März 2019 · Alexander Sängerlaub

Der blinde Fleck digitaler Öffentlichkeiten

Warum sich Desinformationskampagnen in sozialen Netzwerken kaum systematisch analysieren lassen

Stiftung Neue Verantwortung

Think Tank für die Gesellschaft im technologischen Wandel

Executive Summary

Auch bei der EU-Parlamentswahl im Mai besteht Sorge vor Desinformationskampagnen, die online stattfinden und so Einfluss auf den Wahlprozess nehmen könnten. Derzeit ist es fast unmöglich, systematische Untersuchungen über die Verbreitungswege, die beteiligten Akteure und die Reichweite von Desinformationskampagnen innerhalb der Social-Media-Plattformen anzustellen. Wir können derzeit kaum systematisch untersuchen, wie viel Desinformation sich auf welchen Social-Media-Plattformen befinden, wie sehr das Design der Plattformen zur Verbreitung von Falschinformationen beiträgt, wer Desinformation systematisch verbreitet oder wie stark der Einfluss auf unsere politischen Debatten ist. Dies ist nicht nur ein Problem für Fact-Checking-Organisationen oder Behörden, die versuchen, gegen Hetze oder Propaganda vorzugehen. Es ist auch ein enormes Problem für politische Entscheidungsträger:innen, die derzeit in vielen Ländern aber auch supranational, versuchen, wirksame Maßnahmen zum Schutz demokratischer digitaler Öffentlichkeiten zu entwickeln.

Der Zugriff auf die dafür benötigten Daten von Facebook, Twitter, YouTube, Instagram und anderen Plattformen ist derzeit stark eingeschränkt, sodass Untersuchungen der digitalen Öffentlichkeit quasi blind sind. Die benötigten Daten werden von den Plattformen zwar erhoben, sie stehen der Öffentlichkeit aber nicht zur Verfügung. Dabei geht es explizit nicht um die privaten Daten von Nutzer:innen, sondern um die Daten von Beiträgen, die in sozialen Medien öffentlich geteilt werden – etwa von Firmen, Nachrichtenseiten, Politiker:innen oder Influencer:innen.

Für die systematische Analyse solcher öffentlichen Beiträge sind unterschiedliche Datenpunkte relevant. Dazu gehören unter anderem technische Daten (wer hat den Beitrag wann erstellt), Reichweitendaten (wie vielen Nutzer:innen wurde der Beitrag angezeigt, wie viele Accounts haben mit dem Beitrag interagiert, indem er geliket oder geteilt wurde), Suchergebnisdaten (über welche Stichwörter kann man den Beitrag in Suchmaschinen oder auf der Plattform selbst finden), Daten darüber, ob der Beitrag als Werbeanzeige geschaltet wurde und wie viel Geld dafür ausgegeben wurde sowie Angaben dazu, ob der Beitrag nachträglich geändert wurde.

Alle diese Daten können derzeit nur bedingt systematisch eingesehen werden. Dies gilt für Facebook, Twitter, Instagram und YouTube, die in Bezug auf Nutzer:innenzahlen zu den größten sozialen Netzwerken in Deutschland zählen und die damit auch relevant für die öffentliche Meinungsbildung sind.

Sie alle stellen nur sehr vereinzelt Schnittstellen bereit, die aber keine umfassenden systematischen Analysen ermöglichen.

Wie problematisch es ist, wenn Daten darüber fehlen, welche Social-Media-Desinformationskampagnen kursieren, wie sie sich verbreiten und welche Akteure an ihrer Verbreitung beteiligt sind, zeigten die Vorfälle in Chemnitz im Sommer 2018, bei denen ein junger Mann getötet wurde. In der Folge gab es Ausschreitungen durch Rechtsextreme, die sich auch online organisiert und durch Desinformation Stimmung gemacht haben. Versucht man, diese Desinformationskampagnen systematisch zu analysieren, stößt man je nach Art des Beitrags schnell an die Grenzen. Reine Textbeiträge können auf manchen Plattformen zumeist noch durch die Suchfunktion gefunden werden, etwa über das Stichwort "Chemnitz". Inhalte in Bild- und Videodaten lassen sich aber so nicht finden, obwohl es durch automatische Texterkennung technisch möglich wäre. Doch auch die Suchfunktion selbst ist problematisch: Die Reihenfolge, in der Ergebnisse angezeigt werden, birgt das Potential, zur Verbreitung von Desinformation beizutragen. Ob ein Beitrag bestimmten Nutzer:innen als Werbung angezeigt wird, um die Sichtbarkeit zu erhöhen, ist auf Facebook derzeit nicht systematisch nachvollziehbar.

Der Zugang zu relevanten öffentlichen Daten der Social-Media-Plattformen muss verbessert werden, damit Forschung und Politik in Sachen Desinformation nicht länger im Dunkeln tappen. Die wenigen Datenschnittstellen, die etwa Google oder Facebook auf freiwilliger Basis für politische Werbung im Vorfeld von Wahlen anbieten, reichen nicht aus. Benötigt werden Datenzugänge, die sowohl das Echtzeit-Monitoring von sozialen Netzwerken ermöglichen, als auch den Zugriff auf Daten aus der Vergangenheit. Erst wenn diese Daten vorliegen, können Regierungen und Behörden wirksame Maßnahmen gegen Desinformation entwickeln und überprüfen. Gleichzeitig würde dies auch die Wissenschaft in die Lage versetzen, fundierte Studien zu den Herausforderungen unserer digitalen Öffentlichkeit zu erstellen.

Inhalt

1. Blinde Flecken, viele Fragen

Die Demokratisierung unserer Öffentlichkeit, an der jede und jeder heute mehr denn je teilhaben kann, ist ein großer Gewinn der demokratischen Öffentlichkeiten des 21. Jahrhundert. Doch es gibt eine Kehrseite: Populismus, Hate Speech, Propaganda und Desinformation im Netz sind ebenso zu einem festen Bestandteil digitaler Öffentlichkeiten geworden. Politik, Regulierungsbehörden, Fact-Checking-Organisationen und viele zivilgesellschaftliche Organisationen haben reagiert und versuchen derzeit mit verschiedenen Mitteln, digitale Desinformation zu bekämpfen. Doch um solche Initiativen sinnvoll und zielgerichtet umsetzen und evaluieren zu können, ist es unabdingbar, belastbare Informationen über das Ausmaß, Ausprägungen und Akteuren von Desinformationskampagnen zu erlangen: Wie viel und welche Formen von Desinformation befinden sich auf den Social-Media-Plattformen? Welche außer- und innerstaatlichen Akteure versuchen mit welchen Methoden Einfluss auf Debatten zu nehmen? Wie sehr trägt das Design der Plattformen zur Verbreitung von Falschinformationen bei? Wie erfolgreich ist das Fact-Checking? Wer kreiert Desinformation, wer teilt sie, welche Rolle spielt der klassische Journalismus und was bedeutet das alles für die Güte unserer Debatten in der Demokratie?

