens führt allerdings zu hohen Kosten und stellt damit einen Eingriff in die Wirtschaftsfreiheit der Plattformen dar. Die Massnahme muss deshalb den Anforderungen von Art. 36 BV genügen, wobei unter anderem die Eignung zu berücksichtigen ist sowie Eingriffszweck und Eingriffswirkung in Relation zu setzen sind. Der Zweck der prozeduralen Vorgaben besteht nicht in der Bekämpfung von Desinformation, sondern in der Gewährleistung der Meinungsfreiheit der Nutzer:innen. Prozedurale Vorgaben können wohl dazu beitragen, dass Nutzer:innen oder Inhalte nicht willkürlich von einer Plattform entfernt werden. Da die Plattformen jedoch meist rein interne Beschwerdeverfahren vorsehen, ist die Unparteilichkeit der Beschwerdeinstanz nicht gewährleistet. Es wäre deshalb sinnvoll, die Plattformen einer externen Streitbeilegungsinstanz zu unterwerfen; dabei könnte es sich um eine (bestehende) gerichtliche oder um eine aussergerichtliche Instanz handeln. Je nach Umfang der auferlegten prozeduralen Vorgaben und dem damit verbundenen Aufwand stellt sich zudem die Frage, ob solche Vorgaben dem Erfordernis der Verhältnismässigkeit im engeren Sinn genügen würden. Diesem Aspekt könnte Rechnung getragen werden, indem die prozeduralen Vorgaben nur für Plattformen gelten, die eine bestimmte Anzahl von Nutzer:innen haben oder sonst über eine grössere Reichweite verfügen.

Die Veröffentlichung von Angaben zu entschiedenen Fällen kann zusätzliche Transparenz schaffen und kann der Überprüfung dienen, wie die verschiedenen Plattformen ihren Pflichten nachkommen bzw. ihre internen Regeln umsetzen.

#### b) Transparenzvorgaben für Plattformen

Das Ranking von Inhalten entscheidet über die Sichtbarkeit und damit über die Wirkung der Meinungsäusserung von Nutzer:innen. Anbieter:innen von Online-Plattformen, die dem DSA unterstehen, müssen deshalb bei der Verwendung von Empfehlungssystemen die wichtigsten Parameter bekannt geben, welche das Ranking des Inhalts bestimmen.

Denkbar wäre eine solche Transparenzpflicht auf sämtliche Rankings von Inhalten auszuweiten. Dabei kann es allerdings nicht um die Offenlegung der Algorithmen gehen, die regelmässig als Geschäftsgeheimnisse zu qualifizieren sind und von den (allermeisten) Nutzer:innen ohnehin nicht verstanden werden. Erforderlich ist nur, aber immerhin, dass für die Nutzer:innen Transparenz bezüglich der «logic involved» besteht, in dem die für das Ranking relevanten Parameter offengelegt werden. Diese Transparenz sollte auch die Kriterien umfassen, die für das Schalten von Werbung relevant sind. Solche Transparenzpflichten könnten sich grundsätzlich an den datenschutzrechtlichen Vorgaben zu den automatisierten Einzelfallentscheiden und den damit verbundenen Auskunftspflichten orientieren (siehe dazu oben D.II.2.i.).

Bei den Transparenzvorgaben stellt sich das Problem der Abwägung zwischen dem Schutz der Geschäftsgeheimnisse der Plattformen auf der einen und den Interessen der Nutzer:innen und der Gesellschaft an Transparenz auf der anderen Seite. Eine gesetzliche Verpflichtung müsste deshalb mit unbestimmten Rechtsbegriffen operieren und bspw. eine «angemessene» und «adressatengerechte» Information verlangen. Auch hier muss zudem der Zweck des Eingriffs – der Schutz der Betroffenen vor willkürlicher Behandlung durch die Plattformen – in einem vernünftigen Verhältnis zum Eingriff in die Wirtschaftsfreiheit der Plattformen stehen. Mit Blick auf die Verhältnismässigkeit dürfte auch hier eine Begrenzung der Verpflichtung auf Plattformen mit einer bestimmten Anzahl von Nutzer:innen oder einer gewissen Reichweite sinnvoll sein.

c) Recht auf Zugang für Nutzer:innen

Am Prozess der öffentlichen Meinungsbildung sollten möglichst viele Personen beteiligt sein. Zentral ist namentlich, dass sich alle Menschen und Organisationen, die sich an einem Diskurs beteiligen möchten, auch tatsächlich an diesem beteiligen können. Dadurch soll sichergestellt werden, dass gegenüber Desinformation korrektive Informationen in die öffentliche Meinungsbildung einfließen. Zu diesem Zweck könnte ein «Recht auf Zugang» von Nutzer:innen zu (sehr grossen) Plattformen oder zumindest eine Begrenzung der Möglichkeiten des Ausschlusses von Nutzer:innen von diesen Plattformen eingeführt werden. Einen solchen faktischen Kontrahierungszwang gibt es bereits im Bereich der staatlichen Leistungsverwaltung (z.B. Art. 12 Personenbeförderungsgesetz, PBG – Pflicht der SBB zum Abschluss eines Beförderungsvertrages mit den Passagieren oder Art. 11 FMG, Pflicht des marktbeherrschenden Anbieters zur Interkonnektion). Weil die staatliche Tätigkeit nicht unter dem Schutz der Wirtschaftsfreiheit steht, sind solche gesetzlichen Pflichten zum Vertragsabschluss weitgehend unproblematisch.

Anders sieht es bei einem Kontrahierungszwang im Privatrecht aus. Das Kartellrecht erklärt in Art. 7 Abs. 2 lit. a KG die Verweigerung von Geschäftsbeziehungen durch ein marktbeherrschendes Unternehmen als unzulässig, soweit dafür keine sachlichen Gründe («legitimate business reasons») vorliegen. Eine Zugangspflicht für Private lässt sich vorliegend aber nicht auf Kartellrecht abstützen, weil diese Normen dem Schutz von Wettbewerb dienen.

Der Ursprung des kartellrechtlichen Kontrahierungszwangs ist im Persönlichkeitsrecht (Art. 28 ZGB) zu finden (Klett, S. 161 ff.). Damit liesse sich das «Recht auf Zugang» möglicherweise auf den privatrechtlichen Persönlichkeitsschutz abstützen; eine formelle Verankerung des Zugangsrechts könnte sich dann allenfalls erübrigen. Diese Lösung scheint allerdings wenig praktikabel, weshalb eine ausdrückliche gesetzliche Regelung zu begrüssen wäre.

#### d) Anwendung des Sachgerechtigkeitsgebots

Mit Blick auf die Bedeutung der (grossen) Plattformen für die öffentliche Meinungsbildung wäre es denkbar, diese zur inhaltlichen Neutralität oder zur Sachgerechtigkeit zu verpflichten. Als mildere Massnahme könnten die (grossen) Plattformen zumindest dazu verpflichtet werden, Diskurse nicht durch up- und downranking zu verzerren. Alle diese Massnahmen sind allerdings grundsätzlich problematisch, weil sie in die Wirtschaftsfreiheit der Plattformen eingreifen.

Für Radio und Fernsehen ist in Art. 92 Abs. 2 BV und in Art. 4 Abs. 2 RTVG das Sachgerechtigkeitsgebot verankert (siehe dazu vorne C.II.7.). Nach Art. 4 Abs. 2 RTVG müssen redaktionelle Sendungen mit Informationsgehalt Tatsachen und Ereignisse sachgerecht darstellen, sodass sich das Publikum eine eigene Meinung bilden kann. Ansichten und Kommentare sind als solche kenntlich zu machen. Eng damit verbunden ist das Vielfaltsgebot, wonach die Vielfalt der Ansichten angemessen zum Ausdruck zu bringen ist (Weber & Nobel, S. 229). Ein Grund für die inhaltlichen Vorgaben des Gesetzgebers ist die Relevanz der Leitmedien für die Meinungsbildung (Botschaft «Radio- und Fernsehen», S. 1598.)

Vor diesem Hintergrund kann man sich fragen, ob nicht auch Social-Media-Plattformen einem Sachgerechtigkeitsgebot unterworfen werden sollten, weil sie in gleicher Weise wie Radio und Fernsehen die Weiterverbreitung von audiovisuellen Inhalten zulassen, denen ein besonderes Potenzial für Meinungsmanipulation beigemessen wird (Müller & Schefer, S. 485; Gujic, S. 189). Denkbar wäre eine Pflicht zur inhaltlichen Neutralität oder Sachgerechtigkeit für (sehr grosse) Plattformen, um einen breiten Diskurs zu gewährleisten. Eine solche «Neutralitätspflicht» könnte z.B. dadurch umgesetzt werden, dass Plattformen verboten wird, bestimmte Themen von der Diskussion auf der Plattform auszuschliessen (z.B. LGBTQ) oder Diskurse durch systematisches Up- oder Downgrading von bestimmten Themen oder Inhalten zu beeinflussen.

Im Ergebnis würde das Sachgerechtigkeitsgebot die Umsetzung von mehreren, bereits vorstehend analysierten Massnahmen erfordern: Das Löschen von Inhalten dürfte nur für (klar) widerrechtliche Inhalte erlaubt sein, es müsste zumindest minimal begründet werden und dürfte nur nach vorheriger Information der Betroffenen erfolgen; Nutzer:innen dürfen zudem nur bei wiederholten schweren Verstössen ausgeschlossen werden und sie müssten vorher über einen drohenden Ausschluss informiert werden und die Möglichkeit haben, sich dazu zu äussern. Bei nicht widerrechtlichen Inhalten käme sodann nur ein Downgrading infrage.

Eine Ausweitung des Sachgerechtigkeitsgebots auf Social-Media-Plattformen wird in der Lehre meist abgelehnt (Masmajan, S. 95 und S. 96; Schefer & Cueni, S. 49) nicht zuletzt mit der Begründung, dass dieses Gebot insb. auf die im Rundfunk bestehende Frequenzknappheit zurückzuführen ist. Dies ist bei der längst etablierten Verbreitung der Inhalte der klassischen Medien über die Infrastruktur des Internets allerdings kein Problem mehr. Aus heutiger Sicht

kann das Sachgerechtigkeitsgebot deshalb nur noch mit der Meinungsbildungsrelevanz der klassischen Medien begründet werden; eine solche Meinungsbildungsrelevanz können auch gewissen Plattformen haben. Der Auferlegung eines Sachgerechtigkeitsgebots steht allerdings entgegen, dass Plattformen keine eigenen Inhalte verbreiten und die auf der Plattform verbreiteten Inhalte entweder nicht kontrollieren können (Messenger) oder nach geltendem Recht in der EU und nach h.L. in der Schweiz nicht kontrollieren müssen (Social Media). Die Einführung eines Sachgerechtigkeitsgebots macht deshalb im Ergebnis keinen Sinn.

#### e) Werbeverbot für problematische Inhalte

Verschiedentlich wird diskutiert, dass die Werbebeschränkungen, die sich gewisse Plattformen freiwillig auferlegt haben, auf Stufe der Ko-Regulierung verankert werden sollten.

In der Umsetzung stellt sich eine Einschränkung der Werbung aber als schwierig dar. Erstens bedingt dies eine Instanz (z.B. Trusted Flagger), die Desinformation identifiziert und den Wahrheitsgehalt von Nachrichten bewertet. Zweitens greifen solche Vorgaben in relevantem Umfang in die Wirtschaftsfreiheit der Plattformen ein – und dies im Kern ihres Geschäftsmodells. Und drittens sind die wirtschaftlichen Interessen der Plattformen direkt betroffen, was die Anreize mindert, in dieser Frage aktiv zu werden.

Wir halten deshalb die Selbstorganisation der Plattformen im Bereich der Werbung auch für die Schweiz für den sinnvolleren Weg als eine staatliche Regulierung.

#### f) Zugangsrecht zu Daten

Um die empirischen Grundlagen zur Erfassung des Phänomens der Desinformation zu stärken, wäre ein verbesserter Zugang zu (anonymisierten) Daten von Social-Media-Plattformen und Messengerdiensten zu Forschungszwecken äusserst wichtig. Initiativen der Selbstregulierung waren bislang wenig erfolgreich (u.a. Social Science One von Meta/Facebook). Es ist denn auch kaum davon auszugehen, dass Plattformen künftig im erforderlichen Umfang freiwillig Zugang zu den relevanten Daten gewähren werden, zumal die Herausgabe von Daten mit Aufwand verbunden ist und Dritten einen Einblick in (allenfalls problematische) Praktiken der Plattformen vermittelt. Ohne konkrete Vorgaben des Gesetzgebers werden Plattformen kaum bereit sein, Forschenden im erforderlichen Umfang Zugang zu ihren Daten zu gewähren. Es sollte deshalb geprüft werden, wie ein Recht auf Zugang von Forschenden im Rahmen einer Ko-Regulierung umgesetzt werden könnte und ob ein solches Recht auch anderen Interessierten gewährt werden könnte, bspw. investigativen Journalist:innen und/oder Nicht-Regierungsorganisationen (NGO), deren Zweck in der Bekämpfung von Desinformation besteht.

