NJ Neue Justiz

neue-justiz.nomos.de

Beilage 1 | 2024
Jahrgang 78
Seiten B1–B40

Zeitschrift für Anwalts- und Gerichtspraxis

Prof. Dr. Johannes Weberling (Hrsg.)

20. Frankfurter Medienrechtstage 2024

„Strategien gegen Desinformation und Propaganda“

Desinformation und Propaganda haben das Potenzial, das Vertrauen in staatliche Einrichtungen zu zerstören, die gesellschaftliche Trennung zu verschärfen und politisches Engagement zu behindern. In den letzten Jahren haben die COVID-19-Pandemie und der russische Angriffskrieg in der Ukraine die Verbreitung von Desinformation verstärkt. Als Reaktion auf diese Herausforderung haben mehrere Staaten Strategien entwickelt, um dieses Problem anzugehen. Die Europäische Union hat wichtige Vorschriften, darunter den Digital Services Act, eingeführt und arbeitet aktiv an weiteren Initiativen wie dem Media Freedom Act. Die demokratischen Gesellschaften müssen Strategien entwickeln, um die Verbreitung von Desinformation und Propaganda zu bekämpfen und gleichzeitig die Grundsätze der Meinungsfreiheit und des unabhängigen Journalismus zu wahren.

Bei den 20. Frankfurter Medienrechtstagen vom 17. – 18. Januar 2024 an der Europa-Universität Viadrina Frankfurt (Oder) diskutierten deshalb auf Einladung des Studien- und Forschungsschwerpunkts Medienrecht der Juristischen Fakultät der Europa-Universität Viadrina, der Konrad-Adenauer-Stiftung und der Südosteuropa-Gesellschaft mit Unterstützung der Märkischen Oderzeitung 70 Journalisten, Vertreter von Nichtregierungsorganisationen sowie Forscher und Studierende Strategien gegen Desinformation und Propaganda. Resultat der Vorträge und Diskussionen war, dass Desinformation hat zwar mit dem völkerrechtswidrigen Angriffskrieg der russischen Föderation auf die Ukraine seit dem 24. Februar 2022 stark zugenommen habe, man sich aber schon seit der ebenso völkerrechtswidrigen Annexion der Krim durch Russland im Jahr 2014 in einem hybriden Informationskrieg befinde. Insbesondere vor dem Hintergrund des Superwahljahrs 2024 sei Desinformation eine ernstzunehmende Gefahr für die Demokratie und die Gesellschaft, da Desinformation das Vertrauen der Menschen als zentrale Voraussetzung für ein konstruktives Zusammenleben in einer Gesellschaft angreife.

Die 20. Frankfurter Medienrechtstage haben aus internationaler und interdisziplinärer Perspektive die ernstzunehmenden Gefahren von Desinformation identifiziert, zugleich aber interdisziplinäre und moderne Strategien aufgezeigt, durch deren Implementierung Desinformation und Propaganda wirksam bekämpft werden könnten. Um diesen vielfältigen Vorschlägen und Anregungen eine breitere und nachhaltigere Öffentlichkeit zu verschaffen, werden neben einem ausführlichen Bericht über den Verlauf und die wesentlichen Resultate der 20. Frankfurter Medienrechtstage die dort gehaltenen Vorträge in dieser Beilage, die der NJ 10/2024 beigefügt wird, veröffentlicht.

Nomos

Impressum

Neue Justiz: Zeitschrift für Anwalts- und Gerichtspraxis (NJ)
ISSN 0028-3231

Redaktion:
RA Prof. Dr. Johannes Weberling (V.i.S.d.P.), RA Carsten Herlitz, RA Dr. Malte Nieschalk, RAin Dr. Katrin Raabe, Susanne Weberling M.A.

Verlag und Gesamtverantwortung für Druck und Herstellung:
Nomos Verlagsgesellschaft mbH & Co. KG
Waldseestr. 3-5
76530 Baden-Baden
Telefon: 07221/2104-0
Telefax 07221/2104-27
www.nomos.de

Geschäftsführer: Thomas Gottlöber
HRA 200026, Mannheim

Beilage 1/24 78. Jahrgang, Seiten B 1 - B 40

NJ Neue Justiz

Zeitschrift für Anwalts- und Gerichtspraxis

20. Frankfurter Medienrechtstage 2024 „Strategien gegen Desinformation und Propaganda“

Martin Stöhr, Frankfurt (Oder)*

Am 17. und 18. Januar 2024 fanden an der Europa-Universität Viadrina in Frankfurt (Oder) die 20. Frankfurter Medienrechtstage zum Thema „Strategien gegen Desinformation und Propaganda“ statt, die vom Studien- und Forschungsschwerpunkt Medienrecht der Juristischen Fakultät der Europa-Universität Viadrina in bewährter Zusammenarbeit mit der Konrad-Adenauer-Stiftung und der Südosteuropa-Gesellschaft und mit der Unterstützung der „Märkischen Oderzeitung“ veranstaltet wurden. *Prof. Dr. Claudia Weber*, Leiterin der Zweigstelle Frankfurt (Oder) der Südosteuropa-Gesellschaft e.V., *Prof. Dr. Eduard Mühle*, Präsident der Europa-Universität Viadrina und *Christoph Plate*, Leiter des Medienprogramms Südosteuropa der Konrad-Adenauer-Stiftung e.V. eröffneten die Frankfurter Medienrechtstage, hoben die Aktualität des Themas hervor und wünschten den Teilnehmern einen produktiven Austausch und konstruktive Diskussionen. Desinformation und das Ziel ihrer Emittenten, politische Wahlentscheidungen zu beeinflussen, ist vor dem Hintergrund des Superwahljahres 2024, inklusive drei deutscher Landtagswahlen und der Wahl des Europäischen Parlaments, aktueller denn je. Es bedarf einer sorgfältigen wissenschaftlichen Betrachtung von Desinformation und Strategien gegen eben jene. Hierfür wurden hochkarätige Experten und Wissenschaftler aus Afrika, Südosteuropa, England und Deutschland aus den Bereichen Rechtswissenschaft, Psychologie, Wirtschaft, Journalismus und Politik geladen.

Dr. Pavel Usvatov, Leiter des Rechtsstaatsprogramms Südosteuropa der Konrad-Adenauer-Stiftung mit Sitz in Bukarest, begann mit einer juristischen Betrachtung der Desinformation und Regulierung von Medien. So unterschied er die Desinformation von der Fehl-/Misinformation anhand des subjektiven Elements des Täuschungsvorsatzes.[1]

Misinformation sei laut *Dr. Usvatov* eine falsche, fehlerhafte oder dekontextualisierte Information, die jedoch ohne vorsätzliche Täuschungs- oder Irreführungsabsicht verbreitet werde. Als Beispiel wurde die Krisenberichterstattung aufgeführt, in der es den Journalisten aufgrund von Feuergefechten, abiotischer Umweltumstände oder eingeschränkter Kommunikationsmöglichkeiten häufig nicht möglich sei,

* Der Autor ist cand. jur und Absolvent des Schwerpunktbereichs 7 „Medienrecht“ der Juristischen Fakultät der Europa-Universität Viadrina Frankfurt (Oder).

1 Vgl. *Usvatov*, Wahrheit, Vertrauen und Transparenz: Medien und ihre Regulierung, NJ 2024, B 5.

Informationen sofort und umfangreich zu verifizieren, den Gesamtkontext zu erschließen oder die Quellen genauestens zu überprüfen, was dazu führen könne, dass unabsichtlich ungenaue Informationen veröffentlicht werden würden. Dies sei bei Krisenberichterstattung nicht völlig vermeidbar.

Im Gegensatz zu der Fehl-/Misinformation stelle Desinformation die absichtliche Verbreitung falscher oder irreführender Inhalte dar, um die Rezipienten bewusst zu täuschen.

Dr. Usvatov erläuterte die Vielfältigkeit und Komplexität von Desinformation anhand von fünf Erscheinungsformen:

1. Die intendierte Dekontextualisierung, bei der eine korrekte Information in einem inkorrekten Kontext präsentiert wird, wodurch die Wirkung der Information beeinflusst werden soll.
2. Die bewusste Verbreitung falscher Informationen, bei der absichtlich vollständig unwahre und fiktive Informationen verbreitet werden.
3. Die manipulative politische Werbung, die strategisch genutzt wird, um gezielt Einfluss auf das Wahlverhalten zu nehmen.
4. Pseudojournalismus, der vorliegt, wenn sich der Pseudojournalist nicht an die journalistischen Sorgfaltspflichten hält oder bei der Präsentation desinformativer Inhalte etablierte Nachrichtenquellen imitiert werden, um die vermeintlich Glaubwürdigkeit der Inhalte künstlich zu verstärken.
5. Propaganda, bei der staatliche Akteure die eigene oder ausländische Bevölkerung strategisch manipulieren, um politische Vorteile zu generieren.

Anschließend thematisierte *Usvatov* die Regulierung von Desinformation und hob hervor, dass die Grenzen der Regulierung in Art. 5 GG lägen. Art. 5 GG sei mit seiner verfassungsrechtlichen Garantie, nach der jeder seine Meinung frei äußern und sich unterschiedlicher Medien bedienen darf, eines der höchsten Güter der Verfassung. Die essenzielle Bedeutung der Meinungs- und Medienfreiheit habe Deutschland nach zwei Diktaturen auf deutschem Boden auf dem harten Weg lernen müssen. Speziell vor dem historischen Hintergrund starker Restriktion der Meinungs- und Pressefreiheit zur Zeit des Dritten Reichs und der DDR müsse eine demokratische Gesellschaft auch schwierige Diskussionen und zu einem gewissen Grad auch Desinformation und Misinformation aushalten.

Der Referent erläuterte die subsidiäre Staffelung der präventiven Maßnahmen gegen Desinformation im deutschen Recht.

Das mildeste Mittel seien die Transparenzvorgaben wie z.B. die Impressumspflicht nach § 8 BbgPG, § 5 TMG und § 18 MStV.

Die Korrektur von desinformativen Inhalten stelle die nächste präventive Maßnahme mit höherer Eingriffsintensität dar, welche aus unterschiedlichen Rechtsquellen wie aus der Selbstverpflichtung des Pressekodex, vor allem aber aus den in den Landespressegesetzen und dem Medienstaatsvertrages geregelten Gegendarstellungsansprüchen folge.

Die Tilgung der Desinformation stelle die ultima ratio der präventiven Maßnahmen dar, wofür wiederum verschiedene Rechtsgrundlagen zur Verfügung stünden wie z.B. der quasinegatorische Unterlassungsanspruch nach §§ 823, 1004 BGB analog oder der wettbewerbsrechtliche Beseitigungs- und Unterlassungsanspruch nach § 8 UWG. Zudem ermögliche § 109 MStV, dass die zuständige Landesmedienanstalt ein Angebot bei Verstößen gegen die journalistische Sorgfaltspflicht vorübergehend oder dauerhaft untersagen könnte.

Dr. Usvatov führte aus, dass es als eingriffsintensivste repressive Reaktion auf desinformative Äußerungen strafrechtliche Normen gäbe, obwohl kein allgemeiner Straftatbestand für die Verbreitung von Desinformationen existiere. Normen des Strafrechts, die als allgemeines Gesetz gem. Art. 5 II GG die Kommunikationsgrundrechte nach Art. 5 I GG beschränkten, sind z.B. die Strafbarkeit der üblen Nachrede gem. § 186 StGB, der Verleumdung nach § 187 StGB, des Betrug gem. § 263 StGB, der Volksverhetzung im Sinne der Leugnung des Holocaust gem. § 130 III StGB oder der Störung des öffentlichen Friedens durch das Vortäuschen des Vorliegens einer Straftat nach § 126 II StGB.

Im zweiten Panel, welches globale Desinformations- und Propagandastrategien thematisierte, stellten die Referenten heraus, dass Russland und China die beiden Hauptakteure globaler Desinformations- und Propagandakampagnen seien.

Prof. Justine Limpitlaw von der Universität Witwatersrand (Johannesburg, Südafrika) stellte Desinformations- und Propagandastrategien auf dem afrikanischen Kontinent vor.[2]

Russland fokussiere sich auf die Einflussnahme südafrikanischer Medien und die Verbreitung russischer Narrative, da Südafrika eine zentrale politische Figur auf und für den afrikanischen Kontinent sei und Russland sich durch den Einfluss in Südafrika mehr Macht auf dem gesamten Kontinent verspreche. Die Agenda sei klar: Russland wolle Südafrika weg von einer liberalen, rechtsstaatlichen Demokratie hin zu einem autoritären Staat führen, es näher an die BRICS-Staaten und insbesondere Russland binden und mehr vom Westen lösen, der der Haupthandelspartner Südafrikas sei.

Allerdings nehme China noch größeren Einfluss auf Afrika. Dies erfolge laut *Limpitlaw* zum einen mit der China typischen Strategie der Kreditgewährung für teure Projekte, die derart konzipiert seien, dass sie bei Fälligkeit kaum beglichen werden können und infol-

2 Vgl. *Limpitlaw*, Navigieren in der Rechtslandschaft: Die Bedeutung der vergleichenden Analyse im Medienrecht der Länder südlich der Sahara, NJ 2024, B 11.

gedessen China Eigentum an wichtiger Infrastruktur und damit Einfluss in fremden Ländern gewinne. Auf diese Weise erwarb China beispielsweise im Zuge der teuren Transformation von analogem zu digitalem Fernsehen in Sambia Eigentum an einem öffentlich-rechtlichen Sender, was China großen Einfluss und Kontrolle in Sambia sichere.

Zur Erweiterung der medialen Einflusssphäre Chinas in Afrika verbreite China eigene Medien, insbesondere Fernsehsender in Afrika und stelle sie im Gegensatz zu europäischen Sendern kostenlos zur Verfügung, was natürlich dazu führe, dass der kostenlose chinesische Medieninhalt inklusive Propaganda vermehrt konsumiert werde. Auch stellen chinesische Nachrichtenagenturen Nachrichten für afrikanische Sender kostenlos zur Verfügung.

Justine Limpitlaw verdeutlichte damit, wie wichtig die finanzielle Unabhängigkeit der Medien für die Unabhängigkeit der Berichterstattung ist und adressierte das für China typische Vorgehen, die prekäre finanzielle Situation für die Generierung eigener politischer Vorteile auszunutzen.

Ferdinand Gehringer, politischer Berater für interne Sicherheit und Cybersicherheit der Konrad-Adenauer-Stiftung, lenkte den Blick auf die Desinformations- und Propagandastrategien von Russland und China, die Deutschland und die Europäische Union betreffen und hielt fest, dass weder Deutschland noch die EU derzeit ausreichend darauf vorbereitet seien, die langfristig angelegten Desinformationskampagnen seitens der autoritären Regime zu bekämpfen.[3] Er betonte, dass Russland und China bereits langjährig den öffentlichen Diskurs infiltrieren und ihre Narrative verbreiten. *Gehringer* verdeutlichte, dass das Hauptziel der russischen und chinesischen Desinformation in Europa darin bestehe, die freiheitlichen Demokratien zu destabilisieren. Dies geschehe, indem sie durch gezielte und manipulative Meinungsbeeinflussung verschiedener gesellschaftlicher und politischer Gruppen, durch die Verbreitung ihrer Narrative und die Skizzierung einer unwahren, „alternativen“ Realität, die Polarisierung und die Spaltung der demokratischen Gesellschaft forcieren. Zudem soll die faktenbasierte Realität als Fundament der Gesellschaft geschwächt werden. Ein wichtiges Instrument hierfür sei die gezielte Meinungsmanipulation in den sozialen Medien durch angeheuerte menschliche Internetkommentatoren oder „Social Bots“.

Die Inszenierung eines bestimmten dominanten Meinungsklimas stelle die optimale Voraussetzung für Meinungsbeeinflussung dar. Aus der Theorie der „Schweigespirale“ folge, dass Menschen aus Angst vor sozialer Ausgrenzung dazu tendieren, ihre eigenen Meinungen und Überzeugungen zu verschweigen oder sie gar anzupassen, wenn diese von der als dominierend wahrgenommenen öffentlichen Meinung abweichen.

In der von *Christoph Plate* moderierten anschließenden Diskussion stellten beide Referenten heraus, dass im Kampf gegen Desinformation zunächst die Medienkompetenz und die Kommunikation bezüglich der Gefahr von Desinformation gefördert werden solle und die Bildung und Aufklärung der Gesellschaft noch vor der rechtlichen Regulierung, die zweifelsohne wichtig für die Bekämpfung von Desinformation sei, erfolgen solle.

Im Zuge dessen hob *Ferdinand Gehringer* eine Strategie Estlands im Kampf gegen Desinformation hervor. Die Strategie bestünde darin, dass das estnische Verteidigungsministerium jeden Freitag die Bevölkerung über aktuelle Gefährdungslagen informiert. Dazu zählen auch russische Desinformationskampagnen, die seit dem Beginn des russischen Angriffskriegs auf die Ukraine zugenommen haben. Die regelmäßige Kommunikation und Aufklärung über die Gefahr von Desinformation sei ein Schlüsselelement für die Bekämpfung eben jener.

Gleicher Auffassung war *Dr. Rakoen Maertens*, Universität Oxford, der im nächsten Panel die psychologische Wirkung von Desinformation beleuchtete.[4] *Dr. Maertens* erklärte, dass nach der „Inokulationstheorie“ Menschen durch Aufklärung und Kommunikation psychologisch gegen Desinformation „geimpft“ werden könnten. Zum einen durch die möglichst frühe und konkrete Warnung vor aufkommender Desinformation und zum anderen durch die Konfrontation mit „abgeschwächter“ Desinformation, die aber sofort widerlegt würde, wodurch ein Bewusstsein für die eigene Anfälligkeit von Desinformation geschaffen werde und dadurch „mentale Antikörper“ gebildet werden würden.

Dr. Maertens erläuterte aus psychologischer Perspektive, warum sich Desinformation schneller verbreitet als faktenbasierte Information. Dies läge an der intendierten emotional mobilisierenden und manipulativen Gestaltung von Desinformation und der menschlichen Natur. Desinformation verbreite sich schneller als die Wahrheit, da Menschen dazu tendieren, sich eher mit negativen Inhalten zu beschäftigen. Man habe einen evolutionären Vorteil, wenn man sich über potenzielle Gefahren informiere. Desinformation mit emotional mobilisierendem Inhalt ziehe mehr Aufmerksamkeit auf sich und verbreite sich somit schneller.

Vor dem höchst aktuellen Hintergrund digitaler Desinformation und der Dynamik ihrer Verbreitung erklärte *Maertens* zudem das Phänomen des „Illusory Truth Effect[5]“, nach dem Rezipienten Aussagen, auch

3 Vgl. *Gehringer*, Einblicke in globale Desinformations- und Propagandastrategien, NJ 2024, B 18.

4 Vgl. *Maertens*, Die Psychologie der Misinformation - Verstehen und Abwehren von Misinformationen durch eine verhaltensorientierte Sichtweise und Künstliche Intelligenz, NJ 2024, B 24.

5 Deutsch: Wahrheitseffekt.

wenn es sich dabei um Desinformation handele, mehr Glaubwürdigkeit beimessen, wenn sie diese wiederholt wahrnehmen. Die rapide und massive Verbreitung von Desinformation, die durch KI möglich ist, verstärke diesen Effekt.

Der nachfolgende Referent *Dr. Christopher Nehring*, Gastdozent für Desinformation, Medien und Geheimdienste des Medienprogramms Südosteuropa der Konrad-Adenauer-Stiftung an der Universität Sofia, hob durch den Hinweis darauf, dass das Weltwirtschaftsforum Desinformation durch KI als größtes globales Risiko eingestuft hat, die Aktualität und Relevanz der von den Frankfurter Medienrechtstagen abgedeckten Themen hervor und thematisierte die Gefahren, die von KI in Bezug auf Desinformationen ausgingen.[6] So stellte *Nehring* dar, dass durch große generative KI-Modelle Desinformationen schneller, einfacher und kostengünstiger erstellt werden können und perspektivisch deshalb eine quantitative Zunahme von Desinformation sicher sei. Zudem gäbe es KI-Modelle, die die Qualität von desinformativen Inhalten stark verbessert haben. Er verwies hierzu auf die täuschend echt wirkenden Bild-, Video- und Audio-Deepfakes und beschrieb anhand von aktuellen Beispielen die Gefahr, die durch ihre realitätsverzerrende Wirkung für den politischen Meinungsbildungsprozess bestehe. Der Referent skizzierte das Gefahrenpotenzial von KI-Modellen, welches aufgrund des bahnbrechenden technischen Potenzials der Technologie bis hin zu einem Zustand führen könne, indem nicht mehr zwischen der reellen Abbildung der Realität und völlig synthetisch generierten und fiktiven Szenerien unterschieden werden könne.

In der anschließenden von *Ralitsa Stoycheva*, wissenschaftliche Mitarbeiterin des Medienprogramms Südosteuropa der Konrad-Adenauer-Stiftung in Sofia, moderierten Diskussion, wurde die Perspektive auf die KI komplettiert, indem auch ihr Potenzial für die Desinformationsbekämpfung erörtert wurde. So erklärte *Prof. Dr. Martin Grothe*, Geschäftsführer der complexium GmbH, dass die Fähigkeit der KI, riesige Datensätze auswerten und analysieren zu können, eine entscheidende Rolle für „Digital Listening“ spiele.[7] Dabei werden soziale Medien oder Internetplattformen gescreent und Aktivitäten, Diskussionen und Stimmungen analysiert, um Gefahren und Sicherheitsrisiken aus dem digitalen Datenstrom herauszufiltern, sie als solche frühzeitig zu identifizieren, um Desinformationskampagnen oder Anschläge wirksam verhindern zu können.

Dr. Maertens ergänzte, dass KI wissenschaftliche Informationen verständlich vermitteln und somit einen wertvollen Beitrag für die Gesellschaft und die Bekämpfung von Desinformation leisten könne.

Während die Referenten die Chancen und Risiken von KI durchaus unterschiedlich gewichteten, bestand Konsens darüber, wie wichtig qualitativ hochwertiger und seriöser Journalismus in Zeiten von strategischen Desinformationskampagnen sei. Dieser sichere die freie Meinungsbildung und manipulationsfreie Wahlentscheidungen der Bürger als Souverän, was die unentbehrliche Schlüsselrolle freier Qualitätsmedien für eine lebendige Demokratie belege.

Als Vertreter des freien Qualitätsjournalismus berichteten *Claus Liesegang*, Chefredakteur der „Märkische Oderzeitung“, und die stellvertretende Chefredakteurin der „Märkische Oderzeitung“ *Beate Bias* über ihre praktische Tätigkeit in der Redaktion und erklärten den Prozess hinter der journalistischen Entlarvung von Desinformation.

Essenzieller Teil ihrer Strategie gegen Desinformation sei es, dass von jedem der angestellten Journalisten während der Recherche Faktenchecks durchgeführt werden und die journalistische Arbeit durch die strenge Einhaltung der journalistischen Sorgfaltspflichten geprägt sei. Hohe Standards an die journalistische Sorgfalt führten zur Qualität des Mediums. Zudem setze die „Märkische Oderzeitung“ im Vergleich zu anderen Medienhäusern auf eine hundertprozentige Festanstellung der beschäftigten Journalisten und führe ständig Schulungen bezüglich der Entlarvung von Desinformation, der Wahlprogramme und dem politischen Geschehen durch, um insbesondere populistische Desinformation schnell und belegbar aufdecken zu können.

Am zweiten Tag der Frankfurter Medienrechtstagen skizzierten *Alexander Kachamov* aus Bulgarien, *Codruta Simina* aus Rumänien, *Orest Dabija* aus der Republik Moldau, *Hyrije Mehmeti* aus dem Kosovo und *Dragan Sekulovski* aus Nordmazedonien in zwei von *Prof. Dr. Claudia Weber* und Rechtsanwalt *Emil Georgiev* aus Sofia moderierten Panels die Situation von Desinformation und ihrer Bekämpfung in der Medienlandschaft Südosteuropas. Sie schilderten, dass die südosteuropäischen Länder aufgrund ihrer geopolitischen Lage primär von russischen Desinformationskampagnen attackiert werden würden, die darauf abzielen, pro-russische Narrative zu fördern und die russische Einflusssphäre in der Region zu festigen.

Zudem berichteten die Referenten darüber, dass sich die südosteuropäische Medienlandschaft zwar von gelenkten Staatsmedien weg hin zu einem deutlich liberalisierten Mediensystem entwickele, es jedoch immer noch äußerst einflussreiche, regierungsnahe und unkritische Medien gäbe, weshalb die Lage der Medien- und Pressefreiheit in Südosteuropa unverändert als prekär einzustufen sei. Dieser Umstand erschwere die Umsetzung einer in den Panels diskutierten Strategie gegen Desinformation, die aus Identifikation, Analyse und Kommunikation von Desinformation bestehe.

---

6 Vgl. Nehring / ŁAbuz, Deepfakes und KI-Manipulationen in Demokratie und Recht: Gefahren und Lösungen, NJ 2024, B 30.

7 Vgl. Grothe, Desinformation und Digital Listening, NJ 2024, B 35.

Prof. Dr. Johannes Weberling, Koordinator des Schwerpunktbereichs Medienrecht an der Europa-Universität Viadrina Frankfurt (Oder) und Initiator der nun seit 22 Jahren stattfindenden Frankfurter Medienrechtstage fasste am Ende der Tagung als Resümee zusammen, dass gezielte Desinformation zwar mit dem völkerrechtswidrigen Angriffskrieg der russischen Föderation auf die Ukraine am 24. Februar 2022 stark zugenommen habe, man sich aber schon seit der ebenso völkerrechtswidrigen Annexion der Krim durch Russland im Jahr 2014 in einem hybriden Informationskrieg befinde. Insbesondere vor dem Hintergrund des Superwahljahrs 2024 sei Desinformation eine ernstzunehmende Gefahr für die Demokratie und die Gesellschaft, da Desinformation das Vertrauen als zentrale Voraussetzung für ein konstruktives Zusammenleben in einer Gesellschaft attackiere. Das Vertrauen in eine faktenbasierte Realität als Konsens des gesellschaftlichen Diskurses, Vertrauen in die Medien, Vertrauen in die gesellschaftliche Kommunikation und auch das Vertrauen in die Politik, welches jedoch nie völlig vorbehaltlos und unkritisch sein dürfe. Die Strategien gegen Desinformation müssten wie Desinformation an sich vielseitig und interdisziplinär sein. So plädierte *Weberling* für die Förderung von Medienkompetenz an Schulen und an eine offene, aufgeklärte und direkte Kommunikation des Problems seitens der Politik. Er appellierte aber auch an die Eigenverantwortung der Zivilgesellschaft, Desinformation und Propaganda entschieden entgegenzutreten. Es zeige sich wieder, welch konstituierende Rolle unabhängiger, qualitativer und sorgfältiger Journalismus für die Demokratie habe und dieser durch entsprechende Fortbildungen und Festanstellungen gesichert und gestärkt werden müsse. Auch gesetzliche Regulierungen bezüglich digitaler Desinformationen seien wichtig, jedoch hätten die Vorträge und Diskussionen im Rahmen 20. Frankfurter Medienrechtstage verdeutlicht, dass das dynamische und interdisziplinäre Problem der Desinformation nicht wirksam durch Regulierung oder gar Überregulierung gelöst werden könne.

Die 20. Frankfurter Medienrechtstage haben aus internationaler und interdisziplinärer Perspektive umfangreich die ernstzunehmenden Gefahren von Desinformation identifiziert und zugleich interdisziplinäre und moderne Strategien aufgezeigt, durch deren Implementierung Desinformation und Propaganda wirksam bekämpft werden können.

Wahrheit, Vertrauen und Transparenz: Medien und ihre Regulierung

Dr. Pavel Usvatov, Bukarest*

„… der Französische Botschafter [… hat] an Seine Majestät den König [F.W. v. Preußen] noch die Forderung gestellt, …, dass S. Maj. der König sich für alle Zukunft verpflichte, niemals wieder seine Zustimmung zu geben [zur Kandidatur eines Hohenzollern auf den spanischen Thron]. Seine Maj. der König hat es darauf abgelehnt, den Franz. Botschafter nochmals zu empfangen, und demselben durch den Adjutanten vom Dienst sagen lassen, dass S. Majestät dem Botschafter nichts weiter mitzuteilen habe."

Norddeutsche Allgemeine Zeitung v. 13. Juli 1870 (sog. „Emser Depeche")

Wie das Beispiel aus dem Epigraph zeigt, begann die Instrumentalisierung der Medien nicht erst mit dem Eintreten in das Zeitalter des Internets und der „social media". Auch Bismarck nutzte die Macht der Presse für eigene Zwecke, in diesem Fall zur Provokation Frankreichs durch Veröffentlichung des sehr stark verkürzten Telegramms aus Ems. Die Übersetzung des Wortes „Forderung" als eine bloße Bitte des französischen Botschafters, «il a exigé», und der Amtsbezeichnung „Adjutant vom Dienst" (Oberleutnant oder Major) als «adjutant» (franz. Hauptfeldwebel) durch die französische Nachrichtenagentur Havas verstärkte nochmals die Wirkung in Frankreich. Seine Öffentlichkeit empfand diese nun sehr undiplomatisch klingende Zurückweisung der französischen „Bitte" durch Wilhelm als Demütigung und Ehrverletzung. Am 19. Juli 1870 reagierte Frankreich mit einer Kriegserklärung an Preußen und den Norddeutschen Bund.[1]

An diesem Beispiel lässt sich erkennen, welch weitreichenden Auswirkungen der Missbrauch der Presse und der Medien allgemein auf die Bevölkerung und die öffentliche Meinung haben kann. Im Zeitalter der Informationsgesellschaft begegnen wir täglich einer massiven Informationsflut, die nicht mehr jeder ohne weiters bewältigen kann. Insbesondere in den sog. „Sozialen Medien" ist der Anteil ungenauer, unwahrer, manipulierter und irreführender Informationen in

* Der Autor ist Leiter des Rechtsstaatsprogramms Südosteuropa (RSP SOE) der Konrad-Adenauer-Stiftung e.V. (KAS) mit Sitz in Bukarest.