Das sind relevante Fragen, auf die derzeit nur schwer Antworten zu finden sind. Denn derzeit ist es für Politik, Wissenschaft, Journalismus (darunter auch die Fact-Checking-Organisationen), NGOs weitestgehend unmöglich, solche Fragen umfassend zu beantworten. Was innerhalb sozialer Netzwerke passiert, entzieht sich derzeit an vielen Stellen einer systematischen Analyse. Es fehlen schlichtweg die Zugänge zu wichtigen, öffentlichen Daten. Ohne solche Daten können höchstens – falls überhaupt – stichprobenartige Eindrücke darüber gewonnen werden, was innerhalb digitaler Öffentlichkeit sozialer Netzwerke kursiert.

Wie sich dieser Blindflug auf die mediale und politische Öffentlichkeit auswirkt, kann man anhand der Vorgänge im August 2018 in Chemnitz gut erkennen. Hier hat die politische Öffentlichkeit inzwischen durch aufwendige Recherchen von Fact-Checking-Organisationen und Journalist:innen zumindest eine ungefähre Ahnung davon, wie hier die Plattformen als Transmitter und Katalysatoren zur Verbreitung von Desinformation beigetragen haben. Gerüchte, Lügen und falsche Tatsachen zu den Tathergängen verbreiteten sich über YouTube, Facebook und Twitter wie Lauffeuer und hatten ihren Anteil an der starken Mobilisierung rechtextremer Kräfte auf den Straßen von Chemnitz. Diesen Einzelfall möchten wir nutzen, um zu zeigen, welche Daten

es eigentlich bräuchte, um Desinformationskampagnen systematisch zu beobachten. Denn demokratische politische Öffentlichkeiten sind nur funktionstüchtig, wenn alle teilnehmenden Akteure auch beobachten können, was in diesen Öffentlichkeiten passiert, solange sie eine relevante Rolle für unsere digitalen Öffentlichkeiten in der Mediennutzung spielen.[1]

Insbesondere Facebook hat sich dabei in den letzten Jahren unter dem Druck der Öffentlichkeit nicht etwa weiter geöffnet und Zugriffe auf ihren hohen Datenschatz zur digitalen Öffentlichkeit ermöglicht. Im Gegenteil wurden nach dem Skandal um Cambridge Analytica Datenschnittstellen geschlossen, sodass die Beobachtung nun noch schlechter möglich ist, als sie es ohnehin schon war.[2] Dass durch Cambridge Analytica private Daten von Nutzer:innen illegal weiterverbreitet wurden, sollte generell nicht als Vorwand für Social-Media-Unternehmen fungieren, öffentliche Daten unter Verschluss zu halten. Dazu gehören eben nicht die persönlichen Informationen wie Privatnachrichten oder sensible personenbezogene Daten, sondern Postings die als Werbeanzeigen geschaltet werden, auf YouTube als Video hochgeladen, in Bildform über Facebook, Twitter oder Instagram als Meme geshared und geliked oder sogar von Medien aufgenommen und weiterverbreitet werden, weil diese die Herkunft bestimmter Information zuweilen nur ungenau prüfen.[3]

Damit lösen soziale Netzwerke eine tiefgehende Disruption auf unsere Öffentlichkeiten aus. Propaganda, Falsch- und Desinformation können sich zum Teil ungehindert verbreiten[4] und es gibt kaum Möglichkeiten, dieses Phänomen unabhängig und datenbasiert zu untersuchen.

Dieses Papier zeigt am Beispiel der Vorfälle in Chemnitz, was für öffentliche Daten derzeit für die Bekämpfung von Desinformaton – das heißt für wissenschaftliche Forschung, Fact-Checking-Organisationen und Journalist:innen fehlen und welche Probleme dies für Politik und Gesellschaft erzeugt. Es geht, auch das soll deutlich werden, dabei nicht um die privaten oder perso-

1 Reuters Institute (2018): Digital News Report 2018. http://media.digitalnewsreport.org/wp-content/uploads/2018/06/digital-news-report-2018.pdf?x89475

2 Axel Bruns (25.04.2018): Facebook Shuts the Gate after the Horse Has Bolted, and Hurts Real Research in the Process. https://medium.com/@Snurb/facebook-research-data-18662cf2cacb

3 Alexander Sängerlaub/Miriam Meier/Wolf Dieter-Rühl (2018): Fakten statt Fakes. Verursacher, Verbreitungswege und Wirkungen von Fake News im Bundestagswahlkampf 2017, https://www.stiftung-nv.de/de/publikation/fakten-statt-fakes-verursacher-verbreitungswege-und-wirkungen-von-fake-news-im

4 Soroush Vosoughi, Deb Roy, Sinan Aral (2018): The spread of true and false news online. In Science: 03/2018, Vol. 359, http://science.sciencemag.org/content/359/6380/1146.full

nenbezogene Daten der Nutzer:innen, sondern nur um Daten, die eigentlich ohnehin öffentlich sind oder es sein sollten.

Darüber hinaus plädiert das Papier dafür, einen kontrollierten Datenzugang zu schaffen, damit Zivilgesellschaft, Wissenschaft und Gesetzgeber eine Möglichkeit erhalten, die Probleme in den Plattformen besser zu verstehen und öffentlich nachzeichnen können.

Im letzten Kapitel werden neue und bisherige Ansätze, auch der Plattformen, vorgestellt, um zu zeigen, wie der Zugang konkret in der Anwendung aussehen kann.[5] Es ist dringend Zeit, die zahllosen öffentlichen Daten, die Plattformen wie Facebook ohnehin erheben, der Gesellschaft einfach zugänglich zu machen, damit wir die Vorgänge in den Plattformen besser verstehen.

2. Daten, Datenschnittstellen und öffentlich vs. privat

Viele Daten in den sozialen Netzwerken sind öffentlich einsehbar – weil sie per Definition öffentlich sind: seien es Tweets bei Twitter, Beiträge von Spiegel Online bei Facebook oder Videos von alternativen Medienanbietern, wie RT Deutsch, bei YouTube. Über die Suchfunktionen der Plattformen, wie Twitter, Instagram, Facebook oder YouTube, genauso über Suchmaschinen, wie Google, lassen sich diese Inhalte auffinden. Doch nur weil sich Inhalte auffinden lassen, sind sie noch lange nicht systematisch für die Wissenschaft oder Fact-Checking-Organisationen analysierbar. Hierfür bedarf es Daten-

5 Eine Möglichkeit wäre hier ein ähnliches Transparenzangebot zu machen, wie mit dem Wahlwerberegister. Diese wurden von Facebook und Google zur Europawahl angekündigt, sind aber bisher noch nicht veröffentlich worden. Näheres dazu in Kap. 4. Wahlwerberegister. Facebook nennt dieses Tool „Werbebibliothek“, diese ist (Stand: 08.03.2019) bisher für die USA, Großbritannien, Indien und Brasilien verfügbar: https://www.facebook.com/ads/archive/?active_status=all&ad_type=political_and_issue_ads&country=ALL, Google nennt seine Übersicht „Politische Werbung bei Google“, diese enthält bisher (Stand 08.03.2019) nur die USA: https://transparencyreport.google.com/political-ads/overview

schnittstellen[6], die Daten beispielsweise für Social Media Intelligence Tools[7] zugänglich machen. Diese Tools ermöglichen es, die Daten aufzubereiten, auszuwerten, graphisch darzustellen, aggregierbar oder kategorisierbar zu machen. Auch gibt es Daten, die für die Öffentlichkeit zwar relevant wären, aber bisher nicht sichtbar sind, weil die Plattformen sie zwar erheben, aber unter Verschluss halten. Dazu gehören beispielsweise Daten zum Erfolg des Fact-Checkings als Maßnahme gegen Desinformation.