### 3. Selbstregulierung (insb. Branchenregulierung)

Die Selbstregulierung der Plattformen in Zusammenarbeit mit der EU (insb. Verhaltenskodex für Desinformation 2022) erscheint sinnvoll. Soweit ersichtlich, werden diese Vorgaben derzeit auch für die Schweiz umgesetzt. Die freiwillige Umsetzung der europäischen Vorgaben in der

Schweiz steht allerdings allein im Ermessen der Plattformen und erfolgt (wohl) primär aus Gründen der Effizienz. Es erscheint deshalb sinnvoll, die Umsetzung der wichtigsten Vorgaben auch in der Schweiz durch geeignete Massnahmen sicherzustellen.

Denkbar wären die folgenden Massnahmen:

a) Verstärktes Aktivwerden des Presserates

Anders als die schweizerische Lauterkeitskommission (SLK), deren Zuständigkeit vom Inhalt («kommerzielle Kommunikation») abhängt, bezieht sich die Zuständigkeit des Presserates auf journalistisch tätige Personen. In diesem Sinne ist der Presserat auch für Inhalte auf Social Media und für öffentlich zugängliche Inhalte auf Messengerdiensten zuständig, wenn diese von Journalist:innen erstellt wurden.

Ein verstärktes Aktivwerden des Presserates bzw. eine Ausdehnung seiner Zuständigkeit auf weitere Personen erscheint wenig sinnvoll, weil Desinformation kaum von Journalist:innen verbreitet wird. Denkbar wäre zwar, dass der Presserat seine Zuständigkeit weniger von der Person und der Aufmachung von Inhalten (bspw. redaktionelle Gestaltung) und mehr von der Reichweite einer Person abhängig macht. Dazu wäre aber erforderlich, dass die betroffenen Personen sich freiwillig zur Einhaltung von journalistischen Regeln verpflichten, was wenig wahrscheinlich ist.

b) Verstärktes Aktivwerden der Lauterkeitskommission

Die schweizerische Lauterkeitskommission (SLK) ist nur für die kommerzielle Kommunikation zuständig. Sie kann deshalb nur aktiv werden, wenn Desinformation einen kommerziellen Bezug hat, bspw. bei einer Missachtung des Trennungsgebots in sozialen Medien, in Messengerdiensten und in alternativen Medien oder auch bei irreführender oder unwahrer Werbung.

c) Schaffung eines Digitalrates

Die Plattformen könnten zusammen mit weiteren Unternehmen ein Selbstregulierungsorgan schaffen, das analog zum Presserat gewisse gemeinsame Richtlinien und Standards definiert und unter Umständen auch durchsetzt.

Da Desinformation oft über verschiedene Kanäle verbreitet wird, könnte ein Digitalrat der gemeinsamen und kanalübergreifenden Bekämpfung von Desinformation dienen. Hierzu könnten sich die Beteiligten auf gemeinsame Grundsätze einigen, sinnvollerweise unter Einbezug der Nutzer:innen. Dies hätte den Vorteil, dass problematische Inhalte auf allen Plattformen gleichbehandelt und gelöscht oder wegen Downranking kaum noch zur Kenntnis genommen und nicht mehr in relevantem Mass weiterverbreitet würden. Ein Digitalrat könnte auch als Beschwerdeinstanz dienen, die Beschwerden gegen das Sperren oder Löschen von Inhalten oder Nutzer:innen behandelt (siehe dazu vorn, C.IV.2.). Im Gegensatz zu den internen

Beschwerdeinstanzen der Plattformen könnte man dem Digitalrat je nach konkreter Ausgestaltung zumindest eine gewisse Unabhängigkeit attestieren. Umgekehrt hätte ein plattformübergreifendes Sperren oder Löschen von Nutzer:innen oder Inhalten auch viel einschneidendere Wirkungen für die Verfasser:innen und Verbreiter:innen.

Wichtig wäre vor diesem Hintergrund die Beachtung der kartellrechtlichen Schranken, insb. besteht bei einer brancheninternen Zusammenarbeit das Problem des wettbewerbswidrigen Informationsaustausches. In welcher Form und mit welchen Schranken die Schaffung eines Digitalrates im Rahmen des geltenden Kartellgesetzes möglich ist, wäre noch vertieft zu prüfen. Jedoch bestünden sowohl im europäischen mit der Freistellung wie auch im schweizerischen Recht mit der ausnahmsweisen Zulassung nach Art. 8 KG Möglichkeiten, um kartellrechtlich heikle Kooperationen im Einzelfall als zulässig erklären zu lassen.

#### 4. Selbstorganisation

Social-Media- und Messengerdienste-Anbieter:innen sind bei der Gestaltung ihrer internen Regelwerke, insb. der sog. «Community Standards» grundsätzlich frei. Namentlich sind diese Anbieter:innen nicht an die Grundrechte gebunden und haben deshalb einen recht grossen Gestaltungsspielraum. Zahlreiche Community Standards sehen zur Bekämpfung von Desinformation bereits Massnahmen vor, etwa die Möglichkeit, problematische Inhalte zu melden und löschen zu lassen oder Accounts (temporär) zu sperren. Forschende haben auch die Akzeptanz dieser Massnahmen geprüft und gezeigt, dass amerikanische Teilnehmende bspw. eher die Löschung von Inhalten als das Sperren von Accounts bevorzugen (Kozyreva et al., 2023).

Wie bei der Selbstregulierung können auch bei der Selbstorganisation nur Vorschläge gemacht werden, von denen angenommen werden darf, dass sie von den Plattformen freiwillig umgesetzt werden. Kann dies nicht erwartet werden, kommen nur eine staatliche Regulierung oder eine Ko-Regulierung infrage. Dass die Plattformen selbst entscheiden, wo und wie sie ansetzen, ist ein grundlegendes Problem der Selbstorganisation. Die Macht der Plattformen und ihr Einfluss auf den öffentlichen Diskurs wird nochmals gestärkt, wenn man weite Felder der Selbstregulierung und Selbstorganisation überlässt. Aus einer normativen Perspektive ist es zudem problematisch, wenn private Unternehmen darüber entscheiden (müssen), welche Inhalte und welche Akteur:innen im öffentlichen Diskurs zugelassen werden. Hinzu kommt, dass die grossen US-Plattformen die Normen in ihrem Herkunftsland (auch aus rechtlichen Überlegungen) stärker in die Ausgestaltung ihrer Selbstorganisation einfliessen lassen. Facebook bzw. Meta löscht bspw. Darstellungen von Nacktheit äusserst rigide, während mit Verweis auf die Meinungsfreiheit auch teilweise radikale politische Meinungen zugelassen werden (Ibrahim, 2017). Trotz problematischer Aspekte sind Massnahmen der Selbstorganisation wichtige Elemente, um die Verbreitung von Desinformation plattformspezifisch zu ver-

hindern bzw. zu reduzieren. Neben den bereits bestehenden sind vor allem die folgenden Massnahmen denkbar oder, soweit bestehend, auszubauen:

a) Veränderung der plattformeigenen Infrastruktur

Forschungsergebnisse zeigen, dass die Infrastruktur der sozialen Medien zur Verbreitung von Desinformation beitragen kann. Aufgrund des sozialen Kontexts werden vermehrt soziale Bedürfnisse (bspw. nach Zugehörigkeit oder Anerkennung) bei den Nutzer:innen aktiviert, die dazu führen können, dass weniger auf die Richtigkeit von Informationen als auf die soziale Bewertung geachtet wird (vgl. Rathje et al., 2023, Epstein et al., 2023). Eine Veränderung der Interaktionsmöglichkeiten durch andere Schaltflächen («Like» vs. «Trust», vgl. Globig et al., 2023), regelmässige Betonungen der Wichtigkeit der Richtigkeit von Information sowie ein Unterbrechen des Teilens durch kurze Erinnerungen daran, den Wahrheitsgehalt zu prüfen (Pennycook & Rand, 2022) könnten helfen, die Verbreitung von Desinformation zu reduzieren.

b) Sperren von Nutzer:innen in sozialen Medien (Deplatforming)

Das Sperren von Nutzer:innen bzw. ihrer Accounts auf Social-Media-Plattformen und Messengerdiensten ist eine weitreichende Massnahme, welche die Meinungsfreiheit der Verfasser:innen und Verbreiter:innen beschränkt. Weil die gesperrten Nutzer:innen dadurch vom Kommunikationsprozess auf der jeweiligen Plattform ausgeschlossen werden, sollte sie nur restriktiv und nach mehrmaligem oder schwerem Verstoss gegen die Community Standards gesperrt werden. Die Wirkung der Massnahme ist zudem fraglich. Denn Nutzer:innen können in der Regel anonym und schnell neue Accounts erstellen. Eine Ausnahme bilden Accounts mit hoher Reichweite (z.B. von sog. Influencer:innen, prominenten Personen oder professionellen Organisationen), die sich nicht ohne weiteres unmittelbar ersetzen lassen. Prominentes Beispiel ist die Sperrung des Twitter-Accounts des ehemaligen US-Präsident Donald Trump. Dieses sog. Deplatforming hat aber oftmals den Effekt, dass sich die Problematik auf andere, weniger öffentliche Plattformen (u.a. Messengerdienste) verlagert. Bereits heute gibt es bei fast allen Plattformen gewisse minimale Verfahrensgarantien, die vor dem Ausschluss von Nutzer:innen eingehalten werden. So können sich die Nutzer:innen bspw. in der Regel vor dem Sperren dazu äussern. Die Gewährung solcher Rechte ist zu begrüessen, auch wenn sie nur auf Ebene der Selbstorganisation bestehen. Als mildere Massnahme wäre denkbar, nur die Kommentar-, Like- und Post-Möglichkeiten der fehlbaren Nutzer:innen zu sperren, diese aber weiterhin als passive Nutzer:innen auf der Plattform zu belassen, sodass sie weiterhin auf Informationen zugreifen können.

c) Löschung und Sperrung von Inhalten

Weniger weit als das Sperren von Nutzer:innen geht das Löschen und Sperren von Inhalten. Diese Massnahmen werden schon länger mit Bezug auf strafrechtlich relevante Inhalte eingesetzt. Weil sie die betroffenen Inhalte faktisch unsichtbar machen, sind diese Massnahmen nur im Einzelfall ein zulässiges Mittel, insb. bei rechtswidrigen Inhalten. Heikel ist das Löschen

und Sperren von Inhalten, die nur (aber immerhin) manipuliert sind und/oder offensichtlich nicht der Wahrheit entsprechen. Das gilt auch, wenn das Verbreiten von Inhalten eine Gefahr für die Rechtsgüter Dritter schafft. Liegt kein Verstoss gegen geltendes Recht vor, muss das Löschen und Sperren von Inhalten auf eine hinreichend klare Grundlage in den Community Standards gestützt werden können. In diesem Bereich kommt den Plattformen grosse Macht und entsprechend viel Verantwortung zu, zumal sie über «Wahrheit» und «Unwahrheit» entscheiden und mit ihren Entscheidungen die öffentliche Meinungsbildung beeinflussen. Neben dem Löschen und Sperren von problematischen Inhalten könnten auch ein Blockieren oder Verwischen von mit diesen Nachrichten assoziierten Bildern vorgenommen werden, die neben diesen Inhalten angezeigt werden, um diese Inhalte weniger glaubwürdig zu machen (vgl. Forschungsergebnisse in Bezug auf den Truthiness-Effekt, Newman & Zhang, 2021). Gerade bei Situationen, in denen nach einer ersten noch eine intensivere Prüfung der Inhalte vorgenommen wird, könnte eine derartige (vorübergehende) Massnahme den weiteren Einfluss von Inhalten reduzieren.

#### d) Downranking von Inhalten

Das Downranking hat eine ähnliche Wirkung wie das Sperren oder Löschen von Inhalten. Es ist zwar weniger invasiv, weil es nur dazu führt, dass die als problematisch beurteilten Inhalte weniger gut gefunden und damit weniger stark verbreitet werden, kann aber im Ergebnis einen ähnlichen Einfluss auf die öffentliche Meinungsbildung haben. Anders als beim Löschen oder Sperren lassen sich beim Downranking verschiedene Eingriffstiefen verwirklichen, die von weniger gut auffindbar, nur auffindbar mit konkreter Suche bis zu fast nicht mehr auffindbar reichen.

Im Gegensatz zum Sperren und Löschen können die Nutzer:innen ein Downranking oft nicht erkennen, kaum nachweisen und damit auch nicht beanstanden. Meist spielen mehrere Faktoren eine Rolle für das Ranking; die Unglaubwürdigkeit eines Inhaltes ist dabei nur ein mögliches Kriterium. Dass die fehlende Glaubwürdigkeit zum Downranking eines Inhalts führt, ist kaum zu beanstanden. Wichtig ist allerdings, dass das Vorhandensein des Kriteriums für die Nutzer:innen transparent ist, die Glaubwürdigkeit einer Aussage durch hinreichend zuverlässige Verfahren beurteilt und die Unglaubwürdigkeit nicht leichthin angenommen wird.