1 Für dieses Beispiel spielt es nur eine nachgeordnete Rolle, dass Frankreich diesen Anlass nur zu bereitwillig aufgriff.

den letzten Jahren stetig gewachsen. Solche Informationen verbreiten sich um ein Vielfaches schneller als „echte" Inhalte.[2] Dieser Entwicklung entgegenzutreten und zugleich die Freiheitsrechte einer demokratischen Bürgergesellschaft zu gewährleisten ist kein einfaches Unterfangen.

I. Information, Wahrheit, Vertrauen und Transparenz

Für eine juristische Auseinandersetzung und Erörterung der Regulierungsmöglichkeiten ist eine Definition der Begriffe normalerweise unerlässlich. Eine abgrenzungsfähige Definition einiger zentraler Begriffe im Bereich der Desinformation ist jedoch nur schwer möglich, weshalb auch die ergänzenden Umstände außerhalb der Begriffe selbst einbezogen werden müssen.

Das beginnt bereits beim Terminus *Information* selbst, für den die Meisten zwar wie beim physikalischen Begriff „Masse" ein „intuitives Gefühl" oder „Vorstellung" haben, es aber selbst Experten nicht genau definieren können.[3] Der Duden definiert den Begriff eher eng u. a. als „Unterrichtung über eine bestimmte Sache" oder schlicht als Synonym für „Auskunft".[4] In unterschiedlichen Wissenschaftsbereichen existieren verschiedene Definitionen: Wirtschaftswissenschaftlich soll es „derjenige Anteil einer Nachricht" sein, „der für den Empfänger neu ist",[5] in der Bibliothekwissenschaft sind es z. B. „die von den verschiedenen Medien übertragenen Inhalte".[6] Auch in der Informatik gibt es zahlreiche Versuche einer Definition, deren Auflistung den Rahmen dieses Beitrags sprengen würde.[7] Juristen haben sich ebenfalls daran versucht: Im Unternehmensrecht sind Informationen im Bereich der Geschäftsgeheimnisse beispielsweise „Tatsachen, Daten, Umstände sowie Vorgänge im weitesten Sinne".[8] Und in Bezug auf die Informationsfreiheit aus Art. 5 GG heißt es, Information ist ein „Vorgang oder Ergebnis eines Vorgangs […], in dessen Rahmen ein kognitives System einen Sachverhalt interpretiert und dadurch sein Verhalten oder seinen Zustand ändert".[9] Dabei sei im Unterschied zu „Daten", die immer körperlich vorlägen, für Informationen „Unkörperlichkeit" maßgeblich.[10] *Information* kann folglich alles einschließen, was ein Mensch – und inzwischen wohl auch eine Maschine – wahrnehmen und verstehen kann.

Über die Definition von *Wahrheit* (z. B. „wahre" Information) streiten Philosophen schon seit jeher. Während der Duden den Begriff teilweise mit einer Tautologie zu definieren versucht – „wirklicher, wahrhafter Sachverhalt oder Tatbestand; Übereinstimmung einer Aussage mit der Sache, über die sie gemacht wird"[11] – findet sich im juristischen Bereich keine eindeutige Definition. Immerhin wird zwischen einer *absoluten* und einer *subjektiven* Wahrheit abgegrenzt, z.B. im Rahmen des § 138 Abs. 1 ZPO (Wahrheitspflicht im Zivilprozess),[12] und im Strafrecht findet sich der Begriff der „Nichterweislichkeit" als Kehrseite der Erweislichkeit einer Tatsache (also der Wahrheit), die für eine Strafbarkeit nach § 186 StGB (Üble Nachrede) erforderlich ist.[13] Wann etwas als „erweislich" gilt, bleibt indessen unausgesprochen, es wird wohl auch auf die Intuition vertraut. Im vorliegenden Zusammenhang könnte deshalb eine Definition aus dem Cambridge Dictionary wohl am ehesten weiterführen: "The real facts about a situation, event or person" und "The actual fact or facts about a matter",[14] was impliziert, dass es sich um real existierende Tatsachen handeln muss. Doch diese Definition stößt auf ihre Grenzen, wenn es um neueste Sachverhalte beispielsweise in virtuellen Umgebungen wie einem Metaverse, einem Computerspiel oder um „Deep Fakes" geht. Im Zusammenhang mit Information sollte deshalb hinzugefügt werden, dass die Quellen für die Beurteilung des Wahrheitsgehalts einer Information essenziell sind (s.u. Vertrauen und Transparenz). Außerdem dürfte *Steinbach* zuzustimmen sein: „Empirie ist anerkanntermaßen nicht etwas schlicht Existierendes, kein einfaches Abbild der Realität. Zweifel an der Unterscheidbarkeit von Tatsachen und Meinungen sind schon angezeigt, weil Realitätswiedergabe stets als ein sprach- und beobachtungsabhängiger situativer Vorgang konstruiert wird."[15]

Das *Vertrauen* ist ein sehr subjektiver Begriff, im alltäglichen Verständnis ist es ein „festes Überzeugtsein von der Verlässlichkeit, Zuverlässigkeit einer Person oder Sache",[16] also eine Art Glaube: „*the belief*

2 *Vosoughi/Roy/Aral*, Science, 2018 Vol 359, Issue 6380, pp. 1146-1151, DOI: 10.1126/science.aap9559; Maertens, NJ 2024, B 24, B 25 f.

3 Beispiel aus *Blieberger/Burgstaller/Schild*, Informatik: Grundlagen, 2002, S. 15.

4 *https://www.duden.de/rechtschreibung/Information* (letzter Abruf am 3. September 2024).

5 Gabler Wirtschaftslexikon, *https://wirtschaftslexikon.gabler.de/definition/information-40528* (letzter Abruf am 3. September 2024).

6 *Saur*, Bibliothekarisches Grundwissen, 2016, S. 6.

7 *Kuhlen/Seeger/Strauch*, Grundlagen der praktischen Information und Dokumentation, 2004, Bd. 1, S. 683 ff.

8 *Keller* in: *Keller/Schönknecht/Glinke*, Geschäftsgeheimnisschutzgesetz, 1. Aufl. 2021, § 2 GeschGehG, Rn. 15.

9 *Grabenwarter* in: *Dürig/Herzog/Scholz*, GG-Kommentar, 102. EL 2023, Art. 5 Abs. 1 GG, Rn. 1002.

10 *Grabenwarter* (Fn. 9), Rn. 1000.

11 *https://www.duden.de/rechtschreibung/Vertrauen* (letzter Abruf am 3. September 2024).

12 *Groh/Werner* in: *Weber*, Rechtswörterbuch, 31. Ed. 2023, Wahrheitspflicht.

13 *Valerius* in: v. *Heintschel-Heinegg*, BeckOK StGB, 59. Ed. Stand 01.11.2023, § 186 StGB, Rn. 18.

14 *https://dictionary.cambridge.org/de/worterbuch/englisch/truth* (letzter Abruf am 3. September 2024).

15 *Steinbach*, Meinungsfreiheit im postfaktischen Umfeld, JZ 2017, 654 Fn. 15 m. w. N.

16 *https://www.duden.de/rechtschreibung/Vertrauen* (letzter Abruf am 3. September 2024).

that you can trust someone or something".[17] In der Medien- und Kommunikationswissenschaft wird der Quelle eine entscheidende Rolle zugeschrieben, denn in einer Welt, in der das Wissen einer Person nur sehr begrenzt ist, kommt es unter anderem auf die Vertrauenswürdigkeit (auch Glaubwürdigkeit) der Informationsressource an. Denn der Empfänger kann den Wahrheitsgehalt einer Information in der Regel nicht selbst überprüfen.[18] Fehlt das Vertrauen, so wird auch an eine Wahrheit nicht geglaubt. Im Bereich der Medien i. w. S. ist die Quelle der Information entscheidend dafür, ob jemand auf ihre Wahrheit vertraut. Es geht hier, wie im Finanzrecht, nicht um die abstrakte Information, sondern um ein Vertrauen in die einem bestimmten Urheber zugeschriebene Information.[19]

Damit schließt sich der Kreis zur *Transparenz*, in der Alltagssprache „Durchschaubarkeit, Nachvollziehbarkeit".[20] Eine juristische Definition findet sich zu diesem sprachlich wohl gut abgegrenzten Begriff nicht, es gibt jedoch zahlreiche Regelwerke mit Bestimmungen dazu, wie die Transparenz zu gewährleisten ist: die Auskunftsansprüche der Medien gegenüber dem Staat nach Art. 5 Abs. 1 Satz 2 GG, dem Medienstaatsvertrag und den Landespressegesetzen,[21] die Informationsfreiheitsgesetze, das UWG, das Geldwäschegesetz, Regelungen zu Finanzprodukten, der neue Digital Services Act der EU (insb. Art. 15, 24, 27, 39 und 42) und natürlich der Medienstaatsvertrag (MStV). In §§ 85 und 93 MStV werden Medienplattformen und Anbietern von Medienintermediären umfassende Aufklärungspflichten, beispielsweise hinsichtlich der Medienauswahl, Sortierung, Anordnung und Hervorhebung von Inhalten u. v. m. auferlegt. Es geht darum, nicht nur die Quellen nachvollziehbar und offen zu kommunizieren, sondern auch die Entscheidung verständlich zu machen, eine bestimmte Information in einer bestimmten Weise zu veröffentlichen, und Interessen offenzulegen. Eine komplette Offenlegung der Quellen ist im journalistischen Bereich allerdings oft nicht möglich.

II. Fehlinformation und Desinformation

Nicht jede ungenaue oder unrichtige Information stellt zugleich eine Desinformation dar. Eine Abgrenzung zu „bloßen" Fehlinformation ist besonders dann wichtig, wenn es um die Entscheidung geht, ob und mit welchen Maßnahmen auf sie reagiert werden soll. Ebenfalls von Bedeutung für diese Entscheidung sind die Herkunft der Information, ihre Urheber, Medien und Intermediäre (Vermittler, z.B. Suchmaschinen oder soziale Netzwerke).

1. Fehlinformation

Juristen scheinen sich bisher noch nicht auf eine eindeutige Definition des Begriffs „Fehlinformation" (engl. *misinformation*) geeinigt zu haben. Die EU-Kommission definiert in ihrem „Aktionsplan für Demokratie" unter der Überschrift „Bekämpfung von Desinformation" Fehlinformation als „falsche oder irreführende Inhalte, die ohne vorsätzliche Schädigungsabsicht weitergegeben werden, deren Auswirkungen jedoch schädlich sein können, z. B. wenn Personen falsche Informationen gutgläubig an Freunde und Familienangehörige weitergeben".[22] Diese Definition kann mit einem Vorbehalt für juristische Zwecke fruchtbar gemacht werden.[23] Die Abgrenzung von Fehlinformation und Desinformation über Vorsatz erscheint grundsätzlich sinnvoll: Es wird auf die Schädigungsabsicht[24] abgestellt. Durch das Erfordernis von potenziell schädlichen Auswirkungen wird der Begriff zudem begrenzt und von folgenlosen Falschinformationen abgegrenzt. Es stellt sich aber die Frage, ob ein Inhalt noch als Fehlinformation und nicht schon als Desinformation qualifiziert werden sollte, wenn der Urheber ohne Schädigungsabsicht, aber bewusst falsch informiert und die falschen Inhalte schädliche Auswirkungen auslösen (können). Mit Blick auf mögliche Sanktionen gegenüber dem Urheber sollte unterschieden werden, ob er falsche oder irreführende Inhalte bewusst (hier ließe sich auch nach Vorsatzform abstufen) veröffentlicht oder im guten Glauben an ihre Richtigkeit.

Eine weitere Schwierigkeit dürfte darin bestehen zu bestimmen, wann eine Auswirkung falscher Inhalte „schädlich sein kann". Neben der Beurteilung des Inhalts selbst wird auf ein unbestimmtes Merkmal abgestellt. Es ist zwar möglich, auch das Gefährdungspotential miteinzubeziehen (wie z. B. im Strafrecht). Bei Medieninhalten kann oft aber nur darüber spekuliert werden, welche Auswirkungen sie haben werden. Eine Gefahrenprognose ist äußerst schwierig.

Schließlich ist unklar, ob ungenaue Informationen und Dekontextualisierung von dieser Definition erfasst werden. Ungenaue Informationen sind nicht immer falsch oder irreführend an sich, sondern haben

17 *https://dictionary.cambridge.org/de/worterbuch/englisch/trust* (letzter Abruf am 3. September 2024).

18 Ausführlich *Kohring*, Vertrauen in Journalismus, 2004, S. 82 ff.

19 *Mülbert/Sajnovits*, Vertrauen und Finanzmarktrecht, ZfPW 2016, S. 7.

20 *https://www.duden.de/rechtschreibung/Transparenz* (letzter Abruf am 3. September 2024).

21 *Weberling* in *Ricker/Weberling*, Handbuch des Presserechts, 7. Auflage 2021, 18. Kap., Rn. 1 ff.

22 COM (2020) 790 final DE, S. 22: *https://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX:52020DC0790* (letzter Abruf am 3. September 2024).

23 Sie wird aber z. T. kommentarlos übernommen, s. z.B. *Brorsen/Falk*, Neue Compliance-Pflichten nach dem Digital Services Act, MMR 2024, 33.

24 Die tautologische Formulierung „vorsätzliche Schädigungsabsicht" im Dokument scheint auf einem Übersetzungsfehler zu beruhen, denn in der englischsprachigen Fassung heißt es nur „without harmful intent", COM (2020) 790 final EN, S. 18.

erst im Kontext Auswirkungen auf ihre Empfänger (s. das Beispiel der Emser Depesche, bei der Dekontextualisierung noch hinzukommt). Dekontextualisierung lässt sich überhaupt nicht an der Information selbst festmachen, sondern entfaltet nur durch ihre Platzierung oder Einbindung in andere Inhalte ihr Potential zur Irreführung (z. B. falsche Datierung oder Ortsangabe eines Bildes oder Videos). Die Definition sollte also um diese beiden Begriffe ergänzt werden.

2. Desinformation

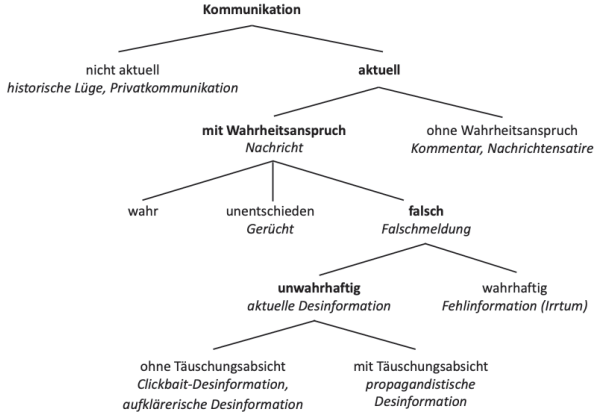
Für den Begriff *Desinformation* existieren einige juristische Definitionsversuche. *Moura Vicente* z.B. bezeichnet sie – ähnlich wie die EU-Kommission[25] und *EU High Level Expert Group on fake news and online disinformation*[26] – als „Erstellung, Präsentation und Verbreitung nachweislich falscher oder irreführender Informationen, die aus Gewinnstreben oder in der Absicht, die Adressaten zu täuschen, erstellt werden und den öffentlichen Interessen schaden können".[27] *Reinhardt* fasst sich kürzer: „Bewusste Täuschung, insbesondere durch Manipulation authentischer Informationen, die sich von schlichten Falschmeldungen, irreführenden Behauptungen und ‚Fake News' unterscheidet".[28]

Im Unterschied zur Fehlinformation wird hier deutlich, dass der Urheber sich der Falschheit oder Irreführung der Inhalte bewusst ist und gerade diese einsetzt, um seine Ziele zu erreichen. Ob diese Definitionen jedoch für eine rechtssichere Handhabung und Regulierung ausreichen, erscheint zweifelhaft. Anders als im Fall der Fehlinformation wird eine Schädigungsabsicht nicht gefordert, die Definition von *Moura Vicente* wird zudem auf die darin genannten zwei Fälle des Vorsatzes (Gewinnstreben und Täuschungsabsicht) begrenzt. Damit lägen auf andere Ziele gerichtete Absichten (z. B. aufklärerische Information, Selbstprofilierung ohne Täuschungs- oder Gewinnerzielungsabsicht oder politischer Vorteil, wie in der Definition der EU-Kommission enthalten) außerhalb der Definition. Umfassender scheint insoweit *Reinhardts* Ansatz zu sein, der schlicht jede bewusste Täuschung als Desinformation charakterisiert, wenn sie keine – unbewusste – Fehlinformation ist. Schwierigkeiten dürfte aber die Abgrenzung zu sog. „Fake News" bereiten, deren Begriffsinhalt bis heute nicht eindeutig definiert ist und in einigen Fällen vom Begriff „aktuelle Desinformation" absorbiert wurde.[29]

Um die Handhabung zu vereinfachen, könnte allein auf den Vorsatz hinsichtlich der Falschheit der Information an sich („Unwahrhaftigkeit")[30] und Schadenspotenzial für die Öffentlichkeit oder auch Einzelne abgestellt werden. Es dürfte keinen legitimen Grund für eine vorsätzliche Veröffentlichung bewusst falscher Informationen geben, mit der möglichen Ausnahme einer Gefahrenabwehr (ggf. Rechtfertigungsgrund).

Nützlich erscheint bei der Abgrenzung das Schema von *Zimmermann/Kohring*:[31]

Abbildung 1: Definitionskriterien aktueller Desinformation



Abschließend sollen noch einige praxisrelevante Arten von Desinformation erwähnt werden, die bei der Regulierung bzw. Maßnahmen bedacht werden sollten, deren Einordnung beizeiten Schwierigkeiten bereitet. Dazu gehören Dekontextualisierung wahrer Informationen, manipulative Werbung, Pseudojournalismus und Propaganda. Dabei handelt es sich nicht immer um objektiv falsche Inhalte, der Schaden wird jedoch durch andere Handlungen angerichtet:

Dekontextualisierung: Objektiv wahre Informationen oder ihre Teile werden sinnverzerrend aus einem Kontext gerissen oder in einen neuen Kontext gesetzt und erhalten dadurch einen anderen, irreführenden Inhalt. Hier könnte der Fall der Emser Depeche eingeordnet werden: Wichtige Teile des Originaltelegramms wurden entfernt und der Inhalt absichtlich verkürzt, um die intendierte Wirkung – eine den Krieg unterstützende öffentliche Meinung – zu erreichen.

Manipulative Werbung: Bestimmte Teile einer Information werden besonders hervorgehoben, um Aufmerksamkeit auf sich zu ziehen (z.B. Clickbait). Auch verschiedene Arten von Targeting können dafür eingesetzt werden.

Pseudojournalismus: Es werden von vorneherein erfundene, falsifizierte oder selbstthematisierende In-

25 "[D]isinformation is false or misleading content that is spread with an intention to deceive or secure economic or political gain and which may cause public harm", COM (2020) 790 final EN, S. 18.

26 "Disinformation [...] includes all forms of false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit", Final report of the High Level Expert Group on fake news and online disinformation (2018), S. 11 li. Sp.

27 ***Moura Vicente***, Schutz vor Desinformation im Internet, MMR 2023, 261.

28 ***Reinhardt*** in: ***Kipker***, Cybersecurity, 2. Aufl. 2023, Kap. 20.4, Rn. 13.

29 ***Zimmermann/Kohring***, „Fake News" als aktuelle Desinformation, Medien & Kommunikationswissenschaft, 11/2018, S. 526 ff.

30 Zum Begriff s. ***Zimmermann/Kohring*** (Fn. 29), S. 535.

31 ***Zimmermann/Kohring*** (Fn. 29), S. 531.

halte oder „inszenierte Pseudoereignisse“[32] veröffentlicht. Dabei werden unrechtmäßig etablierte Namen oder Marken verwendet oder ihnen stark ähnelnde Designs zur Täuschung der Zielgruppe eingesetzt.

Propaganda: Vorsätzlich „täuschende Falschmeldungen, deren Urheber dezidiert politische Ziele verfolgen.[33] Diese hegen die Absicht, die Vorstellungen der Rezipienten mittels unwahrer Behauptungen zu manipulieren, um deren Meinungen, Einstellungen und Handeln in eine bestimmte Richtung zu lenken“.[34]

Möglich sind auch Kombinationen verschiedener Desinformationstechniken.

III. Regulierung

Die Regulierung von Fehl- und Desinformation wirft eine Reihe verfassungsrechtlicher Fragen auf, die besonders in Deutschland auf Grund seiner historischen Erfahrungen mit Zensur und Propaganda hohe Brisanz haben. Diese betreffen insbesondere die Balance zwischen dem Schutz der Kommunikationsgrundrechte gemäß Art. 5 GG sowie Art. 10 Abs. 1 EMRK und anderen rechtlichen, gesellschaftlichen und politischen Interessen. Art. 5 Abs. 1 S. 1 GG schützt primär die Kundgabe subjektiver Werturteile („Meinung“), während isolierte wahre Tatsachenbehauptungen wegen ihrer Objektivität keine Meinungsäußerungen darstellen und nur ausnahmsweise in den Schutzbereich fallen sollen.[35] Bewusst unwahre Tatsachenbehauptungen genießen nach Rechtsprechung des BVerfG keinen Schutz,[36] was durchaus deutlich kritisiert wird.[37] Auch die Pressefreiheit gem. Art. 5 Abs. 1 S. 2 GG, die die gesamte Pressearbeit umfasst, schützt nur „nach bestem Wissen wahrheitsgemäße Informationen“.[38] Jedoch fällt nicht schon jede unüberlegte, gutgläubig geäußerte oder nachlässig recherchierte Tatsachenbehauptung aus dem Schutzbereich der Kommunikationsgrundrechte.[39] Stattdessen könnten Fehlinformationen z. B. unter die Kunstfreiheit des Art. 5 Abs. 3 GG fallen oder ein hinzunehmendes „erlaubtes Risiko“ darstellen,[40] z. B. Fehler in einer (Live-)Berichterstattung über eine unübersichtliche Lage, um eine „Lähmung der Medien“ zu vermeiden,[41] oder im Feuer eines Meinungskampfes.[42] Eine „maßgebliche Schranke wird der Kunstfreiheit aber durch die verfassungsrechtlich gewährleistete funktionierende staatliche Ordnung gesetzt, welche die Effektivität des Grundrechtsschutzes überhaupt erst sicherstellt“.[43] Gerade die Aufrechterhaltung einer solchen Ordnung dürfte neben Persönlichkeitsrechten in vielen Fällen von Fehl- und Desinformation das kollidierende Verfassungsgut sein, das bei der Frage der Verhältnismäßigkeit im Zusammenhang mit einer Regulierung zu berücksichtigen sein wird. Das ist umso wichtiger, als das Abstellen auf den „Wahrheitsgehalt“ einer Information und damit die Bestimmung der Tatsachenqualität nicht immer zu einem eindeutigen Ergebnis führen wird: „Jeder tatsachenbezogene Bericht einer Person durchläuft zwangsläufig einen subjektiven Filter, in dem die vielfältigen Darstellungsformen ausgewählt werden, das heißt was erwähnt, was weggelassen wird, was ausführlich, was ausgeschmückt zur Sprache kommt“.[44]

1. Präventive Maßnahmen

Präventive Maßnahmen im Bereich der Kommunikationsgrundrechte sind problematisch, weil die Grenze zur nach Art. 5 Abs. 1 Satz 3 GG untersagten (Vor)Zensur nicht überschritten werden darf. Einschränkende Maßnahmen „vor der Herstellung oder Verbreitung eines Geisteswerkes“ sind unzulässig.[45] Das Zensurverbot ist nur dann nicht berührt, wenn die Maßnahmen nachträglich einsetzen.[46] Dass diese Lösung angesichts des teilweise hohen Gefährdungspotenzials von Desinformation für manche unbefriedigend erscheint, muss angesichts der kaum zu überschätzenden Bedeutung der Kommunikationsgrundrechte (u. a. deren sog. „Polizeifestigkeit“[47]) in einer demokratischen Gesellschaft hingenommen werden.

Möglich und längst geltende Rechtslage sind jedoch Maßnahmen, die nicht die Inhalte selbst betreffen, sondern die Publikationsmedien. Dabei werden durch sie nicht die Veröffentlichungsmöglichkeiten selbst eingeschränkt, sondern Transparenz gewährleistet. Dazu zählen bspw. Informations- und Impressums-

32 *Hohlfeld*, Vom Informations- zum Pseudojournalismus, Communication Socialis, Bd. 36 Nr. 3 (2003), S. 239.
33 *Jowett & O'Donnell*, 2012, 24.
34 *Zimmermann/Kohring* (Fn. 29), S. 536.
35 Das BVerfG wendet eine weite Definition an, die Mischformen in den Schutzbereich einbezieht, z. B. wenn Wertung und Tatsache nicht getrennt werden können und „der tatsächliche Gehalt ... in den Hintergrund tritt“, vgl. BVerfG, Beschl. v. 22. Juni 1982 - 1 BvR 1376/79, BVerfGE 61, 1, 9.
36 BVerfG (Fn. 35.), BVerfGE 61, 1, 8; Beschl. v. 10. November 1998 - 1 BvR 1531/96, BVerfGE 99, 185; ***Mafi-Gudarzi,*** Desinformation: Herausforderung für die wehrhafte Demokratie, ZRP 2019, S. 67; zur Kritik daran s. mit vielen Nachweisen *Steinbach* (Fn 15), S. 656 ff.
37 S. mit vielen Nachweisen *Steinbach* (Fn. 15), S. 656 ff.
38 *Grabenwarter* (Fn. 9), Rn. 272.
39 *Grimm*, Die Meinungsfreiheit in der Rechtsprechung des Bundesverfassungsgerichts, NJW 1995, 1699.
40 *Rühl*, Tatsachen - Interpretationen - Wertungen, 1998, S. 251.
41 BVerfG, Beschl. v. 3. Juni 1980 - 1 BvR 797/78, BVerfGE 54, 208, 220 = NJW 1980, 2072; ***Rühl***, Tatsachenbehauptungen und Wertungen, AfP 2000, 17, 22.
42 „Weder die ‚Vergiftung des geistigen Klimas‘ und schon gar nicht die ‚Beunruhigung, die die geistige Auseinandersetzung im Meinungskampf mit sich bringt‘, könnten ein Eingriffsgrund [in die Meinungsfreiheit] sein“, *Steinbach* (Fn. 15), S. 658.
43 LG Mannheim Urt. v. 2. März 2020 – 15 Ns 806 Js 10181/18, BeckRS 2020, 62758, Rn. 115; ***Burghart*** in: ***Leibholz/Rinck***, Grundgesetz, 79. Lieferung 2019, Art. 5 GG, Rn. 1056 m w. N.
44 *Steinbach* (Fn. 15), S. 655 m. w. N.
45 *Bethge* in: *Sachs*, GG, 9. Aufl. 2021, GG Art. 5 Rn. 129 ff.
46 *Richter,* Das NetzDG – Wunderwaffe gegen „Hate Speech“ und „Fake News“ oder ein neues Zensurmittel?, ZD-Aktuell 2017, 05623.
47 Vgl. *Ricker* in *Ricker/Weberling* (Fn. 21), 10. Kap., Rn 4.

pflichten, u.a. § 18 Abs. 1 und 2 MStV, §§ 5 und 6 TMG oder Kennzeichnungspflichten, z.B. § 8 Abs. 3, § 18 Abs. 3, § 22 Abs. 1 MStV. Darüber hinaus gelten von vorneherein Sorgfaltspflichten, die sich aus dem Pressekodex, den Pressegesetzen der Länder und z.B. § 19 MStV ergeben.

2. Repressive Maßnahmen

In den Bereich der repressiven Maßnahmen sind Korrekturpflichten und – als ultima ratio – Löschung falscher Inhalte einzuordnen, wenngleich die dadurch erreichte Verhinderung ihrer weiteren Verbreitung auch einen präventiven Charakter hat. Die Pflicht zur Gegendarstellung ergibt sich neben § 20 MStV aus den Pressegesetzen des Länder, kann aber auch zivilrechtlich begründet werden. Die schärfste Reaktion ist die Anordnung der Löschung bestimmter als falsch beurteilter Inhalte z.B. gem. §§ 10 Abs. 1 TMG; §§ 109 Abs. 1 S. 2, § 51 Abs. 1, § 6 und 19 Abs. 1 MStV; § 3 NetzDG[48] oder Durchsetzung des Anspruchs auf Löschung, z.B. gem. §§ 1004, 823 ff. BGB (analog)[49] oder §§ 8, 5 UWG.[50] Abgesehen von eher eindeutigen, die Persönlichkeitsrechte betreffenden und i. d. R. gerichtlich überprüften Fälle im zivilrechtlichen Bereich sind Löschungen nach öffentlichrechtlichen Rechtsgrundlagen zuweilen verfassungsrechtlich problematisch, weil sie die Meinungsfreiheit unverhältnismäßig einschränken können.[51] Der am 17. Februar 2024 in Kraft getretene *Digital Services Act (DSA)*, der ähnliche Pflichten und Sanktionen wie das NetzDG vorsieht (insb. Art. 9 DSA), muss sich derselben Kritik stellen. Ein großes Problem besteht u. a. darin, dass nicht staatliche Stellen und erst recht nicht Gerichte über die Rechtswidrigkeit der Inhalte entscheiden, sondern diese Entscheidung Privaten übertragen wird (Plattformen, Intermediäre etc.), wobei auch offen ist, wie bspw. „offensichtlich rechtswidrige" und einfach „rechtswidrige" Inhalte unterschieden werden sollen.[52]

Ein repressives Handeln gegen Fehl- und Desinformation ist schließlich im Falle des Vorliegens der üblichen Voraussetzungen der Tatbestände des Ordnungs- und des Strafrechts möglich. So enthält z.B. § 115 MStV einen umfassenden Ordnungswidrigkeiten-Katalog. Das Strafrecht bietet ein weitreichendes Instrumentarium an Delikten, darunter im Bereich der Gefährdung demokratischer Ordnung die §§ 86, 86 a, 90, 90 a StGB, des Gemeinschaftsfriedens die §§ 111, 140, 166 StGB, insb. § 126 Abs. 2 StGB (Vortäuschung einer schweren rechtswidrigen Tat)[53] und § 130 StGB (Volksverhetzung),[54] sowie Delikte gegen die persönliche Ehre wie § 185 StGB (Beleidigung), § 186 StGB (Üble Nachrede)[55] und §§ 187 f. StGB (Verleumdung). Auch insoweit ist über § 3 NetzDG und Art. 9 DSA die Anordnung der Löschung der strafbaren Inhalte möglich. Die hier erfassten Fälle erscheinen verfassungsrechtlich weniger problematisch, weil die Kriterien eindeutiger und zudem gerichtlich einfacher zu überprüfen sind bzw. die Löschung erst nach einem gerichtlichen Urteil oder einer Anordnung erfolgt.