Die verschiedenen Typen von Daten wollen wir in diesem Kapitel vorstellen, um deutlich zu machen, welche Daten für die politische Öffentlichkeit von den Plattformen zugänglich gemacht werden sollten. Das anschließende Kapitel zur „Klaviatur der Desinformation" im konkreten Anwendungsfall zu Chemnitz soll noch einmal die Dringlichkeit der Zugänglichkeit dieser Daten unterstreichen sowie das vielleicht etwas abstraktere *Kap. 2* durch Beispiele unterfüttern.

Welche Daten sind relevant und öffentlich?

Soziale Netzwerke erheben eine Vielzahl von nutzerspezifischen Verhaltensdaten, die Aufschluss über die politische Einstellung, die sexuelle Orientierung oder anderweitig grundprivate Daten einzelner Personen geben. All

6 **Datenschnittstellen**, bzw. „Anwenderprogrammierschnittstellen" (englisch API für application programming interface) ermöglichen die Anbindung von einer Software an die andere. Der Zugriff auf die Datenschnittstellen ermöglicht es wissenschaftlicher Forschung, Daten einfach zu erheben, um diese dann auszuwerten. **Social-Media-Intelligence-Tools** nutzen diese Anbindung, um auch ihr grafisches Interface mit den Plattformdaten zu verbinden, sodass die Anwendung und Analyse der Daten auch ohne Programmierkenntnisse möglich sind. Auch haben die Schnittstellen rechtlich Bedeutung, da eine öffentlich verfügbare oder eingekaufte Anbindung an eine Schnittstelle das Arbeiten mit den zur Verfügung gestellten Daten erlaubt – anders als beim Scrapen (also dem illegalen Abschöpfen von Daten von außen, das durch die meisten Terms of Service der Plattformen untersagt ist). Die Plattformen wiederum haben den Vorteil, dass sie innerhalb der Schnittstellen bestimmen können, welche Daten sie freigeben und damit eine bessere Übersicht über das eigene Datenmanagement haben.

7 Am Markt gibt es zahlreiche ähnliche Tools zu dem von uns 2017 benutzten Tool Talkwalker wie BuzzSumo, Pulsar, Meltwalter, Crimson Hexagon, complexium, Alto und andere. Sie bieten alle auch **Social Media Intelligence** an, also die systematische Analyse von Daten aus sozialen Netzwerken. Nicht immer sind diese – wie bei Talkwalker – gepaart mit Daten aus klassischen Medienangeboten. All diese Tools sind allerdings in ihrer Funktionstüchtigkeit auch davon abhängig, welche Schnittstellen die jeweiligen Plattformen (wie Facebook, Twitter, etc.) für den Datenzugriff gewähren. Fast alle Tools unterliegen kommerziellen Interessen, ähnlich wirkmächtige frei verfügbare Tools, auf die einfach zugegriffen werden kann, existieren unseres Wissens derzeit nicht. Der Nachteil ist, dass diesen Tools selten eine (sozial-)wissenschaftliche Methodologie zugrunde liegt. Will heißen: Oft sind die Daten nur mit Vorsicht zu genießen, wenn eine fundierte Grundausbildung im quantitativ empirischen Arbeiten fehlt (die nicht alle kommerziellen Tools automatisch mitliefern) oder das Wissen um Kontexte und Funktionslogik digitaler Öffentlichkeit.

diese sehr persönlichen und privaten Daten, die auch eine Rolle beim Cambridge-Analytica-Skandal spielten, sind ausdrücklich *nicht* gemeint, wenn es um – für die Öffentlichkeit relevante – öffentliche Daten zu politischen Debatten innerhalb der sozialen Netzwerke und den damit verbundenen Fragen zu Desinformation, Hate Speech oder alternativen Medienechokammern geht.

Welche Daten relevant wären, soll an einem einfachen Beispiel aus Facebook verdeutlicht werden:

Abbildung 1: Ein öffentlicher Post der Seite „RT Deutsch“ vom 21.01.2019.

Abb. 1 zeigt einen Beitrag der Seite „RT Deutsch“ auf Facebook. In dem Beitrag ist ein Video von Donald Trump angehängt, welches auch über den Beitrag verlinkt wird. Der Beitrag selbst ist keine Desinformation, zeigt aber ein Framing, das auch von Donald Trump gern bemüht wird: „Mainstream-Medien glaubt keiner mehr“. Würde man – egal mit welchem Social-Media-Intelligence-Tool – nach „Trump“ für einen gewissen Zeitraum suchen, um beispielsweise herauszufinden, wie die Debatte in den sozialen Netzwerken über den US-Präsidenten geführt wird, würde dieser Beitrag nicht gefunden werden, weil derzeit die dafür nötige Datenschnittstelle fehlt (bzw. von Facebook nach Cambridge Analytica geschlossen wurde[8]).

Der Beitrag der Seite erschien am Montag, dem 21.01.2019, auf der Facebook-Seite der alternativen Nachrichtenplattform *RT Deutsch*, die selbst immer wieder Desinformation teilt und vom russischen Staat mitfinanziert wird.[9] Diese Seite ist *öffentlich*, das heißt, dass ihre Inhalte beispielsweise über Suchmaschinen außerhalb von Facebook oder via Suchfunktion innerhalb der Plattform zugänglich sind und jede:r Nutzer:in die Inhalte in der Plattform sehen kann. Damit sind die Daten dieser Seite aber noch lange nicht systematisch analysierbar: Dafür braucht es Datenschnittstellen, welche die Daten in sogenannte Rohdaten für andere Softwareanwendungen überführen, mit deren Hilfe zum Beispiel statistische Analysen, grafische Aufbereitungen oder thematische Suchen durchgeführt werden können. Welche Daten stecken alle in diesem Beitrag?

8 Der einzige Ausweg wäre derzeit RT Deutsch als Seite „zu beobachten“. D.h. das die Seiten, die für die Analyse interessant wären, vorher definiert werden müssen, anstatt das Facebook einfach nach Keywords über die Datenschnittstelle durchsucht werden kann. Gerade aber bei der Logik von sozialen Netzwerken und in Bezug auf Desinformation ist a) vorneherein nicht klar, welche Seiten Desinformation teilen werden, b) ob dieser Content innerhalb der Plattform erstellt/hochgeladen wird (wie hier das hochgeladene Video) oder außerhalb der Plattform erstellt und auf den Plattformen nur geteilt wird und c) es für jede öffentliche Seite potentiell möglich ist, einen „viralen Hit“ zu landen, der Aufmerksamkeit generiert. Die Beispiele in Kapitel 3 vertiefen diese Problematik.