#### e) Flagging durch andere Nutzer:innen (User Flagging) // Meldefunktion

Die Möglichkeit zur Prüfung und Meldung problematischer Inhalte durch Nutzer:innen ist die erste Stufe von Massnahmen, die von Social-Media-Plattformen ergriffen wurden. Die Gewährung dieser Möglichkeit ist für gewisse Online-Plattformen nach dem DSA nun Pflicht. Die Meldung ist oft auch Voraussetzung für das Löschen bzw. Sperren von Nutzer:innen und Inhalten und das Downranking von Inhalten. Dass die Meldungen durch die Nutzer:innen vorgenommen werden können, hat den Vorteil, dass dadurch die ganze Breite der Inhalte einer (potenziell) ständigen «Überwachung» unterliegt und Desinformation leichter identifiziert werden kann (vgl. Stein et al., 2023). Ein Problem bei der Meldung durch Nutzer:innen ist

allerdings ihre oft fehlende Expertise zur Beurteilung der Inhalte (Dreyer et al., S. 52). Ein weiteres Problem liegt im Missbrauchspotenzial, zumal die Meldefunktion auch dazu benutzt werden kann, gewisse Nutzer:innen oder Meinungen systematisch aus dem Diskurs auszuschliessen (Dreyer et al., S. 52). Der DSA versucht dem potenziellen Missbrauch dadurch zu begegnen, dass eine (temporäre) Sperrung auch für Nutzer:innen vorgesehen werden kann, die häufig und unbegründet Meldungen erstatten (II.2.c). Das Flagging greift als solches nicht in die Meinungsfreiheit ein; es scheint zudem sogar in heterogenen Gruppen gute Ergebnisse in Bezug auf die Richtigkeit von Informationen erzielen zu können – allerdings ist dies nach der Forschung nicht der Fall, wenn die Nutzer:innen segregiert und polarisiert sind (vgl. Stein et al., 2023).

Die Beurteilung des Flaggings hängt von der Zusammensetzung der Nutzer:innen und davon ab, welche Massnahmen aufgrund des Flagging ergriffen werden. Das Flagging von Inhalten hat zudem den Nebeneffekt, dass Inhalte unter Umständen erst und gerade wegen des Flagging beachtet werden (siehe auch f.) und Nutzer:innen bei Inhalten, die nicht geflaggt wurden, annehmen könnten, dass diese als unproblematisch bzw. wahr eingestuft werden (Pennycook & Rand, 2019).

Das Flagging von Inhalten durch Nutzer:innen ist aus unserer Sicht eine sinnvolle Massnahme, zumal die Plattformen durch die Nutzer:innen beim Sichten der Inhalte unterstützt werden. Zudem haben die Nutzer:innen die Möglichkeit, auf Inhalte hinzuweisen, die besonders (und ev. nur) aus Sicht einer bestimmten Gruppe (bspw. ethnische Minderheiten) problematisch erscheinen. Zudem kann das Flagging die Eigenverantwortung der Nutzer:innen stärken und die Wichtigkeit von Richtigkeit als Norm und Wert verdeutlichen. Wichtig wäre hier allerdings sichtbar zu machen, dass nicht-geflaggte Inhalte nicht automatisch als unproblematisch eingestuft werden und die Anzahl an Meldungen zur Anzahl der Sichtungen in Relation gestellt werden muss. Zudem kann es empfehlenswert sein, das Flagging durch Nutzer:innen mit Downranking zu verbinden, da selbst geflaggte Informationen bei wiederholter Darbietung immer noch eine erhöhte Glaubwürdigkeit haben können (siehe Pennycook & Rand, 2018, Unkelbach & Greifeneder, 2018).

#### f) Flagging durch besonders qualifizierte Personen (Trusted Flaggers)

Um das Problem der fehlenden Expertise der meldenden Nutzer:innen zu entschärfen, kann die Melde- und Flaggingfunktion nur besonders qualifizierten Personen (sog. Trusted Flaggers) zugestanden werden. Mit diesem Ansatz können Plattformen zugleich dem Missbrauchspotenzial des Flaggings durch Nutzer:innen begegnen. Trusted Flaggers sind zwar eher in der Lage, Inhalte richtig zu bewerten; es bräuchte aber sehr viele Trusted Flaggers, um die Flut an Inhalten auf grossen Plattformen zu sichten und zu flaggen. Zudem müsste auch hier sichergestellt werden, dass nichtgeflaggte Inhalte nicht automatisch als unproblematisch erscheinen.

Die Ergänzung des Meldesystems durch Trusted Flaggers ist aus unserer Sicht zu begrüßen. Aufgrund ihrer Expertise werden Trusted Flaggers Inhalte wohl in der Regel zu Recht melden. Jedenfalls ist eine klar höhere Qualität zu erwarten als beim Flagging durch Nutzer:innen. Wenn Meldungen von Trusted Flaggers von einer Plattform präferiert behandelt werden, kann zudem sichergestellt werden, dass problematische Inhalte schnell gelöscht bzw. downgeranked werden können und entsprechend wenig Reichweite erlangen.

g) Automatische Prüfung von Inhalten (Upload Filter)

Das Aufschalten von problematischen Inhalten kann auch mithilfe von sog. Upload-Filtern bekämpft werden, die Inhalte vor der Aufschaltung prüfen. Eine präventive Prüfung von Inhalten ist allerdings ein schwerer Eingriff in die Meinungsfreiheit (Zu den Upload-Filtern im Zusammenhang mit der RL 2019/790 bspw. Durantaye, S. 596). Wenn die Prüfung automatisiert erfolgt, kann der Kontext einer Information meist nicht bewertet werden, was die Gefahr von Fehlentscheidungen stark erhöht. Automatisierte Upload-Filter mögen ein Mittel sein, um das Aufschalten von klar rechtswidrigen und leicht (wieder-)erkennbaren Inhalten zu verhindern (bspw. urheberrechtlich geschützte Musikstücke). Für die Bekämpfung von Desinformation sind sie aber ungeeignet, weil für deren Bewertung oft der Kontext miteinbezogen werden muss. Automatisierte Verfahren lassen sich nach dem heutigen Stand der Technik in der Regel nur (aber immerhin) als ergänzende Massnahme zu Trusted Flagging und User Flagging einsetzen, um eine gewisse Triage der Inhalte zu erreichen und Inhalte wiederzuerkennen, die bereits als problematisch qualifiziert wurden und nun erneut in identischer oder hinreichend ähnlicher Form verbreitet werden.

h) Internes Fact-Checking und Debunking

Fact-Checking ist darauf ausgerichtet, den Wahrheitsgehalt einer Information zu prüfen und gegebenenfalls mit journalistischen Mitteln richtigzustellen, sog. Debunking. Es handelt sich dabei in der Regel um eine an das Flagging oder eine Meldung anschließende Massnahme für Inhalte, die sich nicht ohne weiteres als richtig oder falsch qualifizieren lassen. Die Richtigstellung von zweifelhaften oder problematischen Informationen wurde als Massnahme teilweise als umstritten bewertet, weil unklar war, ob eine Korrektur falscher Informationen tatsächlich zu einer veränderten Einschätzung bei Personen führt. Es wurde bspw. befürchtet, dass die für die Korrektur erforderliche erneute Nennung der Desinformation zu einer erhöhten Aufmerksamkeit in Bezug auf diese führen und bestehende Positionen sich sogar verhärten könnten (Backfire Effekt). Weiterführende Arbeiten konnten sog. Backfire Effekte allerdings nicht immer bestätigen (Ecker et al., 2019; Swire-Thompson et al., 2020), sodass eine klare Empfehlung für das Debunking von Informationen ausgesprochen wurde (Lewandowsky et al., 2020).

Fact-Checking ist oft mit grossem Aufwand verbunden. Dies führt dazu, dass nicht allen Inhalten die gleiche Aufmerksamkeit zukommen kann. Hinzu kommt, dass mit Fact-Checking ein politisches Programm verfolgt werden kann, indem nur (oder vor allem) bestimmte

Aussagen überprüft werden. In Verbindung mit den hohen Kosten der Massnahme scheint das Fact-Checking als weitreichendes Mittel zur Bekämpfung von Desinformation wenig geeignet. Ein Vorteil vertiefter Analysen von Inhalten besteht darin, dass sich damit erkennen lässt, zu welchen Themen Desinformation verbreitet wird. Wenn die Ergebnisse des Fact-Checking veröffentlicht werden, kann auch das Bewusstsein der Nutzer:innen für das Problem der Desinformation erhöht werden.

i) Tagging

Plattformen mit einer Meldefunktion für Desinformation oder problematische Inhalte sehen als Teil ihres Massnahmesystems oft auch die bloße Kennzeichnung problematischer Inhalte vor, sog. Tagging. Die Inhalte sind so weiterhin wahrnehmbar und das Tagging bildet sozusagen Teil des Diskurses, indem Zweifel an der Richtigkeit der Inhalte signalisiert werden. Das Tagging greift damit nicht in relevanter Weise in die Meinungsfreiheit der Verfasser:innen und Verbreiter:innen von Desinformation ein und ist als deutlich mildere Massnahme als das Löschen, Sperren oder Downranking von Inhalten zu qualifizieren. Es eignet sich damit besonders für Inhalte, deren (Un)wahrheitsgehalt nicht offenkundig ist oder nicht zweifelsfrei überprüft werden kann. Durch das Tagging kann das Bewusstsein der Nutzer:innen für problematische Inhalte erhöht werden, was dazu beitragen mag, dass die Inhalte weniger weiterverbreitet werden, zumindest nicht unreflektiert (vgl. Koch et al., 2023). Das Tagging kann aber auch dazu führen, dass die Inhalte stärker verbreitet werden, weil sie besonders interessant erscheinen. Diese Effekte hängen wesentlich von der Ausgestaltung der Algorithmen und von den Nutzer:innen der Plattformen ab. Grundsätzlich ist Tagging eine sinnvolle Massnahme für das Unterbinden der Verbreitung von Desinformation. Die Wirksamkeit wird allerdings bisweilen bezweifelt, weil gewisse Gruppierungen ein Tagging als Beweis für einen (staatlich) manipulierten Diskurs betrachten.

j) Einschränkung der Möglichkeit, Inhalte weiterzuleiten

Durch die Einschränkung der Möglichkeit, Inhalte weiterzuleiten, soll die massenhafte Weiterverbreitung von Desinformation verhindert oder zumindest eingeschränkt werden. Die Verbreitung von Inhalten auf Messengerdiensten aber auch auf Social Media kann durch die Beschränkung der Gruppengrößen und der Möglichkeit zur Weiterleitung von Inhalten eingeschränkt werden. Das wird bereits verschiedentlich umgesetzt (dazu vorne IV.4.b.vii). Über die Effekte dieser Massnahmen ist noch wenig bekannt. Sie hängen vermutlich stark von der Plattform ab. Plattformspezifisch ausgestaltet kann die Einschränkung der Möglichkeit, Inhalte weiterzuleiten, eine sinnvolle Massnahme zur Unterbindung der Verbreitung von Desinformation sein.

Einschränkungen der Weiterleitung könnten durch Verzögerungen oder Unterbrechungen der Weiterleitung ergänzt werden: Durch vorgeschaltete Fragen könnten Nutzer:innen in regelmässigen Abständen gebeten werden, zuerst den Wahrheitsgehalt der Information, die

sie betrachten oder teilen möchten, zu bewerten. Dies könnte dazu führen, dass Nutzer:innen stärker auf die Richtigkeit (statt auf den sozialen Wert bspw. durch den erwarteten Zuspruch von wichtigen anderen Personen) der Information achten, was gemäss Forschung die Unterscheidung von Information und Desinformation beim Teilen von Inhalten verbessert. Diese Massnahmen betreffen nur die Verbreitung von Inhalten; sie schränken die Meinungsfreiheit damit nicht in relevanter Weise ein. Diese Massnahmen betreffen allerdings jede Kommunikation auf der Plattform und sind nicht auf Desinformation oder problematische Inhalte beschränkt. Insofern gehen sie weit über die vorliegende Problemstellung hinaus. Dennoch kann die Beschränkung oder Unterbrechung der Weiterleitung dazu beitragen, das Verbreiten von Desinformation auf einen begrenzten Kreis von Nutzer:innen einzuschränken bzw. generell zu reduzieren.

#### 5. Zivilgesellschaftliches und individuelles Engagement

Auch zivilgesellschaftliche Organisationen haben sich dem Thema Desinformation angenommen. Die Fachstelle für Rassismusbekämpfung (FRB) geht bspw. spezifisch gegen rassistische Desinformationsinhalte vor und fördert Projekte, die sich in diesem Bereich einsetzen. In der Schweiz gibt es, soweit bekannt, keine klassischen Fact-Checking-Organisationen, wie bspw. Correctiv.org oder FactCheck.org, die in Deutschland bzw. international aktiv sind. Factcheck.ch prüft bspw. ausschliesslich Behauptungen zum UN-Migrationspaket auf ihren Wahrheitsgehalt. Journalistische Medien wie Swissinfo, tagesanzeiger.ch oder watson.ch bieten weiteres bereichsspezifisches Fact-Checking an. Im Lichte dessen werden auf zivilgesellschaftlicher und individual Ebene folgende Massnahmen vorgeschlagen:

##### a) Gründung von Fact-Checking-Agenturen

Fact-Checking wird in der Schweiz von Medienorganisationen gemacht, allerdings in einem beschränkten Umfang. Denkbar wäre es, Fact-Checking-Agenturen oder investigative Recherche-Organisationen mit Fokus auf die Schweiz zu gründen oder zu unterstützen. In der Schweiz sind andere Inhalte als problematisch einzustufen als anderswo. Deshalb wären Fact-Checking-Agenturen zu begrüssen, die auf die Prüfung von Inhalten mit Bezug auf die Schweiz fokussieren. Hinter solchen Agenturen könnten bspw. Stiftungen stehen. Denkbar wäre auch eine finanzielle Unterstützung durch den Staat. In diesem Fall müssten aber geeignete Massnahmen ergriffen werden, um die inhaltliche Unabhängigkeit der Fact-Checking-Agenturen vom Staat sicherzustellen.