3. Berücksichtigung der Urheberschaft

Die existierenden Regelungen differenzieren de lege lata nicht nach der Person des Urhebers, sondern stellen auf die Information selbst ab. Für die Beurteilung der Qualität einer Desinformation ist allerdings bedeutsam, von wem der Inhalt stammt. Nur auf diese Weise werden die Absichten nachvollziehbar, die das Vorgehen gegen eine Desinformation womöglich erst rechtfertigen. Insbesondere im Falle von Desinformation, die von einem Staat lanciert wurde (Propaganda), ist die Urheberschaft wichtig: Solche Inhalte fallen entweder nicht in den Schutzbereich der Kommunikationsfreiheiten nach dem Grundgesetz, und/oder sie können wegen eines Verstoßes gegen das völkerrechtliche Nichteinmischungsgebot einfacher unterbunden werden, wie es bspw. bei „Russia Today" erfolgt ist.[56] Aber auch bei anderen Sachverhalten ist die Unterscheidung zwischen Privatpersonen (höchster Schutz v. a. durch Meinungsfreiheit), Medien (Pressefreiheit), politischen Parteien (über Art. 21 GG) und staatlichen Stellen (geringster Schutz nur im Bereich ihrer Aufgabenerfüllung) relevant sowohl mit Blick auf den Schutzbereich der Kommunikationsgrundrechte als auch auf die Beurteilung der Verhältnismäßigkeit der Eingriffe.

---

48 Teile des NetzDG sind mit dem Inkrafttreten des DSA am 17. Februar 2024 überholt, s. ***Kuhlmann/Trute,*** Die Regulierung von Desinformationen und rechtswidrigen Inhalten, GSZ 2022, 116.

49 BGH, Urt. v. 28. Juli 2015 - VI ZR 340/14, MMR 2016, 210, 211 ff.; ***Holznagel,*** Phänomen „Fake News" – Was ist zu tun?, MMR 2018, 20.

50 ***Rehart/Ruhl/Isele*** in: ***Fritzsche/Münker/Stollwerck***, BeckOK UWG, 22. Ed. 01.10.2023, § 5 UWG, Rn. 62 ff., zur Durchsetzung der DSA über §§ 3 a i. V.m. 8 UWG s. ***Gerdemann/Spindler***, Das Gesetz über digitale Dienste (Digital Services Act) (Teil 2), GRUR 2023, 115.

51 Zu verfassungsrechtlichen Bedenken s. ***Müller-Franken***, Netzwerkdurchsetzungsgesetz: Selbstbehauptung des Rechts oder erster Schritt in die selbstregulierte Vorzensur? – Verfassungsrechtliche Fragen, AfP 2018, 1 ff.; ***Papier H.-J.*** in ***Bär/Grädler/Mayr*** (Hrsg.), Digitalisierung im Spannungsfeld von Politik, Wirtschaft, Wissenschaft und Recht, 2018, S. 171 ff.; zu den zivilrechtlichen Aspekten des NetzDG s. ***Peifer***, AfP 2018, 14 ff.

52 ***Kuhlmann/Trute*** (Fn. 48), S. 115; ***Steinbach*** (Fn. 15), S. 659 f.

53 LG Mannheim, LG Mannheim Urt. v. 2. März 2020 – 15 Ns 806 Js 10181/18, BeckRS 2020, 62758 (Blogbeitrag über einen erfundenen Terroranschlag).

54 Meist fehle insoweit eine Aufstachelung zum Hass bei einer Desinformation, ***Mafi-Gudarzi*** (Fn. 36), S. 67.

55 Problematisch sei hier die Feststellung eines Personenbezugs, ***Tschorr***, Wenn der Staat Fake News verbreitet, ZfDR 2021, 381.

56 ***Jamnejad/Wood***, The Principle of Non-intervention, Leiden Journal of International Law, 22(2) 2009, 345 ff.

IV. Fazit: Differenzierung, Zurückhaltung und Bildung

Beim Thema Desinformation und ihre Regulierung zeigt sich so deutlich wie bei kaum einem anderen Gegenstand, wie stark das politisch Gewollte und das (verfassungs-)rechtlich und vielleicht auch demokratisch Zulässige auseinandergehen können. Der Wunsch des Staates, seine Ordnung vor Destabilisierung durch Desinformation zu schützen, ist nachvollziehbar und grundsätzlich legitim. Einen legalen Weg, um die Verbreitung von Desinformation komplett zu unterbinden, gibt es jedoch nicht in einem Verfassungsstaat, der das Recht auf Meinungsfreiheit, ja Kommunikationsfreiheit allgemein garantiert. Bereits eine rechtssichere, verfassungskonforme Entscheidung darüber, ob und welche Arten von Desinformation bekämpft werden sollen, ist schwierig. Wann ist die Schwelle von einer ungenauen zu einer falschen Information überschritten? Dürfen eine Fehlinformation und eine Desinformation gleichbehandelt werden? Welche Rolle spielt der Vorsatz? Wann liegt eine Schädigung oder eine Gefährdung durch eine Fehl- oder Desinformation vor? Und wann wiegt die Schädigung oder die Gefährdung so schwer, dass die Kommunikationsgrundrechte des Urhebers zurücktreten müssen? Und schließlich: Wer darf darüber entscheiden, ob und ab wann ein Inhalt rechtswidrig ist? Bei diesem – nicht abschließenden – Fragenkatalog ist die technische und technologische Seite noch gar nicht berücksichtigt. Sie spielt eine zunehmend relevante Rolle zum einen bei der Definition von Desinformation, z. B. hinsichtlich der Nutzung von Algorithmen im Falle der Online-Information, Einordnung von Deep Fakes und Künstlicher Intelligenz. Zum anderen ist sie auch für die Sanktionierung und Bekämpfung von Desinformation relevant, z. B. bei zeitlich begrenzten (ggf. automatischen) Nutzersperren, algorithmischer Reduzierung der Sichtbarkeit, KI-Faktenprüfung etc.

Die Möglichkeiten des Staates zur Reaktion auf Desinformation sind folglich nicht nur auf Grund des Grundsatzes der „Staatsferne der Medienaufsicht" eng begrenzt. Auch die Umwälzung der grundsätzlich der Gerichtsbarkeit exklusiv zugewiesenen Aufgaben wie der Feststellung der Rechtswidrigkeit und Verhältnismäßigkeit auf Private ist nicht weniger problematisch. Eine Hinterfragung und ggf. Regulierung der Geschäftsmodelle der einflussreichen Intermediäre, z. B. ihre Nutzung von Algorithmen, wäre ein gangbarer, aber in der Realität wohl sehr steiniger Weg. Es muss also verstärkt auf die „Soft Power" gesetzt werden, die vor allem in der Vermittlung von Medienkompetenz an die Bevölkerung besteht. Die Medienkonsumenten müssen befähigt werden, falsche, dekontextualisierte, manipulierte bzw. manipulative Inhalte zu erkennen und bspw. durch eigene Recherche zu überprüfen. Insoweit steht die Gesellschaft aber noch ganz am Anfang.

Navigieren in der Rechtslandschaft: Die Bedeutung der vergleichenden Analyse im Medienrecht der Länder südlich der Sahara

Prof. Justine Limpitlaw, Johannesburg*

Dieser Beitrag ist ein diskursiver, historischer Blick auf ein zwanzigjähriges Projekt der Konrad-Adenauer-Stiftung (KAS) zur Unterstützung, Entwicklung und Aktualisierung von Handbüchern, die die Medienrechtslandschaft in einer Reihe von Ländern des südlichen und östlichen Afrikas in der Zeit nach der Verfassungsrechtsreform in den 1990er Jahren detailliert beschreiben. Ziel ist es, die Auswirkungen dieser demokratiefördernden Maßnahmen zur Unterstützung des Journalismus, der Medien und der Medienrechtsreform im Allgemeinen zu reflektieren.

Die Anfänge

Anfang der 2000er Jahre traf ich die erste Direktorin des neu ins Leben gerufenen Medienprogramms in Subsahara-Afrika der Konrad-Adenauer-Stiftung (KAS). Sie war entsetzt über den Mangel an Material zum Medienrecht in dieser Region. Zu dieser Zeit war ich als Rechtsanwältin tätig, aber auch als Dozentin in Teilzeit an der juristischen Fakultät der Universität Witwatersrand in Johannesburg, Südafrika, und unterrichtete unter anderem Medienrecht sowie Rundfunkrecht und -regulierung.

* Die Autorin, BA LLB LLM. Electronic Communications Lawyer. ist Honorary Adjunct Professor beim LINK Centre der Universität Witwatersrand, Johannesburg. Der Beitrag wurde in die deutsche Sprache übersetzt von Frau Margarita Hamann, Studien- und Forschungsschwerpunkt Medienrecht der Juristischen Fakultät der Europa-Universität Viadrina Frankfurt (Oder). Alle Internet-Quellen wurden zuletzt abgerufen am 6. September 2024.

Wir trafen uns und sie brachte die Idee ein, Handbücher für das Südliche Afrika zu erstellen – eine Reihe von Werken mit Kapiteln zu jedem Land, die Journalisten, Studenten, Anwälten, Richtern, Akademikern, Medienbesitzern und politischen Entscheidungsträgern helfen sollten, das rechtliche und regulatorische Medienumfeld der einzelnen Länder zu verstehen. Ursprünglich wollten wir die Stimmen von Medienschaffenden/Journalisten in den verschiedenen Ländern einbeziehen, aber dies erwies sich als weniger nützlich als ursprünglich angenommen, da die Kommentare offensichtlich subjektiv waren und, wie sich herausstellte, die Journalisten selbst oft nur sehr wenig über die geltenden Gesetze und Vorschriften in ihren eigenen Ländern wussten, eben weil so wenig über das Medienrechtsumfeld geschrieben worden war und die Gesetze und Vorschriften nur schwer zugänglich waren.

Die erste Ausgabe

Für die erste Reihe von Handbüchern haben wir uns mit einigen Ländern der Entwicklungsgemeinschaft des südlichen Afrikas (SADC) befasst. Ich stellte ein Team aus Studenten (anfangs) und angehenden Rechtsanwälten (später) zusammen, das bei der Suche nach relevanten Artikeln, Kopien von Gesetzen, Verordnungen und Rechtsprechung helfen sollte. Ziel war es, einen umfassenden Überblick über die Medienlandschaft in jedem Land zu gewinnen. Die erste Reihe von Handbüchern wurde im A5-Format veröffentlicht und war recht dünn, da sie sich hauptsächlich auf Recherchen am Schreibtisch stützte. Dennoch boten sie einen vergleichenden Blick auf eine Reihe von Schlüsselthemen:

- Erfahrungen von Journalisten
- Verfassungsrechtlicher Schutz der Medienfreiheit
- Gesetzgebung für die Medien (Printmedien und Rundfunk)
- Verhaltenskodizes für die Medien (sowohl staatliche als auch selbstregulierende)
- Verordnungen, die die Medien betreffen sowie
- Rechtsprechung mit Auswirkungen auf die Medien.

In diesen ersten Bänden[1] wurden elf SADC-Länder behandelt: Botswana, die Demokratische Republik Kongo (DRC), Lesotho, Malawi, Mosambik, Namibia, Südafrika, Swasiland, Tansania, Sambia und Simbabwe. Ich hatte das Glück, einen mosambikanischen Jurastudenten an einer anderen örtlichen Universität zu finden, der mir bei den Gesetzen Mosambiks (auf Portugiesisch verfasst) helfen konnte, und einen kongolesischen Anwaltskandidaten, der mir bei den Gesetzen der Demokratischen Republik Kongo (auf Französisch verfasst) half.

Die von 2003 bis 2006 veröffentlichten Bücher fanden großen Anklang – Journalisten, Redakteure, Aktivisten für Medienfreiheit, Medienanwälte, Studenten und Akademiker wünschten sich eindeutig, in einem einzigen Text etwas über ihr eigenes medienrechtliches Umfeld lesen und in der Lage sein zu können, auf Material über andere SADC-Länder zuzugreifen. Die Bücher wurden auch online verfügbar gemacht, was ihre Reichweite erheblich vergrößerte.

Die Bücher wurden in Kinshasa (Demokratischen Republik Kongo) vorgestellt, wobei drei Aspekte besonders hervorstachen:

Zunächst wurden wir auf der Fahrt vom Flughafen in die Stadt fast verhaftet, als ich vom Auto aus einige Fotos von Straßenständen machte. Ich hatte vergessen, dass man dort eine Genehmigung zum Fotografieren braucht, selbst für Touristenschnappschüsse.

Zweitens waren die Journalisten, die eingeladen waren, über die Veranstaltung zu berichten, nicht glücklich darüber, dass sie für ihre Teilnahme nicht in bar bezahlt wurden. Eine deutliche Erinnerung an die Unsitte des so genannten "Brown Bag Journalism", die viele Medien auf dem Kontinent plagt – wo Journalisten schlecht bezahlt werden und ihr Einkommen "aufbessern", indem sie dafür bezahlt werden, über bestimmte Ereignisse zu berichten.

Drittens, und vielleicht am amüsantesten, war die Rede des Ehrengastes bei der Eröffnung, des für Medienangelegenheiten zuständigen Ministers. Er stand auf und sagte dem Publikum, dass er auf dem Weg zum Veranstaltungsort noch überlegte, ob er mir gratulieren oder mich verhaften sollte! Offensichtlich war er verärgert über die scharfe Kritik an der Medienlandschaft in der DRK.

Die zweite Ausgabe

Im Jahr 2008 trat der nächste Direktor des KAS-Medienprogramms an mich heran, um eine weitere Reihe zu verfassen, diesmal jedoch eine viel umfassendere, maßgebliche Ausgabe, die weit über eine

1 Vgl. Volume 1, https://www.humanrightsinitiative.org/programs/ai/rti/international/laws_papers/malawi/Study%20on%20media%20freedom%20in%204%20SADC%20-%20Mandela%20Institute.pdf;
Volume 2,
https://www.kas.de/c/document_library/get_file?uuid=8e39c554-2c4c-140f-ef31-a8fb64c7a2f3&groupId=252038;
Volume 3,
https://www.kas.de/c/document_library/get_file?uuid=af37ebec-eef1-9a82-739c-e7559fbf3baf&groupId=252038;
Mozambique,
https://www.kas.de/c/document_library/get_file?uuid=ca3e5d52-e367-f74d-0a05-193211f54f75&groupId=252038;
Democratic Republic of Congo,
https://www.kas.de/en/web/medien-afrika/einzeltitel/detail/-/content/medienrechtshandbuch-fuer-dr-kongo-aktualisiert1.

Schreibtischrecherche hinausging. Dies erforderte eine neue Methode: die Beschaffung von Kopien aller relevanten Verfassungen/Gesetze/Verordnungen/Kodizes/Fälle, um ein Höchstmaß an Vollständigkeit zu gewährleisten. Es war klar, dass kein Außenstehender in der Lage sein würde, alle erforderlichen Materialien zu beschaffen. Also machte ich mich daran, in jedem Land Juristen zu finden, die sämtliches juristisches Material zusammentrugen und die ausgearbeiteten Kapitelentwürfe auf ihre Richtigkeit hin überprüften.

Für die Praktiker in den Industrieländern ist dies eine relativ einfache Aufgabe, da alle derartigen Materialien online frei verfügbar sind. Nicht so in Entwicklungsländern oder am wenigsten entwickelten Ländern. In der Demokratischen Republik Kongo zum Beispiel war kein einziges Material elektronisch verfügbar. Alle Materialien mussten in Papierform gekauft oder von den Anwälten in ihren Kanzleien fotokopiert werden. Der in der DRK ansässige Anwalt, den wir unter Vertrag genommen hatten, lebte in Lubumbashi, aber die staatliche Druckerei – der einzige Ort, an dem Kopien der staatlichen Amtsblätter erworben werden konnten – befand sich in Kinshasa – über zweitausend Kilometer entfernt, und die Straße ist schlecht (geschätzte Reisezeit – 38 Stunden). Das Zusammentragen der Materialien in der DRK war eine Herkulesaufgabe, und ich war äußerst dankbar, als ich schließlich ein DHL-Paket mit wertvollen Ausdrucken aller rechtlichen Materialien erhielt.

Die Durchsicht der Materialien war offenbar auch eine Herausforderung – in den meisten Ländern waren (damals) keine konsolidierten Fassungen von Gesetzen verfügbar, so dass man mühsam die Änderungen der Gesetze durchgehen und sicherstellen musste, dass aufgehobener Text geändert und neue Bestimmungen eingefügt wurden. Es kostete Tage und Wochen an Arbeit, nur um ein Gesetz zur Überprüfung zusammenzustellen. Besonders schwierig war es, dies in einer Fremdsprache zu tun. Die Übersetzung der Materialien dauerte noch viel länger, und ich verließ mich in hohem Maße auf einen kongolesischen Jurastudenten, der sich bereit erklärt hatte, beim Kapitel über die Demokratische Republik Kongo zu helfen – wir saßen tage-, wochen- und monatelang Seite an Seite, während wir die juristischen Materialien durcharbeiteten und das Kapitel über die Demokratische Republik Kongo auf Englisch entwarfen, und anschließend übersetzte er es ins Französische, um es in der Demokratischen Republik Kongo zu verwenden.

In den Jahren 2012 und 2013 wurde schließlich ein umfangreiches zweibändiges Handbuch[2] zum Medienrecht veröffentlicht, das in Bezug auf Umfang und Tiefe der behandelten Themen viel umfassender war als die ersten SADC-Handbücher zum Medienrecht. Aber es gab auch Lücken: Da ich keinen lusophonen Assistenten finden konnte, war es mir nicht möglich, ein Kapitel über Mosambik (oder Angola) zu verfassen, und auch die Inseln im Indischen Ozean haben wir nicht berücksichtigt. Wir haben die zehn Länderkapitel nach einer bestimmten Vorlage verfasst und versucht, sie so umfassend wie möglich zu gestalten. Die behandelten Themen waren:

- Medien und die Verfassung
- Medien und Gesetzgebung
  - Wie Gesetzgebung gemacht wird
  - Gesetze, die Folgendes regeln: Printmedien, Rundfunkmedien, staatliche Rundfunkanstalten, Filmproduktion, Medienschaffende und die Verbreitung von Rundfunksignalen.
  - Zensurgesetze
  - Abhörgesetze (für Themen wie Telefonüberwachung usw.)
  - Gesetze, die die Medien unterstützen, wie z. B. Zugang zu Informationen, Schutz von Informanten, Transparenzanforderungen und ähnliches
- Vorschriften, die die Medien betreffen
- Selbstregulierung, einschließlich Verhaltenskodizes und Mechanismen zur Durchsetzung der Selbstregulierung
- Rechtsprechung mit Auswirkungen auf die Medien.

Die Handbücher gingen jedoch über ein sehr viel ausführlicheres und umfassenderes Kapitel zu jedem Land hinaus. Wir begannen damit, Materialien normativer Art aufzunehmen. Wir haben drei erste Kapitel für das Handbuch entwickelt, in denen die wichtigsten normativen Grundsätze eines freien Medienumfelds erläutert werden, die als Maßstab für die Bewertung des Engagements eines Landes für demokratische Medienpraktiken und die Reform des Medienrechts dienen sollen.

Und so enthält dieses zweibändige Werk drei neue Eingangskapitel:

- Die Rolle der Medien und die Pressefreiheit in der Gesellschaft
- Merkmale eines demokratischen Medienumfelds
  - Grundsätze eines demokratischen Medienumfelds
  - Grundsätze einer demokratischen Rundfunkregulierung
- Medienrecht: Fallstricke und Schutzmöglichkeiten für die Medien

2 Vgl. Volume 1, https://www.kas.de/c/document_library/get_file?uuid=6282e973-1437-2d89-db7c-0acebb127fa5&groupId=252038;Volume 2, https://www.kas.de/c/document_library/get_file?uuid=c399ed08-7ff5-f7cf-685f-cbdf0fce920c&groupId=252038.

Das Kapitel über die Rolle der Medien und die Pressefreiheit in der Gesellschaft befasste sich mit den konstitutiven Gründen für das Recht auf freie Meinungsäußerung, mit der zentralen Bedeutung dieses Rechts für den Einzelnen und mit der Tatsache, dass die Rechte auf Gleichheit und Würde grundlegend für die Anerkennung des Rechts auf freie Meinungsäußerung sind, und wie wichtig das Recht auf freie Meinungsäußerung wiederum für die Entwicklung und Sicherung der Autonomie des Einzelnen und seiner Persönlichkeit ist. Das Kapitel fasst auch die wichtigsten internationalen Rechtsinstrumente zusammen, aus denen sich die konstitutiven Begründungen ableiten, darunter die Allgemeine Erklärung der Menschenrechte[3], der Internationale Pakt über bürgerliche und politische Rechte[4] und die Afrikanische Charta der Menschenrechte und Rechte der Völker[5]. Anschließend werden die instrumentellen Begründungen für das Recht auf freie Meinungsäußerung erörtert, darunter die Suche nach der Wahrheit auf dem Ideenmarkt und die Tatsache, dass es für die Demokratie von entscheidender Bedeutung ist. Anschließend wird die Rolle der Medien in der Gesellschaft erörtert, einschließlich der allgemeinen Rolle der Presse und ihrer spezifischeren Rolle als Wächter, Aufdecker, Erzieher der Öffentlichkeit, Befürworter von Demokratie und guter Regierungsführung und Katalysator für Demokratie und Entwicklung. Den Abschluss bildet eine Diskussion über die Bedeutung der Rundfunkmedien.

Das Kapitel über die Merkmale eines demokratischen Medienumfelds listet eine Reihe von internationalen Instrumenten auf, die demokratische Grundsätze für die Medien- und Rundfunkregulierung festlegen, und nennt die zehn wichtigsten Grundsätze der demokratischen Medienregulierung, nämlich:

- Freiheit der Presse und anderer Medien
- Unabhängige Medien
- Vielfalt und Pluralismus in den Medien
- Professionelle Medien
- Schutz der Vertraulichkeit von Quellen
- Zugang zu Informationen
- Verpflichtung zu Transparenz und Rechenschaftspflicht
- Engagement für die öffentliche Debatte und Diskussion
- Verfügbarkeit von lokalen Inhalten
- Sicherstellung, dass die Staaten ihre Werbemacht nicht zur Beeinflussung von Inhalten einsetzen.

Außerdem werden acht Schlüsselprinzipien einer demokratischen Rundfunkregulierung erörtert, nämlich:

- Nationale Rahmenbedingungen für die Regulierung des Rundfunks müssen gesetzlich verankert werden
- Unabhängige Regulierung des Rundfunks
- Pluralistische Rundfunkumgebungen sehen drei Ebenen des Rundfunks vor: öffentlich-rechtliche, kommerzielle und gemeinnützige.
- Öffentlich-rechtliche im Gegensatz zu staatlichen Rundfunkdiensten
- Verfügbarkeit von Bürgerrundfunk
- Gerechte, faire, transparente und partizipative Lizenzierungsverfahren, auch in Bezug auf die Frequenzen
- Universeller Zugang zu Rundfunkdiensten und gerechter Zugang zur Signalverteilung und anderen Infrastrukturen
- Regulierung von Rundfunkinhalten im öffentlichen Interesse.

Das Kapitel "Fallstricke und Schutzmaßnahmen für die Medien" befasste sich mit Gesetzen, die die Verbreitung bestimmter Formen der Meinungsäußerung durch die Medien rechtmäßig regeln oder verbieten, darunter

- Lizenzvergabe und Regulierung von Rundfunk und Kino
- Schutz des guten Rufs
- Schutz der Rechte anderer im Allgemeinen
- Schutz der Privatsphäre
- Regulierung der Obszönität und Schutz von Kindern und der Sittlichkeit
- Propaganda für den Krieg
- Hassreden oder diskriminierende Äußerungen
- Schutz der nationalen Sicherheit oder der territorialen Integrität
- Krieg oder Ausnahmezustand
- Schutz der öffentlichen Ordnung oder Sicherheit
- Schutz der öffentlichen Gesundheit
- Aufrechterhaltung der Autorität und Unparteilichkeit der Justiz
- Verhütung von Straftaten
- Verhinderung der Weitergabe von vertraulichen Informationen.

Sie hat auch die Arten von Gesetzen identifiziert, die die Medien bei der Erfüllung ihrer Aufgaben aktiv behindern, nämlich diejenigen, die:

- den Marktzugang in unverhältnismäßiger Weise beschränken
- eine Vorabzensur vorsehen
- die Rechte des Einzelnen (insbesondere von Amtsträgern) über das Recht der Öffentlichkeit auf Information stellen

3 https://www.ohchr.org/en/human-rights/universal-declaration/translations/german-deutsch?LangID=ger.
4 https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights.
5 https://au.int/sites/default/files/treaties/36390-treaty-0011_-_african_charter_on_human_and_peoples_rights_e.pdf.

- nicht den international anerkannten Gründen für die Einschränkung der folgenden Arten der Meinungsäußerung entsprechen:
  - Anstößige Materialien
  - Propaganda für den Krieg oder Hassreden
  - Bedrohung der nationalen Sicherheit, der territorialen Integrität, der öffentlichen Ordnung und der Strafverfolgung
  - Untergrabung der Judikative
- Verhängung des Ausnahmezustands auf unbestimmte Zeit
- Kriminalisierung von Verleumdungen.

Im weiteren Abschnitt werden Beispiele für die Arten von Gesetzen genannt, die die Medien bei der Wahrnehmung ihrer verschiedenen Aufgaben unterstützen, darunter:

- Verfassungen mit einer Bill of Rights und anderen medienbezogenen Schutzbestimmungen
- Gesetze, die den international anerkannten Standards für eine demokratische Medien- und Rundfunkregulierung entsprechen.
- Gesetze, die den international anerkannten Standards für die Einschränkung von Veröffentlichungen oder Sendungen entsprechen.
- Gesetze über den Zugang zu Informationen
- Gesetze zum Schutz von Whistleblowern und/oder zur Korruptionsbekämpfung.
- Gesetze zur Einrichtung unabhängiger Stellen, die im öffentlichen Interesse handeln, wie Ombudsleute, ein Public Protector, eine Menschenrechtskommission usw.

Außerdem haben wir am Ende von Band Zwei ein neues Schlusskapitel eingefügt: Medienrecht in der Region – Wohin führt der Weg? In diesem Kapitel wurde aus der Vogelperspektive bewertet, inwieweit die Länder die zehn Grundprinzipien einer demokratischen Medienordnung und die acht Grundprinzipien einer demokratischen Rundfunkordnung einhalten. Außerdem wurde dargelegt, was internationale Organisationen zur Förderung der Medienfreiheit auf dem Kontinent unternommen haben. Schließlich wurden die wichtigsten Herausforderungen für die Medienfreiheit in der SADC-Region aufgezeigt, darunter

- Die Zensur als Erbe des Kolonialismus
- Gesetze zur Registrierung von Medien
- Restriktive Rundfunkgesetze
- Strafrechtliche Verleumdungsgesetze
- Beleidigungsgesetze
- Überholte Obszönitätsgesetze
- Überzogene Gesetze gegen Aufruhr und andere Sicherheitsgesetze.

Die viel umfassendere Behandlung des medienrechtlichen Umfelds der behandelten Länder wurde von Anfang an sehr gut aufgenommen. Das zweibändige Werk, das von der KAS kostenlos zur Verfügung gestellt und mit einer Creative-Commons-Lizenz[6] online zugänglich gemacht wurde, war im südlichen Afrika weit verbreitet und wurde rasch als Standardwerk auf Universitätsebene übernommen und in den Redaktionen der Region verteilt.

Ich wohnte dem Prozess gegen Bheki Makhubu bei, den Herausgeber und Verleger des Magazins The Nation, einer Monatszeitschrift, die kritische Artikel über die politische und soziale Landschaft von Swasiland (eSwatini)[7] veröffentlichte. Er fühlte sich nicht wohl und hörte nur halb zu, als unsere Delegation, ein Team südafrikanischer Menschenrechtsanwälte, die den Prozess beobachteten, uns vorstellte. Als er jedoch meinen Namen hörte, sagte er: "Justine Limpitlaw? Das ist alles Ihre schuld!". Er erklärte, er habe das Kapitel über Swasiland gelesen und sei begeistert gewesen. Er sagte: "Also dachte ich: 'Veröffentlichen und verdammt sein' und ich habe es veröffentlicht und jetzt bin ich verdammt!"

Bei einer anderen Gelegenheit flog ich vom Flughafen in Lusaka, Sambia, ab, als der Beamte der Einwanderungsbehörde kerzengerade aufsaß, als er meinen Namen in meinem Reisepass sah. "Justine Limpitlaw?", fragte er – und es ist nicht unbedingt etwas Gutes, wenn ein Einwanderungsbeamter deinen Namen erkennt. Als ich das bejahte, erzählte er mir, dass er nebenbei Jura studiere und eines seiner Fächer Medienrecht sei, dass das Handbuch zu seinen Pflichtlektüren gehöre und dass er das Kapitel über Sambia fertiggelesen habe.