9 Bundeszentrale für politische Bildung (28.06.2016): Propaganda und Desinformation. Ein Element "hybrider" Kriegführung am Beispiel Russland, http://www.bpb.de/apuz/232964/propaganda-und-desinformation?p=all

Technische Daten:	Absender: RT Deutsch, Uhrzeit/Datum: Montag, 21.01.2019, 15:40
Bilddaten:	im Beispiel keine[10]
Videodaten:	Das Video über Donald Trump, welches untertitelt ist. Dieser Text könnte wiederum durchsuchbar gemacht werden, um Schlagworte aus dem Video ebenfalls auffindbar zu machen.
Suchergebnisdaten:	Bei welchen „Schlagworten" taucht dieser Beitrag innerhalb der Facebook-Suchfunktion an welcher Position auf? Mit anderen Worten: Wenn sich Nutzer:innen zu Trump über Facebook informieren, erhalten sie zuerst Vorschläge von Medien wie RT Deutsch oder von bspw. Qualitätsmedien wie Tageszeitungen oder öffentlich-rechtlichen Angeboten?
Performance/ Reichweite:	Shares, Likes & Comments = „Engagement" als Social-Media-Währung für „aktive Reichweite" sowie die Aufrufzahlen (bisher nur für Videos verfügbar). Facebook weiß auch, bei welchen Zielgruppen dieses Video erfolgreich war. Diese Daten könnten aggregiert auch für die Öffentlichkeit verfügbar gemacht werden.
Metadaten:	Nicht im Beitrag sichtbar, aber theoretisch könnte dieser Beitrag von einer Fact-Checking-Organisation (für Deutschland: Correctiv) geprüft worden sein, ob die Inhalte im Video faktisch richtig sind. Facebook hat diese Informationen für jeden Post, sie sind jedoch bisher für niemanden systematisch erfassbar.

10 Siehe für den Zugang zu Bilddaten, wenn es sich mit Bilder inklusive Text handelt, Kapitel 3.

11 Dieses Sichtbarmachen von Änderungen – hier ist Facebook vorbildlich – sollte ein genereller Standard in der digitalen Öffentlichkeit werden, um Leser:innen über Veränderungen an Texten zu informieren. Oft werden im Nachhinein auf „Nachrichten"-Seiten wie BILD oder Epoch Times Textelemente geändert und Falschinformationen im Nachhinein ausgebessert, ohne das dies für die Nutzer:innen der Seiten sichtbar gemacht wird. Zahlreiche Beispiele hierfür finden sich auch in unserer Bundestagswahlstudie.

Editierte Inhalte: Facebook bietet die Option an, den Bearbeitungsverlauf eines Beitrages für jede:n Nutzer:in anzeigen zu lassen (Abb. 2). Hier lässt sich nachverfolgen, ob und wie Beiträge im Nachhinein verändert wurden. Twitter „löst" dieses Problem, indem Tweets nach der Veröffentlichung nicht mehr bearbeitet werden können.[11] Instagram oder YouTube bieten keine solche Funktion an.

Gelöschte Inhalte: Tools wie *Talkwalker* können durch die Speicherung von Daten durch regelmäßige Schnittstellenabfragen auch Beiträge sehen, die später gelöscht wurden. Gerade beim Thema Desinformation, gibt es immer wieder kritische Beiträge, die nachträglich von ihren Verbreiter:innen gelöscht werden – ohne, dass es für die Nutzer:innen kenntlich wäre. Über archivierte Datenbanken sind diese Beiträge manchmal weiter auffindbar.

Werbeanzeigen: Dieser Beitrag könnte auch von RT Deutsch als Werbeanzeige geschaltet worden sein. Die Werbebibliothek von Facebook[12] müsste diese Werbung dann anzeigen und würde dadurch transparent machen, *wer*, *wie viel Geld* ausgibt, um *wen* damit auf der Plattform zu erreichen.

12 Die Werbebibliothek von Facebook ist ein Transparenztool, das politische Werbeanzeigen für alle durchsuchbar und sichtbar macht. Die Werbebibliothek wird dezidiert in Kap. 4 vorgestellt.

Abbildung 2: Bearbeitungshistorie eines Facebook-Posts der BILD-Zeitung (der Case stammt aus unserer Studie "Fakten statt Fakes", in der ersten Version (08.07.) werden Falschinformationen geteilt [auf dem Bild ist weder der Täter zu sehen, noch stand zur Debatte, ob ein Beamter sein Augenlicht verliert]) – deutlich wird hier, wie die Bildzeitung falsche Informationen editiert und entfernt, die Änderungen aber selbst im Post nicht kenntlich macht. Problematisch, wenn ein Großteil der Reichweite auf die Falschinformation (08.07.) entfällt – die spätere(n) Bearbeitung(en) (10.07., 12.07.) werden dabei kaum registriert. Facebooks Editionshistorie ist hier sehr hilfreich, die Chronologie von Desinformation herzustellen, ist aber so oder so auch kein Teil der früher verfügbaren Schnittstellen-Daten (es handelt sich um einen Zufallsfund).

Aus der Summe solcher Daten lassen sich am Ende konkrete Fragen beantworten, die beispielsweise in Bezug auf Desinformationskampagnen von großer Bedeutung sind:

Wer teilt Desinformation? Wie viele Menschen erreichen Desinformation? Wurden diese durch Fact-Checking-Organisationen geprüft? Wie ist das Verhältnis zwischen Desinformation und korrekten Nachrichten zum Thema XY? Werden Desinformation auch über Werbeanzeigen geteilt und wenn ja, wie viel Geld wird darin von wem investiert, um wen damit zusätzlich zu erreichen, der u.U. diesen Seiten auf den Plattformen gar nicht folgt? Wurden falsche In-

halte nachträglich bearbeitet? Wie werden Desinformation in sozialen Netzwerken kommentiert und vielleicht sogar debunked? Etc.

Einige dieser Fragen – aber auch nicht alle – konnten wir bspw. in unserer Studie zur Bundestagswahl 2017 zumindest für die dort aufgeführten Fallbeispiele beantworten.[13] Konkrete Informationen zur methodischen Erhebung sind zudem in einem separaten Methodenpapier zusätzlich veröffentlicht.[14] Allerdings wären die dort beschriebenen Erhebungen derzeit nicht möglich, da, gerade bei Facebook, wichtige Datenschnittstellen (u.a. im August 2018) geschlossen wurden.[15] Andere Teile dieser Daten (wie z.B. die Metadaten zum Fact-Checking oder die Untertitel innerhalb des Videos) sind bisher noch gar nicht verfügbar gewesen. Auch das systematische Erforschen von Memes und deren Verbreitung ist bisher kaum möglich.