##### b) Melden unglaubwürdiger Inhalte durch Nutzer:innen

Wie bereits ausgeführt (siehe dazu vorn, C.II.5), können Nutzer:innen selbst gegen Desinformation vorgehen, indem sie den Plattformen unglaubwürdige Inhalte melden. Denkbar

ist auch eine Meldung der Nutzer:innen an die SLK (siehe dazu vorn, C.III.3.) oder den Presserat (siehe dazu vorn, C.II.2.).

### **III. Indirekte Massnahmen**

Mindestens so wichtig wie direkte Massnahmen gegen das Verfassen und Verbreiten sind indirekte Massnahmen zur Reduktion der Wirkung von Desinformation. Bürger:innen müssen Desinformation erkennen und damit umgehen können. Die verstärkte Thematisierung von Desinformation hat hier massgebliche Vorteile: Eine erhöhte Konfrontation mit der Problematik kann dazu führen, dass Menschen aufmerksamer und ein Stück weit «immuner» gegenüber Desinformation werden, wie die aktuelle Forschung zeigt (Lewandowsky & van der Linden, 2021). Die Analyse der bisherigen Forschung zeigt aber auch, dass zu wenig wissenschaftlich gesicherte Erkenntnisse vorliegen, um die kausalen Auswirkungen von Desinformation hinreichend erfassen und beurteilen zu können. Das gilt auch (und gerade) für die Schweiz. Die empirischen Grundlagen müssen deshalb verbessert werden. Nach dem heutigen Stand der Erkenntnis vermögen die folgenden indirekten Massnahmen den Einfluss von Desinformation zu reduzieren:

#### **1. Teilnahme an Europäischen Gremien und Initiativen**

In Europa bestehen zahlreiche Gremien und Initiativen, die sich mit Desinformation im weitesten Sinn befassen. In einigen dieser Gremien ist die Schweiz als Beobachterin dabei, in anderen wäre eine Mitarbeit oder Teilnahme hilfreich.

Zu denken ist z.B. an eine Teilnahme am Kooperationsnetzwerk für Wahlen.

#### **2. Steigerung der Medien- und Digitalkompetenz**

Der Medien- und Digitalkompetenz im weitesten Sinn kommt eine zentrale Rolle bei der Steigerung der Resilienz gegenüber Desinformation zu. Diese Kompetenzen sind Voraussetzung für das Erkennen und den kritischen Umgang mit Desinformation. Einige Organisationen, bspw. Pro Juventute ([www.projuventute.ch/de/eltern/medien-internet](http://www.projuventute.ch/de/eltern/medien-internet)), der Schweizerische Verlegerverband ([www.schweizermedien.ch/politik/testimonial/fake-news](http://www.schweizermedien.ch/politik/testimonial/fake-news)), das Bundesamt für Sozialversicherung ([www.jugendundmedien.ch](http://www.jugendundmedien.ch)), das Bundesamt für Bevölkerungsschutz (<https://blog.alertswiss.ch/de/rubriken/ Gefahren/wie-sie-desinformation-erkennen/>) und wissenschaftliche Organisationen wie das Forschungszentrum Öffentlichkeit und Gesellschaft ([www.iqesonline.net/bildung-digital/checknews/lernumgebungen/](http://www.iqesonline.net/bildung-digital/checknews/lernumgebungen/)), haben bereits Initiativen zur Förderung der Medienkompetenz mit spezifischen Angeboten zu Desinformation realisiert. Weitere Massnahmen zur Steigerung des Wissens über und der Kompetenz im Umgang mit Desinformation können sich durch empirisch überprüfte Online-Spiele oder Desinformations-Impf-Videos umsetzen lassen. Diese wirken präventiv und steigern nachweislich die

individuelle Sensitivität und Kompetenzen der jeweiligen Person (siehe auch «Prebunking»; vgl. Roozenbeek et al., 2022, van der Linden, 2023).

Solche Initiativen sind zu begrüßen. Sie sollten weiter gefördert werden und möglichst breit abgestützt sein.

#### a) Steigerung der Medien- und Digitalkompetenz bei jüngeren Zielgruppen

Bei den jüngeren Zielgruppen können die Medien- und Digitalkompetenz in der Schule ausgebildet werden. Sinnvoll wäre die Etablierung der Desinformations-Thematik als Pflichtinhalt im Lehrplan. Im Lehrplan 21 erhält der Bereich «Medien und Informatik» mehr Gewicht und wird bereits ab Stufe Kindergarten etabliert. Der übergeordnete Lehrplan enthält auch Massnahmen zum Aufbau von inhaltlicher Medienkompetenz, also bspw. dazu, wie Informationen gesucht, genutzt und überprüft werden (<https://v-fe.lehrplan.ch>). Bei den Kompetenzbereichen im Modul «Medien und Informatik» ist unter anderem vorgesehen, dass «Schülerinnen und Schüler Medien und Medienbeiträge entschlüsseln, reflektieren und nutzen können». Die inhaltliche Medienkompetenz konkurriert allerdings in der Umsetzung mit der Vermittlung von anderen Kompetenzen im Bereich der Informatik, unter anderem mit dem Programmieren, dem Verständnis für Algorithmen und dem Umgang mit Daten. In vielen Angeboten der öffentlichen Hand dominiert zudem die Vorbeugung von Verhaltenssucht, Cybermobbing und Übergriffen im Netz im Umgang mit digitalen Medien (z.B. im Kanton Zürich). Die inhaltliche Medienkompetenz steht bislang weniger im Fokus, die Themen Desinformation oder «Fake News» werden nur punktuell aufgegriffen.

Mit Blick auf jüngere Zielgruppen wird zudem diskutiert, dass viele Massnahmen im Bereich Desinformation zu spät ergriffen werden (i.d.R. Sekundarschule, Gymnasien), obwohl Kinder und Jugendliche bereits viel früher mit digitalen Medien in Kontakt kommen. Die Einführung des Bereichs Medien und Informatik ab Stufe Kindergarten ist deshalb zu begrüßen. In der Ausgestaltung ist darauf zu achten, dass die inhaltliche Medienkompetenz und der Bereich der Desinformation ausreichend gewichtet werden. Erste Projekte setzen da bereits an und bieten auf Jugendliche ausgerichtete Informationen zum Thema Desinformation an (z.B. Pro Juventute); andere stellen Lehrpersonen Unterrichtseinheiten zur Vermittlung von inhaltlicher Medienkompetenz zur Verfügung (Check News).

#### b) Steigerung der Medien- und Digitalkompetenz bei älteren Zielgruppen

Viele Initiativen richten sich explizit an ein junges Zielpublikum. Zielführend wären aber auch Angebote, welche die breite Bevölkerung und vor allem auch ältere Menschen ansprechen. Denn gerade ältere Menschen verfügen oftmals über weniger Digitalkompetenzen, was mit dazu beiträgt, dass sie in Studien als anfälliger für Desinformation identifiziert werden. Da diese Zielgruppe zunehmend auch Social Media und Messengerdienste, allen voran Whatsapp, nutzt, ist sie auch verstärkt dem Risiko ausgesetzt, dort auf Desinformation zu

treffen. Prospektiv wäre es daher wichtig, auch diese Menschen für Desinformation auf den Plattformen zu sensibilisieren und im Umgang damit zu befähigen. Denkbar wäre, verstärkt Organisationen einzubinden, die sich für ältere Zielgruppen engagieren (z.B. Pro Senectute).

c) Steigerung der Medien- und Digitalkompetenz bei Personen mit Migrationshintergrund  
Bislang weiss man wenig über die Bedeutung und Wirkung von Desinformation bei Personen mit Migrationshintergrund, gerade auch in der Schweiz. Ein Grund liegt darin, dass diese Gruppen für die Forschung gar nicht oder nur sehr schwer zu erreichen sind. Weitere Forschung dazu wäre wichtig. Bei der Bekämpfung von Desinformation sind unter anderem Sprachbarrieren zu berücksichtigen. Wünschenswert wären Angebote, die sich der Problematik der Desinformation aus der Perspektive dieser Personen annehmen. Auch hier sollten Organisationen eingebunden werden, die sich im Bereich der Migration engagieren.

### 3. Stärkung des Informationsjournalismus

Weiter sind auch Massnahmen zu begrüessen, die den Informationsjournalismus in der Schweiz stärken. Damit soll der Zugang zu qualitativ hochwertiger Information für die Bevölkerung sichergestellt werden. Der öffentliche Rundfunk spielt dabei eine zentrale Rolle, weil er eine hohe Qualität liefert und viele Menschen erreicht. Dieser Service public sollte aufrechterhalten werden und mit Blick in die Zukunft auch online über Webseiten und Social-Media-Angebote zur Verfügung stehen. Weiter ist der Zugang zu Qualitätsmedien für ein jüngeres Zielpublikum sicherzustellen, etwa durch (teilweise bereits bestehende) vergünstigte Abonnements für Schüler:innen und Studierende.

### 4. Gründung einer (kostenlosen) Medienplattform

Denkbar wären verschiedene Ausgestaltungsformen. Der Staat oder ein Zusammenschluss von bestehenden Medien könnte eine Infrastruktur zur Verfügung stellen, welche von Medien, Journalist:innen und anderen Kommunikator:innen mit Inhalten befüllt werden könnte. Mit Blick auf die Sicherstellung der Qualität könnte der Zugang auf diejenigen Medien und Journalist:innen und anderen Kommunikator:innen beschränkt werden, die sich zur Einhaltung journalistischer Grundsätze, namentlich der Grundsätze des Presserates verpflichten. Möglich wäre auch, dass die Plattform durch die Medien selbst zur Verfügung gestellt wird.

Der Zugang müsste für alle Personen kostenlos sein und es müssten auf einer Plattform Inhalte von mehreren Medien zur Verfügung gestellt werden. Auf diese Weise könnte man der Bevölkerung einen besseren Zugang zu qualitativ hochwertiger Information ermöglichen.

\* \* \* \* \*

## E. Glossar

**AGB:** Allgemeine Geschäftsbedingungen (AGB) sind im Unterschied zu einer Individualabrede alle für eine Vielzahl von Verträgen vorformulierten Vertragsbedingungen, die eine Vertragspartei der anderen Vertragspartei bei Abschluss eines Vertrages stellt.

**Debunking:** Beim Debunking wird Desinformation durch korrigierende oder den Inhalt widerlegende Informationen ergänzt. Die zur Widerlegung notwendige Information wird in der Regel durch Fact-Checking-Organisationen zur Verfügung gestellt.

**Deplatforming:** Das Sperren von Nutzer:innen bzw. ihrer Accounts auf Social-Media-Plattformen und Messengerdiensten.

**Desinformation:** Als Desinformation wird das gezielte Verbreiten von Falschinformationen mit bezeichnet, dessen Ziel es ist, die Gesellschaft, einzelne Gruppen oder Einzelpersonen im Sinn politischer oder wirtschaftlicher Interessen zu täuschen. Selten bezeichnet „Desinformation“ auch die beschriebene Falschinformation selbst.

**Downranking:** Downranking wird auf Social-Media-Plattformen als Massnahme gegen die Verbreitung von Desinformation eingesetzt. Indem gewisse Inhalte weniger oder in einem Feed weiter unten angezeigt werden, kann faktisch eine ähnliche Wirkung erreicht werden wie durch eine Löschung.

**DSA (Digital Services Act):** Das Gesetz über digitale Dienste (Digital Services Act) ist eine Verordnung der EU, die unter anderem Haftungs- und Sicherheitsvorschriften für digitale Plattformen, Dienste und Produkte schaffen und den digitalen Binnenmarkt vollenden soll. Neben der Fortentwicklung geltender Regeln enthält der DSA auch neue Durchsetzungsvorschriften und neue Anforderungen für Online-Plattformen.

**Fact-Checking:** Als Fact-Checking wird ein journalistisches Konzept bezeichnet, das die Aussage einer oder mehrerer Personen anhand von recherchierbaren und nachprüfbaren Fakten systematisch überprüft. Dabei werden mündliche oder schriftliche Aussagen oder audiovisuelle Inhalte den recherchierten Fakten gegenübergestellt.

**Flagging:** Als Flagging wird der Mechanismus zur Meldung anstössiger Inhalte an die Betreiber:in einer Online-Plattform bezeichnet. Das Flagging kann durch andere Nutzer:innen (**User Flagging**) oder durch besonders qualifizierte Personen (**Trusted Flaggers**) erfolgen.