Die dritte Ausgabe – Ostafrika

Noch im selben Jahr (2014) bat mich die KAS, die Handbücher erheblich zu erweitern und ein Handbuch für Ostafrika zu realisieren. Das zweibändige Werk wurde 2017 veröffentlicht und umfasste sechs ostafrikanische Länder: Burundi, Eritrea, Äthiopien, Kenia, Ruanda und Uganda. Leider war es aufgrund der Aufstände in Dschibuti und Somalia nicht möglich, Juristen zu finden, die in der Lage waren, die für die Überprüfung der beiden Länder erforderlichen medienrechtlichen Materialien aufzuspüren.

Ich hatte auch das große Glück, einen eritreischen Anwalt im Exil zu finden, der Kontakte zu Regierungsvertretern hatte und der mir Kopien aller eritreischen Materialien besorgen konnte. Besonders bemerkenswert war der Mut der Anwälte aus repressiven Ländern, die mir unter großer Gefahr für sich selbst Materialien für ein Handbuch zur Verfügung stellten, dass diese Länder kritisieren würde. Die Anwälte, die mich in Burundi und Äthiopien unterstützten, baten darum, im Danksagungsteil des Handbuchs nicht na-

6 https://creativecommons.org/
7 Damals hieß Eswatini noch so.

mentlich erwähnt zu werden, da sie Repressalien befürchteten. Das Handbuch wurde in Kenia vorgestellt und fand ebenfalls großen Anklang.

2019 habe ich auf Einladung der KAS beim Global Media Forum der Deutschen Welle einen Vortrag über die Handbücher gehalten. Eines Abends ging ich zurück zu meinem Hotel und lief mit einem anderen Delegierten aus Afrika zusammen. Wir stellten uns vor, und er erzählte mir, er sei einer der sogenannten Zone 9 Blogger[8] und sagte, das Kapitel über Äthiopien im Handbuch Medienrecht in Ostafrika habe ihm und anderen in Äthiopien Mut gemacht, dass die Menschen darauf achteten, was in dem Land passiere, und dass das Kapitel bestätigt habe, dass ihr Engagement für den Grundsatz der Meinungsfreiheit richtig sei.

Die vierte Ausgabe – Südliches Afrika

Im Jahr 2018 wurde ich beauftragt, eine zweite Auflage (in Wirklichkeit eine dritte Auflage) des Handbuchs Medienrecht für das südliche Afrika[9] zu beginnen, das 2021 in drei Bänden veröffentlicht wurde.

Während die Struktur des Handbuchs gleichgeblieben ist, haben wir nun Materialien zur Regulierung des Internets aufgenommen. So wurde beispielsweise in Kapitel 2, Kennzeichen eines demokratischen Medienumfelds, ein völlig neuer Abschnitt über die Grundsätze einer demokratischen Internetregulierung aufgenommen:

- Internetzugang und Erschwinglichkeit
- Freiheit der Meinungsäußerung und Informationsfreiheit im Internet
- Versammlungs- und Vereinigungsfreiheit im Internet
- Das Recht auf Privatsphäre, Anonymität, Schutz personenbezogener Daten und Freiheit von Überwachung im Internet
- Sicherheit, Stabilität und Widerstandsfähigkeit des Internets
- Demokratische Internet-Verwaltung unter Beteiligung mehrerer Interessengruppen
- Gerechte Verteilung der Einnahmen aus dem Internet.

Auch für diese Ausgabe konnten wir wieder portugiesische Juristen finden, die uns bei der Zusammenstellung und Übersetzung der Mosambik-Materialien halfen (leider konnten wir aber immer noch keinen angolanischen Juristen finden). Außerdem konnten wir zum ersten Mal zwei Inselstaaten, nämlich Mauritius und die Seychellen, in das Handbuch aufnehmen, wodurch sich die Gesamtzahl der Länderkapitel des südlichen Afrikas auf vierzehn erhöht hat. Auch hier haben wir eine französische Übersetzung des Kapitels über die Demokratische Republik Kongo und eine portugiesische Übersetzung des Kapitels über Mosambik erstellt.

Die Bedeutung und die Auswirkungen der Handbücher

In einigen Fällen sind die Handbücher die einzige öffentlich zugängliche Quelle über die Medienrechtslandschaft des jeweiligen Landes. Es stimmt zwar, dass immer mehr afrikanische Länder ihre Gesetze online zur Verfügung stellen, aber nur sehr wenige veröffentlichen konsolidierte Gesetze (in vielen Fällen haben wir die Konsolidierung für die Handbücher von Hand vorgenommen). Fast keines veröffentlicht Rechtsprechung und Vorschriften online, und wenn doch, dann meist hinter Bezahlschranken, die für den Durchschnittsbürger unerreichbar sind.

Auch wenn man zahlreiche akademische Artikel über den einen oder anderen Aspekt des medienrechtlichen Umfelds eines Landes finden kann, gibt es nur wenige Quellen, die einen detaillierten Überblick über die Mediengesetze eines Landes bieten, geschweige denn eine vergleichende regionale Betrachtung, wie sie in den Handbüchern vorgenommen wird.

Besonders wichtig ist, dass die KAS dafür gesorgt hat, dass die Handbücher in Papierform und in elektronischer Form zum kostenlosen Herunterladen im Internet zur Verfügung stehen und dass sie in der Amtssprache des jeweiligen Landes verfügbar sind, d. h. in Französisch für die Demokratische Republik Kongo und Burundi und in Portugiesisch für Mosambik.

Außerdem sind die Handbücher nicht nur deskriptiver Natur, indem sie die medienrechtliche Situation in einem bestimmten Land zum Zeitpunkt der Erstellung darlegen, sondern sie sind auch normativ – sie helfen Journalisten, Redakteuren, Studenten, Akademikern, Anwälten und Akteuren der Zivilgesellschaft zu verstehen, wie ein demokratisches medienrechtliches Umfeld (Print, Rundfunk und Online) aussehen sollte. Dies hilft den Menschen in den verschiedenen Ländern, z. B. Richtlinienentwürfe, Regelungen und Gesetzen in Kenntnis der Grundsätze zu kommentieren, um die es geht.

Die Handbücher wurden in einer Reihe von Ländern als universitäres Lehrmaterial und auch in Online-Kursen verwendet. So wurden sie beispielsweise in dem Online-Kurs "Media Freedom and Freedom of Expression in Africa"[10] verwendet, der vom LINK

8 https://cpj.org/awards/zone-9-bloggers-ethiopia/.
9 https://www.kas.de/en/web/medien-afrika/einzeltitel/detail/-/content/media-law-handbook-for-southern-africa-second-edition.
10 https://www.wits.ac.za/linkcentre/certificate-courses/media-freedom-and-freedom-of-expression-in-africa-jeanette-minnie-course-2018-21/.

Centre der Universität Witwatersrand geleitet und auf der EdX-Plattform angeboten wurde. Mehr als 10.000 Studierende aus 22 afrikanischen Ländern nahmen an dem Kurs teil. EdX berichtet, dass der Kurs die größte Vielfalt an afrikanischen Studierenden aller Kurse auf seiner Plattform aufwies.

Die Handbücher haben dazu geführt, dass europäische Förderorganisationen Schwerpunktbereiche für die Unterstützung der Medienfreiheit in Subsahara-Afrika festgelegt haben. Sie wurden auch von internationalen Organisationen wie der UNESCO genutzt, die medienbezogene Unterrichtsmaterialien in einer Publikation mit dem Titel Teaching Media Policy in Africa[11] entwickelt hat, die zusammen mit dem Namibian Media Trust veröffentlicht wurde und in der achtmal auf die Handbücher verwiesen wird.

Die vielleicht wichtigste langfristige Auswirkung der Handbücher sind die Veränderungen, die wir bei den Gesetzen und Praktiken in den Ländern südlich der Sahara beobachten können. Diese sind natürlich weder einheitlich noch konsistent (oft gibt es zwei Schritte vorwärts, drei Schritte zurück), aber einige Fortschritte können mit den Handbüchern in Verbindung gebracht werden. Im Jahr 2020 wurde ich beispielsweise von einer Nichtregierungsorganisation angesprochen, die das äthiopische Justizministerium bei der Ausarbeitung einer Reihe neuer kommunikationsbezogener Gesetze unterstützt. Sie erwähnten ausdrücklich das äthiopische Kapitel des Handbuchs zum Medienrecht für Ostafrika, als sie mich baten, ihre Gesetzesentwürfe zu überprüfen, um aufzuzeigen, wo die Gesetzesentwürfe nicht den normativen Grundsätzen entsprachen, die in den ersten Kapiteln des Handbuchs dargelegt waren.

In Südafrika war ich an den zivilgesellschaftlichen Bemühungen beteiligt, den öffentlich-rechtlichen Rundfunk, die South African Broadcasting Corporation, davor zu schützen, wieder zu einer staatlichen Rundfunkanstalt zu werden. Das Anwaltsteam, das den Fall vor Gericht vertrat, machte ausgiebig Gebrauch vom Medienrechtshandbuch für das Südliche Afrika, um die wichtigsten Grundsätze für den öffentlich-rechtlichen Rundfunk darzulegen, die im Wesentlichen vom Obersten Gerichtshof in der Rechtssache SOS Support Public Broadcasting Coalition und Andere gegen SABC und Andere[12] übernommen wurden.

Es überrascht nicht, dass die afrikanischen Regierungen afrikanische Texte als Orientierungshilfe in politischen und rechtlichen Fragen bevorzugen. Folglich ist die Möglichkeit, Einfluss auf den Wortlaut kontinentaler Verträge und Erklärungen usw. zu nehmen, von enormer Bedeutung, um die nationale Übernahme demokratischer Grundsätze der Medienregulierung zu fördern. Anfang 2019 wurde ich (aufgrund der Handbücher) zu einem zivilgesellschaftlichen Workshop eingeladen, der vom Sonderberichterstatter für Meinungsfreiheit und Zugang zu Informationen der Afrikanischen Kommission für Menschenrechte und Rechte der Völker, Herrn Lawrence Murugu Mute, geleitet wurde. Er hatte Kopien der Handbücher auf seinem Tisch liegen! Im November 2019 wurde die Grundsatzerklärung zur Meinungsfreiheit und zum Zugang zu Informationen in Afrika[13] von der Kommission angenommen. Darin finden sich Formulierungen, an deren Ausarbeitung ich maßgeblich beteiligt war und für die ich mich im Rahmen von Stellungnahmen der Zivilgesellschaft eingesetzt habe, insbesondere in Bezug auf den öffentlich-rechtlichen Rundfunk (Grundsatz 13: Öffentlich-rechtliche Medien), nämlich

- Die Führungskräfte der öffentlich-rechtlichen Medien werden vom Verwaltungsrat ernannt und sind diesem gegenüber rechenschaftspflichtig. – Grundsatz 13(2)
- Die redaktionelle Unabhängigkeit der öffentlich-rechtlichen Medien ist zu gewährleisten. – Grundsatz 13(3)
- Die öffentlich-rechtlichen Medien werden in einer Weise angemessen finanziert, die sie vor unzulässigen Eingriffen schützt. – Grundsatz 13(3).

Die Rolle der Konrad-Adenauer-Stiftung (KAS):

Ganz einfach: Keines der Handbücher (drei Versionen der Handbücher für das südliche Afrika und das Handbuch für Ostafrika) wäre ohne die administrative und finanzielle Unterstützung der KAS möglich gewesen.

Seit mehr als 20 Jahren finanziert das Medienprogramm der KAS die Handbücher: die Zusammenstellung der Gesetze, die Honorare der Anwälte und Übersetzer in den unterschiedlichen Ländern, die Honorare des Autors und der Redakteure, die Gestaltung, den Satz und den Druck der Bücher und die Kosten für das Hosting der Online-Versionen sowie die Einführung der Bücher in vier verschiedenen Ländern im Laufe der Jahre: Südafrika, die Demokratische Republik Kongo, Kenia und Mosambik.

Das war eine gewaltige Investition.

Meiner Ansicht nach hat diese Art von finanzieller und langfristiger Unterstützung zu einer Reihe von Handbüchern geführt, die nicht nur an sich als Publikationen wichtig und nützlich sind, sondern auch direkt und messbar zur Verbesserung der Medienrechtslandschaft in Afrika südlich der Sahara als Ganzes beigetragen haben.

11 https://unesdoc.unesco.org/ark:/48223/pf0000379923.
12 https://www.saflii.org/za/cases/ZAGPJHC/2017/289.html.
13 https://www.chr.up.ac.za/images/researchunits/dgdr/documents/ati/Declaration_of_Principles_on_Freedom_of_Expression_ENG_2019.pdf.

Einblicke in globale Desinformations- und Propagandastrategien

Rechtsanwalt Ferdinand Gehringer, Berlin*

I. Einleitung

In einer Ära, in der Informationen in Sekundenschnelle um den Globus zirkulieren, sind Desinformation und Propaganda zu mächtigen Werkzeugen geworden. Diese Phänomene sind nicht neu, doch ihre Reichweite und Wirkungskraft haben durch die Digitalisierung und die Verbreitung sozialer Medien eine neue Dimension erreicht. Studien zeigen, dass manipulierte Informationen sich sechsmal schneller verbreiten als echte Nachrichten, und dass wir uns möglicherweise bald in einer Welt befinden könnten, in der mehr Falschinformationen als echte Nachrichten konsumiert werden.[1]

Desinformation, bei der es sich um absichtliche Verbreitung falscher oder irreführender Informationen handelt,[2] und Propaganda, die gezielte Beeinflussung der öffentlichen Meinung, stellen erhebliche Herausforderungen für Gesellschaften weltweit dar. Staatliche und nichtstaatliche Akteure agieren zunehmend professioneller und nutzen ausgeklügelte Strategien, um Narrative zu verbreiten und bestehende gesellschaftliche Spaltungen zu vertiefen. Dieser Kampf der Narrative, durch Informationsoperationen (InfoOps) ausgeführt, wird sowohl militärisch in Kriegszeiten als auch als hybride Einflussnahme in Zeiten, in denen kein kriegerischer Zustand besteht, eingesetzt, mit dem Ziel, Demokratien zu destabilisieren.[3]

Der Kampf gegen diese Narrative dreht sich nicht nur um die Unterscheidung zwischen Wahrheit und Lüge, sondern auch um die Durchsetzung vorherrschender Werte. **Autoritäre Regime wie China und Russland machen dabei keinen Unterschied zwischen Cyberoperationen (wie Cyberangriffe über Phishing Mails oder Schadsoftware) und Informationsoperationen im Informationsraum (wie Desinformationskampagnen).**

Sie verfolgen eine doppelte Strategie: Einerseits versuchen sie, Unentschlossene zu manipulieren und die Reihen bestehender Unterstützer für ein Narrativ zu schließen, andererseits zielen sie darauf ab, Gegner gegeneinander auszuspielen und so die Polarisierung zu erhöhen.

Dieser Beitrag liefert einen kurzen Überblick über die historischen Wurzeln von Desinformation und Propaganda, beschreibt aktuelle Strategien von Staaten und zeigt die globalen Unterschiede auf. Zudem werden die rechtlichen Rahmenbedingungen und die Herausforderungen bei der Bekämpfung von Desinformation und Propaganda beleuchtet. Ziel ist es, nicht nur die Bedrohungen aufzuzeigen, sondern auch mögliche Lösungsansätze für Präventionsstrategien zu liefern.

II. Historischer Abriss

Die Nutzung von Propaganda und Desinformation hat eine lange Geschichte, die bis in die Antike zurückreicht. Bereits im alten Rom und Griechenland wurden gezielte Informationen verbreitet, um die öffentliche Meinung zu beeinflussen und ideologische Ziele zu erreichen. Ein berühmtes Beispiel aus der Antike ist die Nutzung von Propaganda durch Julius Caesar, der seine militärischen Erfolge in den Gallischen Kriegen in seinen "Commentarii de Bello Gallico" darstellte,[4] um seinen politischen Werdegang zu begünstigen.

Im 20. Jahrhundert erlebte die Propaganda eine neue Blütezeit, insbesondere während der beiden Weltkriege. Im Ersten Weltkrieg nutzten die Kriegsparteien Plakate, Zeitungen und Filme, um die Moral der Bevölkerung zu stärken und den Feind zu dämonisieren. Ein bekanntes Beispiel ist das britische Propagandaplakat mit dem Slogan "Your Country Needs You", das zur Rekrutierung von Soldaten aufrief.[5]

Der Zweite Weltkrieg brachte eine noch intensivere Nutzung von Propaganda mit sich. Die nationalsozialistische Propaganda unter Joseph Goebbels in Deutschland und die alliierten Propagandakampagnen in den USA und Großbritannien sind eindrucksvolle Beispiele dafür, wie Medien genutzt wurden, um die öffentliche Meinung zu formen und politische

---

* Der Autor ist sicherheitspolitischer Berater bei der Konrad-Adenauer-Stiftung e.V. (KAS), zugelassener Rechtsanwalt und zertifizierter Mediator. Er berät Politiker zu Fragestellungen der Cybersicherheit, Kritischer Infrastrukturen, Hybrider Bedrohungen und Informationsmanipulation.

1 Study: False News spreads faster than the thruth, MIT Management Sloan School, 08. März 2018, abrufbar unter: https://mitsloan.mit.edu/ideas-made-to-matter/study-false-news-spreads-faster-truth [zuletzt abgerufen am 3. September .2024].

2 Bundesamt für Verfassungsschutz, Desinformation als Mittel gezielter Einflussnahme fremder Staaten, abrufbar unter: https://www.verfassungsschutz.de/SharedDocs/hintergruende/DE/spionage-und-proliferationsabwehr/desinformation.html [zuletzt abgerufen am 3. September 2024].

3 *Aro*, Desinformation als Waffe – Über einen Krieg, den Russland seit Jahren führt, APuZ, 08. Juli 2022, abrufbar unter: https://www.bpb.de/shop/zeitschriften/apuz/krieg-in-europa-2022/510257/desinformation-als-waffe/ [zuletzt abgerufen am 3. September 2024].

4 *Caesar*, The Gallic Wars, übersetzt von W. A. McDevitte und W. S. Bohn, abrufbar unter: https://classics.mit.edu/Caesar/gallic.html [zuletzt abgerufen am 3. September 2024].

5 Kitchener: The most famous pointing finger, BBC News, 4. August 2014, abrufbar unter: https://www.bbc.co.uk/news/magazine-28642846 [zuletzt abgerufen am 3. September 2024].

Ziele zu erreichen. Die Nazis nutzten Radio, Filme und Zeitungen, um ihre Ideologie zu verbreiten und dadurch die Bevölkerung zu mobilisieren.[6]

Auch nach dem Zweiten Weltkrieg und während des Kalten Krieges setzten sowohl die USA als auch die Sowjetunion umfangreiche Kommunikationsstrategien ein. Die Verbreitung von Informationen und Desinformationen wurde zu einem zentralen Element der geopolitischen Auseinandersetzung. Die USA nutzten beispielsweise den Radiosender "Voice of America", um ihre Botschaften in die Sowjetunion und andere kommunistische Länder zu senden.[7]

Mit dem Aufkommen des Internets und der sozialen Medien hat die Verbreitung von Desinformation eine neue Dimension erreicht. Die Möglichkeiten, Informationen schnell und weitreichend zu verbreiten, haben es staatlichen und nichtstaatlichen Akteuren ermöglicht, ihre Propagandastrategien zu verfeinern und zu erweitern. Heute sehen wir eine zunehmende Professionalisierung dieser Akteure.

III. Strategien und Techniken

In der modernen Informationslandschaft haben Desinformations- und Propagandastrategien eine neue Dimension erreicht. Die Digitalisierung und die Verbreitung sozialer Medien haben es ermöglicht, dass sich manipulierte Informationen sich schneller und weiterverbreiten lassen, als je zuvor.

1. Digitale Medien und soziale Netzwerke

Soziale Netzwerke wie Facebook, X und Instagram sind zentrale Plattformen für die Verbreitung von Desinformation. Diese Plattformen ermöglichen es, Informationen schnell und weitreichend zu verbreiten, oft ohne ausreichende Überprüfung der Fakten. Algorithmen, die darauf ausgelegt sind, Inhalte mit hoher Interaktionsrate zu priorisieren, begünstigen die Verbreitung sensationeller und oft falscher Informationen. Zudem können Desinformationen durch gezielte Werbeanzeigen verbreitet werden, die spezifische Zielgruppen ansprechen und so die Reichweite erhöhen. Anknüpfungspunkte sind hierbei psychologische Effekte und Emotionen.[8]

2. Deepfakes und künstliche Intelligenz

Die Entwicklung von Deepfake-Technologien hat die Möglichkeiten der Desinformation erheblich erweitert. Deepfakes sind KI-generierte Videos, Text- oder Audiodateien, die täuschend echt wirken und dazu verwendet werden können, falsche Aussagen oder Handlungen von Personen darzustellen. Diese Technologie wird zunehmend genutzt, um politische Gegner zu diskreditieren oder falsche Narrative zu verbreiten.[9] Darüber hinaus ermöglicht es Künstliche Intelligenz, große Mengen an Daten zu analysieren und gezielt Desinformationskampagnen zu steuern.

3. Automatisierte Accounts und Bots

Automatisierte Accounts, auch bekannt als Bots, spielen eine zentrale Rolle bei der Verbreitung von Desinformation. Diese Bots können große Mengen an Inhalten in kurzer Zeit verbreiten oder Inhalte liken und kommentieren und so den Eindruck erwecken, dass bestimmte Meinungen weit verbreitet sind und die Aufmerksamkeit von echten Usern auf diese Inhalte lenken. Sie übernehmen die Aufgabe, gezielt bestimmte Begriffe massenhaft zu „liken", um deren Sichtbarkeit in den Algorithmen der Suchmaschinen zu erhöhen und auf Social-Media-Plattformen maximale Aufmerksamkeit zu erzielen. Bots werden oft eingesetzt, um Hashtags zu trendigen Themen zu erstellen oder um falsche Informationen zu verbreiten. Sie können auch dazu verwendet werden, Diskussionen zu manipulieren, indem sie sich an dem digitalen Meinungsaustausch beteiligen, und dadurch die öffentliche Meinung beeinflussen.

4. Troll-Fabriken

Für Desinformationskampagnen werden oft sogenannte Troll-Fabriken eingesetzt. Diese Strukturen verbreiten irreführende Inhalte und Memes. Sie zielen darauf ab, die Abfolge von Beiträgen, Kommentaren und Interaktionen auf Websites und sozialen Medien zu beeinflussen.

Ein Beispiel für den Grad der Professionalisierung und der Bedeutung, die einzelne Staaten, wie Russland, diesen Strukturen beimessen, ist die Troll-Fabrik in St. Petersburg, auch Internet Research Agency (IRA) genannt. Die Organisation gleicht einem herkömmlichen Unternehmen.[10] Die Mitarbeiter arbeiten im Schichtbetrieb mit Arbeitszeiterfassung. Die IRA infiltriert soziale Netzwerke und Online-Com-

6 Manipulation of the media in the nazi era, Deutsches Buch- und Schriftenmuseum, abrufbar unter: https://mediengeschichte.dnb.de/DBSMZBN/Content/EN/MassMedia/10-medienmanipulation-im-ns-en.html [zuletzt abgerufen am 3. September 2024].

7 *Johnson, A. R., & Parta, R. E.* (Eds.), Cold War Broadcasting: Impact on the Soviet Union and Eastern Europe, 2010, Central European University Press, S. 16 ff.

8 Social Media Algorithms Distort Social Instincts and Fuel Misinformation, Neuroscience News, abrufbar unter: https://neurosciencenews.com/social-media-behavior-misinformation-23752/ [zuletzt abgerufen am 3. September 2024].

9 *F. Gehringer, Dr. C. Nehring, M. Labuz*, Der Einfluss von Deep Fakes auf Wahlen – Legitime Besorgnis oder bloßer Alarmismus?, 2024, KAS-Monitor, S. 2.

10 *Dr. M. Saletta and R. Stearne*: Understanding Mass Influence - A case study of the Internet Research Agency as a contemporary mass influence operation, S. 20.

munities langfristig, um Vertrauen aufzubauen und dann gezielt falsche Informationen zu verbreiten.[11]

5. Kanäle

Aber auch herkömmliche Fernsehsender und offizielle Webseiten sind ein beliebtes Instrument und elementarer Bestandteil von Desinformationsstrategien. Kanäle, die dem Staat zugerechnet werden können, lassen sich grundlegend in drei Kategorien einteilen: Regierungswebsites, staatlich kontrollierte Kanäle und staatlich zurechenbare Kanäle. Beispiele für Russland und China sind die offizielle Regierungswebsite der chinesischen Regierung (gov.cn) und die Website des Präsidenten der Russischen Föderation (kremlin.ru). Staatlich kontrollierte Kanäle umfassen China Central Television (CCTV), die chinesische Nachrichtenagentur Xinhua und die chinesische Zeitung People's Daily, der russische Kanal Russia Today (RT), die russische Nachrichtenagentur TASS und der Fernsehsender Channel One.

Zu den staatlich zurechenbaren Kanälen gehören die Global Times und CGTN (China Global Television Network) in China und Sputnik News, eine Nachrichtenagentur, die oft die Positionen der russischen Regierung widerspiegelt, und NTV, ein Fernsehsender, der oft regierungsfreundliche Inhalte verbreitet, in Russland.

Im Rahmen einer Untersuchung wurde festgestellt, dass bei 100 verschiedenen Desinformationskampagnen der beiden Länder, 616 digitale Kanäle bei der Verbreitung oder Verstärkung beteiligt waren. Etwa 40 Prozent dieser Kanäle konnten entweder China oder Russland zugerechnet werden, während 60 Prozent der Kanäle keiner staatlichen Struktur zugeordnet werden konnten. Vor allem China und Russland investieren viel in ausländische Einflussnahme durch "eigene Medien" und planen das Aufgreifen der Meldung durch andere Kanäle mit in ihren Strategien ein.

IV. Globale Perspektiven

Der Informationsraum ist regional unterschiedlich. Er setzt sich jeweils aus verschiedenen Narrativen, ideologischen Vorstellungen zusammen. Es werden unterschiedliche Mittel, Medien und Kanäle eingesetzt. Demnach variieren Desinformationsstrategien stark nach der Region und den spezifischen Zielen der Akteure. In autoritären Regimen wie Russland und China werden Desinformationskampagnen oft zentral gesteuert und zielen darauf ab, die öffentliche Meinung sowohl im Inland als auch im Ausland zu beeinflussen.[12] Innerhalb westlicher Demokratien hingegen werden Desinformationskampagnen häufig dezentraler gesteuert und können von einer Vielzahl von Akteuren, einschließlich ausländischer staatlicher und nichtstaatlicher Gruppen, durchgeführt werden.

1. Russland

Russland agiert strukturiert und semi-systematisch im Informationsraum. Eine zentrale Taktik ist der sogenannte „Spiegeleffekt", bei dem Kritik an Russland auf andere projiziert wird. Ein Beispiel hierfür ist das Massaker in Butscha, bei dem behauptet wurde, die ukrainische Armee sei verantwortlich und nicht die russischen Truppen. Russlands Kommunikationsstrategie basiert auf vier Säulen: Offizielle Vertreter des russischen Staatsapparats verbreiten gezielt Desinformationen, während staatlich kontrollierte Medien die gewünschten Narrative weiterverbreiten. Ein Desinformationsökosystem, zu dem Organisationen wie die IRA gehören, spielt ebenfalls eine zentrale Rolle. Die russischen Nachrichten- und Geheimdienste sind ebenfalls in sozialen Medien aktiv und schaffen Netzwerke falscher Identitäten.[13]

Moskau erreicht Menschen in anderen Ländern durch eine gezielte Auswahl von Zielgruppen, die anfällig für bestimmte Narrative sind, wie zum Beispiel NATO-Skepsis („Die NATO hat Russland eingekreist, sie ist der Aggressor, und deshalb mussten wir uns verteidigen").[14] Ein weiteres Element ist die ständige Wiederholung bestimmter Narrative, wie dem vom „Nazi-Regime" in Kiew.[15] In südlichen Ländern verbreitet Russland zudem Narrative über die Getreideblockade oder den Ukraine-Krieg, um die öffentliche Meinung zu beeinflussen. Es wird erwartet, dass Russland bald den Fokus auf die USA und den dortigen Wahlkampf legen könnte.

Die oberste Priorität liegt grundsätzlich auf der Dominanz des Informationsraums im eigenen Land, die zweite auf der Ukraine und die dritte auf größere westliche Staaten.[16]

2. China

China verfolgt langfristige Ziele, um das globale Narrativ über das Land zu verändern und US-Interessen

11 *R. DiResta, Dr. K. Shaffer, B. Ruppel, D. Sullivan, R. Matney, R. Fox, Dr. J. Albright und B. Johnson*, The Tactics & Tropes of the Internet Research Agency, 2019, DigitalCommons@University of Nebraska – Lincoln, Oktober 2019, S. 84, abrufbar unter: https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1003&context=senatedocs [zuletzt abgerufen am 3. September 2024].

12 *J. Kurlantzick*, How Russia and China Learn From Each Other on Disinformation, ECFR, 2022, abrufbar unter: https://www.cfr.org/blog/how-russia-and-china-learn-each-other-disinformation [zuletzt abgerufen am 3. September 2024].

13 *Dr. M. Saletta and R. Stearne* (Fn. 11), S. 5 ff.

14 Offene Konfrontation: https://taz.de/Nato-und-Russland/!5810304/ [zuletzt abgerufen am 3. September 2024].

15 *M. Durm*, Rechtsextreme Bewegungen in der Ukraine – Was ist dran an Putins Nazi-Vorwurf?, SWR2 Wisse, 10. April 2024, https://www.swr.de/swrkultur/wissen/rechtsextreme-bewegungen-in-der-ukraine-was-ist-dran-an-putins-nazi-vorwurf-swr2-wissen-2024-04-11-100.html [zuletzt abgerufen am 3. September 2024].