Inwieweit die Enthüllungen zu Cambridge Analytica durch die britische Zeitung The Guardian im Sommer 2018 bei der Schließung wichtiger Datenschnittschnellen durch Facebook eine Rolle gespielt haben mögen, kann hier nur gemutmaßt werden. Die Zeitschrift T3N dokumentiert: „Die Media-Solutions-API umfasst mehrere API, mit denen Tools für Medienpartner entwickelt werden können. Künftig werden API zur öffentlichen Inhaltssuche auf Seiteninhalte und öffentliche Posts auf bestimmten verifizierten Profilen beschränkt. Nachdem bereits verschiedene Tools dieser Familie deaktiviert wurden, folgen wegen geringer Nutzung am 1. August auch die API für Topic Search, Topic Insights, Topic Feed und Public Figure.“[16]

Als „Ersatz“ wurde das reduzierte „Einhängen von Kanälen“ weiterhin ermöglicht. Das heißt, dass sich thematische Suchen über die API mittels Tools insofern durchführen lassen, wenn man vorher die zu durchsuchenden Kanäle benennt. Im Falle der Messung von Desinformation ist es jedoch utopisch, bereits im Vorhinein abzusehen, welche Kanäle wann und wie welche Desinformation teilen werden. Diese Definition im Vorhinein einer Messung festzulegen widerspricht der Natur der Sache, wie wir im Folgekapitel zu Chemnitz auch noch einmal veranschaulichen werden. Die Anzahl der be-

13 siehe Fußnote 4.

14 Wolf-Dieter Rühl (12/2017): Measuring Fake News – Die Methode. https://www.stiftung-nv.de/sites/default/files/fake_news_methodenpapier_deutsch.pdf

15 Offizielle Ankündigung von Facebook vom 02. Juli 2018 zur Schließung der Schnittstellen: https://newsroom.fb.com/news/2018/07/a-platform-update/

16 T3N (03.07.2018): „Facebook kündigt neue API-Einschränkungen für Apps an“, URL: https://t3n.de/news/facebook-kuendigt-neue-1092496/

obachtbaren Kanäle wird zudem auch noch durch einige Anbieter der Social-Media-Intelligence-Tools begrenzt.

In der Summe der Maßnahmen lässt sich dadurch etwa ein Viertel der Daten aus der von uns untersuchten Desinformation in der gesamten digitalen Öffentlichkeit nicht mehr systematisch beobachten, verglichen mit den Daten, die wir bei der Bundestagswahl erheben konnten – allein durch das Schließen der Facebook-Schnittstellen. Heißt: 27 % der von uns erhobenen Daten zu Desinformation zur Bundestagswahl stammen originär aus Facebook[17] und lassen sich derzeit nicht beobachten. Durch diese Verzerrung und den Wegbruch der Datenlage war es insofern auch methodisch unverantwortlich, beispielsweise eine systematische Analyse zu Desinformation während der bayerischen Landtagswahl durchzuführen – wie es unser Projektteam eigentlich vorhatte. Methodische Alternativen zu diesem Verfahren, die auch rechtlichen Rahmenbedingungen standhalten (also: kein Scrapen), gibt es derzeit unseres Wissens nach nicht.

Unabhängig davon, ist die Analyse von Social-Media-Daten ohnehin mit hohen organisatorischen und finanziellen Aufwendungen verbunden[18], da öffentlich zugängliche Daten, inklusive Datenschnittstellen nicht existieren. Oder die Wissenschaft auf extrem teure Social-Media-Intelligence-Tools angewiesen ist, deren Datenbasis oft schwer nachvollziehbar ist. Das liegt auch daran, dass diese Tools vor allem zu kommerziellen Zwecken eingesetzt werden (Marketing, Marktbeobachtung, Krisenkommunikation).

Abbildung 3 fasst nun noch einmal zusammen, welche Daten, die für Journalismus, Wissenschaft, NGOs oder andere Forschungsorganisationen einsehbar und systematisch analysier- und durchsuchbar sein müssten.

17 Berechnungen unseres Research-Partners Unicepta zur Folge.

18 Die Folge davon ist, dass es in der Wissenschaft einen gigantischen Überfluss an Twitter-Studien gibt, weil die Datenpolitik des Konzerns wesentlich offener und transparenter ist. Twitter hat allerdings in kaum einem Land die gleiche Bedeutung für digitale Öffentlichkeiten, wie Facebook. Die Übertragung von wissenschaftlichen Erkenntnissen von einer Plattform zur anderen ist jedoch aus mehreren Gründen problematisch: unterschiedliches Design sowie die Funktionslogik der Plattformen spielen genauso eine Rolle wie verschiedene Zielgruppen, welche die jeweiligen Plattformen nutzen.

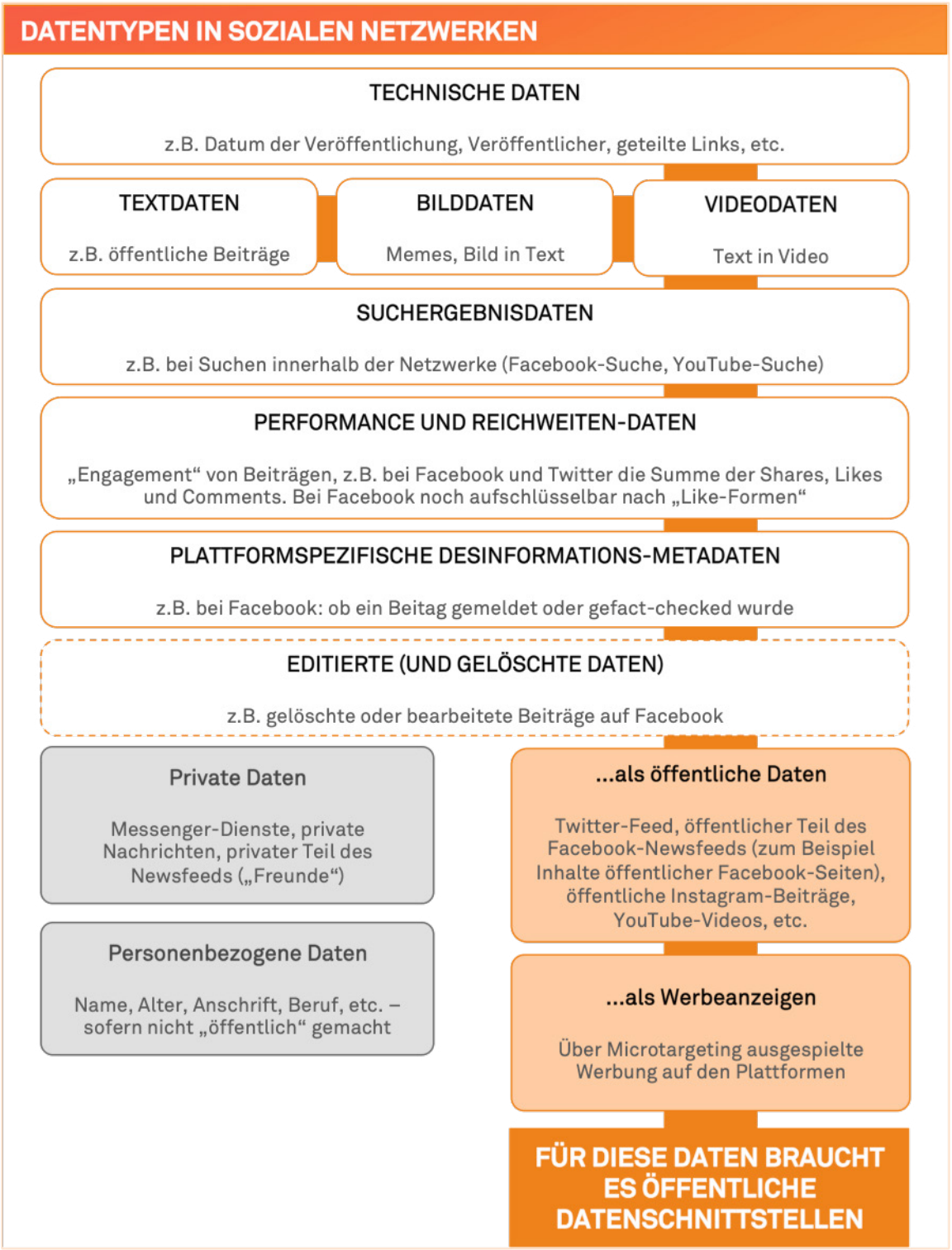

Abbildung 3: Datentypen innerhalb sozialer Netzwerke, die bspw. für das Nachvollziehen von Desinformationskampagnen für die Öffentlichkeit von Bedeutung wären.