**Governance:** Governance ist nach Mayntz (2004, 66) die «Gesamtheit aller nebeneinander bestehenden Formen der kollektiven Regelung gesellschaftlicher Sachverhalte: von der institutionalisierten zivilgesellschaftlichen Selbstregelung über verschiedene Formen des Zusam-

menwirkens staatlicher und privater Akteure bis hin zu hoheitlichem Handeln staatlicher Akteure».

**Internet-Troll:** Als Internet-Troll bezeichnet man eine Person, die im Internet vorsätzlich mit kontroversen Kommentaren eine Diskussion entfachen oder absichtlich Menschen provozieren will. Eine Troll-Farm oder Troll-Armee ist eine organisierte Gruppe von Internet-Trollen, die versucht, sich in politische Meinungsbildungs- und Entscheidungsprozesse einzumischen.

**Ko-Regulierung:** Kombination von staatlicher und nicht-staatlicher Regulierung.

**Media Literacy:** Media Literacy (Medienkompetenz) bezeichnet die Fähigkeit, Medien und ihre Inhalte den eigenen Zielen und Bedürfnissen entsprechend sachkundig zu nutzen.

**Messengerdienste:** Messengerdienste oder Kurznachrichtendienste sind Programme oder Apps, mit denen User:innen Texte, Bilder und Videos in Echtzeit verschicken können. Die bekanntesten Messengerdienste in der Schweiz sind WhatsApp, Telegram, Snapchat oder Facebook Messenger, Threema und Signal.

**Micro-Targeting:** Der Begriff bezeichnet eine aus den USA stammende Kommunikationsstrategie, die vor allem im Bereich der politischen Kommunikation und im Marketing verwendet wird. Sie zielt darauf, die begrenzten Ressourcen der Öffentlichkeitsarbeit möglichst effizient und effektiv einzusetzen und Nutzer:innen anhand von gesammelten Daten mit gezielten und personalisiertem Inhalt zu erreichen.

**Online-Plattform:** Gemäss DSA bezeichnet der Begriff Onlineangebote, die sich nicht auf das Speichern von Daten beschränken (wie etwa Hosting- und Clouddienste), sondern Daten auch aufbereiten, sie verbreiten und eventuell auch priorisieren und auswählen (i.d.R. durch Algorithmen) und mitunter darüber entscheiden, welche Inhalte wem angezeigt werden. In der Regel wird der Begriff als Sammelbegriff für Social-Media-Plattformen (z.B. Facebook, Instagram, TikTok oder Twitter), Messengerdienste (z.B. Whatsapp oder Telegram) und Videoportale (z.B. Youtube oder Vimeo) verstanden.

**Selbstorganisation:** Regeln, die sich die Akteur:innen selbst geben, etwa durch vertragliche Regelwerke, wie Community Standards

**Selbstregulierung:** Unter Selbstregulierung sind Regeln zu verstehen, die sich verschiedene Akteur:innen einer Branche selbst geben.

**Similarity-Attraction Paradigma:** Dieses Paradigma erklärt, dass Menschen dazu neigen, sich zu Menschen hingezogen zu fühlen, die ihnen ähnlich sind. Diese Ähnlichkeiten bestehen unter anderem in gemeinsamen Wertvorstellungen und Überzeugungen, Einstellungen und kulturellen Hintergründen.

**Social Bot:** Social Bots sind Computerprogramme, die in sozialen Netzwerken oder Onlineforen automatisiert Inhalte produzieren und verbreiten. Die meisten Bots sind so programmiert, dass sie Inhalte (z.B. Tweets) mit bestimmten Stichworten (z.B. Hashtags) oder Inhalte von spezifischen Accounts weiterverbreiten. Andere Bots werden so programmiert, dass sie online verfügbare Inhalte nach definierten Regeln zusammenstellen und veröffentlichen. Beispiele sind Accounts, die regelmässig Wetterdaten oder Sportresultate veröffentlichen. Manchmal werden Bots so programmiert, dass sie versuchen, unter Nachahmung menschlichen Verhaltens, mit menschlichen Nutzer:innen zu interagieren.

**Social Media:** Social Media sind digitale Medien bzw. Plattformen, die es Nutzer:innen ermöglichen, sich im Internet zu vernetzen, sich also untereinander auszutauschen und mediale Inhalte einzeln, in einer definierten Gemeinschaft oder offen in der Gesellschaft zu erstellen, zu diskutieren und weiterzugeben.

**Soziale Identität:** Als soziale Identität wird derjenige Teil des Selbstkonzepts einer Person bezeichnet, der sich aus dem Wissen über die Mitgliedschaft in einer sozialen Gruppe (oder in mehreren sozialen Gruppen) und aus dem Wert sowie der emotionalen Bedeutung dieser Mitgliedschaft ableitet. Wenn Menschen als Gruppenmitglieder handeln, haben sie ein Bedürfnis nach positiver sozialer Identität und sind deshalb motiviert, die Eigengruppe positiv von relevanten Fremdgruppen abzugrenzen.

**Tagging:** Als Tagging wird die Kennzeichnung von Inhalten bezeichnet, die aufgrund einer Meldung (Flagging) von Nutzer:innen oder von Trusted Flaggern erfolgen kann. Durch das Tagging soll erreicht werden, dass sich die Nutzer:innen mit einem kritischeren Blick mit den Inhalten auseinandersetzen.

**Very Large Online Platforms (VLOP):** Mit diesem Begriff werden im DSA Online-Plattformen bezeichnet, die mehr als 45 Mio. aktive Nutzer:innen pro Monat in der EU haben.

## F. Literatur

- Abokhodair, N., Yoo, D., & McDonald, D. W. (2015). Dissecting a social botnet: Growth, content and influence in Twitter. In Proceedings of the 18th ACM conference on computer supported cooperative work & social computing, 839-851.
- Adams, Z., Osman, M., Bechlivanidis, C., & Meder, B. (2023). (Why) Is Misinformation a Problem? *Perspectives on Psychological Science*, 174569162211413. <https://doi.org/10.1177/17456916221141344>
- Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211–236.
- Allcott, H., Gentzkow, M., & Yu, C. (2019). Trends in the diffusion of misinformation on social media. *Research and Politics*, 6(2), 211–236. <https://doi.org/10.1177/2053168019848554>.
- Arechar, A. A., Allen, J., Berinsky, A. J., Cole, R., Epstein, Z., Garimella, K., Gully, A., Lu, J. G., Ross, R. M., Stagnaro, M. N., Zhang, Y., Pennycook, G., & Rand, D. G. (n.d.). *Understanding and Combatting COVID-19 Misinformation Across 16 Countries on Six Continents*. Preprint/publication in preparation.
- Auer A., Malinverni G., Hottelier Mic (2023). *Droit constitutionnel suisse*, Volume 1. 3. Auflage Bern 2013.
- Bail, C. A., Guay, B., Maloney, E., Combs, A., Hillygus, D. S., Merhout, F., ... & Volfovsky, A. (2020). Assessing the Russian Internet Research Agency's impact on the political attitudes and behaviors of American Twitter users in late 2017. *Proceedings of the national academy of sciences*, 117(1), 243-250.
- Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, 33(2), 122–139. <https://doi.org/10.1177/0267323118760317>
- Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 US Presidential election online discussion. *First monday*, 21(11-7).
- Besson, M. & Boillet, V. (2020). The guarantee of political rights in view of misinformation. In: Baume S. & Boillet V. & Martenet V., *Misinformation in referenda*, London 2020.
- Besson, S. (2003). Les obligations positives de protection des droits fondamentaux. Un essai en dogmatique comparative. *Revue de droit suisse* 2003/1 49 – 96.
- Biaggini, G. (2017). *BV Kommentar: Bundesverfassung der schweizerischen Eidgenossenschaft*. 2. Auflage Zürich 2017.
- Born, Ch. (2021). «Medienkampagne»: Problematische Ausdehnung des Persönlichkeitsschutzes, Risiken und Nebenwirkungen des Hirschmann II-Urteils des Bundesgerichts, in: *medialex* 10/2021, 7. Dezember 2021, abgerufen unter:

- [www.medialex.ch/2021/12/07/medienkampagne-problematische-ausdehnung-des-persoentlichkeitsschutzes/#post-6611-endnote-2](http://www.medialex.ch/2021/12/07/medienkampagne-problematische-ausdehnung-des-persoentlichkeitsschutzes/#post-6611-endnote-2), zuletzt besucht am 23. März 2022
- Borucki, I., & Meisberger, A. (2019). Dealing and dancing with Bots: Der Umgang der Parteien mit disruptiven Phänomenen im Bundestagswahlkampf. In: Korte, K.R., Korte & J., Schoofs. (Hrsg.) Die Bundestagswahl 2017 (pp. 343-362). Springer VS, Wiesbaden.
- Bowles, J., Larreguy, H., & Liu, S. (2020). Countering misinformation via WhatsApp: Preliminary evidence from the COVID-19 pandemic in Zimbabwe. *PloS one*, 15(10), e0240005.
- Brady, W. J., Gantman, A. P., & van Bavel, J. J. (2020). Attentional capture helps explain why moral and emotional content go viral. *Journal of Experimental Psychology: General*, 149(4), 746–756. <https://doi.org/10.1037/xge0000673>
- Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., van Bavel, J. J., & Fiske, S. T. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences of the United States of America*, 114(28), 7313–7318. <https://doi.org/10.1073/pnas.1618923114>
- Brennen, J.S., Simon, F.M., & Nielsen, R.K. (2021). Beyond (Mis)Representation: Visuals in COVID-19 Misinformation. *The International Journal of Press/Politics*, 26(1), 277-299. <https://doi:10.1177/1940161220964780>
- Bröhm, Alexandra (04.06.21) Fake News gegen das Impfen, Bund.
- Bundesrat (2021). Die Sicherheitspolitik der Schweiz. Bericht des Bundesrates vom 24. November 2021. Abgerufen unter: <https://www.vbs.admin.ch/de/sicherheit/sicherheitspolitik/sicherheitspolitische-berichte/sicherheitspolitischer-bericht-2021.detail.document.html/vbs-internet/de/documents/sicherheitspolitik/sicherheitspolitische-berichte/2021/Sicherheitspolitik-Schweiz-Bericht-Bundesrat-2021.pdf.html>
- Buser D. (2015). Gibt es Grenzen der Einflussnahme Privater in Abstimmungskampagnen?. Jusletter 18 Mai 2015.
- Byrne, D. (1961). Interpersonal attraction and attitude similarity. *Journal of Abnormal and Social Psychology*, 62, 713–715. <https://doi.org/10.1037/h0044721>
- Byrne, D. (1971). *The attraction paradigm*. Academic Press.
- Campaign for Accountability (2017). How Google Makes Millions Off of Fake News. Abgerufen unter: <https://campaignforaccountability.org/work/how-google-makes-millions-off-of-fake-news/>
- Center for Countering Digital Hate (2021). The Disinformation Dozen. Why platforms must act on twelve leading online anti-vaxxers. Abgerufen unter: <https://counterhate.com/research/the-disinformation-dozen/>

- Chia, S. C., Lu, F., & Gunther, A. C. (2022). Who fact-checks and does it matter? Examining the antecedents and consequences of audience fact-checking behaviour in Hong Kong. *International Journal of Press/Politics*. <https://doi.org/10.1177/19401612221142439>
- Cueni, R. (2019). Falsche und irreführende Informationen, ex ante 2019/1, 3 ff.
- Dan, V., Paris, B., Donovan, J., Hameleers, M., Roozenbeek, J., van der Linden, S., & von Sikorski, C. (2021). Visual mis- and disinformation, social media, and democracy. *Journalism & Mass Communication Quarterly*, 98(3), 641-664.
- Donatsch, A., Lieber, V., Summers, S., Wohlers W. (Hrsg.) (2020). Kommentar zur Schweizerischen Strafprozessordnung, 3. Aufl., Zürich – Basel – Genf 2020. Zit. SK StGB-Verfasser:in
- Dreyer, S., Stanciu, E., Potthast, K.C., & Schulz, W. (2021). Desinformation, Risiken für rechtliche Schutzziele, Regulierungslücken und zieladäquate Gegenmassnahmen. Gutachten im Auftrag der Landesanstalt für Medien NRW. Abgerufen unter <https://www.hans-bredow-institut.de/de/publikationen/desinformation-risiken-regulierungsluecken-und-adaequate-gegenmassnahmen>
- Dröge, C. (2003). Positive Verpflichtungen der Staaten in der Europäischen Menschenrechtskonvention.
- Ecker, U. K. H., Lewandowsky, S., Jayawardana, K., & Mladenovic, A. (2019). Refutations of equivocal claims: No evidence for an ironic effect of counterargument number. *Journal of Applied Research in Memory and Cognition*, 8(1), 98–107. <https://doi.org/10.1037/h0101834>
- Egelhofer, J. L. & Lecheler, S. (2019). Fake news as a two-dimensional phenomenon: A framework and research agenda. *Annals of the International Communication Association*, 43(2), 97–116. <https://doi.org/10.1080/2380898>
- Egli, P. & Rechsteiner, D. (2017). Social Bots und Meinungsbildung in der Demokratie, *AJP* 2017, 249-258.
- Ehrenzeller, B., Egli P., Hettich, P., Hongler, P., Schindler B., Schmid S., Schweizer R. (Hrsg.) (2023). Die schweizerische Bundesverfassung. St. Galler Kommentar. Zit. SG BV-Verfasser:in
- Eisenegger, M. (2021). Dritter, digitaler Strukturwandel der Öffentlichkeit als Folge der Plattformisierung. In M. Eisenegger, M. Prinzing P. Ettinger & R. Blum (Hrsg.), *Digitaler Strukturwandel der Öffentlichkeit. Historische Verortung, Modelle und Konsequenzen* (S. 17–38). Springer VS
- Elías, C., & Catalan-Matamoros, D. (2020). Coronavirus in Spain: Fear of 'official' fake news boosts WhatsApp and alternative sources. *Media and Communication*, 8(2), 462-466.
- Equoy, D. (2018). La responsabilité pénale des fournisseurs de services Internet, Etude à la lumière des droits suisse, allemand et français, Bern 2016.