16 *Dr. M. Saletta and R. Stearne* (Fn. 11), S. 18.

zu untergraben. Ein Beispiel ist die sogenannte "50 Cent Army", bei der bezahlte Internetnutzer Diskurse manipulieren.[17] Zudem nutzt China ein Netzwerk koordinierter Accounts auf verschiedenen Plattformen, um verdeckte Propaganda zu verbreiten. Meta hat zwei bedeutende Fälle aufgedeckt, bei denen chinesische Desinformationsnetzwerke entfernt wurden. Im ersten Fall wurden 107 Konten, 36 Seiten und sechs Gruppen auf Facebook sowie 35 Konten auf Instagram gelöscht, die Desinformationen über Taiwan, Südafrika, Japan, Zentralasien und die Uiguren verbreiteten. Im zweiten Fall identifizierte Meta 7704 Konten, 954 Facebook-Seiten und 15 Gruppen, die Nutzer in Taiwan, den USA, Australien, dem Vereinigten Königreich, Japan und chinesischsprachige Nutzer weltweit anvisierten.[18]

China nutzt vier Haupttaktiken, um das Bewusstsein und die Wahrnehmung von Ereignissen bei der jeweiligen Zielgruppe zu beeinflussen. Erstens, durch die Verheimlichung von Informationen werden gezielt solche Ereignisse verbreitet, die Chinas Position unterstützen, um das Verständnis der Zielgruppe zu beeinflussen.

Zweitens, im Rahmen des Diskurswettbewerbs wird die kognitive Struktur der Zielgruppe unbewusst zugunsten Chinas geformt, oft durch „Trolling" in sozialen Medien. Drittens, durch Meinungsmaskierung werden zu bestimmten Zeitpunkten und Ereignissen gezielt öffentliche Meinungen geschaffen, um den Diskurs zu diktieren. Schließlich wird durch Informationsblockaden, einschließlich technischer Angriffe und Blockaden, die Interpretation und das Agenda-Setting monopolisiert.

V. Rahmenbedingungen und Regelwerke

Die Bekämpfung von Desinformation und Propaganda ist eine überregionale und globale Herausforderung, die koordinierte Anstrengungen und umfassende rechtliche Rahmenbedingungen erfordert.

1. EU-Verhaltenskodex und Digital Services Act

Auf EU-Ebene gibt es verschiedene Initiativen, die darauf abzielen, Desinformation und Propaganda zu bekämpfen.

Ein bedeutendes Beispiel ist der EU-Verhaltenskodex zur Bekämpfung von Desinformation, der 2018 eingeführt und 2022 verschärft wurde. Dieser Kodex verpflichtet Online-Plattformen, aktiv gegen Desinformation vorzugehen, indem sie beispielsweise die Transparenz politischer Werbung erhöhen und die Verbreitung von manipulierten Informationen eindämmen. Der Kodex hat sich als wirksames Instrument zur Förderung von Transparenz und Rechenschaftspflicht erwiesen und bietet einen strukturierten Rahmen für die Überwachung und Verbesserung der Strategien der Plattformen zur Bekämpfung von Desinformation.[19]

Der EU-Aktionsplan gegen Desinformation, der auch 2018 entwickelt wurde, zielt darauf ab, die Zusammenarbeit innerhalb der EU-Institutionen und Mitgliedstaaten zu verstärken. Er umfasst Maßnahmen zur Erkennung, Analyse und Aufdeckung von Desinformation sowie zur Sensibilisierung der Öffentlichkeit. Im März 2019 wurde zudem ein Frühwarnsystem eingerichtet, das in Echtzeit und rund um die Uhr alle Mitgliedstaaten vor Desinformationskampagnen warnen kann. Dafür wurden 27 nationale Kontaktstellen eingerichtet, die sich auch inhaltlich austauschen.[20]

Eine weitere bedeutende Initiative ist die Kampagne EUvsDisinfo, die von der East StratCom Task Force im Europäischen Auswärtigen Dienst durchgeführt wird. Diese Initiative zielt darauf ab, Desinformationskampagnen aus Russland aufzudecken und aktiv zu widerlegen.[21] Ergänzend dazu verfügt die Europäische Beobachtungsstelle für digitale Medien (EDMO) über Drehkreuze, die alle EU-Länder abdecken und die Zusammenarbeit zwischen verschiedenen Akteuren fördern, um Desinformation effektiv zu bekämpfen.[22]

Ein weiteres wichtiges Instrument ist der Digital Services Act (DSA) der Europäischen Union, der darauf abzielt, die Verbreitung illegaler Inhalte, einschließlich Desinformation, zu bekämpfen. Der DSA verpflichtet große Online-Plattformen, Maßnahmen zur Erkennung und Entfernung von Desinformation zu ergreifen und regelmäßige Berichte über ihre Bemühungen vorzulegen. Der DSA trat am 16. November

17 *L. Joyce Lok-Teng*, „50-Cent-Armee" in Aktion: Desinformation aus China, dlf, 07. März 2023, abrufbar unter: https://www.deutschlandfunk.de/50-cent-armee-in-aktion-desinformation-aus-china-1-dlf-b6df2473-100.html [zuletzt abgerufen am 3. September 2024].

18 *S. Zurier*, Meta takes down 'Spamouflage' – a prolific Chinese disinformation campaign, SC Magazine, 29. August 2023, abrufbar unter: https://www.scmagazine.com/news/meta-takes-down-spamouflage-a-prolific-chinese-disinformation-campaign [zuletzt abgerufen am 3. Sepetember 2024]

19 Siehe hierzu Verhaltenskodex 2022 für Desinformation, abrufbar unter: https://digital-strategy.ec.europa.eu/de/policies/code-practice-disinformation [zuletzt abgerufen am 3. September 2024].

20 Siehe hierzu Gemeinsam gegen Desinformation, abrufbar unter: https://www.bundesregierung.de/breg-de/schwerpunkte/umgang-mit-desinformation/eu-desinformation-1875918 [zuletzt abgerufen am 3. September 2024].

21 Siehe hierzu Countering Disinformation, abrufbar unter: https://www.eeas.europa.eu/eeas/questions-and-answers-about-east-stratcom-task-force_en [zuletzt abgerufen am 3. September 2024].

22 Siehe hierzu Europäisches Netzwerk gegen Desinformation, abrufbar unter: https://www.bundesregierung.de/breg-de/aktuelles/netzwerk-gegen-desinformation-2153606 [zuletzt abgerufen am 3. September 2024].

2022 in Kraft und ist seit dem 17. Februar 2024 anwendbar.[23]

Der DSA verfolgt mehrere Ziele: Er verpflichtet digitale Dienstleister, insbesondere große Online-Plattformen, zu mehr Transparenz und Rechenschaftspflicht. Plattformen müssen beispielsweise regelmäßige Berichte über ihre Maßnahmen zur Bekämpfung von Desinformation und illegalen Inhalten vorlegen. Zudem stärkt der DSA die Rechte der Nutzer, indem er Maßnahmen gegen sogenannte „dark patterns" einführt, die Nutzer zu ungewollten Entscheidungen verleiten könnten. Mechanismen zur Beschwerde und Streitbeilegung werden ebenfalls eingeführt. Große Plattformen und Suchmaschinen, die monatlich mindestens 45 Millionen aktive Nutzer erreichen, sind verpflichtet, Risikoanalysen durchzuführen und Maßnahmen zur Risikominimierung zu ergreifen.[24]

Der DSA hat das Potenzial, die Verbreitung von Desinformation und illegalen Inhalten erheblich zu reduzieren. Durch die Einführung strengerer Transparenz- und Rechenschaftspflichten werden Plattformen gezwungen, proaktiver gegen schädliche Inhalte vorzugehen. Erste Berichte deuten darauf hin, dass die neuen Regelungen bereits zu einer verbesserten Moderation und Entfernung illegaler Inhalte geführt haben.[25]

Trotz der umfassenden Regelungen gibt es noch einige Herausforderungen und potenzielle Regelungslücken. Die Wirksamkeit des DSA hängt stark von der Durchsetzung durch die Mitgliedstaaten ab. Da der DSA für alle digitalen Dienste gilt, die im EU-Binnenmarkt tätig sind, unabhängig davon, ob sie in der EU ansässig sind oder nicht, ist eine effektive grenzüberschreitende Zusammenarbeit erforderlich. Dies könnte in der Praxis schwierig sein. Zudem muss der DSA kontinuierlich an neue technologische Entwicklungen angepasst werden, um wirksam zu bleiben. Insbesondere die rasante Entwicklung von KI-Technologien und neuen Formen der Desinformation stellen eine Herausforderung dar.[26] Es bleibt abzuwarten, wie effektiv die nationalen Behörden die neuen Vorschriften umsetzen und überwachen werden.

In Deutschland wurde das Netzwerkdurchsetzungsgesetz (NetzDG) eingeführt, das soziale Netzwerke verpflichtet, offensichtlich rechtswidrige Inhalte innerhalb von 24 Stunden zu entfernen. Das Gesetz zielt darauf ab, die Verbreitung von Hassrede und Desinformation zu reduzieren. Das NetzDG wird nun mehr durch den DSA der Europäischen Union faktisch ersetzt, da der DSA umfassendere Regelungen und Pflichten für Online-Plattformen einführt.[27]

Die Bundesregierung hat das Digitale-Dienste-Gesetz auf den Weg gebracht, um die nationalen Vorschriften an die neuen europarechtlichen Vorgaben des DSA anzupassen.

Zusätzlich hat die Bundesregierung eine ressortübergreifende Arbeitsgruppe zur Bekämpfung hybrider Bedrohungen, einschließlich Desinformation, eingerichtet. Diese Gruppe konzentriert sich auf die Erkennung und Abwehr von Desinformation im Zusammenhang mit dem russischen Angriffskrieg gegen die Ukraine.

VI. Lösungsansätze

Die Verbreitung von Falschinformationen kann das Vertrauen in demokratische Institutionen untergraben, gesellschaftliche Spaltungen vertiefen und die öffentliche Meinung manipulieren. Um diesen Bedrohungen wirksam zu begegnen, sind umfassende Strategien und Maßnahmen erforderlich

1. Detektion und Reaktion

Die Erkennung und Bekämpfung von Desinformation erfordern eine Kombination aus technologischen, analytischen und kommunikativen Maßnahmen.

Zu den wichtigsten Methoden gehören die Faktenprüfung, bei der unabhängige Faktenchecker eine zentrale Rolle spielen. Plattformen wie Correctiv und Faktencheck.org bieten umfassende Analysen und Richtigstellungen.[28]

Künstliche Intelligenz (KI) ist ebenfalls ein bedeutendes Werkzeug, da KI-Systeme große Datenmengen analysieren und Muster erkennen können, die auf Desinformation hinweisen. Algorithmen zur Textanalyse und Bilderkennung helfen dabei, Falschinforma-

23 Siehe hierzu Gestaltung der digitalen Zukunft Europas, abrufbar unter: https://digital-strategy.ec.europa.eu/de/policies/digital-services-act-package [zuletzt abgerufen am 3. September 2024].

24 Vgl. Das Gesetz über digitale Dienste, abrufbar unter: https://www.bundesregierung.de/breg-de/themen/digitalisierung/gesetz-ueber-digitale-dienste-2140944 [zuletzt abgerufen am 3. September 2024]; siehe hierzu auch ***Möller-Klapperich***, NJ 2024, 193, 197 f.

25 DSA: Welche Regulierung bringt der Digital Services Act für Plattformbetreiber?,IHK Ratgeber, abrufbar unter: https://www.ihk-muenchen.de/de/Service/Recht-und-Steuern/Internetrecht/dsa-regulierung-von-plattformen/ [zuletzt abgerufen am 3. September 2024].

26 *T. Hartmann*, EU-Kommission vor großen Herausforderungen bei Umsetzung vom DAS, EURACTIV, 30. August 2023, abrufbar unter: https://www.euractiv.de/section/innovation/news/eu-kommission-vor-grossen-herausforderungen-bei-umsetzung-vom-dsa/ [zuletzt abgerufen am 3. September 2024].

27 Siehe hierzu Sicher im Netz unterwegs, abrufbar unter: https://www.bundesregierung.de/breg-de/themen/digitalisierung/digitale-dienste-gesetz-2250526#:~:text=Das%20Digitale%2DDienste%2DGesetz%20trat,Digitale%2DDienste%2DGesetz%20geregelt [zuletzt abgerufen am 3. September 2024].

28 *U. Jonas , S. Timmermann*, Kampf gegen Desinformation: Maßnahmen von Tech-Konzernen greifen zu kurz, 13. Juni 2023, abrufbar unter: https://correctiv.org/in-eigener-sache/2023/06/13/kampf-gegen-desinformation-massnahmen-von-tech-konzernen-greifen-zu-kurz/ [zuletzt abgerufen am 3. September 2024].

tionen automatisch zu identifizieren und herauszufiltern.[29]
Ein weiterer wichtiger Aspekt ist die Transparenz und Zusammenarbeit. Plattformen wie Facebook und X arbeiten zunehmend mit Regierungen und unabhängigen Organisationen zusammen, um Desinformation zu bekämpfen. Diese Zusammenarbeit umfasst den Austausch von Daten und die Entwicklung gemeinsamer Strategien. Bisher reichen die personellen und finanziellen Investitionen der Plattformen nicht aus. Ihnen wird aufgrund ihres großen Wirkungsgrades eine besondere Rolle zuteil.
Durch nachträgliche Maßnahmen (Debunking) können falsche Informationen korrigiert und die Wahrheit wiederhergestellt werden.

2. Prävention

Bildung und Aufklärung sind von entscheidender Bedeutung, um die Resilienz gegen Desinformation zu stärken. Zu den wichtigsten Maßnahmen gehört die Förderung der Medienkompetenz in Schulen und Bildungseinrichtungen. Diese Programme sollen das kritische Hinterfragen von Quellen und das Erkennen von Falschinformationen lehren.[30]
Doch nicht nur der Umgang mit Medien und deren Inhalte ist entscheidend, vielmehr bedarf es auch einer umfassenden Einbettung in geopolitische Zusammenhänge. Die Akteure und deren Strategien und Ziele sollten hierbei ebenfalls beleuchtet werden. Problematisch hingegen ist die Ausbildung und Schulung der Lehrkräfte, damit diese das nötige Wissen in den Bildungseinrichtungen vermitteln können. Zugleich könnten Influencer Aufklärungsarbeit leisten und als Multiplikatoren fungieren. Die Bundesregierung könnte eine „Social-Impact-Initiative" gründen und so die Attraktivität für ein Engagement im Rahmen der Aufklärungsarbeit zu steigern. An besonders aktive Influencerinnen und Influencern wird dann jährlich ein Social-Impact-Award verliehen. Weitergehend wird vor allem die ältere Generation nicht mehr über Bildungseinrichtungen erreicht. Über altersgerechte Informationsvermittlung auf Marktplätzen, vor Supermärkten, in Kirchengemeinden oder auf Dorffesten kann diese Zielgruppe besser erreicht werden. In dieser Hinsicht ist noch großer Spielraum.
Außerdem kann TV- und Radiowerbung zu mehr Aufklärung führen. Regierungen und NGOs können öffentliche Kampagnen starten, um die Bevölkerung über die Gefahren von Desinformation aufzuklären und ihnen Werkzeuge an die Hand zu geben, um diese zu erkennen.

3. Einsatz technologischer Lösungen

Auch technologische Entwicklungen spielen eine zentrale Rolle im Kampf gegen Falschinformationen. Zu den bedeutendsten technologischen Lösungen zählen automatisierte Erkennungssysteme, die mithilfe von Algorithmen und KI-Systemen Desinformationen in Echtzeit erkennen und markieren können. Diese Systeme analysieren Texte, Bilder und Videos auf Anzeichen von Manipulation. Verifikationstools wie InVID und Factmata unterstützen Nutzer dabei, die Echtheit von Bildern und Videos zu überprüfen und sind besonders wertvoll, um Deepfakes und andere manipulierte Medien zu identifizieren.[31] Zudem implementieren soziale Netzwerke Mechanismen, um die Verbreitung von Desinformationen zu verlangsamen. Dazu gehören Warnhinweise und die Reduzierung der Sichtbarkeit verdächtiger Inhalte.

VII. Fazit

Die Bekämpfung von Desinformation und Propaganda ist eine der größten Herausforderungen unserer Zeit. In einer zunehmend vernetzten Welt, in der Informationen blitzschnell verbreitet werden können, sind die Auswirkungen von Falschinformationen tiefgreifend und weitreichend. Sie können das Vertrauen in demokratische Institutionen untergraben, gesellschaftliche Spaltungen vertiefen und die öffentliche Meinung manipulieren. Autoritäre Regime wie China und Russland nutzen Desinformation als strategisches Werkzeug, um ihre geopolitischen Ziele zu erreichen und die Stabilität demokratischer Gesellschaften zu untergraben. Ihre professionellen und gut koordinierten Kampagnen zeigen, wie wichtig es ist, internationale und nationale Maßnahmen zur Bekämpfung von Desinformation zu verstärken.

Um den Bedrohungen wirksam zu begegnen, bedarf es einer Kombination aus technologischen, bildungsbezogenen und rechtlichen Maßnahmen. Die Erkennung und Bekämpfung von Desinformation erfordert den Einsatz fortschrittlicher Technologien wie KI und automatisierter Erkennungssysteme. Gleichzeitig ist die Förderung der Medienkompetenz und die Aufklärung der Bevölkerung von entscheidender Bedeutung, um die Resilienz der Gesellschaft gegen Falschinformationen zu stärken.

Letztlich erfordert die Bekämpfung von Desinformation und Propaganda eine gemeinsame Anstrengung von Regierungen, Technologieunternehmen, Bildungseinrichtungen und der Zivilgesellschaft. Die Strategien der feindlichen Akteure sind klar, auch auf diese gilt es sich im Besonderen einzustellen. Nur durch koordinierte und umfassende Maßnahmen können die Integrität unserer Informationslandschaft geschützt und die Werte unserer demokratischen Gesellschaften bewahrt werden.

---

29 *M. Wiegmann*, Die rettende Lösung gegen Desinformation?, tagesschau.de, 06. März 2023, abrufbar unter: https://www.tagesschau.de/faktenfinder/kontext/ki-gegen-desinformation-101.html [zuletzt abgerufen am 3. September 2024].

30 Siehe hierzu bspw. Medienkompetenzen stärken: „Stop Fake News!", abrufbar unter: https://www.schulministerium.nrw/medienkompetenzen-staerken-stop-fake-news [zuletzt abgerufen am 3. September 2024].

31 Siehe hierzu InVID Verification Plugin, abrufbar unter: https://www.invid-project.eu/tools-and-services/invid-verification-plugin/ [zuletzt abgerufen am 3. September 2024].

Die Psychologie der Misinformation – Verstehen und Abwehren von Misinformationen durch eine verhaltensorientierte Sichtweise und Künstliche Intelligenz

Dr. Rakoen Maertens, Oxford*

I. Die Psychologie hinter Künstlicher Intelligenz[1] und Misinformation

Um Künstliche Intelligenz grundlegend verstehen zu können, ist es sinnvoll, ein einen Schritt zurück zu gehen und zunächst den Menschen als solches zu verstehen. Denn ein Großteil von KI basiert auf neuronalen Netzen und damit auch auf menschlichem Verhalten. Oftmals werden die großen Sprachmodelle an menschlichen Texten, an menschlicher Sprache und an von Menschen produzierten Inhalten trainiert.

Psychologen befassen sich mit dem menschlichen Verhalten. Sie setzen sich mit der Fragestellung auseinander, was Menschen einerseits tatsächlich tun, was sie andererseits denken und welche Verbindung zwischen diesen beiden Aspekten besteht. Derartige Fragestellungen mögen Psychologen. Denn es zeigt diese klare Verbindung zwischen dem Glauben an bestimmte Dinge wie z.B. „weil meine Vorgesetzten gesagt haben, dass 5G alle möglichen Probleme verursacht" und einem dazu entsprechenden Verhalten, nämlich dem Verüben von Brandanschlägen auf Mobilfunktürme. An dieser Stelle ist ein eindeutiger Zusammenhang zu bejahen. In den meisten Fällen ist das jedoch weniger offensichtlich. Manchmal verhalten sich Menschen, ohne wirklich zu verstehen, warum sie das plötzlich tun, oder sie glauben an bestimmte Dinge, aber sie handeln nicht entsprechend. Und genau diese Verbindung zwischen diesen beiden Aspekten ist höchst spannend.

1. Misinformationen im psychologischen Sinne

Wenn ich als Psychologe von Misinformation[2] spreche, dann meine ich das stets unabhängig von der Absicht. Für Psychologen sind Misinformationen sämtliche Informationen, die falsch oder irreführend sind, ob sie nun beabsichtigt sind oder nicht. Das unterscheidet sich von der juristischen Definition. Wenn ich also von Misinformation spreche, dann beziehe ich mich auf jede Art davon. Darüber hinaus unterscheiden wir in der Psychologie zwischen expliziten und impliziten Misinformationen. Es gibt also zwei Haupttypen, mit denen wir uns beschäftigen. Und wir entwickeln verschiedene Interventionen, um ihnen zu begegnen.

a. Arten von Misinformationen

Explizite, also eindeutige Misinformationen sind diejenigen, die man mit wissenschaftlichen Methoden als wahr oder falsch beweisen kann. Typischerweise handelt es sich dabei um Behauptungen wie die „Die Erde ist flach." Das kann man nachprüfen und verneinen. Aber das ist nur der kleinere Teil von Misinformationen.

Bei den ganz überwiegenden impliziten Misinformationen handelt es sich um Aussagen, die entweder wahr oder falsch oder auch solche sind, die zwar buchstäblich wahr, aber dennoch irreführend sein könnten. Ein Beispiel für eine derartige Aussage ist „Menschen sind aufgrund des Covid-19-Impfstoff gestorben. Es ist gefährlich ihn einzusetzen." Wenn jemand zu einer Familie gehört, in der ein Familienmitglied an Blutgerinnseln, die durch den AstraZeneca-Impfstoff verursacht wurden, gestorben ist, dann könnte derjenige behaupten, dass dies eine wahre Aussage sei. Schaut man allerdings auf die Gesellschaft als Ganzes, könnte es wesentlich gefährlicher sein, den Impfstoff nicht einzusetzen als ihn einzusetzen. Das sind also irreführende Aspekte.

Eine dritte Art von Misinformationen sind Inhalte, die so generiert werden, dass wir nicht genau wissen, ob sie aus einer echten Quelle stammen oder nicht, also nicht-authentische Informationen. Ein Beispiel für solche Inhalte ist ein vollständig von einer KI generierte Bild von Papst Franziskus in einer Kapuzenjacke, dassehr realistisch aussieht und daher viral ging.[3]

---

* Der Autor ist Post-Doc am Institut für Experimentelle Psychologie der Universität Oxford. Sein Vortrag auf den 20. Frankfurter Medienrechtstagen wurde für die Veröffentlichung übersetzt und bearbeitet von Margarita Hamann, Studien- und Forschungsschwerpunkt Medienrecht der Juristischen Fakultät der Europa-Universität Viadrina Frankfurt (Oder). Alle Internet-Quellen wurden zuletzt am 6. September 2024 abgerufen.

1 Im Nachfolgenden weitgehend als „KI" abgekürzt.

2 Vgl. Usvatov, NJ 2024, B XX ff.

3 Vgl. https://www.zdf.de/nachrichten/panorama/prominente/papst-daunenjacke-fake-ki-kuenstliche-intelligenz-100.html.

b. Arbeitsweise von Psychologen

Wir betrachten Inhalte aus drei unterschiedlichen Blickwinkeln, man kann auch von drei Dimensionen sprechen: Genauigkeit, Manipulierbarkeit und Authentizität. Diese Betrachtungsweise ist für das Verständnis und den Umgang, aber auch für die Zukunft von KI wichtig. Wenn wir als Psychologen Menschen darin schulen, Misinformationen zu erkennen, schauen wir uns oft die verschiedenen Techniken an, mit denen Menschen desinformiert oder fehlinformiert werden. In meiner Tätigkeit beschäftige ich mich mit Diskreditierung, Appell an Emotionen, polarisierende Sprache, Nachahmung anderer sowie Verschwörungstheorien und Trolling.

c. Arbeitsweise des menschlichen Verstandes

Zunächst einige Bemerkungen zum Kern der Psychologie. Vereinfacht ausgedrückt, arbeitet der Verstand mit zwei Systemen. Das erste ist das System, das wir 95 % der Zeit nutzen. Es ist im Grunde unser „intuitives System", mit dem wir die meisten unserer Entscheidungen treffen.[4] Schließlich setzen wir uns nicht wirklich hin und denken rational darüber nach, was wir gerade tun. Wir entscheiden uns zwar, ob wir uns hinsetzen und die Beine übereinanderschlagen wollen oder nicht, aber wir denken nicht wirklich aktiv darüber nach. Es ist ein automatisches System. Wenn wir allerdings das rationale Denken von Menschen aktivieren, sind wir ziemlich gut darin, nur tun wir es nicht oft und das ist auch Teil des Problems. Deshalb versuchen viele psychologische Interventionen, die Aktivierung von System zwei zu erreichen, wenn es wichtig ist. Wir wissen bereits, dass sich Misinformationen schneller verbreiten als echte Nachrichten. Ein wichtiger Grund dafür ist auch, dass Menschen dazu neigen, sich mehr mit negativen Inhalten zu beschäftigen, weil das von evolutionärem Wert ist. Wenn man überleben will, will man ganz genau wissen, ob das eigene Haus in Flammen steht oder nicht. Auch außerhalb sozialer Medien erhalten Misinformationen, negative Informationen und vor allem Informationen mit moralischem Inhalt, also emotionale Inhalte, mehr Aufmerksamkeit und verbreiten sich daher auch schneller. Teils ist das ein Problem der Algorithmen, wie auch das der menschlichen Natur. Es ist nicht so, dass wir negative Informationen mögen. Fragt man danach, würden die meisten Menschen positive Informationen vorziehen. Dennoch beschäftigen wir uns mehr mit negativen Informationen. Diese Dynamik spiegelt sich auch in den Algorithmen auf den unterschiedlichen Plattformen und den Gewinnen wieder, die mit dem Engagement erzielt werden können.

d. Der „rationale" Verstand an einem Beispiel

Ein weiteres Beispiel für System zwei in Aktion, aber auch, um die Schwierigkeiten bei der Erforschung von Verhaltensweisen und Einstellungen ist die Auseinandersetzung bezüglich der Teilnehmerzahlen bei der Amtseinführung von Donald Trump 2017 im Vergleich zur Amtseinführung von Barack Obama.[5]

Sobald man den Teilnehmern in den USA erklärte, dass auf dem einen Bild die Menschenmenge bei der Amtseinführung von Donald Trump und auf dem anderen Bild die Menschenmenge bei der Amtseinführung von Barak Obama sei, wird man feststellen, dass einige Trump-Wähler behaupten würden, dass auf dem Bild zur Amtseinführung von Trump eindeutig mehr Leute zu sehen seien. Hierbei handelt es sich um ein konkretes Beispiel mit Trump und Obama, aber auch umgekehrt wird man ähnliche Auswirkungen feststellen können. Es ist also keinesfalls so, dass die Trump-Wähler extrem leichtgläubig seien. Jeder von uns ist leichtgläubig und jeder drückt sich gerne aus. Dringt man tiefer in das Problem ein, wird es noch interessanter. Einige Leute ändern plötzlich ihre Meinung und beginnen tatsächlich zu glauben, dass irgendwo an den Seiten versteckt noch mehr Menschen sind, die nicht auf den ersten Blick erkennbar sind. Andere glauben es zwar nicht wirklich, aber würden trotzdem sagen, dass auf dem einen Bild mehr Menschen abgebildet sind. Diese Menschen tun dies, weil es ihnen einen sozialen Vorteil verschafft, solche Dinge zu sagen. Das dokumentiert die häufige Diskrepanz zwischen dem Verhalten und dem tatsächlichen Glauben.

2. Kognitive Informationsüberlastung und Mechanismen des implizierten Gedächtnisses

Wir werden mit Informationen überflutet. Wir bezeichnen das als Informationsüberlastung, die zur kognitiven Überlastung führt. Wie können wir also das zweite System, das Energie für das rationale Denken benötigt, bei der schieren Menge an Informationen wirklich weiter nutzen? Erhalten wir zu viele Informationen, greifen wir auf automatische, instinktive Verhaltensweisen zurück. Vieles hängt dann von der Bestätigungstendenz ab, also der Suche nach der Bestätigung dessen, was wir bereits wissen, einer motivierten Argumentation, die sogar Dinge diskreditiert, die nicht mit dem eigenen Weltbild übereinstimmen, weil es zu viel Aufwand bedeutet, all die neuen Informationen zu integrieren, die nicht mit der eigenen Wahrnehmung übereinstimmen.

Es gibt auch weitere Mechanismen, die problematisch sind, wie z. B. die Wahrheitsseffekte. Stellen Sie sich vor, Sie wollen grüne Äpfel kaufen, die gesund, vitamin – und ballaststoffreich sind. Und nun stellen Sie sich ein Bananenhandelsunternehmen vor, dass

4 Vgl. https://neurofied.com/thinking-fast-slow-down-system-1-and-2/.

5 Vgl. https://www.spiegel.de/politik/ausland/amtseinfuehrung-von-donald-trump-fotos-wurden-von-behoerden-bearbeitet-a-1227163.html.

den Verkauf von Bananen steigern und keine Äpfel verkaufen will. Sie könnten eine Anzeige schalten oder auch eine Misinformationskampagne starten, in der es hieße, Äpfel seien giftig oder gefährlich. Die meisten von uns würden sagen, dass das völliger Unsinn sei und sich weigern, daran zu glauben. Wiederholt man diese Aussage allerdings immer wieder in verschiedenen Quellen und an verschiedenen Orten, wird man das nächste Mal beim Einkauf möglicherweise geneigt sein, eher eine Banane als einen Apfel zu kaufen. Man wüsste zwar, dass es Fake News sind, doch hätte man ein Bild von giftigen Äpfeln im Kopf. Müsste man dann eine Entscheidung treffen, würden manche nachgeben oder sogar anfangen dran zu glauben. Aus einer von fünf Personen, würden es durch die ständige Wiederholung plötzlich zwei von fünf werden. Wenn man etwas nur oft genug wiederholt, gilt es nach und nach als Norm. Das ist ein einfacher Wiederholungseffekt.