3. Exkurs: Klaviatur der Desinformationsdaten – das Beispiel Chemnitz

In diesem Kapitel soll noch einmal die Bedeutung der Zugänglichkeit dieser Daten und Datenschnittstellen für die Öffentlichkeit verdeutlicht werden.[19] Dafür ziehen wir die Vorfälle in Chemnitz im August 2018 heran, da hier Desinformation in den sozialen Netzwerken einen großen Anteil an dem Ausmaß der Proteste und deren Organisation hatten. Viele Einzelheiten der Vorgänge in den sozialen Netzwerken sind nicht etwa der Transparenz von Facebook und Co. zu verdanken, sondern der investigativen Arbeit von Journalist:innen und Fact-Checking-Organisationen. Dabei müsste es nicht nur Aufgabe, sondern auch im eigenen Interesse der sozialen Netzwerke sein, solche Desinformationskampagnen zu unterbinden, damit Verschwörungstheorien, Lügen und Propaganda nicht auf einen fruchtbaren Nährboden fallen und massiv verbreitet werden. Letztlich sind soziale Netzwerke nicht immer alleiniger Auslösungsort, auch unseriöse Medienanbieter und schlecht gemachter Journalismus spielen zuweilen eine Rolle[20], allerdings haben die Plattformen eine Verantwortung als Transmitter und Katalysator von Desinformation.

Ausschreitungen in Chemnitz

Nach dem gewaltsamen Tod Daniel H.s in Chemnitz, in der Nacht vom 25. zum 26. August 2018, kam es in den Folgetagen dort zu Ausschreitungen durch Rechtsextreme und Neonazis, die im Anschluss in einer medialen Debatte bis hin zu Einlassungen des damaligen Leiters des Bundesamts für Verfassungsschutz Hans-Georg Maaßen endeten, ob es sich dabei um „Hetzjagden“ handle. Dabei sind viele Desinformation über soziale Netzwerke geteilt worden, unter anderem um Menschen zu mobilisieren, an Protestzügen teilzunehmen. Unter die Demonstrant:innen mischten sich auch Rechtsextreme und Neonazis, die sich über soziale Netzwerke verabredet hatten. Die vielen geteilten Desinformation haben wahrscheinlich zur Mobilisierung beigetra-

19 Wer bereits überzeugt ist, kann mit Kapitel 4 fortfahren.

20 Alexander Sängerlaub/Miriam Meier/Wolf Dieter-Rühl (2018): Fakten statt Fakes. Verursacher, Verbreitungswege und Wirkungen von Fake News im Bundestagswahlkampf 2017, https://www.stiftung-nv.de/de/publikation/fakten-statt-fakes-verursacher-verbreitungswege-und-wirkungen-von-fake-news-im

gen und dafür gesorgt, dass die Polizei mit der Masse der Demonstrierenden und Randalierenden überfordert war, wie der Spiegel aufgearbeitet hat.[21]

Die größte Reichweite erlangten dabei, sofern ohne Datenschnittstellen überhaupt rekonstruierbar, drei konkrete Falschaussagen: 1. der Getötete wollte angeblich deutsche Frauen beschützen, 2. es hätte noch ein zweites Mordopfer gegeben, 3. der Täter hätte den Mord vorher über soziale Netzwerke angekündigt. Darüber hinaus gab es noch eine Vielzahl weiterer nicht korrekter Aussagen, wie eine angebliche Reisewarnung zu Sachsen aus Kanada. Alle Vorfälle und Desinformation sind sehr sorgfältig und umfangreich von journalistischen Medien aufbereitet worden.[22] Der Sinn der gezielt verbreiteten Falschinformation: „...um Fakten kümmern sich die Urheber der zahlreichen Falschmeldungen nicht. Ihre Fake News sind Mittel zum Zweck: ‚Die Wut soll gezielt geschürt werden – und bei den geplanten Demonstrationen vom Netz auf die Straße getragen werden.'", schreiben Patrick Gensing und Karolin Schwarz im *Faktenfinder.*[23]

Im nun Folgenden geht es weniger um eine Rekonstruktion der Ereignisse, sondern viel mehr um die unterschiedlichen Formen der Desinformation sowie der Datenlogik dahinter. Diese orientieren sich nach wie vor an der Logik von *Abb. 3*.

Textdaten (öffentliche Beiträge)

Öffentliche Beiträge sind Beiträge von Seiten, denen jede:r Nutzer:in folgen kann – unabhängig ob nun bei Facebook, Twitter oder Instagram. Deren Beiträge erscheinen dann beispielsweise im News Feed der Nutzer:innen. Die plattformübergreifende Suche nach konkreten Beiträgen mittels Suchbegriffe via Tools wie Talkwalker – und sofern es sich nur um Text handelt – ist derzeit nur bei Twitter und YouTube möglich. Bei Instagram sind die Suchmöglichkeiten eingeschränkt (derzeit ist nur die Suche nach „Hashtags" möglich). Ebenso bei Facebook funktioniert die Suche über die Datenschnittstelle nur noch dann, wenn der Kanal vorher zur Analyse ausgewählt wurde (womit kanalunabhängige Themenanalysen nicht möglich sind). Ganz

21 SPIEGEL/SPIEGEL ONLINE (22.12.2018): Ausschreitungen in Chemnitz. Wie Politik und Polizei versagten, URL: http://www.spiegel.de/plus/ausschreitungen-in-chemnitz-wie-politik-und-polizei-versagten-a-00000000-0002-0001-0000-000161498502 (Paywall)

22 Ausführliche Hintergründe finden sich dazu beispielsweise bei Patrick Gensing und Karolin Schwarz, die für den Faktenfinder von Tagesschau.de berichteten: „Das Trauerspiel von Chemnitz" (30.08.2018), URL: https://faktenfinder.tagesschau.de/inland/chemnitz-fakes-trauerspiel-101.html

23 ebd.

konkret am Beispiel: Der Beitrag der *AfD Erzgebirge (Abb. 4)* würde nur bei Facebook mittels Tools über die API gefunden werden können, wenn der Kanal „AfD Erzgebirge“ vorher als zu beobachtend ausgewählt werden würde. Bei der Verbreitung von Desinformation ist es aber nahezu unmöglich, vorher auszuwählen, welche Kanäle diese verbreiten oder teilen werden.