- Epstein, Z., Sirlin, N., Arechar, A., Pennycook, G., & Rand, D. (2023). The social media context interferes with truth discernment. *Science Advances*, 9, eabo6169. <https://doi.org/10.1126/sciadv.abo6169>
- Ferrara, E. (2017). Disinformation and social bot operations in the run up to the 2017 French presidential election. *First Monday*, 22(8). <https://doi/10.5210/fm.v22i8.8005>
- Franta, B. (2021). Early oil industry disinformation on global warming. *Environmental Politics*, 30(4), 663-668.
- Freelon, D., & Wells, C. (2020). Disinformation as political communication. *Political Communication*, 37(2), 145–156. <https://doi.org/10.1080/10584609.2020.1723755>
- Freitas Melo, P. D., Vieira, C. C., Garimella, K., Melo, P. O., & Benevenuto, F. (2019). Can WhatsApp counter misinformation by limiting message forwarding? *International conference on complex networks and their applications*, 372-384.
- Gallwitz, F., & Kreil, M. (2021). The Rise and Fall of Social Bot Research. Available at SSRN 3814191.
- Geiser, Th. & Fountoulakis, Ch. (2018). Basler Kommentar zum Zivilgesetzbuch I, Art. 1 – 452 ZGB, Basel 2018. Zit. BK ZGB-Verfasser:in
- Gerdemann, S. & Spindler, G. (2023). Das Gesetz über digitale Dienste (Digital Services Act), Teil 2, GRUR 2023, 115 – 125.
- Gill, H., & Rojas, H. (2020). Chatting in a mobile chamber: effects of instant messenger use on tolerance toward political misinformation among South Koreans. *Asian Journal of Communication*, 30(6), 470-493.
- Global Disinformation Index (2019). Cutting the Funding of Disinformation: The Ad-Tech Solution. Abgerufen unter: <https://www.disinformationindex.org/research/2019-5-1-cutting-the-funding-of-disinformation-the-ad-tech-solution/>
- Globig, L. K., Holtz, N., & Sharot, T. (2022). Changing the incentive structure of social media platforms to halt the spread of misinformation (Preprint). <https://doi.org/10.31234/osf.io/26j8w>
- González-Bailón, S., & Lelkes, Y. (2022). Do social media undermine social cohesion? A critical review. *Social Issues and Policy Review*, 17(1), 155-180. <https://doi.org/10.1111/sipr.12091>
- Graeupner, D., & Coman, A. (2017). The dark side of meaning-making: How social exclusion leads to superstitious thinking. *Journal of Experimental Social Psychology*, 69, 218-222.
- Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, 5(1), eaau4586.
- Guess, A., Nyhan, B., & Reifler, J. (2018). *Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 U.S. presidential campaign.*

<http://www.ask-force.org/web/Fundamentalists/Guess-Selective-Exposure-to-Misinformation-Evidence-Presidential-Campaign-2018.pdf>

- Gunjic, I. (2020). Die schweizerische Demokratie und «Fake News», in: Meyer J., Zurkinder N., Staffel L.(Hrsg.), Recht und Innovation, Zürich/St. Gallen 2020.
- Ha, L., Andreu Perez, L., & Ray, R. (2021). Mapping Recent development in scholarship on Fake news and misinformation, 2008 to 2017: Disciplinary contribution, topics, and impact. *American Behavioral Scientist*, 65(2), 290–315. <https://doi.org/10.1177/0002764219869402>
- Häfelin, U., Haller, W., Keller H., Turnherr D. (2020). Schweizerisches Bundesstaatsrecht, 10. Auflage, Zürich 2020.
- Haenschen, K. (2022). The conditional effects of microtargeted facebook advertisements on voter turnout. *Political Behavior*. <https://doi.org/10.1007/s11109-022-09781-7>
- Heglich, S., & Janetzko, D. (2016). Are social bots on Twitter political actors? Empirical evidence from a Ukrainian social botnet. In Proceedings of the Tenth International AAAI Conference on Web and Social Media (ICWSM 2016), 579–582.
- Heimgartner, S. (2022). Strafbarkeit dschihadistischer Propaganda, AJP 2022, 1217 – 17 41.
- Heinemann, A. & Althaus, B. (2015). Posten, Liken, Sharen – Urheberrecht in sozialen Netzwerken. Jusletter 12. Oktober 2015.
- Herrero-Diz, P., Conde-Jiménez, J., & Reyes de Cózar, S. (2020). Teens' motivations to spread fake news on WhatsApp. *Social Media+ Society*, 6(3), 2056305120942879.
- Hilty, R. & Arpagaus, R. (2013). Basler Kommentar zum Gesetz gegen den Unlauteren Wettbewerb (UWG), Basel 2013. Zit. BK UWG-Verfasser:in
- Holzer, S., & Sengl, M. (2020). Quelle gut, alles gut? Glaubwürdigkeitsbeurteilung im digitalen Raum. In R. Hohlfeld, M. Harnischmacher, E. Heinke, L. S. Lehner, & M. Sengl (Hrsg.), Fake News und Desinformation: Herausforderungen für die vernetzte Gesellschaft und die empirische Forschung (S. 155 – 178). Nomos. <https://doi.org/10.5771/9783748901334-179>
- Howard, P. N., & Kollanyi, B. (2016). Bots, #StrongerIn, and #Brexit: Computational Propaganda during the UK-EU Referendum. SSRN Electronic Journal. <https://doi/10.2139/ssrn.2798311>
- Humprecht, E. (2019). Where 'fake news' flourishes: a comparison across four Western democracies, *Information, Communication & Society*, 22(13), 1973-1988, <https://doi.org/10.1080/1369118X.2018.1474241>
- Humprecht, E., Esser, F., & Van Aelst, P. (2020). Resilience to online disinformation: A framework for cross-national comparative research. *The International Journal of Press/Politics*, 25(3), 493-516.

- Humprecht, E., Esser, F., Aelst, P. V., Staender, A., & Morosoli, S. (2021). The sharing of disinformation in cross-national comparison: analyzing patterns of resilience. *Information, Communication & Society*, 1-21.
- Hohlfeld, R. (2020). Wahr oder falsch? Eine empirische Untersuchung zur Wahrnehmung von »Fake News« und echten Nachrichten in der politischen Kommunikation. In R. Hohlfeld, M. Harnischmacher, E. Heinke, L. S. Lehner, & M. Sengl (Hrsg.), *Fake News und Desinformation: Herausforderungen für die vernetzte Gesellschaft und die empirische Forschung* (S.155 – 178). Nomos. <https://doi.org/10.5771/9783748901334-179>
- Hohlfeld, R., Harnischmacher, M., Heinke, E., Lehner, L., & Sengl, M. (2020). *Fake News und Desinformation: Herausforderungen für die vernetzte Gesellschaft und die empirische Forschung*. Nomos Verlag.
- Holt, K., Ustad Figenschou, T., & Frischlich, L. (2019). Key dimensions of alternative news media. *Digital Journalism*, 7(7), 860-869.
- Hughes, H. C., & Waismel-Manor, I. (2021). The Macedonian fake news industry and the 2016 US election. *PS: Political Science & Politics*, 54(1), 19-23.
- Ibrahim, Y. (2017). Facebook and the Napalm Girl: reframing the iconic as pornographic. *Social Media+ Society*, 3(4), 2056305117743140.
- Keller, T. R., & Klinger, U. (2019). Social bots in election campaigns: Theoretical, empirical, and methodological implications. *Political Communication*, 36(1), 171-189.
- Kiener, R. (2020). Grundrechte in der Bundesverfassung, V. Teil: Grund- und Menschenrechte. In: Diggelmann, O. & Hertig, M. & Schindler, B. (Hrsg.), *Verfassungsrecht in der Schweiz* Bd. II, Zürich 2020.
- Kiener, R., Kälin, W., Wyttenbach, J. (2018). *Grundrechte*, 3. Auflage, Bern 2018.
- Klett K., Vertragsfreiheit und Kontrahierungszwang, *BJM* 2005, 161 ff.
- Koch, T. K., Frischlich, L., & Lermer, E. (2023). Effects of fact-checking warning labels and social endorsement cues on climate change fake news credibility and engagement on social media. *Journal of Applied Social Psychology*. <https://doi.org/10.1111/jasp.12959>
- Kozyreva, A., Wineburg, S., Lewandowsky, S., & Hertwig, R. (2022). Critical Ignoring as a Core Competence for Digital Citizens. *Current Directions in Psychological Science*, 32(1), 81–88. <https://doi.org/10.1177/09637214221121570>
- Kozyreva, A. I., Herzog ID, S. M., Lewandowsky, S., Hertwig, R. I., Lorenz-Spreen, P. I., Leiser, M. I., & Reifler, J. (2023). Resolving content moderation dilemmas between free speech and harmful misinformation. *Proceedings of the National Academy of Sciences of the United States of America*, 120(7), e2210666120. <https://doi.org/10.1073/pnas>
- Kühling, J. (2021). «Fake News» und «Hate Speech» - Die Verantwortung der Medienintermediäre zwischen neuen NetzDG, MStV und Digital Services Act, ZUM 2021.

- Kuhlmann, S. & Trute, H.H. (2022). Die Regulierung von Desinformation und rechtswidrigen Inhalten nach dem neuen Digital Services Act. *Zeitschrift für das gesamte Sicherheitsrecht* 2022. 115 – 123.
- Lawson, M. A., Anand, S., & Kakkar, H. (2023). Tribalism and Tribulations: The Social Costs of Not Sharing Fake News. *Journal of Experimental Psychology: General*. <https://doi.org/10.1037/xge0001374>
- Lehner, L. S. (2020, September). Social Media–Vox populi oder Forum der Agitatoren? Elementare Aufgaben-, Wirkungs- und Problemfelder sozialer Netzwerke im syrischen Bürgerkrieg. In R. Hohlfeld, M. Harnischmacher, E. Heinke, L. S. Lehner, & M. Sengl (Hrsg.), *Fake News und Desinformation: Herausforderungen für die vernetzte Gesellschaft und die empirische Forschung* (S.155 – 178). Nomos. <https://doi.org/10.5771/9783748901334-179>
- Lewandowsky, S., Cook, J., Ecker, U. K. H., Albarracín, D., Amazeen, M. A., Kendeou, P., Lombardi, D., Newman, E. J., Pennycook, G., Porter, E., Rand, D. G., Rapp, D. N., Reifler, J., Roozenbeek, J., Schmid, P., Seifert, C. M., Sinatra, G. M., Swire-Thompson, B., van der Linden, S., ... Zaragoza, M. S. (2020). *The debunking handbook 2020*. <https://doi.org/10.17910/b7.1182>
- Lewandowsky, S., & van der Linden, S. (2021). Countering misinformation and fake news through inoculation and prebunking. *European Review of Social Psychology*, 32(2), 348–384. <https://doi.org/10.1080/10463283.2021.1876983>
- Linville, D. L., & Warren, P. L. (2020). Troll factories: Manufacturing specialized disinformation on Twitter. *Political Communication*, 37(4), 447-467.
- Löber, L. & Rossnagel, A. (2020). Desinformation aus der Perspektive des Rechts. In Steinebach M., Bader K., Rinsdorf L., Krämer N., Rossnagel A., *Desinformation aufdecken und bekämpfen. Interdisziplinäre Ansätze gegen Desinformationskampagnen und für Meinungspluralität*. Baden – Baden 2020.
- Lyons, B. A., Merola, V., & Reifler, J. (2020). How bad is the fake news problem? The role of baseline information in public perceptions. In R. Greifeneder, M. E. Jaffé, E. J. Newman, & N. Schwarz (Eds.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 11–26). Routledge.
- Johann, M., & Wagner, J. (2020). Neue Debatte, altes Dilemma? Die Herausforderungen des Phänomens ‚Fake News‘ für die Unternehmenskommunikation. In R. Hohlfeld, M. Harnischmacher, E. Heinke, M. Sengl, & L. S. Lehner (Hrsg.), *Fake News und Desinformation: Herausforderungen für die vernetzte Gesellschaft und die empirische Forschung* (S. 99–116). Nomos. <https://doi.org/10.5771/9783748901334-97>
- Mäder, L. (2022). In russischen Telegram-Kanälen taucht ein gefälschtes Schweizer Plakat auf – und verbreitet sich im Internet. <https://www.nzz.ch/technologie/ein-gefaelschtes->