II. Interventionen gegen Misinformationen

Wenn man sich die Interventionen ansieht, um Misinformationen entgegenzuwirken, können wir mit dieser Information etwas auf der Zeitachse vor, während oder nach der Exposition tun. Normalerweise denken wir bei der Bekämpfung von Misinformationen lediglich an die Zeit nach der Exposition. Es sei denn, wir sprechen über Medienkompetenztraining, das betrifft die Zeit vor der Exposition. Aber wenn man über solche Maßnahmen nachdenkt und darüber, ob sie funktionieren, muss man berücksichtigen, dass einige der Maßnahmen zu Veränderungen von Überzeugungen führen, während andere Veränderungen im Online-Verhalten bewirken. Bei wiederum anderen geht es um Verhaltensänderungen im wirklichen Leben und diese sind nicht perfekt miteinander verknüpft. Man muss sich dann ernsthaft fragen, ob es sich dann um eine gute Intervention handelt, wenn man zwar davon überzeugt wird, dass der Impfstoff gut für einen ist, sich aber dennoch weniger häufig impfen lässt. Ist eine Intervention wirklich sinnvoll, wenn man zwar mehr an die Wahrheit glaubt, sich das Verhalten aber trotz dessen verschlechtert? Man muss also berücksichtigen, was tatsächlich eine gute Intervention ist und sich fragen, was die Folgen für jede dieser Dimensionen sein können. Es gibt unterschiedliche Studien in denen sich beispielsweise abzeichnet wie eine Intervention dazu führt, dass die Menschen einerseits skeptischer gegenüber Fake News, aber gleichzeitig auch skeptischer gegenüber echten Nachrichten werden. Ist das dann als eher positiv oder negativ zu beurteilen? Bei Interventionen gilt es also, solche interessanten Dilemmas zu lösen.

1. „Debunking" – die Entlarvung von Misinformationen

Bei der Entlarvung von Misinformationen, dem „Debunking", einer der häufigsten Interventionen, kann eine ganze Menge schiefgehen. Es gibt verschiedene Arten potenzieller Rückschlageffekte.[6] Wiederholt man Mythen zu oft, könnte es passieren, dass den Mythen mehr Glauben geschenkt wird. Denn echte Nachrichten könnten so verwirrend und schwierig oder einfach nicht mit dem Weltbild vereinbar sein, dass man Fake News vorzieht. Selbst wenn eine Sache kurios erscheint, kann es immer noch passieren, dass Fake News einen Einfluss auf jemanden haben wie am Beispiel mit den vermeintlich giftigen Äpfeln veranschaulicht. Natürlich gibt es Mittel und Wege, Fake News sehr effizient zu entlarven. Beispielhaft sei das etwa 15 Seiten lange Debunking-Handbuch 2020 genannt.[7]

2. „Accuracy nudge" – Anregung zum Nachdenken

Die andere Lösung könnte darin bestehen, schon während der Veröffentlichung etwas zu unternehmen. Betrachtet man X (früher:„Twitter") oder diverse andere Plattformen, wird dies bereits angestrebt, wenn gefährliche Schlagzeilen geteilt werden sollen. Dabei wird der Nutzer gefragt, ob der Inhalt wirklich geteilt werden soll und die Information auch vertrauenswürdig ist. Das bezeichnen wir als einen „accuracy nudge". Die Idee dahinter ist, das rationale System zum Nachdenken zu motivieren. Das Vorgehen funktioniert bis zu einem gewissen Grad, aber die Wirkung lässt auf lange Sicht oft schnell nach.

3. Inokulationstheorie – bewusste Schadensprävention

Und schließlich folgt mein Forschungsbereich: die Schadensverhütung. Meine Dissertation befasste sich mit dem Thema der mcguierischen Inokulationstheorie.[8] *William J. McGuire* hat versucht, herauszufinden, warum Kriegsgefangene von der Propaganda der jeweils anderen Seite überzeugt waren. Dieses Phänomen hat McGuire auf beiden Seiten gleichermaßen beobachtet und es zunächst nicht verstanden. Er fragte sich, was dafür ursächlich und ob das US-amerikanische Bildungssystem nicht schon ausreichend gut war. Während seiner Forschung fand er heraus, dass wir mit den Misinformationen, die auf uns zukommen, in der Regel nicht vertraut sind. Wenn man mit solchen Informationen nicht vertraut ist, hat man auch keine Gegenargumente entwickelt. So fängt man unter Umständen an, die ursprünglichen Erstinformationen zu glauben. Denn Menschen neigen dazu, Informationen eher zu glauben, als sie zu widerlegen,

6 Sog. „Backfire-Effekt" bzw. Bumerang-Effekt.

7 Vgl. https://skepticalscience.com/translationblog.php?n=4886&l=6.

8 *Maertens*, The Long-Term Effectiveness of Inoculation Against Misinformation, https://www.repository.cam.ac.uk/handle/1810/344848, https://doi.org/10.17863/CAM.92273.

wenn sie diese zum ersten Mal wahrnehmen. Eine Möglichkeit, Menschen mit einem „Impfsystem" zu schützen, besteht also darin, ihnen bewusst zu machen, dass sie für Angriffe durch Misinformationen anfällig sind, und ihnen zahlreiche Beispiele für die Art von Misinformationen zu präsentieren, mit der sie in Zukunft konfrontiert sein könnten. Wenn sie dann auf ein konkretes, ähnliches Beispiel stoßen, hätten sie dieses interne Verteidigungssystem zu ihrer Verfügung.

a. Beispiel für eine textbasierte Intervention

Auch wir haben die Möglichkeit textbasierte, spielerische sowie videobasierte Interventionen mit Google durchführen. Ein kurzes Beispiel: Ich beschäftige mich viel mit Misinformationen über den Klimawandel. Der wissenschaftliche Konsens zum Klimawandel liegt bei einer siebenundneunzigprozentigen Wahrscheinlichkeit, dass es sich um eine vom Menschen verursachte, globale Erwärmung handelt. Wenn man sich nun den wahrgenommenen wissenschaftlichen Konsens ansieht, dann sind es in den Vereinigten Staaten 72 %. Wir könnten an dieser Stelle versuchen, dies zu korrigieren und uns die Frage stellen, was passiert, wenn man über den tatsächlichen wissenschaftlichen Konsens von 97 % aufklären würde. Die gute Nachricht ist, die Verbreitung echter Nachrichten hilft tatsächlich. Die Verbreitung von Misinformationen dagegen verschlimmert den negativen Effekt zusätzlich. Keinerlei Wirkung erzielt man, wenn man falsche und echte Nachrichten nebeneinander präsentiert, denn sie heben sich gegenseitig auf. Warnt man allerdings nach der Konfrontation mit echten Nachrichten darüber, dass man einer falschen Petition über den Klimawandel ausgesetzt sein könnte, die versucht hinsichtlich des Konsens zu verwirren und begründet man anschließend, aus welchen Gründen diese nicht richtig sei und kommt dann einen weiteren Monat später erneut auf die Befragten zu und setzt sie einer Misinformation aus, wird die Wirkung fast vollständig neutralisiert. Man kann also durchaus präventive Gegenbotschaften kreieren. Es ist zu hoffen, dass KI künftig bei der Erstellung derartiger Botschaften hilft.

b. Spielerische Interventionen

Selbstverständlich müssen wir das Ganze noch ein wenig ausbauen. Und so entwickeln wir auch einige Spiele, die sich vermehrt auf den Schutz auf technischer Ebene sowie dem Schutz gegen emotionale Sprache und allgemeine oder polarisierende Sprache konzentrieren. „Getbadnews.com"[9] ist eines solcher Spiele. Es handelt sich um ein 15-minütiges Spiel bei dem man etwa sechs Techniken erlernt. Es hat bereits ungefähr zwei Millionen Spieler. „Cranky Uncle"[10] ist ein weiteres Spiel, das von einem anderen Entwickler stammt, der sich unter andrem gezielt mit Fehlinformationen zum Klimawandel beschäftigt. Beide Spiele eignen sich als hervorragende und moderne Maßnahmen zur Förderung der Medienkompetenz. Auch „Spot the Troll"[11] reiht sich hier ein. Das Spiel wurde von einem Kollegen von mir, Jeff Lees, entwickelt. Lees befasst sich mit authentischen und gefälschten Konten auf der Plattform X (früher: „Twitter"). Bei dem Spiel handelt es sich um ein einfaches Quiz mit sieben Fragen, ob ein echter oder ein gefälschter Twitter-Account vorliegt. Ich kann garantieren, dass es fast unmöglich sein wird, die Höchstpunktzahl zu erreichen. ZahlreicheExperten auf dem Gebiet der Misinformationen haben drei oder vier von sieben Punkten erreicht.

III. Der Einsatz von Künstlicher Intelligenz in meiner Forschung

1. Eine Anekdote aus meiner lehrenden Tätigkeit

Ich begann, mich für künstliche Intelligenz zu interessieren als ich 2019 über GPT2 las, zu einer Zeit als es noch nicht in aller Munde war. Dabei habe ich mir verschiedene Dinge angeschaut. Wenn ich beispielsweise an den Universitäten Oxford oder Cambridge dazu lehre, beginne ich dies oft mit Bildgenerationen. Wenn sie beispielsweise die Wortabfolge „University of Cambridge überschwemmt echtes Foto" in DALLE-2[12] eingeben, erhalten sie ein ziemlich realistisches Bild von einem Cambridge das unter Wasser steht. Genau ein solches, fotorealistisches Bild habe ich in meinen Studenten gezeigt. Ich werde nie vergessen, wie ich diese Vorlesung zum ersten Mal hielt und wie mich ein paar Wochen darauf einer der Studenten bei einer Supervisionssitzung fragte, wo ich das Foto eines überfluteten Cambridge gefunden hätte. Nach meiner Antwort, dass es sich um ein mittels KI generierten Bildes handle, war der Student sichtlich überrascht, dass es kein echtes Foto war, glaubte mir allerdings. Das lag wahrscheinlich daran, dass ich den fotorealistischen Aspekt in die Generierungsbedingungen eingefügt habe. Auf diese Weise habe ich tatsächlich jemanden falsch informiert, definitiv unabsichtlich. Zum Glück hatte das keine allzu großen Auswirkungen, es ist allerdings ein passendes Beispiel dafür, was man tun und wie gefährlich das potenziell werden könnte. Ein anderes Beispiel ist ein veröffentlichtes Deepfake von Wolodymyr Selenskyj zu erwähnen, in der er die Ukrainer zur Kapitulation auffor-

9 Vgl. https://www.lernen-digital.nrw/arbeitshilfen/games-im-unterricht-get-bad-news-unterrichtsmaterial.

10 Vgl. https://crankyuncle.com/.

11 Vgl. https://spotthetroll.org/start; siehe hierzu auch Lees et al., The Spot the Troll Quiz game increases accuracy in discerning between real and inauthentic social media accounts, PNAS nexus, Volume 2, Issue 4, April 2023https://academic.oup.com/pnasnexus/article/2/4/pgad094/7083320.

12 Vgl. chttps://openai.com/dall-e-2/

dert.[13] Da Inhalte jeglicher Art generiert werden können, können sie zur falschen Zeit, am falschen Ort nachteilige und schwerwiegende Folgen haben.

2. Ähnlichkeiten zwischen Künstlicher Intelligenz und neuronalen Netzwerkmodellen

Alles begann also im Februar 2019, als ich eine Publikation über das angekündigte GPT2 las. Was ich dabei besonders interessant fand war, dass ich bereits in meinem Grundstudium neuronale Netzwerksysteme verwendet hatte. Im Zusammengang mit der Lehre zum Gehirn lernte ich wie man menschliches Verhalten aus der Perspektive der Computational Neuroscience (Theoretische Neurowissenschaft) verstehen kann. Denn letztendlich sind beide KI oder zumindest die tiefen neuronalen Netzwerkmodelle der KI in ihrer Funktionsweise dem menschlichen Gehirn sehr ähnlich. Man könnte sagen, dass die Neuronen - daher auch der Name neuronales Netz - auf der neuronalen Funktionsweise des Gehirns basieren. Sowohl über das Gehirn als auch über die KI kann man viel lernen, besonders durch gegenseitiges Lernen untereinander. So schreiben meine Studenten bereits Aufsätze mittels KI während andere Studenten offenbar Apps entwickeln, die KI-geschriebene Aufsätze erkennen sollen. Es ist also ein regelrechtes Wettrüsten im Gange.

3. Entwicklung des MIST

a. Einsatz von Künstlicher Intelligenz bei der Erstellung eines psychologischen Tests

Ich habe mir überlegt, wie ich GPT-2 für meine Forschung nutzen könnte, um Bausteine für meine psychologischen Umfragen zu erstellen. Als das funktionierte, fragte ich mich, ob man auch einen psychologischen Test erstellen könne, bei dem die Teilnehmer zwischen echten Nachrichten und Fake News unterscheiden sollen. Sie wurden gebeten, KI-generierte irreführende Nachrichten und relativ neutral gelesene, echte Schlagzeilen zu erkennen und ob anhand dessen eine Art übergreifendes psychologisches Konstrukt deutlich würde und wie es aussehen könnte. Das ist auch die große Frage, mit der ich mich in den letzten Jahren intensiv befasse. Eine weitere Frage ist, ob es unterschiedliche Fähigkeiten zur Unterscheidung zwischen Wahrheit und Irreführung benötigt. Wir haben das zwar als explizite und implizite Misinformation definiert. Aber erfordern beide wirklich die gleiche psychologische Fähigkeit – all das fragte ich mich im Rahmen meiner Forschung. Ich fand heraus, dass es sich dabei um zwei Seiten derselben Medaille handelt. Wenn man also gut darin ist, Wahrhaftigkeit zu erkennen, ist man in der Regel auch gut darin, Irreführung zu erkennen. Trotzdem sind es voneinander leicht abweichende Fähigkeiten.

Also habe ich diesen Test entwickelt, den „Misinformation Susceptibility Test“ (Misinformations-Anfälligkeitstest), kurz MIST.[14] Ich bat GPT-2 zunächst mir 20.000 irreführende Überschriften zu generieren. Es dauerte wenige Sekunden, das fand ich bereits ziemlich beeindruckend. Dann habe ich das Ganze gesteigert, in dem ich GPT-2 einige Beispiele für Misinformationen nannte und aufforderte hundert weitere vom gleichen Typ, welche die gleichen irreführenden Strategien verwenden, zu erstellen. Tatsächlich kamen Hunderte weitere in Sekundenschnelle hinzu, das war im Jahr 2019. Diese Sprachmodelle haben sich seitdem dramatisch verbessert. Um es an einem, meiner Ansicht nach brillantem, Beispiel zu veranschaulichen: „Die UN: Mehr als 50 % der Weltbevölkerung werden 2070 von einer Hungersnot betroffen sein“. Viele Leute, die das lesen, würden zustimmen und sagen, dass es sich bei dieser Schlagzeile um vertrauenswürdige Nachrichten handelt. Dabei hat GPT-2 sie erzeugt. Das zweite, eher witzige und offensichtlichere Beispiel. Das gefällt mir, denn es ist ein phonetisches Wortspiel mit meinem Namen Rakoen (ra’kun): „Eilmeldung: Waschbären (engl. racoon(s)) sind jetzt die beliebtesten Haustiere in New York City.“ Die wenigsten Leute würden das wirklich glauben und falls schon, dann sind diese Personen wahrscheinlich ziemlich einfach zu täuschen. Aber das Schöne ist, dass man wirklich einen Test mit verschiedenen Schwierigkeitsgraden erstellen kann.

Bei der Entwicklung des MIST wählten wir ein paar hundert Artikel aus den Quellen aus, die von Media Bias/Fact Check (MBFC)[15] als am wenigsten voreingenommen und als sehr sachlich eingestuft worden waren. Sie sollten uns als echte Nachrichten dienen und die von GPT-2 generierten als Fake News. Anschließend haben wir Tausende von Menschen mit Hunderten von Schlagzeilen konfrontiert und sie gefragt, ob es sich dabei um Fake News oder echte Nachrichten handelt. Dann haben wir mit Hilfe von maschinellem Lernen eine Auswahl von nur 20 Schlagzeilen getroffen. Hier kommt dieses Quiz[16] ins Spiel, mit dem man sich selbst in seinem MIST abfragen kann. Im Grunde bekommt man seine Punktzahl für die allgemeine Urteilsfähigkeit, für die Erkennung von echten Nachrichten und von gefälschten Nachrichten und ob man generell eher skeptisch oder eher aufgeschlossen ist. Das ist der erste Test dieser Art. Es ist nicht der perfekte Test, aber es ist ein Beispiel dafür, wie man KI einsetzen kann, um Misinformatio-

13 Vgl. https://www.spiegel.de/netzwelt/web/meta-loescht-fake-video-das-wolodymyr-selenskyj-falsche-worte-in-den-mund-legt-a-5600045c-8057-4359-bd31-ee02c6e585d5.

14 Vgl. Maertens et al., The Misinformation Susceptibility Test (MIST): A psychometrically validated measure of news veracity discernment, 29. Juni 2023, https://link.springer.com/article/10.3758/s13428-023-02124-2.

15 Vgl. https://mediabiasfactcheck.com/.

16 https://yourmist.streamlit.app/.

nen zu generieren und um psychologische Tests zu entwickeln.
Interessant ist, dass die Top-Nachrichtenquellen von denjenigen, die eine hohe Punktzahl auf dem MIST haben, Nachrichten von AP (Asociated Press), Reuters und anderen sind. Noch interessanter ist, dass die Teilnehmer, die auf dem MIST schlecht abschneiden, Nachrichten von Snapchat, Truth Social, WhatsApp, TikTok konsumieren und gänzlich auf Rundfunk verzichten. Es bestätigt weiterhin die Ansicht, dass ein gewisses Risiko von neuen Plattformen ausgeht und dass das schlechte Abschneiden der jüngeren Personen möglicherweise damit zusammenhängt.

b. Einsatz des MIST in der Praxis am Beispiel der Impfquote im Vereinigten Königreich

Wir können den MIST auch nutzen, um die Anfälligkeit für Misinformationen in einem Land zu ermitteln. So haben wir uns die Impfquote im Vereinigten Königreich angesehen und eine Karte, die zeigt, wie gut die Menschen im Vereinigten Königreich echte Nachrichten erkennen. Man erkennt also die verschiedenen Regionen und wie gut sie in der Lage sind, Fake News zu erkennen.Wir haben einige geografische Untersuchungen durchgeführt, um Risikoregionen zu ermitteln. Anschließend haben wir versucht die Quote für die Auffrischungsimpfung/Booster-Quote vorherzusagen und konnten so sehr vorläufige Ergebnisse erzielen. Die Ergebnisse müssen noch verifiziert werden. Aber wir haben festgestellt, dass die Erkennung von echten Nachrichten nicht so sehr mit der Inanspruchnahme von Impfungen zusammenhängt, wohl aber mit der Erkennung von FakeNews. Das ist eine interessante Betrachtungsweise: unterschiedliche Dimensionen, unterschiedliche Verhaltensergebnisse.

c. AOT-Variable (Actively Open-minded Thinking) - aktiv aufgeschlossenes Denken

Schließlich habe ich in all meinen Studien, die wir in fünfeinhalb Jahren durchgeführt haben, eine überraschende Variable gefunden, die immer wieder als bester Prädiktor für die Widerstandsfähigkeit gegenüber Misinformationen auftaucht. Ich hatte das überhaupt nicht erwartet, da ich immer davon ausging, dass man dazu eine gute Bildung, ein Medienkompetenztraining sowie eine gewisse Skepsis braucht und darüber hinaus ein rationaler Denker sein muss. Doch dann fand ich die „AOT-Variable“ (Actively Openminded Thinking), nämlich das aktiv aufgeschlossene Denken. Das AOT ist durchweg ein besserer Prädiktor als das rationale Denken. Im Grunde besagt es, dass man offen für neue Perspektiven sein muss und in der Lage sein, seine eigene Meinung zu aktualisieren, bescheiden sein, was den eigenen Intellekt und die eigene Wahrnehmung angeht. Dann ist die Wahrscheinlichkeit geringer, dass man Verschwörungstheorien Glauben schenkt. Das klingt etwas widersprüchlich. Denn in dem Fall könnte man sich auch fragen, ob man dann auch glaube müsse, dass die Erde flach sei. Man geht schließlich davon aus, dass ein Mensch sehr aufgeschlossen sein muss, um eine solche Theorie zu glauben. Tatsächlich ist das aber so, dass die Menschen, die solche Ansichten glauben, in der Tat zu engstirnig sind, um ihre Meinung zu ändern. Man könnte es theoretisch glauben, aber würde man sich die Beweise anschauen, müsste man doch erkennen, dass das völlig unhaltbar ist. Man muss die Fähigkeit besitzen, seine eigenen Überzeugungen zu aktualisieren und dabei scheint das AOT extrem wichtig zu sein. Wir versuchen immer noch herauszufinden, ob man das AOT fördern und auf einfache Weise steigern und ob man eingreifen oder dafür bessere Plattformen schaffen kann. Damit gilt es sich in Zukunft zu beschäftigen.

IV. Ein paar Nuancen zum Abschluss

Zum Abschluss noch ein paar Vorbehalte und Nuancen. Ich bin Psychologe. Daher betrachte ich oftmals die individuelle Ebene. Aber ich muss gestehen, dass viele dieser Maßnahmen, wie Debunking, Prebunking usw., zwar funktionieren, aber deren Auswirkungen letztlich recht gering sind. Wir müssen uns also die Frage stellen, ob dies wirklich die beste Lösung ist und ob wir nicht mehr Änderungen auf Systemebene vornehmen müssen. Ich spreche an dieser Stelle über Algorithmen. Es gibt Teams, die mit Algorithmentypen experimentiert und sich überlegt haben, wie sie den zivilisierten Dialog verbessern könnten. Sie fanden auch Möglichkeiten dazu. Aber die Schlussfolgerung war final dieselbe: sie würden es nicht einführen, da auf diese Weise das Engagement auf den Plattformen zurückgehen würde. Es gibt also Teams und Unternehmen, die sich tatsächlich mit Lösungen befassen und denen das auch gelingen könnte. Ich habe diese Lösungsansätze bereits bei Demos in Aktion gesehen. Die Frage ist nur, wie wir sie aktivieren können und die Betreffenden davon überzeugen können, diese zu implementieren. Wir müssen also ernsthaft nachdenken, wie wir Anreize dafür schaffen können. Ein letztes Beispiel ist VTaiwan, das ich persönlich sehr interessant finde. Die Idee dahinter war, einen zivilisierten Dialog über politische Themen zu führen. Das Interessante an dieser Plattform war allerdings, dass sie die Konsensbildung und die Einigung zwischen den Menschen in den Vordergrund gestellt hat. Die ersten Ergebnisse und Versuche waren sehr positiv. Es schien tatsächlich zu einem positiven Dialog zu führen. Die Plattform wurde jedoch schnell wiedereingestellt, da es nicht viele aktive Nutzer gab. Letzten Endes hat es nicht wirklich funktioniert. Dennoch ist das ein guter Konzeptnachweis für unterschiedliche Arten und Möglichkeiten von Plattformen, die man potenziell ausbauen kann.

Deepfakes und KI-Manipulationen in Demokratie und Recht: Gefahren und Lösungen

Dr. Christopher Nehring, Sofia / Mateusz Łabuz, Krakau/Chemnitz*

Diese Studie untersucht die spezifischen Risiken von KI-generierten Deepfakes für das Rechtssystem und die Justiz. Anhand von Fallstudien sowie Studien internationaler Strafverfolgungs- und Justizbehörden wird dabei gezeigt, dass Deepfakes eine signifikante Bedrohung für die Integrität von Beweismitteln und die Funktionsweise und das Vertrauen in das Justizsystems darstellen können. Anschließend analysiert diese Studie Lösungsmöglichkeiten und Gegenmaßnahmen. Im Zentrum steht dabei die Analyse des KI-Gesetzes der EU (AI Act oder KI-Gesetz)[1] als grundlegender Rechtsakt zur Regulierung von KI-basierten Manipulationen, und seiner Effektivität in der Bekämpfung Deepfake-spezifischer Gefahren. Obgleich Transparenzpflichten hier ein wichtiger Schritt sind, bleibt es zweifelhaft, ob die Regelungen des AI Acts alleine dauerhaft in der Lage sind, den Herausforderungen von Deepfakes zu begegnen. Deshalb zeigt die Untersuchung weitere Gegenmaßnahmen, deren Kernstück eine Erweiterung rechtlicher Maßnahmen um bildungsbezogene und technologische Lösungen ist.

Einleitung, Definition, Fragestellung

Deepfakes, obgleich nicht die einzige Spielart von Manipulation und Fälschungen mithilfe Künstlicher Intelligenz (KI), gelten als eine der großen Gefahren im Zusammenhang mit KI. Das KI-Gesetz der EU (EU AI Act) definiert Deepfakes als „*durch KI erzeugten oder manipulierten Bild-, Ton- oder Videoinhalt, der wirklichen Personen, Gegenständen, Orten, Einrichtungen oder Ereignissen ähnelt und einer Person fälschlicherweise als echt oder wahrheitsgemäß erscheinen würde*“ (AI Act, 2024). Das Gesetz, wie auch die Forschung stuft diese Art der Manipulationen als erhebliche Bedrohung für die Demokratie ein.[2] In vielen Grundsatzdiskussionen, insbesondere im EU-Parlament, stand im Kontext von Deepfakes vor allem ihr Potential für Desinformation, Fälschung und politischer Einflussnahme und Manipulation im Vordergrund. Dies reduziert jedoch die komplexe Problematik der vielfältigen Einsatz- und Angriffsmöglichkeiten durch Deepfakes. Von Anfang an seit ihrem Aufkommen im Jahr 2017 sind sie untrennbar mit anderen schädlichen, manipulativen und bösartigen Einsatzmöglichkeiten verbunden, allen voran zur Herstellung nicht einvernehmlicher Pornografie.[3] Trotzdem gelang es Gesetzgebern und Strafverfolgungsbehörden weltweit nicht, diese Risiken entsprechend anzugehen und wirksame Gegenmaßnahmen zu entwickeln und umzusetzen.

Insbesondere die dynamische Entwicklung im Bereich von KI-Technologien, aber auch ein fehlendes Problembewusstsein mögen dafür die Ursachen gebildet haben. In der jüngsten Vergangenheit jedoch hat generative Künstliche Intelligenz (KI) viele Lebensbereiche revolutioniert. Negative Anwendungen und missbräuchlicher Einsatz, z. B. für Deepfakes, haben an Bedeutung gewonnen. Das Recht hält in der Regel nicht mit der technologischen Entwicklung Schritt, was zu einer verzögerten Umsetzung von Lösungen führt.[4] In einigen Fällen kann diese Verzögerung zu irreversiblen Veränderungen in der Funktionsweise der Gesellschaft führen.

Um dies zu verhindern, untersucht diese Studie das Phänomen von Deepfakes im Hinblick auf ihre spezifischen Risiken für Recht und Justiz. Anschließend wird der breitere Kontext systemischer Gefahren von Deepfakes erläutert, bevor der Ansatz des KI-Gesetzes der EU (AI Act) und die dort vorgesehenen Provisionen für Regulierung und Gegenmaßnahmen analy-

* Der Autor Nehring ist Gastdozent für Desinformation, Medien und Geheimdienste des Medienprogramms Südosteuropa der Konrad-Adenauer-Stiftung e.V. (KAS) an der Universität Sofia, e-mail research_cn@proton.me. Der Autor Łabuz ist ehem. Cyber-Attaché an der polnischen Botschaft in Berlin und Berufsdiplomat im polnischen Außenministerium, Doktorand an der Technischen Universität Chemnitz, Dozent für Cybersicherheit und künstliche Intelligenz an der Universität der Nationalen Bildungskommission in Krakau sowie Dozent für Cybersicherheit an der Päpstlichen Universität Johannes Paul II. in Krakau, e-mail mateuszlabuz@gmail.com

1 Verordnung (EU) 2024/1689 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnung (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/767 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz) v. 13. Juni .2024, ABl. L 2024, 1689.

2 Vgl. ***Vaccari & Chadwick***, Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. Social Media + Society 2020, Vol. 6(1). doi: 10.1177/2056305120903408; ***Huijstee van et al.***, Tackling deepfakes in European policy, European Parliamentary Research Service. Brüssel 2021; ***Maham & Küspert***, Governing General Purpose AI, Stiftung Neue Verantwortung. Berlin 2023; hierzu und zur KI-Verordnung der EU s. a. ***Möller-Klapperich***, NJ 2024,337, 338 f.

3 Vgl. ***Chesney & Citron***, Deepfakes: A Looming Challenge for Privacy, Democracy, and National Security. California Law Review 2019, Vol. 107(18). S. 1753 - 1820.