Abbildung 4: Öffentlicher Beitrag der AfD Erzgebirge vom 26. August 2018. Der Hinweis, dass die Opfer „Frauen helfen wollten“ ist eine Falschinformation. Bisherigen Erkenntnissen nach handelte es sich um einen Raubüberfall. Auf vielen rechtsextremen Seiten wurde auch unterschlagen, dass das Opfer Deutsch-Kubaner ist.

Sonderfall: Textdaten als öffentliche Beiträge privater Personen

Auch Privatpersonen können ihre Beiträge auf öffentlich stellen, sodass sie von Nutzer:innen geteilt werden können. Dieses Konzept des „öffentlich privaten Beitrags“ gibt es so nur bei Facebook. Auch mithilfe der nun geschlossenen Schnittstelle war es nicht möglich, diese Posts systematisch zu analysieren. Rechte Influencer:innen, aber auch Privatleute, nutzen ihre privaten Facebook-Accounts zur Verbreitung von Falschinformationen, indem sie ihre Beiträge öffentlich stellen.

Hierfür zwei Beispiele (Abb. 5 & 6). Abb. 6 zeigt den privaten Account von Franz P., der ein hochgeladenes Video des rechtsnationalen Rappers Chris Ares, „mit kurzem Draht zur Identitären Bewegung“[24], teilt.

24 Stuttgarter Nachrichten (29.08.2018): Wie YouTube mit rechter Hetze umgeht, URL: https://www.stuttgarter-nachrichten.de/inhalt.ausschreitungen-in-chemnitz-wie-youtube-mit-rechter-hetze-umgeht.50e2546c-505b-45b6-95be-b80d2b5b3d9d.html

Abbildung 5: So ziemlich alles ist an diesem öffentlich geteilten Post dieses privaten Nutzers falsch, außer der Nachricht, dass Daniel H. erstochen wurde. Problematisch auch, dass hier ein Bild des Opfers, wie auch einer weiteren Frau, geteilt wird – ohne auf die Privatsphäre des Opfers Rücksicht zu nehmen. Der Post wurde allein knapp 20.000 mal geteilt. EDIT: Der Name des Nutzers und die Gesichter der Personen auf dem Foto wurden nachträglich von uns verpixelt.

Ein weiterer Sonderfall ist das Kommentieren von Privatnutzer:innen von öffentlichen Beiträgen – diese Kommentare werden innerhalb der Plattform öffentlich. Hier könnte für den Zugriff der Wissenschaft die Namen der Kommentierenden einfach gelöscht werden – wichtiger wären ohnehin nur die Inhalte der Kommentare zu öffentlichen Beiträgen, um zum Beispiel zu sehen, ob in den Kommentarspalten Hate Speech verbreitet wird, oder Desinformation – auch das kommt vor – durch Nutzer:innen debunked werden.

Abbildung 6: Dieses durch eine Privatperson öffentlich geteilte Video, in dem der rechte Rapper Chris Ares zu den Vorfällen in Chemnitz Desinformation verbreitet, hat extrem hohe Reichweiten. Es wird alleine 111.000 geteilt, dazu kommen 20.000 Likes und 4.000 Kommentare. Facebook gibt eine Reichweite 2,6 Mio. Views an. Zum Vergleich: die erfolgreichste „Fake News", die wir zur Bundestagswahl messen konnten, kommt – mit einer Vielzahl von Akteuren, die sie zusammen geteilt haben – auf ein Engagement von 250.000. EDIT: Name von uns verpixelt.

Videodaten

Daran anschließend kommen wir gleich zu einem der Hauptprobleme: Falschinformationen in Videos, wie dem des rechten Rappers Chris Ares, lassen sich überhaupt nicht auffinden, wenn die Videobeschreibung den Inhalt nicht hergibt. Heißt: Der Inhalt des Videos lässt sich nicht durchsuchen, obwohl es technisch möglich wäre. Wer als Fact-Checker beispielsweise herausfinden will, ob in einem Videobeitrag massiv Desinformation geteilt werden, muss ihn ganz klassisch manuell Sichten. In diesem Fall konnte nur dadurch, dass das Wort „Chemnitz" im Beitrag von Franz P. (*Abb. 6*) fällt, das Video überhaupt aufgefunden werden. Der Inhalt des Videos ist nicht durchsuchbar, dennoch verbreiten sich die vielen darin getroffenen Falschaussagen in hohen Zahlen (das Video hat allein bei Facebook 2,6 Millionen Abrufe):

> *„In dem Clip ruft Chris Ares, der während der Aufnahme in einem Auto sitzt, offen zu den rechten Protesten in Chemnitz auf und verbreitet Unwahrheiten. So sei der Anlass für die Messerattacke, die von Rechten für ihren Aufmarsch instrumentalisiert wurde, angeblich sexuelle Belästigung einer Frau gewesen. Außerdem spricht Ares von zwei Toten. Beides ist laut Polizei und sächsischem Innenministerium falsch."*[25]

Das Video haben allein in den ersten Tagen bei YouTube etwa 435.000 Menschen gesehen. Sowohl Facebook als auch YouTube haben den Account von Chris Ares gesperrt. Das Video wird aber weiterhin von anderen Accounts hochgeladen (Stand: 22.12.2018 existiert noch immer eine Version auf YouTube[26], hochgeladen von einem Dritt-Account, ohne Hinweis, dass die Aussagen in dem Video falsch sind).

Was aber die Durchsuchbarkeit von Aussagen in Videos betrifft, sind die technischen Voraussetzungen dafür eigentlich längst gegeben, wie *Abb. 2* beim RT-Beispiel und der Untertitelung Trumps zeigen: So existiert Software, die Audiospuren in Videos zu Text umwandeln kann. Diese wird sogar von YouTube bereits genutzt, wenn man Videos hochlädt und diese automatisch untertiteln möchte. Damit bestehen bereits die technischen Voraussetzungen für die Durchsuchbarkeit der Sprachinhalte der Videos und ihr mögliches Auffindbarmachen über textbasierte Suchen.

25 ebd.

26 Video von Chris Ares auf YouTube – hochgeladen von einem Drittacount. https://www.youtube.com/watch?v=jxNAipkgSAI&t=21s