[schweizer-plakat-taucht-in-russischen-telegram-kanaelen-auf-und-verbreitet-sich-im-internet-ld.1702736](#)

- Martenet, V. & Dubey, J. (Hrsg.) (2021). Commentaire romand, Constitution fédérale. Bale 2021.
- Masmejan, D. (2020). Débat public en ligne et protection des libertés de communication. Etude réalisée sur mandat de l'Office fédéral de la communication. <https://rsf-ch.ch/wp-content/uploads/2020/06/Lire-l%C3%A9tude.pdf>
- Masmejan, D., Cottier, B., Capt, N. (Hrsg.) (2014). Loi sur la radio-télévision (LRTV), Bern 2014. Zit. LRTV-Verfasser:in
- Mayntz, R. (2004). Governance im modernen Staat. In A. Benz (Hrsg.), Governance—Regieren in komplexen Regelsystemen. Eine Einführung (S. 65–76). VS Verlag.
- McEvily, P. (2022). Energiespar-Kampagne ist fake – eine Spurensuche. <https://www.srf.ch/news/schweiz/aufruf-zur-denunziation-energiespar-kampagne-ist-fake-eine-spurensuche>
- McKay, S., & Tenove, C. (2021). Disinformation as a threat to deliberative democracy. *Political Research Quarterly*, 74(3), 703-717.
- McKee, M. (2017). The tobacco industry: the pioneer of fake news. *Journal of Public Health Research*, 6(1), 878.
- Mourão, R. R., & Robertson, C. T. (2019). Fake news as discursive integration: An analysis of sites that Publish false, misleading, hyperpartisan and sensational information. *Journalism Studies*, 20(14), 2077–2095. <https://doi.org/10.1080/1461670X.2019.1566871>
- Marchal, N., Kollanyi, B., Neudert, L.-M. & Howard, P. N. (2019). Junk News During the EU Parliamentary Elections: Lessons from a Seven-Language Study of Twitter and Facebook. Verfügbar unter [https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2017/09/ComProp\\_GermanElections\\_Sep2017v5.pdf](https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2017/09/ComProp_GermanElections_Sep2017v5.pdf)
- Mossburger, T. (2021). Alles über Telegram: The good, the bad & the ugly. <https://www.br.de/nachrichten/netzwelt/alles-ueber-telegram-the-good-the-bad-and-the-ugly,SdESFKQ>
- Müller, P.J. (1996). Abstinenz und Engagement des Strafrechts im Kampf gegen Ausländerfeindlichkeit, *AJP* 1996, 659 ff.
- Müller, P. J. (2018). Die Verwirklichung der Grundrechte nach Art. 35 BV: Der Freiheit Chancen geben. Bern 2018.
- Müller, P.J. (2020). V. Teil: Grund- und Menschenrechte: Entstehung und Entwicklung der Grundrechte in der Schweiz, in: Diggelmann O. & Hertig. M. & Schindler B., *Verfassungsrecht in der Schweiz Bd. II*, Zürich 2020.
- Müller, P.J. & Schefer, M. (2009). *Grundrechte in der Schweiz*, 4. Auflage, Bern 2009.

- Nachtwey, O., Schäfer, R., & Frei, N. (2020). Politische Soziologie der Corona-Proteste. <https://doi.org/10.31235/osf.io/zyp3f>
- Neudert, L.-M., Kollanyi, B. & Howard, P. N. (2017). Junk news and bots during the German parliamentary election: What are German voters sharing over Twitter? Verfügbar unter [https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2017/09/ComProp\\_GermanElections\\_Sep2017v5.pdf](https://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2017/09/ComProp_GermanElections_Sep2017v5.pdf)
- Newman, N., Fletcher, R., Kalogeropoulos, A., & Nielsen, R. (2019). Reuters institute digital news report 2019 (Vol. 2019): Reuters Institute for the Study of Journalism.
- Newman, E. J., & Zhang, L. (2020). Truthiness: How non-probative photos shape belief. In R. Greifeneder, M. E. Jaffé, E. J. Newman, & N. Schwarz (Eds.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 90–114). Routledge.
- Newton, C., Feeney, J., & Pennycook, G. (2023). On the disposition to think analytically: Four distinct intuitive-analytic thinking styles. *Personality & Social Psychology Bulletin*, 1461672231154886. Advance online publication. <https://doi.org/10.1177/01461672231154886>
- Niggli, A. & Wiprächtiger, H. (2018). Basler Kommentar zum Strafrecht, Basel 2018. Zit. BK StGB-Verfasser:in.
- OECD (2021), 21st-Century Readers: Developing Literacy Skills in a Digital World, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/a83d84cb-en>
- Oehmer, F. (2018). Meinungsfreiheit für Social Bots?. Jusletter vom 29. April 2019.
- Oehmer, F. & Pedrazzi, S. (2020). Was schützt (vor) Social Bots? Vorschläge zur Governance von computergenerierten Softwareagenten im Internet. UFITA: Archiv für Urheber- und Medienrecht, 84(1), 7-35.
- Oehmer, F., Udris, L., & Eisenegger, M. (2020). Zur Wirkung von Tech-Informations-Intermediären auf die (Schweizer) Öffentlichkeit – Chancen und Risiken. Bericht für das Bundesamt für Kommunikation. Zürich. [https://www.bakom.admin.ch/dam/bakom/de/dokumente/bakom/elektronische\\_medien/Zahlen%20und%20Fakten/Studien/bericht-chancen-risiken-intermediaere-2020.pdf.download.pdf/Bericht\\_Chancen\\_Risiken\\_Intermedia%CC%88re\\_310720\\_fin.pdf](https://www.bakom.admin.ch/dam/bakom/de/dokumente/bakom/elektronische_medien/Zahlen%20und%20Fakten/Studien/bericht-chancen-risiken-intermediaere-2020.pdf.download.pdf/Bericht_Chancen_Risiken_Intermedia%CC%88re_310720_fin.pdf)
- Peduzzi, R. (2004). Meinungs- und Medienfreiheit in der Schweiz, Zürich/Basel/Genf 2004.
- Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, 592(7855), 590–595. <https://doi.org/10.1038/s41586-021-03344-2>
- Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, 147(12), 1865–1880. <https://doi.org/10.1037/xge0000465>

- Pennycook, G., & Rand, D. G. (2019). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences of the United States of America*, 116(7), 2521–2526. <https://doi.org/10.1073/pnas.1806781116>
- Pennycook, G., & Rand, D. G. (2022). Accuracy prompts are a replicable and generalizable approach for reducing the spread of misinformation. *Nature Communications*, 13(1), 2333. <https://doi.org/10.1038/s41467-022-30073-5>
- Pereira, A., Harris, E., & Van Bavel, J. J. (2023). Identity concerns drive belief: The impact of partisan identity on the belief and dissemination of true and false news. *Group Processes & Intergroup Relations*, 26(1), 24–47. <https://doi.org/10.1177/13684302211030004>
- Puppis, M. (2010). Media governance: A new concept for the analysis of media policy and regulation. *Communication, Culture & Critique*, 3(2), 134–149. <https://doi.org/10.1111/j.1753-9137.2010.01063.x>
- Rathje, S., Roozenbeek, J., van Bavel, J. J., & van der Linden, S. (2023). Accuracy and social motivations shape judgements of (mis)information. *Nature Human Behaviour*. <https://doi.org/10.1038/s41562-023-01540-w>
- Rathje, S., van Bavel, J., & van der Linden, S. (2021, July 13). Why Facebook really, really doesn't want to discourage extremism. *The Washington Post*. <https://www.washingtonpost.com/politics/2021/07/13/why-facebook-really-really-doesnt-want-discourage-extremism/>
- Robertson, C. E., Pröllochs, N., Schwarzenegger, K., Pärnamets, P., van Bavel, J. J., & Feuerriegel, S. (2023). Negativity drives online news consumption. *Nature Human Behaviour*. <https://doi.org/10.1038/s41562-023-01538-4>
- Roozenbeek, J., van der Linden, S., Goldberg, B., Rathje, S., & Lewandowsky, S. (2022). Psychological inoculation improves resilience against misinformation on social media. *Science Advances*, 8(34), eabo6254. <https://doi.org/10.1126/sciadv.abo6254>
- Rauchfleisch, A., Vogler, D., & Eisenegger, M. (2020). Wie das Coronavirus die Schweizer Twitter-Communitys infizierte. *Jahrbuch Qualität der Medien*, 11, 61-75.
- Rauchfleisch, A. & Kaiser, J. (2020). The false positive problem of automatic bot detection in social science research. *PloS one*, 15(10), e0241045.
- Rauchfleisch, A., & Kaiser, J. (2020). The German far-right on YouTube: An analysis of user overlap and user comments. *Journal of Broadcasting & Electronic Media*, 64(3), 373-396.
- Reuter, M. (2019). Social Bots: Was nicht erkannt werden kann, sollte nicht reguliert werden. Abgerufen unter: <https://netzpolitik.org/2019/social-bots-was-nicht-erkannt-werden-kann-sollte-nicht-reguliert-werden/>
- Riedo, C. & Beglinger, R. (2022). Ehrverletzungsdelikte im Internet– insbesondere auf Facebook, *medialex* 03/22. <https://doi.org/10.52480/ml.22.08>.

- Rogers, R. (2020). Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media. *European Journal of Communication*, 35(3), 7662–7669. <https://doi.org/10.1177/0267323120922066>
- Rosenthal, D. (2013). Wie sich Privatpersonen gegen Verletzungen ihrer Persönlichkeits-Rechte durch Dritte auf Social Media Plattformen wehren können. Bericht im Auftrag des BAKOM. Abgerufen unter [https://www.bakom.admin.ch/dam/bakom/de/dokumente/informationsgesellschaft/verletzungen\\_persoelichkeitsrechteaufsocialmedia.pdf.download.pdf/verletzungen\\_persoelichkeitsrechteaufsocialmedia.pdf](https://www.bakom.admin.ch/dam/bakom/de/dokumente/informationsgesellschaft/verletzungen_persoelichkeitsrechteaufsocialmedia.pdf.download.pdf/verletzungen_persoelichkeitsrechteaufsocialmedia.pdf)
- Rossini, P., Stromer-Galley, J., Baptista, E. A., & Veiga de Oliveira, V. (2020). Dysfunctional information sharing on WhatsApp and Facebook: The role of political talk, cross-cutting exposure and social corrections. *New Media & Society*, 2430–2451. <https://doi.org/10.1177/1461444820928059>.
- Rossini, P., Baptista, É. A., de Oliveira, V. V., & Stromer-Galley, J. (2021). Digital media landscape in Brazil: Political (Mis) information and participation on Facebook and WhatsApp. *Journal of Quantitative Description: Digital Media*, 1. <https://doi.org/10.51685/jqd.2021.015>
- Sängerlaub, A., Meier, M. & Rühl, W.-D. (2018). Fakten statt Fakes: Das Phänomen «Fake News». Verursacher, Verbreitungswege und Wirkungen von Fake News im Bundestagswahlkampf 2017. Verfügbar unter [https://www.stiftung-nv.de/sites/default/Files/snv\\_fakten\\_statt\\_fakes.pdf](https://www.stiftung-nv.de/sites/default/Files/snv_fakten_statt_fakes.pdf)
- Sängerlaub, A. (2020) Fakten versus Fakes: Was Fact-Checking als Massnahme gegen Desinformation leisten kann und was nicht. In R. Hohlfeld, M. Harnischmacher, E. Heinke, M. Sengl, & L. S. Lehner (Hrsg.), *Fake News und Desinformation: Herausforderungen für die vernetzte Gesellschaft und die empirische Forschung* (S. 99–116). Nomos. <https://doi.org/10.5771/9783748901334-97>
- Sarmina, A. (2018). Fact-Checking as Defence Against Propaganda in the Digital Age. In O. Hahn, & F. Stalph (Hrsg.), *Digital Investigative Journalism* (S. 193-202). Palgrave Macmillan.
- Saurwein, F., & Spencer-Smith, C. (2020). Combating disinformation on social media: Multilevel governance and distributed accountability in Europe. *Digital Journalism*, 8(6), 820-841.
- Schefer, M. (2020). Teil: Grund- und Menschenrechte, in: Diggelmann O. & Hertig Randal M. & Schindler B., *Verfassungsrecht in der Schweiz Bd. II. Rechtsstaatlichkeit Grund- und Menschenrechte*, Zürich 2020.
- Schefer, M. & Cueni, R. (2020). Öffentlichkeit im Wandel: Überlegungen aus grundrechtlicher Sicht, Studie zuhanden des BAKOM, Juli 2020. Abgerufen unter:

- [https://www.bakom.admin.ch/dam/bakom/de/dokumente/bakom/elektronische\\_medien/Zahlen%20und%20Fakten/Studien/studie-oeffentlichkeit-im-wandel-2020.pdf.download.pdf/Schefer\\_Cueni\\_%C3%96ffentlichkeit-im-Wandel\\_August.pdf](https://www.bakom.admin.ch/dam/bakom/de/dokumente/bakom/elektronische_medien/Zahlen%20und%20Fakten/Studien/studie-oeffentlichkeit-im-wandel-2020.pdf.download.pdf/Schefer_Cueni_%C3%96ffentlichkeit-im-Wandel_August.pdf).
- Scheufele, D. A. & Krause, N. M. (2019). Science audiences, misinformation, and fake news. *Proceedings of the National Academy of Sciences of the United States of America*, 116(16), 7662-7669. <https://doi/10.1073/pnas.1805871115>
- Schnyder, A. (2014). Einführung in die Dogmatik der AGB nach schweizerischem und europäischem Recht. In: Brunner Alexander/Schnyder Anton K./Eisner-Kiefer Andrea (Hrsg.), *Allgemeine Geschäftsbedingungen nach neuem Schweizer Recht*, Zürich/Basel/Genf 2014.
- Schoop, F., Skinner, B., & Kelen, J. (2021). Das sind die 10 lautesten Schweizer Corona-Kritiker auf Telegram. <https://www.nzz.ch/gesellschaft/bei-corona-skeptikern-sind-gefuehle-wichtiger-als-fakten-ld.1603166>
- Schwaiger, Lisa. (2022). *Gegen die Öffentlichkeit*. Transcript Verlag.
- Shao, C., Ciampaglia, G. L., Varol, O., Yang, K., Flammini, A., & Menczer, F. (2018). The spread of low-credibility content by social bots. Abgerufen unter: <http://arxiv.org/pdf/1707.07592v4>
- Ständer, A., Humprecht, E., Esser, F., Morosoli, S., & Van Aelst, P. (2021). Is sensationalist disinformation more effective? Three facilitating factors at the national, individual, and situational level. *Digital Journalism*, 1-21.
- Stein, J., Frey, V., & van de Rijt, A. (2023). Realtime user ratings as a strategy for combatting misinformation: an experimental study. *Scientific Reports*, 13(1). <https://doi.org/10.1038/s41598-023-28597-x>
- Steinebach, M., Bader, K., Rinsdorf, L., Krämer, N., & Roßnagel, A. (2020). *Desinformation aufdecken und bekämpfen – Interdisziplinäre Ansätze gegen Desinformationskampagnen und für Meinungspluralität*, Baden-Baden.
- Stieglitz, S., Brachten, F., Ross, B., & Jung, A. K. (2017). Do social bots dream of electric sheep? A categorisation of social media bot accounts. arXiv preprint arXiv:1710.04044.
- Stier, S., & Jungherr, A. (2019). Digitale Verhaltensdaten und Methoden der Computational Social Science in der politischen Kommunikationsforschung. In: J. Hoffmann et al. (Hrsg.). *Politik in der digitalen Gesellschaft: zentrale Problemfelder und Forschungsperspektiven*. Transcript, S. 309-325.
- Swire-Thompson, B., DeGutis, J., & Lazer, D. (2020). Searching for the backfire effect: Measurement and design considerations. *Journal of Applied Research in Memory and Cognition*, 9(3), 286–299. <https://doi.org/10.1016/j.jarmac.2020.06.006>

- Tantuco, V. (2021). On Facebook's messaging apps, false information spreads undetected, unchecked. <https://www.disinfo.eu/publications/on-facebooks-private-messaging-apps-harmful-misinformation-spreads-largely-undetected-and-unchecked/>
- Thouvenin, F. (2017). Vergleichs- und Bewertungsdienste: eine Analyse aus Sicht des Wettbewerbsrechts (UWG). In: Thouvenin F. & Weber R.H. Werbung - Online. Zürich: Schulthess Verlag, 131-160.
- Trechsel S. & Pieth M. (2021). Praxiskommentar Schweizerisches Strafgesetzbuch. 4. Auflage Zürich 2021 (zit. PK StGB-Verfasser:in)
- Udris, L., Rivière, M., Vogler, D., & Eisenegger, M. (2022). Reuters Institute Digital News Report 2022 – Länderbericht Schweiz. Abgerufen unter: [https://www.foeg.uzh.ch/dam/jcr:04d547fc-0fc3-42fc-bbf8-cee04a06ebcf/DNR\\_22\\_Schweiz.pdf](https://www.foeg.uzh.ch/dam/jcr:04d547fc-0fc3-42fc-bbf8-cee04a06ebcf/DNR_22_Schweiz.pdf)
- Unkelbach, C., & Greifeneder, R. (2018). Experiential fluency and declarative advice jointly inform judgments of truth. *Journal of Experimental Social Psychology*, 79, 78–86. <https://doi.org/10.1016/j.jesp.2018.06.010>
- Urman, A., & Katz, S. (2022). What they do in the shadows: examining the far-right networks on Telegram. *Information, communication & society*, 25(7), 904-923.
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media+ Society*, 6(1), 2056305120903408.
- van Bavel, J. J., Packer, D. J., Ray, J. L., Robertson, C., & Ungson, N. (n.d.). *How social identity tunes moral cognition*. Preprint.
- van Bavel, J. J., & Pereira, A. (2018). The Partisan Brain: An Identity-Based Model of Political Belief. *Trends in Cognitive Sciences*, 22(3), 213–224. <https://doi.org/10.1016/j.tics.2018.01.004>
- van der Linden, S. (2023). Foolproof. Why misinformation infects our minds and how to build immunity. HarperCollins Publishers.
- Vargo, C. J., Guo, L., & Amazeen, M. A. (2017). The agenda-setting power of fake news: A big data analysis of the online media landscape from 2014 to 2016. *New Media & Society*, 20, 2028–2049. <https://doi/10.1177/1461444817712086>
- Valenzuela, S., Halpern, D., & Araneda, F. (2021). A Downward Spiral? A Panel Study of Misinformation and Media Trust in Chile. *The International Journal of Press/Politics*, 27(2), 353-373. <https://doi.org/10.1177/19401612211025238>
- Valenzuela, S., Halpern, D., Katz, J. E., & Miranda, J. P. (2019). The Paradox of Participation Versus Misinformation: Social Media, Political Engagement, and the Spread of Misinformation. *Digital Journalism*, 1-22. <https://doi/10.1080/21670811.2019.1623701>

- Vogler, D., & Rauchfleisch, A. (2018). #NoBillag auf Twitter: Grabenkämpfe zwischen Gegnern und Befürwortern.
- Vogler, D., Schwaiger, L., Schneider, J., Udris, L., Siegen, D., Marschlich, S., Rauchfleisch, A., Eisenegger, M. (2021). Falschinformationen, Alternativmedien und Verschwörungstheorien – Wie die Schweizer Bevölkerung mit Desinformation umgeht. Bericht für das Bundesamt für Kommunikation.
- Volz, S. (2013). Trennungsgebot und Internet: Ein medienrechtliches Gebot in Zeiten der Medienkonvergenz, Zürich 2013.
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151. <https://doi/10.1126/science.aap9559>
- Vraga, E. K., & Tully, M. (2021). News literacy, social media behaviors, and skepticism toward information on social media. *Information, Communication & Society*, 24(2), 150–166. <https://doi.org/10.1080/1369118X.2019.1637445>
- Waldmann, B., Belser, E.M. & Epiney, A. (Hrsg.) (2015). Bundesverfassung, Basler Kommentar, Basel 2015. Zit. BK BV-Verfasser:in
- Walker, C., O'Neill, S., & De-Wit, L. (2020). Evidence of Psychological Targeting but not Psychological Tailoring in Political Persuasion Around Brexit. *Experimental Results*, 1, E38. <https://doi.org/10.1017/exp.2020.43>
- Walther, S., & McCoy, A. (2021). US extremism on Telegram. *Perspectives on Terrorism*, 15(2), 100-124.
- Wardle, C. (2018). The need for smarter definitions and practical, timely empirical research on information disorder. *Digital Journalism*, 6(8), 951–963. <https://doi.org/10.1080/21670811.2018.1502047>
- Wardle, C., & Derakhshan, H. (2017). Information Disorder: Toward an interdisciplinary framework for research and policy making (DGI(2017)09; Council of Europe Report). Council of Europe report.
- Weber, R.H. (2008). Rundfunkrecht – Kommentar zum Bundesgesetz vom 24. März 2006 über Radio und Fernsehen, Bern 2008.
- Weber, R.H. & Nobel, P. (2021). Medienrecht, 4. Auflage, Zürich 2021.
- Wyss, P (2002). Appell und Abschreckung. Verfassungsrechtliche Beobachtungen zur Versammlungsfreiheit. *ZBI* 2002, S. 393 – 410.
- Zimmermann, F., & Kohring, M. (2020). Mistrust, disinforming news, and vote choice: A panel survey on the origins and consequences of believing disinformation in the 2017 German parliamentary election. *Political Communication*, 37(2), 215–237. <https://doi.org/10.1080/10584609>.

Zimmermann, F., & Kohring, M. (2018). "Fake News" als aktuelle Desinformation: Systematische Bestimmung eines heterogenen Begriffs. *Medien & Kommunikationswissenschaft: M & K*, 66(4), 526-541.

## Materialienverzeichnis

### Europäische Union

Gemeinsame Mitteilung an das Europäische Parlament, den Europäischen Rat, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen, Aktionsplan gegen Desinformation, JOIN(2018) 36 final («Aktionsplan Desinformation»)

Gemeinsame Mitteilung an das Europäische Parlament, den Europäischen Rat, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen, Bericht über die Umsetzung des Aktionsplans gegens Desinformation, JOIN(2019) 12 final, <https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:52019JC0012&from=EN#footnoteref8>, zuletzt besucht am 29. Juni 2023

Verhaltenskodex für Desinformation 2022, (Strengthened Code of Practice on Desinformation), abrufbar unter <https://digital-strategy.ec.europa.eu/en/library/2022-strengthened-code-practice-disinformation> , zuletzt besucht am 29. Juni 2023

Verordnung des Europäischen Parlaments und des Rates über einen Binnenmarkt für digitale Dienste (Gesetz über digitale Dienste, Digital Services Act und zur Änderung RL 2000/31/EG, COM (2020) 825 final)

### Reglemente und Leitlinien von Selbstregulierungsorganen

Geschäftsreglement der schweizerischen Lauterkeitskommission («Geschäftsreglement SLK»), abrufbar unter <https://www.faire-werbung.ch/wp-content/uploads/2021/09/SLK-Geschaeftsreglement-DE-1.1.2019.pdf>, zuletzt besucht 29. Juni 2022

Geschäftsreglement des Presserates («Geschäftsreglement PR»), abrufbar unter [www.presserat.ch/wp-content/uploads/2018/01/Geschäftsreglement-Januar-2018.pdf](http://www.presserat.ch/wp-content/uploads/2018/01/Geschäftsreglement-Januar-2018.pdf), zuletzt besucht am 29. Juni 2023

Erklärung der Pflichten und Rechte der Journalistinnen und Journalisten, («Erklärung»), abrufbar unter [www.presserat.ch/journalistenkodex/richtlinien/](http://www.presserat.ch/journalistenkodex/richtlinien/), zuletzt besucht am 29. Juni 2029

### Verzeichnis der Selbstregulierung

#### Facebook

Facebook Gemeinschaftsstandards, abrufbar unter <https://transparency.fb.com/de-de/policies/community-standards/?source=https%3A%2F%2Fwww.facebook.com%2Fcommunitystandards%2F%3Frefsrc%3Dabout.facebook.com%252Factions%252Fpromoting-safety-and-expression%252F> zuletzt besucht am 29. Juni 2023

## Instagram

Gemeinschaftsrichtlinien, abrufbar unter <https://help.instagram.com/477434105621119/?refsrc=about.facebook.com%2Factions%2Fpromoting-safety-and-expression%2F> zuletzt besucht am 29. Juni 2023.

## Youtube

Youtube Community Guidelines, abrufbar unter <https://www.youtube.com/howyoutubeworks/policies/community-guidelines/> zuletzt besucht am 29. Juni 2023.

## Whatsapp

WhatsApp Nutzungsbedingungen, Zulässige Nutzung unserer Dienste, Stand 4. Januar 2021, abrufbar unter <https://www.whatsapp.com/legal/terms-of-service-eea#terms-of-service-acceptable-use-of-our-services> zuletzt besucht, 29. Juni 2023.

## Threema

Threema Broadcast: Nutzungsbedingungen, Stand 1. März 2022, 4. Nutzung des Vertragsgegenstandes, <https://broadcast.threema.ch/de/nutzungsbedingungen>, zuletzt besucht am 29. Juni 2023

## Telegram

Telegram Privacy Policy, abrufbar unter <https://telegram.org/privacy>  
Telegram Terms of Service, abrufbar unter <https://telegram.org/tos/terms-of-service-for-telegram>, zuletzt besucht am 29. Juni 2023

## TikTok

Community-Richtlinien, Stand März 2023, abrufbar unter <https://www.tiktok.com/community-guidelines/de-de/overview/>, zuletzt besucht am 29. Juni 2023