4 Vgl. ***Olson***, The Double-Side of Deepfakes: Obstacles and Assets in the Fight Against Child Pornography. Georgia Law Review 2022, Vol. 56(2). S. 865 - 892.

siert werden. Abschließend erfolgt ein Ausblick über weitere, nicht rechtliche Gegenmaßnahmen gegen Deepfakes bevor die grundlegenden Schlussfolgerungen dieser Studie zusammengefasst werden. Zentral für diese Untersuchung ist dabei das KI-Gesetz der EU als dem grundlegenden Rechtsakt zur Regulierung von KI und Deepfakes und Ausgangspunkt für Überlegungen zu weiteren Gegenmaßnahmen.[5] Darüber hinaus bedient sich diese Analyse dem Konzept der „epistemischen und Informationsapokalypse", das eine totale Sinnentleerung von Sprache und Information durch die Ununterscheidbarkeit von KI-Inhalten sowie eine Überschwemmung des globalen Informationsraumes mit Fakes und KI-Inhalten bezeichnet.[6]

Deepfakes als Gefahren für Recht und Justiz

Deepfakes haben das Potenzial, auch das Recht, die Jurisprudenz und deren praktische Umsetzung maßgeblich zu beeinflussen.[7] Reale Anwendungsfälle und Beispiele aus der Praxis untermauern dabei die Schädlichkeit von Deepfakes im Kontext der Funktionsweise und Integrität des Justizsystems und der Praxis der Strafverfolgungsbehörden.[8] Viele Risiken hängen insbesondere mit der starken Verbreitung der Technologie (sog. „Demokratisierung") zusammen, d. h. dass der Zugang zu Deepfake-Technologie, die früher allenfalls professionellen Akteuren vorbehalten war, nun massenhaft möglich ist.[9] Gleichfalls befördert werden die Gefahren von Deepfakes durch die zunehmende Qualität der Synthese von Inhalten (z.B. Text, Audio, Bild), die durch KI erzeugt werden. Eine Störung der Grenze zwischen dem, was „falsch" und „wahr" ist bzw. der Ununterscheidbarkeit zwischen beidem, bezeichnet das Konzept der sogenannten „epistemischen Apokalypse".[10] Diese Gefahren durch Deepfakes bedrohen das Justizsystem direkt, aber auch indirekt als Katalysator für negative gesellschaftliche Auswirkungen.

Eine der größten Gefahren für Recht und Justiz bezieht sich auf „die Glaubwürdigkeit und Zulässigkeit von audiovisuellem Filmmaterial als elektronisches Beweismittel".[11] Wenn also audiovisuelle Beweismittel zum Einsatz kommen, wie wird dann sichergestellt, dass diese Inhalte authentisch sind und nicht mittels Deepfake-Technologie erstellt wurden? Die massenhafte Verbreitung von KI sowie die enorme Qualitätssteigerung bedeuten also ganz neue Möglichkeiten für die Fälschung von Beweismitteln. Experten warnen bereits vor einer deutlichen Zunahme von Versuchen der Beweismittelmanipulation mittels moderner KI-Technologien.[12] *Llorente* macht ferner auf die Möglichkeit einer Kompromittierung von Beweisen und einer erheblichen Erweiterung des Materials aufmerksam, dessen Wahrheitsgehalt gründlich analysiert werden muss: „*Die schiere Menge potenziell KI-generierter oder manipulierter Inhalte droht Ermittler und Analysten, die mit der Sichtung audiovisueller Inhalte beauftragt sind, zu überfordern, was möglicherweise zu Justizirrtümern führt*".[13]

Andererseits können KI-Technologien noch leichter und noch besser zur Urkunden- und Verfahrensbeweisfälschung sowie zur Umgehung der biometrischen Sicherheit und Identitätsdiebstahl eingesetzt werden. Bestehende Sprach- und biometrische Sicherheitsmethoden müssen aufgrund der Fähigkeiten von KI in vielen Fällen überarbeitet werden,[14] was gleichfalls für Überprüfungen und Verifikationen von Urkunden und Beweisen allgemein vor Gericht gilt.

Die Schwierigkeit, zuverlässig, glaubhaft und rechtssicher zwischen KI-generieten (synthetischen) und authentischen, von Menschen erstellten, Inhalten zu unterscheiden, hat heute bereits ganz konkrete Konsequenzen in der Strafverfolgung. So zum Beispiel bei der Arbeit der Strafverfolgungsbehörden bezüglich der erschwerten Identifizierung von Opfern von Kinderpornografie.[15] Die Ununterscheidbarkeit von KI-generierten Materialien von solchen, in denen echte Opfer vorkommen, führt dazu, dass die mit der Ver-

5 Vgl. ***Łabuz***, Regulating Deepfakes in the Artificial Intelligence Act. Applied Cybersecurity & Internet Governance 2023, Vol. 2(1). doi: 10.60097/ACIG/162856.2023; ***Romero Moreno***, Generative AI and deepfakes: a human rights approach to tackling harmful content. International Review of Law, Computers & Technology 2024. doi: 10.1080/13600869.2024.2324540. 2024; ***Möller-Klapperich***, NJ 2024, 337 ff.

6 Vgl. ***Schick***, Deep Fakes and the Infocalypse. Octopus Books. Ottawa 2020; ***Fallis***, The Epistemic Threat of Deepfakes. Philosophy & Technology, Vol. 34(4) 2021. S. 623 - 643. doi: 10.1007/s13347-020-00419-2. 2021.

7 Vgl. ***Martsenko***, Influence of artificial intelligence on the legal system. Studia Prawnoustrojowe 2024, Vol. 54. S. 385 - 403. doi: 10.31648/sp.7101.2021; ***Vargas-Murillo et al.***, Transforming Justice: Implications of Artificial Intelligence in Legal Systems. Academic Journal of Interdisciplinary Studies 2024, Vol. 13(2). Doi: 10.36941/ajis-2024-0059.

8 Vgl. ***Europol***, Facing reality? Law enforcement and the challenge of deepfakes, an observatory report from the Europol Innovation Lab. Publications Office of the European Union. Luxemburg 2022; ***Delfino***, The Deepfake Defense - Exploring the Limits of the Law and Ethical Norms in Protecting Legal Proceedings from Lying Lawyers, SSRN Electronic Journal 2023, doi: 10.2139/ssrn.4355140; ***Llorente***, Deepfakes in the Dock: Preparing International Justice for Generative AI. The SciTech Lawyer 2024, Vol. 20(2). S. 28 - 33.

9 Vgl. ***Westerlund***, The Emergence of Deepfake Technology: A Review. Technology Innovation Management Review 2019, Vol. 9 (11). S. 39-52. doi: 10.22215/timreview/1282.

10 Vgl. ***Fallis*** (Fn. 6)

11 Vgl. ***Ciancaglini***, Gibson & Sancho, Malicious Uses and Abuses of Artificial Intelligence. Trend Micro Research, United Nations Interregional Crime and Justice Research Institute, Europol's European Cybercrime Centre 2020.

12 Vgl. Europol (Fn. 8); ***Llorente*** (Fn. 8).

13 Vgl. ***Llorente*** (Fn. 8), S. 32.

14 Vgl. Onfido, Identity Fraud Report 2024. London 2024.

15 Vgl. ***Crawford & Smith***, Illegal trade in AI child sex abuse images exposed, 15. Dezember 2023, https://www.bbc.com/news/uk-65932372; ***Internet Watch Foundation*** (IWF), How AI is being abused to create child sexual abuse imagery, Internet Watch Foundation, Cambridge 2023.

folgung von Pädophilen befassten Einheiten überlastet sind und zusätzliche Investitionen in spezialisierte Software und Humankapital erforderlich sind, um die hohe Anzahl von Fällen zu bearbeiten. Der Einsatz von Ressourcen könnte auch eine bewusste Strategie von Desinformationsakteuren und eine spezifische Form der Beeinträchtigung der Funktionsweise staatlicher Institutionen sein. Diese ist eine sehr reale und wirkungsmächtige Auswirkung des Konzepts der „epistemischen Apokalypse", also der Sinnentleerung von Informationen und Ununterscheidbarkeit zwischen synthetischen und authentischen Inhalten.

Eine andere, gleichfalls systemische, Herausforderung für Recht und Justiz ist die Möglichkeit, sich auf das Manipulationspotenzial von Deepfake-Technologien zu berufen, um sich der Verantwortung für Aussagen und Taten zu entziehen. In der (zumeist sozialwissenschaftlichen) Forschung wird dieses Phänomen als „liar's dividend" (Dividende des Lügners)[16] und im Kontext der Fälschung, Manipulation, Verzerrung oder Untergrabung der Integrität von Beweisen „Deepfake Defence" definiert.[17] Sie bezeichnen zum Beispiel die Möglichkeit, dass Politiker die Verantwortung für Aussagen mit dem Hinweis von sich weisen, „es könnte ja auch nur ein Deepfake gewesen sein".[18] Im Rechtskontext kann hierdurch der Wert, das Vertrauen und die Glaubwürdigkeit von Beweisen und Zeugen beeinträchtigt werden. In der Rechtspraxis wendete beispielswiese Tesla bzw. deren Rechtsvertreter diese Argumentation 2023 in einem Prozess über Todesfälle im Zusammenhang mit dem Programm für autonomes Fahren der Firma an. Angebliche Aussagen von Unternehmenschef Elon Musk über die Sicherheit des autonomen Fahrens wurden vor Gericht zurückgewiesen mit dem Hinweis, dass sich Musk nicht an solche Aussagen erinnern könne und „er, wie viele Personen des öffentlichen Lebens, regelmäßig Opfer von Deepfakes ist, die ihn Dinge sagen oder tun lassen, die er nie gesagt oder getan hat".[19] Fälle wie diese zeigen das Manipulationspotenzial von Deepfakes vor Gericht auf. Mit Hilfe von Deepfake-Technologie ist es also nicht nur möglich, Beweise und Materialien zu fälschen, sondern auch, die Glaubwürdigkeit echter Beweise, Aussagen und Materialien zu untergraben und zu diskreditieren.

Im schlimmsten Fall könnte Deepfake-Technologie auch genutzt werden, um die Erinnerungen einzelner Personen (z.B. Zeugen oder Opfern) zu manipulieren und Geständnisse zu erzwingen. Ein sehr beunruhigendes Ergebnis eines Experiments in der Verhaltensforschung zeigte zum Beispiel, dass „*Teilnehmer, denen gefälschte Videos von sich selbst beim Betrügen in einem Glücksspiel gezeigt wurden, bereit waren, falsche Geständnisse zu unterschreiben, wobei viele plausible Geschichten fabrizierten, um sich selbst zu erklären, warum sie betrogen haben könnten*".[20] Die Manipulation der menschlichen Psyche mit Hilfe der neuen technologischen Möglichkeiten von KI bedeutet daher mannigfache Auswirkungen auf das Funktionieren und die Fundamente von Recht und Justiz. Überall, wo Vertrauen, Wahrheit und Unterscheidbarkeit zentrale Eckpunkte für das Funktionieren sozialer Systeme sind, bedeutet die massenhafte Verbreitung und die Qualitätssteigerung von Deepfakes eine enorme Disruption. Dies kann dazu führen, dass die Autorität des Rechts- und Justizsystems untergraben wird und das Vertrauen der Öffentlichkeit in die Funktionsweise von Institutionen, die die Grundlage jeder gesunden Demokratie bilden, sinkt.

Deepfakes und das KI-Gesetz der EU

Das KI-Gesetz verwendet eine Reihe unterschiedlicher Risikokategorien: Deepfakes wurden als *Systeme mit begrenztem Risiko und besonderen Transparenzpflichten* eingestuft.[21] Durchaus umstritten ist dabei die Frage, ob Deepfakes aufgrund ihres nahezu universell anwendbaren Schadenspotentials, das sich auch in den Schnittstellenbereichen von Justiz und Strafverfolgung zeigt, nicht eher als „systemisches Risiko" hätten eingestuft werden müssen. Systemische Risiken sind mit den KI-Systemen mit allgemeinem Verwendungszweck verbunden, die „*aufgrund ihrer Reichweite oder aufgrund tatsächlicher oder vernünftigerweise vorhersehbarer negativer Folgen für die öffentliche Gesundheit, die Sicherheit, die öffentliche Sicherheit, die Grundrechte oder die Gesellschaft insgesamt erhebliche Auswirkungen auf den Unionsmarkt haben*" (Art. 3, Abs. 65 KI-Gesetz)). Diese Klassifikation bringt begrenzte Rechtsbehelfe mit sich.[22]

Das KI-Gesetz verweist ferner in ErwG 61 auf die möglichen erheblichen Auswirkungen auf die Demokratie, die Rechtsstaatlichkeit, die individuellen Freiheiten sowie das Recht auf einen wirksamen Rechtsbehelf und ein unparteiisches Gericht bei bestimmten KI-Systemen. Diese KI-Systeme wurden als hochriskant eingestuft. Und weiter: „*Um insbesondere den Risiken möglicher Verzerrungen, Fehler und Undurchsichtigkeiten zu begegnen, sollten KI-Systeme, die von einer Justizbehörde oder in ihrem Auftrag dazu genutzt werden sollen, Justizbehörden bei der Ermittlung und Auslegung von Sachverhalten und Rechtsvorschriften und bei der Anwendung des Rechts auf konkrete Sachverhalte zu unterstützen, als hochriskant eingestuft werden*". In diesem Zusam-

16 Vgl. ***Chesney & Citron*** (Fn. 3).
17 Vgl. ***Delfino*** (Fn. 8).
18 Vgl. ***Delfino*** (Fn. 8).
19 Vgl. ***The Guardian***, Elon Musk's statements could be 'deepfakes', Tesla defence lawyers tell court, 15. Dezember 2023, https://www.theguardian.com/technology/2023/apr/27/elon-musks-statements-could-be-deepfakes-tesla-defence-lawyers-tell-court.
20 Vgl. ***Rini & Cohen***, Deepfakes and the Epistemic Backstop. Philosophers' Imprint 2020. Vol. 24.
21 Vgl. ***Łabuz*** (Fn. 5).
22 Vgl. ***Romero Moreno*** (Fn. 5).

menhang beziehen sie sich vor allem auf KI-Systeme zur Unterstützung der Rechtsprechung, also spezielle „Rechts-KIs" und Software-Produkte, die zur Unterstützung von Anwälten, Staatsanwälten oder Gerichten eingesetzt werden. Diese Art der KI-Anwendungen haben ein ganz eigenes Schadens- und Risikopotential, wie z. B. das Auftreten sog. „Halluzinationen", also richtig formulierter, jedoch inhaltlich falscher (vor allem Text-)Inhalte, die innerhalb des Rechtssystems eingesetzt werden und für Verwirrung, falsche Ergebnisse, Benachteiligung und Diskriminierung oder höhere Kosten sorgen können. Ein solcher Beispielfall trat 2023 auf als ein US-amerikanischer Anwalt sein Plädoyer offenbar auf Präzedenzfällen aufbaute, die eine KI-Anwendung „erfunden" hatte.[23]

Die Bedrohungen für Justiz- und Rechtssysteme im Zusammenhang mit Deepfakes werden im KI-Gesetz nur unzureichend dargelegt. Im ErwG 133 wird zu Recht darauf hingewiesen, dass „*eine Vielzahl von KI-Systemen große Mengen synthetischer Inhalte erzeugen kann, bei denen es für Menschen immer schwieriger wird, sie von von Menschen erzeugten und authentischen Inhalten zu unterscheiden. Die breite Verfügbarkeit und die zunehmenden Fähigkeiten dieser Systeme wirken sich erheblich auf die Integrität des Informationsökosystems und das ihm entgegengebrachte Vertrauen aus, weil neue Risiken in Bezug auf Fehlinformation und Manipulation in großem Maßstab, Betrug, Identitätsbetrug und Täuschung der Verbraucher entstehen.*" Ähnlich wie z. B. für Journalismus, Medien und andere Bereiche öffentlicher Kommunikation rückt das KI-Gesetz hier die Gefahr in den Vordergrund, das KI- und von Menschen generierte Inhalte ununterscheidbar werden. Folgen können Vertrauensverlust oder auch zum Beispiel Lücken in der Haftung für Aussagen oder unklare Verantwortlichkeiten – auch innerhalb des Rechtssystems – sein.

Gegenmaßnahmen gegen Deepfakes im KI-Gesetz

Um den Gefahren von Deepfakes rechtlich zu begegnen, schlägt das KI-Gesetz konkrete Lösungen und Gegenmaßnahmen vor, um das Risiko zu verringern, das mit der unangemessenen Verwendung von Deepfakes verbunden ist. Diese allgemeinen Lösungen gelten damit auch für die Gefahren, die Deepfakes für Recht und Justiz darstellen. Realistischerweise jedoch muss erwartet werden, dass diese Regeln für entschlossene, professionelle böswillige Akteure kein Hindernis darstellen und bereits viele technologische und andere Tools zur Verfügung stehen (z.B. Entfernungssoftware für digitale Wasserzeichen),[24] um sie zu umgehen.

In Art. 50 EU KI-Gesetz, der Transparenzpflichten für bestimmte, risikoreiche KI-Systeme einführt, werden zwei grundlegende Schutzbarrieren eingeführt, die aus der Perspektive der gesamten Lieferkette schädlicher Inhalte interpretiert werden können:[25] Beginnend mit der Erstellung des Systems und der Implementierung von Modellen, bis hin zur Erstellung spezifischer Inhalte. Laut KI-Gesetz könnten solche Maßnahmen auf der Einführung geeigneter Markierungen in der Phase der Modellerstellung und -implementierung basieren. In der ersten Phase (Anbieter von KI-Systemen) würden zuverlässige, interoperable, wirksame und belastbare Techniken erforderlich sein, wie z. B. Wasserzeichen, Metadatenidentifizierungen, kryptografische Methoden zum Nachweis der Herkunft und Authentizität des Inhalts, Protokollierungsmethoden oder digitale Fingerabdrücke (ErwG 133 in Bezug auf Arti. 50 Abs. 2 KI-Gesetz). In der zweiten Phase (Betreiber von KI-Systemen) besteht die Verpflichtung, offenzulegen, dass die Inhalte künstlich erzeugt oder manipuliert wurden. Ausnahmen von diesen Regeln sind unter anderem, wenn die Verwendung zur Aufdeckung, Verhütung, Ermittlung oder Verfolgung von Straftaten gesetzlich zulässig ist (Art. 50 Abs. 4 KI-Gesetz). Grundsätzlich sollten also fast alle Formen von Deepfakes entsprechend gekennzeichnet sein, wobei die Form der Kennzeichnung einer Standardisierung unterliegen wird.

Zusätzlich wird im KI-Gesetz eine Art Sicherheitsökosystem durch eine Verknüpfung mit dem im ErwG 136 genannten Gesetz über digitale Dienste (Digital Service Act - DSA) der EU ergänzt. Es bezieht sich auf die „*Pflicht der Anbieter sehr großer Online Plattformen oder sehr großer Online-Suchmaschinen, systemische Risiken zu ermitteln und zu mindern, die aus der Verbreitung von künstlich erzeugten oder manipulierten Inhalten entstehen können, insbesondere das Risiko tatsächlicher oder vorhersehbarer negativer Auswirkungen auf demokratische Prozesse, den gesellschaftlichen Diskurs und Wahlprozesse, unter anderem durch Desinformation*". Dies ist nur eine beispielhafte Liste von Prozessen, deren Hauptzweck die Begrenzung der algorithmischen Verstärkung schädlicher Inhalte ist, die die Hauptquelle für Sichtbarkeit und Interaktion mit KI-generierten Inhalten ist. Es ist jedoch zu beachten, dass diese Maßnahmen in erster Linie auf Deepfakes abzielt, die Teil

23 Vgl. ***Spiegel***, New Yorker Gerichtsposse. Anwalt blamiert sich mit Fake-Fällen aus ChatGPT, 29. Mai 2023, https://www.spiegel.de/netzwelt/apps/new-york-anwalt-blamiert-sich-mit-fake-urteilen-aus-chatgpt-a-8935d1c8-b6c2-4079-8ecd-1cf4c2d33259 2023.

24 Vgl. ***Harris & Norden***, Meta's AI Watermarking Plan Is Flimsy, at Best. Watermarks are too easy to remove to offer any protection against disinformation, IEEE Spectrum, 4. März 2024, https://spectrum.ieee.org/meta-ai-watermarks.

25 Vgl. ***Chowdhury***, AI-fuelled election campaigns are here — where are the rules?. Nature 2024, Vol. 628(237). doi: 10.1038/d41586-024-00995-92024; ***Łabuz***, Deep fakes and the Artificial Intelligence Act—An important signal or a missed opportunity?. Policy & Internet 2024. doi: 10.1002/poi3.406.

politischer Desinformation sind (und nicht etwa auf Deepfakes im Bereich Pornographie, Imageangriffe, Gewalt oder Beweisfälschung).

Weitere Gegenmaßnahmen gegen Deepfakes

Die im KI-Gesetz enthaltenen Vorschläge identifizieren und adressieren nur einen Teil der Herausforderungen und Gefahren, die Deepfakes für unterschiedliche Bereiche darstellen. Dementsprechend etabliert das KI-Gesetz nur ein gesetzliches Minimum, das sich, wie die oben analysierten Provisionen des Gesetzes zeigen, vor allem auf technische Lösungen und Transparenzvorschriften beschränkt. Es ist realistischerweise davon auszugehen, dass entschlossene und professionelle Akteure, egal ob im Bereich der Kriminalität oder politischer Desinformation, sich nicht nur nicht um Transparenzpflichten kümmern, sondern auch personalisierte KI-Modelle verwenden, die somit nicht den Offenlegungsregeln auf Anbieterebene unterliegen. Oder kurz: Professionelle böswillige Akteure entwickeln und betreiben ihre eigene KI-Systeme, deren Produkte nicht entsprechend gekennzeichnet sein werden. Deshalb ist es so wichtig, das Problem von durch KI generierten und manipulierten Inhalte pragmatisch anzugehen und in Bereiche zu investieren, die die Gesellschaft vor der Verbreitung von Desinformation und deren Verstärkung in populären und leicht zugänglichen Quellen schützen können.

Mögliche Lösungsansätze und Gegenmaßnahmen umfassen eine breite Palette von Maßnahmen: Von großer Bedeutung wird in Europa in Zukunft zum Beispiel sein, dass EU, Staaten und nationale Behörden die Umsetzung und Anwendung des 2024 in Kraft getretenen Digital Services Act (DSA) und seiner spezifischen Regelungen zu Desinformation, Hate Speech, Pornographie und anderer schädlicher Inhalte umsetzen. Dies hat gleichfalls enorme Auswirkungen und Strahlkraft auf die Verbreitung KI-generierter schädlicher Inhalte. Eine erste Analyse eines im Februar 2024 von der EU-Kommission veröffentlichten Datensatzes über nach dem DSA gelöschter Inhalte auf Social Media Plattformen ergab beispielsweise, dass alleine 10.000 synthetische Inhalte bereits entfernt wurden; die überwältigende Mehrheit davon betraf „sexuelle Inhalte".[26] Das KI-Gesetz selbst verweist gleichfalls auf den DSA als maßgebliche gesetzliche Grundlage immer dann, wenn es spezifisch um KI-generierte Desinformation geht. Darüber hinaus gehende, die Schwächen und Versäumnisse des KI-Gesetz korrigierende rechtliche Gegenmaßnahmen und Regulierung von Deepfakes sind auf nationaler Ebene zwar möglich, für den Rechtsraum der EU jedoch extrem unwahrscheinlich für die kommenden Jahre. Das Vereinigte Königreich hingegen erließ 2024 beispielsweise ein spezielles Gesetz, dass den Einsatz von Deepfake-Technologie für nicht-einvernehmliche Pornographie unter Strafe stellt.[27]

Bislang fiel zu wenig Beachtung gefunden haben jedoch Ansätze, die Bildung, Aufklärung und die Stärkung menschlicher Fähigkeiten in Hinblick auf Problembewusstsein, Erkennungs- und Überprüfungsfähigkeiten sowie den Umgang mit KI-generierten Inhalten in den Vordergrund stellen. Dies gilt sowohl für die allgemeine Bevölkerung als auch für spezifische Berufsgruppen, wie zum Beispiel im Bereich Recht und Justiz. Die Transparenz- und Kennzeichnungsregeln des KI-Gesetz werden nur begrenzte Auswirkungen haben können, wenn das Publikum und die Rezipienten von KI-Inhalten keine oder nur geringe digitale und KI-Bildung aufweisen. Im Bereich Recht und Justiz kann das beispielsweise bedeuten, dass Richter, Staatsanwälte, Anwälte und andere Rechtshelfer mit niedrigem Grad an KI-Bildung bzw. Problembewusstsein für Deepfakes leichter und öfter auf KI-Fälschungen hereinfallen oder Opfer der „Deepfake Defence" werden können. Ein Auf- und Ausbau von KI-spezifischer Expertise, Schulungen zum Aufbau eines Problembewusstseins sowie Kenntnisse über Lösungs-, Überprüfungs- und Gegenmaßnahmen sind daher auch hier Grundlage für Resilienz und effektive Abwehr schädlicher Deepfakes.

Experten haben in diesem Kontext bereits erste Vorschläge für Akteure des internationalen Rechts und er Justiz ausgearbeitet.[28] Diese sehen z .B. eine Kombination aus dem Aufbau und der Stärkung technischer Fähigkeiten und Lösungen mit dem Aufbau und der Stärkung menschlicher Fähigkeiten in den entsprechenden Institutionen vor. So sollen Institutionen z. B. einen eigenen Pool an Software-Tools zur Überprüfung von Inhalten bereithalten, einen Überblick über sog. „content credentials" aufbauen, digitale Wasserzeichen in Inhalte und Daten einbauen und auslesen oder eigene Applikationen entwickeln. Neben technologischen Möglichkeiten zur Überprüfung und Verifikation oder Sicherung von Inhalten sind die Entwicklung und der Ausbau analytischer und forensischer Expertise an Institutionen ein weiterer wichtiger Punkt. Dazu gehören Kapazitätsaufbau, Schulungen und Weiterbildungen, aber auch die Einbeziehung, Kooperation und der Wissensaustausch mit externen Experten.

Es gibt sogar Konzepte, einen Teil der Verantwortung für die Beurteilung der Zulässigkeit von Beweismitteln auf Rechtsanwälte zu übertragen. In einem solchen Szenario würden sie mithilfe von KI aktiv überprüfen, ob die Beweise nicht gefälscht wurden. „Es

26 Vgl. *Nehring*, 10.000 deleted pieces of AI-content, 15. März 2024, https://www.linkedin.com/posts/christopher-nehring-423b06257_syntheticmedia-aicontent-socialmediaplatforms-activity-7165736889805750272-tVFH?utm_source=share&utm_medium=member_desktop.

27 Vgl. *UK Government*, Press release: Government cracks down on 'deepfakes' creation, 16. April 2024, https://www.gov.uk/government/news/government-cracks-down-on-deepfakes-creation.

28 Vgl. *Llorente* (Fn. 8).

geht darum, die wahrheitssuchende Funktion unserer Gerichte aktiv zu wahren“[29]. Dies würde eine Neudefinition bzw. Erneuerung der Rolle der Anwälte in Gerichtsverfahren als Garanten der Fairness und Vertreter des weitgefassten Rechtssystems erfordern.

Darüber hinaus ist die Etablierung von einheitlichen Standards innerhalb eines Rechtssystems unabdingbar, um Vertrauen zu bewahren und rechtseinheitliche Herangehensweisen zu sichern. Wie, also anhand welcher Software, welcher inhaltlicher Faktoren und welcher Ergebnisse, sollen beispielsweise Beweisstücke vor Gericht forensisch auf ihre Herstellung mit KI untersucht werden? Auf diese vermeintlich simple Frage gibt es bislang keine einheitliche Antwort, geschweige denn etablierte Vorgehensweisen. Als Resultat überprüfen und verifizieren Gerichte, Strafverfolgung, Experten, aber auch private Organisationen unterschiedlich und nach Gutdünken. Die Ergebnisse sind daher weniger verlässlich und auch deutlich leichter angreifbar. US-Experte *Hany Farid* verwies deshalb z. B. darauf, dass Gerichte nicht nur einheitliche Standards für Überprüfung und Verifikation, sondern auch entsprechende Ressourcen (insb. auch ausreichend Zeit) haben müssten.[30]

Fazit

Diese Untersuchung hat gezeigt, dass Deepfakes nicht nur eine technologische Herausforderung darstellen, sondern auch eine ernsthafte Bedrohung für Recht und Justiz darstellen können. Mit der Fähigkeit, nahezu ununterscheidbare Fälschungen von echten Inhalten zu erstellen, haben sie das Potential, Vertrauen in Beweismittel und die Integrität des Rechtssystems grundlegend in Frage zu stellen. Das KI-Gesetz der EU hat zwar wichtige Schritte unternommen, um den Gefahren durch KI-generierte Inhalte zu begegnen, die dort etablierten Regeln und Gegenmaßnahmen, wie z.B. Transparenzpflichten, sind jedoch zu löchrig und unvollständig, um die vielschichtigen und dynamischen Herausforderungen durch Deepfakes alleine zu bekämpfen. Die Anpassung dieses gerade geschaffenen Regulierungsaktes an die dynamische Entwicklung und Anwendung von KI-Technologien wird daher eine Langzeitaufgabe bleiben. Sowohl in der Forschung als auch in der Praxis sollten künftige Anstrengungen unternommen werden, um diesen Herausforderungen durch ein umfassendes Sicherheitsökosystem zu begegnen, das sowohl technologische als auch bildungsbezogene Maßnahmen beinhaltet. Zentral wird dabei die Förderung von digitaler und KI-Bildung sein, um sowohl (juristische) Fachkräfte als auch die allgemeine Bevölkerung besser auf die Erkennung und Handhabung von KI-generierten Inhalten vorzubereiten. Zusammenarbeit und Austausch zwischen Rechts- und Justizsystem, Wissenschaft, Experten und Zivilgesellschaft sowie die Etablierung von einheitlichen Standards ist eine weitere Zukunftsaufgabe in diesem Kontext. Die Implementierung von robusten, anpassungsfähigen und vor allem proaktiven Maßnahmen wird entscheidend sein, um Vertrauen und Integrität von Recht und Justiz im KI-Zeitalter zu schützen.

29 Vgl. *Schlegel*. The Deepfake Dilemma: Legal Safeguards in the Digital Era. 12. September 2024, https://www.judgeschlegel.com/blog/the-deepfake-dilemma-legal-safeguards-in-the-digital-era.

30 Vgl. *Farid*. AI Audio Deepfakes Are Quickly Outpacing Detection, Scientific American, 26. Januar 2024, https://www.scientificamerican.com/article/ai-audio-deepfakes-are-quickly-outpacing-detection/.

Desinformation und Digital Listening

Prof. Dr. Martin Grothe, Berlin*

Einführung: Ansatz von Desinformation

Bekannterweise sind drei Verhaltensregeln ausreichend, damit – zumindest in der Simulation – Vögel einen Schwarm bilden: Fliege in die Richtung deiner Nachbarn, halte die gleiche Geschwindigkeit ein, halte Abstand. Auch in vielerlei anderen Konstellationen führen wenige einfache Regeln zu komplexem Gesamtverhalten.[1]

Ein böswilliger Manipulator könnte nun eine Regel, etwa die der Geschwindigkeitsangleichung, variieren und so das emergente Muster, das zielfördernde Miteinander beeinträchtigen oder sogar zerstören.

Das soziale Miteinander ist durch explizit kodifizierte Regeln, aber eben auch durch implizite Erwartungen an das Verhalten der jeweils anderen Akteure oder Akteursmehrheiten bestimmt. Eine solche lebendige, weitgehend emergente Ordnung wird durch Manipulation der Erwartungen an andere Personen, Gruppen oder Institutionen unter Umständen massiv gestört.

Dies ist der Ansatzpunkt von Desinformationskampagnen: Grundüberzeugungen, Einschätzungen und Erwartungen an das Verhalten anderer Akteure sollen

* Der Autor ist Geschäftsführer der complexium GmbH, Berlin, und Honorarprofessor an der Universität der Künste (UdK) Berlin.

1 Vgl. *Grothe*, Ordnung als betriebswirtschaftliches Phänomen, Gabler 1997.

aufgeweicht werden, um die Stabilität des Gemeinwesens oder die Stellung einzelner Akteure zu schädigen. Damit sind Desinformationskampagnen ein Mittel, um Erwartungsänderungen von individuellen Akteuren und damit in der Folge geänderte Entscheidungen wahrscheinlicher zu machen. Wahlbeeinflussung ist ein Anwendungsfall.

Solche Kampagnen werden erfolgreich sein, wenn es gelingt, eine bestehende Thematik – eine gesellschaftliche Bruchstelle, ein Riss im Fundament der Demokratie oder auch nur eine Unklarheit im Unternehmenskontext - aufzunehmen und durch aktiv eingebrachte Narrative zu einer Emotionalisierung und wunschgemäßer Aktivierung breiterer Kreise zu kommen. Im böswilligen Idealfall führt diese Folge sogar zu einer Mobilisierung gegen jeweils herausgestellte Kampagnenziele.

1 Thematisierung → 2 Emotionalisierung → 3 Aktivierung → 4 Mobilisierung

Der digitale Raum ist auf der einen Seite der prädestinierte Operationsraum für solche Desinformationskampagnen: Kann doch der Angreifer aus dem Verborgenen oder unter falscher Legende agieren, seine Botschaften aber sehr passgenau in die Themen- und Diskussionslandschaft einbringen und Glaubwürdigkeit vorgaukeln.

Auf der anderen Seite erlaubt der Digitalraum aber auch ein sehr weitgehendes Monitoring und damit eine idealerweise frühzeitige Detektion solcher Kampagnen, die zumeist nicht an einzelnen Beiträgen, sondern kontextuell erkannt werden können. So muss die Anforderung sein, entsprechende Kampagnen bereits in der Anlaufphase zu detektieren und nicht erst im Nachgang zu rekonstruieren.

Gleichwohl ist hervorzuheben, dass einerseits motivierte Angreifer mittlerweile sehr versiert vorgehen, andererseits die Detektions- oder gar Abwehrseite noch sehr unausgeprägt aufgestellt ist. Folglich hat dieser Beitrag zum Ziel, aus Sicht der Unternehmenssicherheit die aktuellen Möglichkeiten der kontinuierlichen und frühzeitigen Detektion von Kampagnen oder ähnlichen Bedrohungen im Digitalraum zu skizzieren.

Digital Listening für die Unternehmenssicherheit

Es wird für Unternehmen und insbesondere die Fachfunktion Unternehmenssicherheit immer wichtiger, den digitalen Raum in der Lagebilderstellung proaktiv zu nutzen. So lassen sich Sicherheitsaufgaben nach den Dimensionen "Zuhause/vor Ort – Rest der Welt" und "Physischer Raum – Digitaler Raum" verorten. Das resultierende Bild macht deutlich, dass es drei Quadranten gibt, die durch die Unternehmenssicherheit bereits überwiegend gut bis sehr gut abgedeckt werden: Objekt-/Event-Sicherheit, Reisesicherheit und Cyber-/IT-Sicherheit.

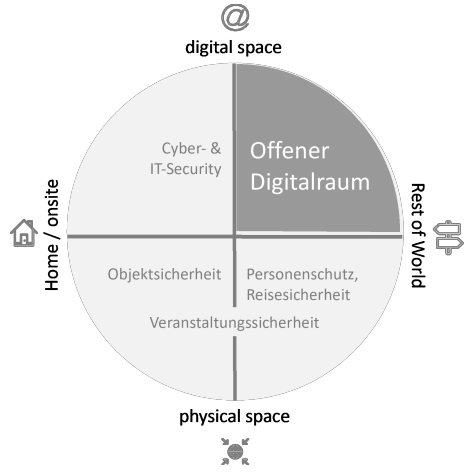


Abbildung 1: Die vier Quadranten der Unternehmenssicherheit (eigene Darstellung)

Aber Quadrant IV, Aktivitäten auf externen digitalen Plattformen, wird noch weitgehend als blinder Fleck behandelt. Was immer großzügiger wird, denn dort werden:

- wahre oder falsche Narrative zu Unternehmen aufgebaut und weitergegeben,
- durchgesickerte Informationen aufgenommen,
- konkrete Vorhaben, Aufrufe und Aktionen von Aktivisten koordiniert,
- gestohlene Daten geteilt, die Kriminelle als Munition nutzen können, um Unternehmen – meist indirekt über Mitarbeiter – anzugreifen,
- unerwünschte Einblicke in die private Sphäre von exponierten Persönlichkeiten preisgegeben.

Der Ansatz, um diese Bedrohungen – im Idealfall so früh wie möglich – zu erkennen, ist das digitale Zuhören: *Digital Listening* spielt eine entscheidende Rolle, um den Grad der unerwünschten Überraschung zu begrenzen. Es liegt auf der Hand, dass die Erkenntnisse aus Quadrant IV dazu beitragen, die Bedrohungslage auch in den anderen drei Quadranten zu reduzieren. Denn grundsätzlich können alle Bedrohungen, die ein Unternehmen, seine Führungskräfte, Mitarbeiter, Standorte und Infrastruktur betreffen, schwache Signale im digitalen Bereich haben. Daher baut die Implementierung von digitalen Aufklärungs- und Analyseprozessen zugleich einen umfassenden Schutzschirm auf.

Aktive Verteidigung überlässt dem Gegner nicht den digitalen Raum. Tatsächlich gibt es keine Entschuldigung dafür, nicht auf die digitalen Quellen zu achten.

So nehmen Bedrohungen durch

- (böswillige) Intentionen,
- (gegnerische) Fähigkeiten sowie
- (zugelassene) Gelegenheiten zu.

Wenn sich Gelegenheiten nun durch digitale Einblicke vervielfachen, Intentionen durch ungestüme, auch ge-

steuerte Emotionalisierung eskalieren und Fähigkeitsanforderungen durch Vereinfachung und Kontrollverlust leichter erreichbar werden, dann steigt der Bedrohungsgrad für Unternehmen und exponierte Persönlichkeiten an.

Unternehmenssicherheit kann dem nur entgegenwirken, indem Gelegenheiten begrenzt, Intentionen frühzeitig erkannt und notwendige Fähigkeiten heraufgesetzt werden. Diese Zielrichtungen lassen sich mit digitaler Früherkennung und Lageverfolgung erreichen. Durch einen Ausbau der digitalen Aufklärung sollen Überraschungsmomente reduziert sowie Vorlauf und Informationsstand für die Unternehmenssicherheit ausgebaut werden. Dafür kann ***Digital Listening*** eingesetzt werden:[2]

Durch weitflächige Aufnahme und systematische Verdichtung von Beiträgen aus dem digitalen Raum sowie intelligente Bildung von Ableitungen werden konsistente Lagebilder aufgebaut und Bedrohungen frühzeitig erkannt.

Dabei können verschiedene Formate eingesetzt werden:

- Ein laufendes Monitoring macht sich zunutze, dass auch ad hoc auftretende Lagen häufig im Vorfeld schwache Signale aufweisen: Inhaltliche Emotionalisierung, aktivierende Aufrufe, Trainingskurse, Anfahrtsplanungen etc. lassen sich detektieren.
- Eine zeitlich abgestimmte digitale Aufklärung im Vorlauf von Veranstaltungen oder kritischen Entscheidungen kann die Planungen und Vorbereitungen von möglichen Gegnerschaften aufzeigen: Eigene Maßnahmen können entsprechend angepasst werden.
- Mit digitalen Sichtbarkeitsanalysen für exponierte Personen wird erkannt, welche unerwünschten Einblicke eine Schutzperson dem böswilligen Dritten gewährt. Möglichkeiten für Anbahnungen und Annäherungen können erkannt und unterbunden werden.

Damit soll unterstrichen werden, dass es beim Einsatz von *Digital Listening* nicht um die Anschaffung eines Software-Tools geht oder um ein Setting in immergleicher Ausrichtung, Intensität und Frequenz, sondern um die agile, passgenaue Gestaltung der nächsten Verteidigungslinie, die derzeit mit menschlicher Expertise und intelligenter Werkzeugunterstützung immer noch am stärksten ist. Mit einer solchen Grundlage erreicht ein Sicherheitsbereich, ob zuständig für ein Unternehmen oder eine exponierte Familie, eine zeitgemäße und perspektivisch belastbare Aufstellung.

Anders als etwa im Marketing oder der Kommunikation ist es für die Unternehmenssicherheit jedoch schwierig, genaugenommen per se unmöglich, potenziell relevante Informationen im Vorfeld zu definieren: Etwas wird gesucht – ein Sicherheitsrisiko, eine Bedrohung – aber es ist noch nicht bekannt, wonach begrifflich gesucht werden sollte. Der Ansatzpunkt einer möglichen Angriffskampagne wird durch den potenziellen Gegner gesetzt.

Diese immanente Herausforderung wird durch den von Donald Rumsfeld geprägten Begriff der "**unknown Unknowns**" beschrieben, also Bedrohungen, von denen noch nicht einmal bekannt ist, dass es sie gibt.[3]

Der Digitalbereich bietet jedoch einen facettenreichen Datenschatz, um eine sicherheitsrelevante Aufklärung und Lageverfolgung umzusetzen. Die Herausforderung, a priori unbekannte schwache Signale zu erkennen, lange bevor sie zu großen Problemen oder bevor ihre Auswirkungen zu Bedrohungen werden, setzt einen strukturierten Prozess und geeignete Werkzeuge voraus.

Analyseprozess durch Intelligence Cycle

Der Digitalraum kann genauso strukturiert analysiert werden wie das analoge Pendant. Um insbesondere auf die Trennung von Datensammlung, -bearbeitung und -analyse hinzuweisen, wird hier der mehrstufige **Intelligence Cycle**[4] als allgemeines Vorgehensmodell dargestellt:

1. Direction

In der ersten Phase wird die Zielsetzung einer Analyse festgelegt. Neben der sach-inhaltlichen Bestimmung muss auch das Ergebnisformat festgelegt werden. Folgende Varianten sind besonders praxisrelevant:

- Tägliches, aber ggf. temporär begrenztes Alerting zu aktuell auftauchenden Bedrohungen,
- wöchentlicher operativer Lagebericht zum Unternehmen, exponierten Personen, Liegenschaften und besonderen Events,
- monatliche Sicherheitsberichte mit Übersichten, Zusammenhängen und allgemeinen aktivistischen Entwicklungen,
- eigenständige Vertiefungsanalysen, sei es zu besonderen Ereignissen, Örtlichkeiten oder Background-Checks.

Sämtliche Formate sind geeignet, um etwaige Desinformationskampagnen aufzunehmen.

2. Data Collection: Daten

Spezifische Vertiefungsanalysen können eine Fokussierung der Datensammlung erfordern. Für eine breite, hypothesenfreie Erschließung sollte jedoch Input

2 Vgl. ausführlich *Grothe*, Digital Listening in Unternehmen, Springer Gabler Wiesbaden 2020.

3 Donald Rumsfeld führte diesen vielfach zitierten Gedanken anlässlich einer Pressekonferenz am 12. Februar 2002 aus. Der Originalwortlauf kann bis heute abgerufen werden, vgl. https://www.youtube.com/watch?v=GiPe1OiKQuk (zuletzt abgerufen am 6. September 2024).

4 https://irp.fas.org/cia/product/facttell/intcycle.htm (zuletzt abgerufen am 6. September 2024).

aus einer möglichst großen Vielzahl der digitalen Plattformen und Formate aufgenommen werden.

Besonders wichtige Quellen sind zumeist die Welt der Telegram-Gruppen und -Kanäle sowie das an Bedeutung verlierende X/Twitter, die Vielzahl der Blogs und nicht durch Suchmaschinen indizierten Noblogs, ebenso das vermutlich aufstrebende BlueSky, Discord, das finstere 4Chan, dessen Anonymität zu rechtsextremen, sexistischen und antisemitischen Inhalten führt, die Vielzahl von spezifischen und mehr oder weniger dunklen Foren mit Raum für den anonymen Austausch, die Mikroblogging-Plattform Gettr aus den USA als neue Heimat für Querdenker und Rechte, Mastodon, das oft übersehene Reddit als ganz eigener Kosmos mit sehr aktiven Nutzern, populäre Soziale Media-Plattformen wie Facebook, Instagram, YouTube und die gefährliche Videoverbreitungsplattform TikTok.

3. Data Processing: Informationen

Die Aufgaben der Datenbereinigung und ggf. Übersetzung sind im Umgang mit großen Datenmengen essenziell. Hier soll auf innovative Ansätze hingewiesen werden: Modelle der Künstlichen Intelligenz können in einer automatisierten Bearbeitung geographische Koordinaten (wo?), zeitliche Bezugsgrößen (wann?) sowie besondere Entitäten (Personen, Organisationen etc.) (wer?) aus den Beiträgen destillieren. Mit einer solchermaßen aufbereiteten Grundlage kann eine Analyse deutlich mächtiger und effizienter ausfallen.

4. Analysis: Intelligence

Im Allgemeinen werden direkte Bedrohungen und Kontexte für Unternehmen, Personen, Lokationen und anstehende Events betrachtet. Es setzt sich zunehmend durch, auch das weitere Umfeld in das Lagebild einzubeziehen: Die allgemeine Sicherheitslage, insbesondere Sabotagefälle, aktivistische Strömungen und neue Operationsweisen sind von Interesse. Aufgabe der Analyse ist es, aus den gefundenen Treffern und relevanten Beiträgen den konkreten Sicherheitsbezug und treffende Ableitungen zu schließen: Intelligence.

5. Presentation

Die Aufbereitung richtet sich nach der Zielsetzung: Ein Warnhinweis, der echtzeitnah zugeleitet wird, muss sofort – auf dem mobilen Device – zu erfassen sein und kann auf Quellenangaben verzichten. Ein vorstandsfähiger Bericht sieht anders aus. Grundsätzlich sollte jede Aufbereitung die Ableitungen und Folgerungen vorrangig darstellen, erst danach die eigentlichen Vorkommnisse und Qualifizierungen. Eine durchgängige Struktur ist unabdingbar, geht es doch nicht mehr darum, einzelne Treffer aufzuzeigen, sondern systematisch entwickelte Ableitungen als Input für Entscheidungsprozesse einzubringen.

6. Dissemination

Die Zuleitung der Ergebnisse und Lagebilder ist der Impuls für die Diskussion der Ableitungen. Zugleich werden Anpassungen für die folgenden Durchgänge des Intelligence Cycles aufgenommen.

Mit einer solchen Ausrichtung ist die digitale Lageverfolgung als lernender Prozess angelegt: Das zu beobachtende Geschehen entwickelt sich, es tauchen neue Vorhaben, Kampagnen, Akteure, Quellen, Begrifflichkeiten und Hashtags auf. Zudem ist die technische Struktur des digitalen Raums konstanter Veränderung ausgesetzt: Von der Relevanz der einzelnen Quellen bis hin zu den Möglichkeiten, die einzelnen Formate automatisiert in eine Analyse einfließen zu lassen.

Analyseinfrastruktur durch OSINT-Lagezentrum

In der Analyse unterstützen Open Source Intelligence (OSINT)-Werkzeuge investigative Prozesse. Diese Werkzeuge weisen zwei bemerkenswerte Aspekte auf: Zum einen sind die zur Analyse benötigten digitalen Daten offen zugänglich und stehen im Prinzip jedem offen, der zumindest Suchmaschinen bereits genutzt hat. Zum anderen existiert eine sehr unübersichtliche Anzahl an Such-Werkzeugen, die sich zudem stetig weiterentwickeln. So bleiben die Quellen oft weiterhin unerschlossen. Für einen tieferen Einblick wird auf weitergehende Literatur verwiesen.[5]

Mit diesen Werkzeugen lassen sich primär einzelne Vertiefungen durchführen, eine kontinuierliche Lageverfolgung bedarf jedoch darüberhinausgehende und verstetigende Prozesse. Als Beispiel wird die Analyseinfrastruktur der **complexium GmbH aus Berlin** herangezogen.[6] Das Lagezentrum arbeitet für Konzernsicherheiten aus den Bereichen Auto, Bank, Chemie, Defense, Energie, Flughafen, Gesundheit und Industrie. In diesen Feldern nutzen Aktivisten, Kritiker und sonstige Gegnerschaften das Internet zur Meinungsmache sowie zur Ankündigung und Durchführung von Vorhaben.

In vielen Fällen werden desinformative Narrative eingesetzt, um eine meinungsbeeinflussende oder gar aktivierende Wirkung gegen die entsprechenden Unternehmen zu erzielen.

5 Vgl. *Grothe* (Fn. 2), S. 32 ff.

6 Weitere Informationen zu Unternehmen und Tätigkeitsgebieten finden sich auf www.complexium.de

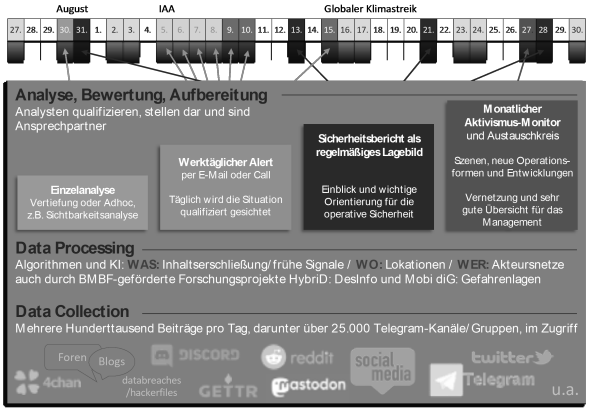


Abbildung 2: Lagezentrum der complexium GmbH (eigene Darstellung)

Das Verständnis von neuer Sicherheit und aktiver Verteidigung setzt dabei konsequent auf *Digital Listening*: Die systematische Früherkennung von Bedrohungen im digitalen Raum.

Täglich werden in diesem Rahmen mehrere Hunderttausend potenziell kritische Beiträge aus relevanten Plattformen eingelesen. Dadurch werden nicht nur Inhalte zu konkreten Suchworten identifiziert, sondern auch Kontexte und herannahende Themen oder Narrative erschlossen, die zu einer Gefährdung werden können. Das interne Entwicklerteam stellt für diese Beitragsflut innovative Werkzeuge für das Analyseteam bereit. Dabei wird die eigene Technologie durch BMBF-geförderte Projekte befruchtet.

Erfahrene Analysten nehmen frühe Signale effizient auf, erstellen qualifizierte Lagebilder, verfolgen Entwicklungen und bringen Klienten im Dialog vor die Lage. Unternehmen erhalten damit kontinuierlich eine bessere Entscheidungsgrundlage sowie mehr Vorwarnzeit für eigene Planungen, Anpassungen und Sicherheitsmaßnahmen.

Erreichbare Transparenz im Digitalraum

Durch analytische Abbildung der zentralen W-Fragen Was, wer, wo wird ein hohes Maß an Transparenz erreicht. Hierbei kommen Methoden aus den Bereichen der Computerlinguistik, Netzwerkanalyse und Künstlichen Intelligenz zum Einsatz.

Früherkennung: Signale, Themen, Cluster

Im Digitalraum impliziert eine Desinformationskampagne, ähnlich wie ein anschwellender Protest, dass bestimmte Begriffe oder Floskeln verstärkt geteilt werden.

Durch computerlinguistische Signifikanzanalyse kann nun echtzeitnah erschlossen werden, welche Terme aus der aktuellen Beitragsflut jeweils häufiger als üblich, d. h. linguistisch signifikant verwendet werden: Diese Metrik führt dazu, dass auch für eher seltene Begriffe eine markante Steigerung in einem Signifikanzranking hochgespült und damit erkennbar wird.

Mit einem solchen Werkzeug wird ein Analyst auf abrupte inhaltliche Veränderungen hingewiesen. Allerdings haben natürlich nicht alle entsprechenden Muster einen sicherheitsrelevanten Hintergrund.

Die Analyseeffizienz wird weiter ausgebaut, wenn durch Social Network Analysis (SNA) aufgezeigt wird, welche dieser signifikanten Terme in einem stärkeren Zusammenhang genannt werden: Inhaltliche Cluster werden deutlich und erleichtern die Einordnung. Es gelingt eine fortlaufende, hypothesenfreie Inhaltserschließung großer Beitragsmengen.

Desinformationskampagnen, die einen bestimmten, aber im Vorhinein unbekannten Aspekt oder Zusammenhang herausstellen, werden damit systematisch transparent.

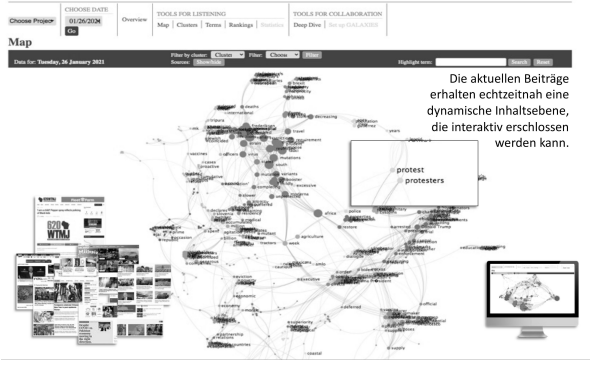


Abbildung 3: Automatisierte Inhaltserschließung mit dem Tool Galaxy (eigene Darstellung)

Telegram: Akteure und Akteursnetze

Ein ganz wesentliches Feld für aktivistische Kommunikation, russische Einflussnahme, aber auch zentraler Verbreitungsraum für Desinformationskampagnen ist Telegram. Mit einer dubiosen Verortung in Dubai geht die Einschätzung einher, dass entsprechende Kommunikation weitgehend im Verborgenen abläuft.

Im Rahmen einer effektiven Frühwarnung für die Unternehmenssicherheit gelingt es mit der vorgenannten Analyseinfrastruktur, über 20.000 relevante Kanäle und Gruppen kontinuierlich auszulesen. Neben der dadurch möglichen Suchfunktion und Inhaltserschließung können mit den Datensätzen umfangreiche Netzwerke nachgezeichnet werden.

Durch die jeweilige Zuordnung, welche Kanäle die aktive Nutzerschaft einer Gruppe noch nutzt, lässt sich beispielsweise eine direkte Nähe zu russischen Inputgebern ableiten, die Einschätzung der Gruppe also entsprechend bewerten.

Zusätzlich lassen sich – durch mathematische Metriken – verschiedene funktionale Rollen in den Verbreitungsnetzwerken erkennen: Etwa Brücken zu anderen Gruppen oder Multiplikatoren.

Erkannte Desinformationskampagnen und andere Emotionalisierungen lassen sich auf Basis dieser Transparenz akteursbezogen zurückverfolgen und idealerweise begrenzen.

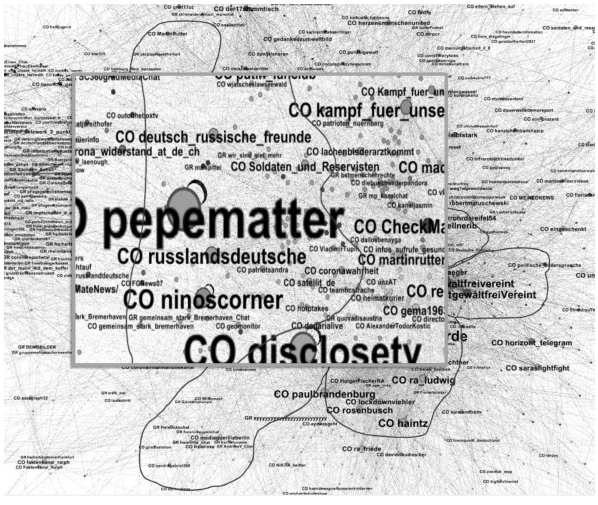


Abbildung 4: Netzwerk von Telegram-Gruppen und -Kanälen aufgrund ihrer Nähe (eigene Erstellung).

Künstliche Intelligenz: Lokationen und andere Entitäten

Für viele Sicherheitsaufgaben stehen bestimmte Lokationen im Zentrum. Durch Desinformationskampagnen oder Demonstrationsaufrufe können Personen zu räumlichen Aktionen bewegt werden: Die jeweilige Nähe solcher Aktionen zu sicherheitsrelevanten Lokationen ist eine überaus wichtige Information.

Modelle der Künstlichen Intelligenz können den kontinuierlichen Beitragsstrom filtern und eine lokale Verortung (Georeferenzierung) von sicherheitsrelevanten Aspekten im regionalen Adressumkreis von Unternehmensstandorten, Wohndomizilen, Liegenschaften, kritischer Infrastruktur, Flughäfen, Veranstaltungsorten etc. leisten.

Auf diese Weise werden (intendierte) Aktionen plastisch auf Land- bzw. Umgebungskarten abgebildet und im Idealfall kontinuierlich fortgeschrieben. Auf ähnliche Weise können – mit anderen KI-Modellen – Entitäten und Zeitangaben aus dem Rauschen herausgefiltert und in eine passende Ordnung gebracht werden.

Fazit

Durch verschiedene Analyseansätze können mehrschichtige Übersichten (Was: Signale, Wer: Akteure, Wo: Lokationen) aus dem kontinuierlichem digitalen Beitragsstrom destilliert werden, die für Analysten echtzeitnah Auffälligkeiten, Entwicklungen, aber auch schlichte Treffer darstellen.

Die Bewertung einer etwaigen Bedrohung, sei es durch eine Desinformationskampagne oder beispielsweise aktivistische Blockadeplanungen, muss zum derzeitigen Stand durch menschliche Analysten mit hinreichendem Kontextwissen vorgenommen werden.

Es empfiehlt sich eine mehrdimensionale Bewertung der jeweiligen Lage. Ein erprobtes Setting nimmt mehrere Dimensionen als **Sicherheitsradar** auf: Aggressivität der Vorhaben/Beiträge, Aktualität/Frequenz, Ansteckungsgefahr, Aktionskraft der Akteure und Anschlussfähigkeit der jeweiligen Thematik. Im Fall des Auftauchens falscher Narrative ist etwa die Beurteilung der Ansteckungsfahr ein wichtiger Faktor.

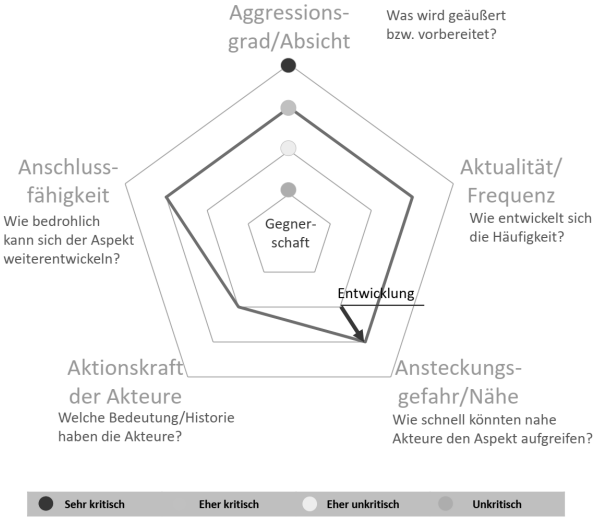


Abbildung 6: Sicherheitsradar für ein Unternehmen (eigene Darstellung)

Eine automatische Detektion von Desinformation ist – auch mit Künstlicher Intelligenz – zum derzeitigen Stand grundsätzlich nicht möglich: So kann auch geschickt selektierte, aber „dekontextualisierte" wahre Information desinformieren.

Ein strukturierter Gegner kann Desinformation als Methode einsetzen, um bestehende Diskussionen aufzunehmen, zu emotionalisieren und auf eine Aktivierung oder gar Mobilisierung von Akteursgruppen hinzuwirken.

Es wurde aber dargestellt, dass sich ungewöhnliche Beitragsinhalte, d.h. signifikante Terme, frühzeitig ausmachen lassen, auffällige Akteure mitunter schnell verorten und sich eine mögliche geographische, personelle oder thematische Nähe herausfiltern lassen kann.

Damit zeichnen sich gewichtige sicherheitsrelevante Entwicklungen bereits a priori ab. Ein Verzicht auf eine solche digitale Aufklärung würde folglich einen unnötigen Blindflug bedeuten. In der konkreten Umsetzung bildet das Zusammenspiel von Bedrohungslage, Sicherheitsbedürfnis und eigenen Kapazitäten damit für die Unternehmenssicherheit den Kontext, um über den Einsatz einzelner Werkzeuge oder die Hinzunahme eines spezialisierten Lagezentrums im Rahmen der digitalen Lageverfolgung zu entscheiden.

Noch steht die Entwicklung erst am Anfang. Um aber in die Lage zu kommen, Desinformationskampagnen wirkungsvoll zu bekämpfen, leistet die frühzeitige Detektion der Desinformation durch *Digital Listening* einen grundlegenden Beitrag.