Bilddaten, Gelöschte Daten

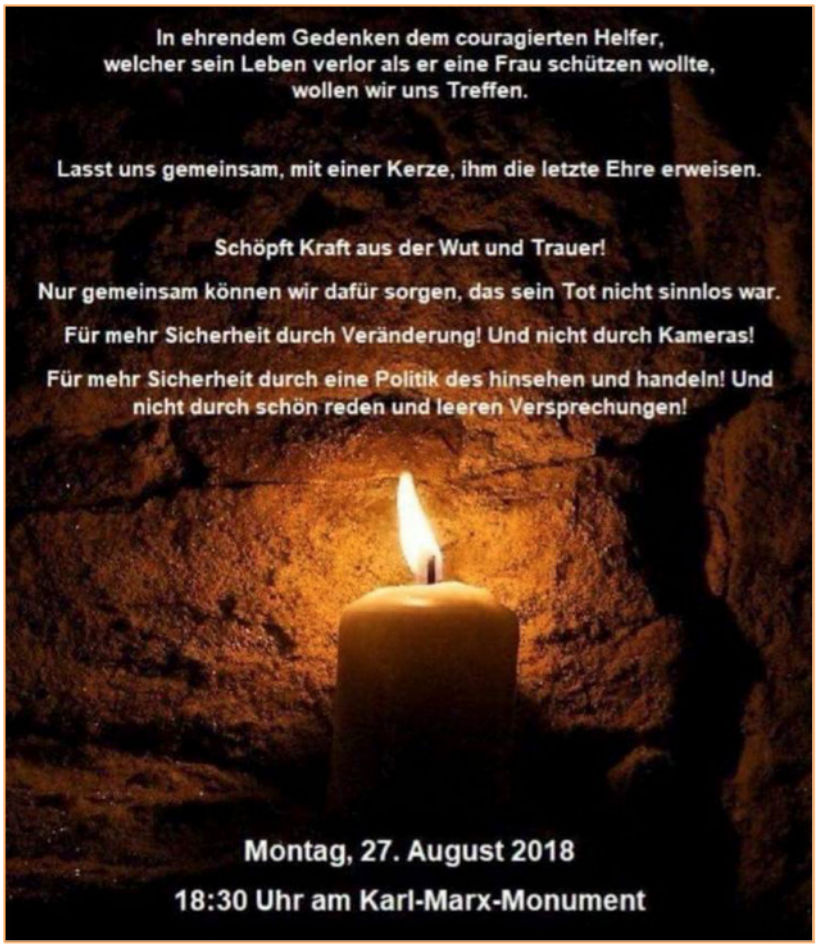

Abbildung 7 und 8: Der Text auf den beiden geteilten Bildern enthält ebenfalls Desinformation. Der Text auf den Bildern ist allerdings nicht durchsuchbar. Auch wenn es technisch möglich wäre.

Bilder sind auf den sozialen Netzwerken als Content besonders erfolgreich – das soziale Netzwerk Instagram ist sogar speziell darauf ausgerichtet, dass Bilder (und auch Videos) geteilt werden. Auch hier gibt es keine uns bekannte Möglichkeit Bilder mit oder ohne Text aufzufinden und zu analysieren.

Manuell bietet Google hierfür die beste Möglichkeit – allerdings muss dafür bereits bekannt sein, welches Bild genau gesucht wird. So ermöglicht es die Bildersuche bei Google, Bilder hochzuladen, zu denen dann Suchergebnisse angezeigt werden.[27] Auch für das Auffinden gelöschter Inhalte kann Google nützlich sein, da die Google-Suchergebnisse nicht jeden Tag vollständig aktualisiert werden. So ließ sich zum Beispiel aufspüren, dass auch der Ac-

27 Auch im Fact-Checking ist diese Funktion von Google sehr hilfreich. Mittels Bilder-Rückwärtssuche lassen sich so zum Beispiel auch gefälschte Bilder aufspüren.

count von *LEGIDA* bei Facebook das Bild geteilt hatte (später wurde dieser Beitrag auf der Seite von *LEGIDA* gelöscht). Auch die rechtsnationale Band *FLAK* hatte dieses Bild auf Facebook geteilt, aber später wieder gelöscht.

Abbildung 9: Die Google-Bildersuche findet das Bild auch bei Facebook, dabei hat LEGIDA das Bild zu dem Zeitpunkt bereits bei Facebook gelöscht. Es erreichte allein knapp 20.000 Likes.

Suchergebnisdaten

Die nächste Datenform, deren systematische wissenschaftliche Analyse bisher nur durch hohen Aufwand möglichst ist, sind Suchergebnisse auf den Plattformen. Sowohl Facebook als auch YouTube stehen hier immer wieder in der Kritik in Bezug darauf, wie sie ihre Suchergebnisse anzeigen, wie *Abb. 10* exemplarisch zeigt. Im Zuge der Auseinandersetzung um die Einlassungen des damaligen Präsidenten des Bundesamts für Verfassungsschutz, Hans-Georg Maaßen[28], haben wir nach Beiträgen auf Facebook darüber gesucht. In den Suchergebnissen zu Videos finden sich gleich vier Videos der Alternative für Deutschland, nur eines vom ZDF schafft es in die Top 5-Suchergebnisse.

28 Zeit Online (07.09.2018): Verfassungsschutzpräsident Maaßen äußert Zweifel an Hetzjagdvorwürfen, URL: https://www.zeit.de/politik/deutschland/2018-09/verfassungsschutz-hans-georg-maassen-chemnitz-hetzjagd

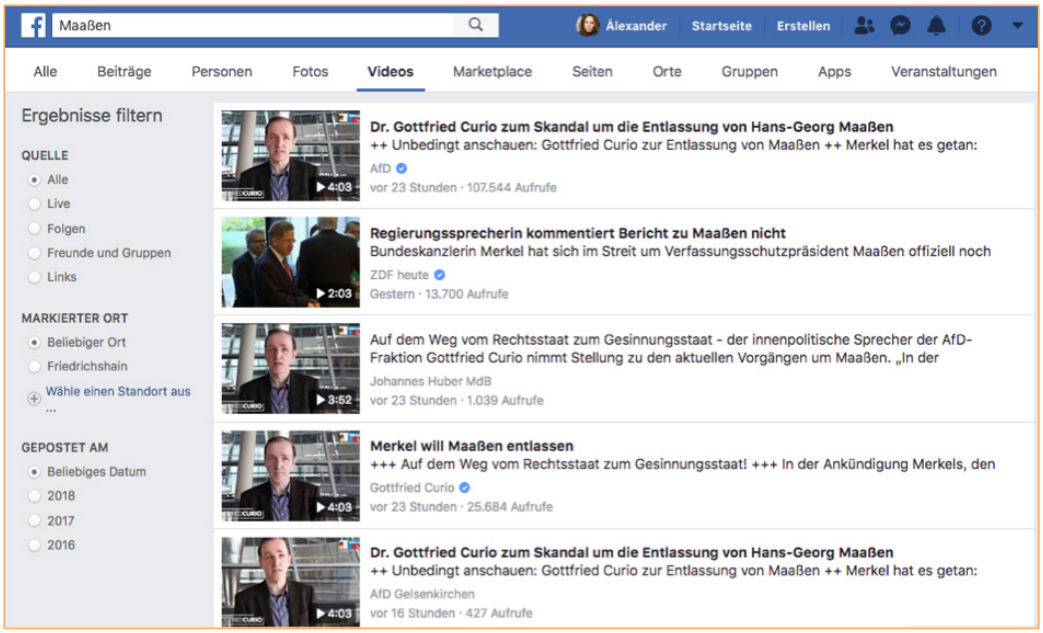

Abbildung 10: Eine relevante Frage ist auch, welche Inhalte bei welchen Suchbegriffen von den Plattformen ausgegeben werden. Sucht man beispielsweise nach „Maaßen“, um die Debatte um den ehemaligen BfV-Chef abzubilden, finden sich fast nur Ergebnisse von der AfD.

Auch der Videoplattform YouTube wird immer wieder vorgeworfen, als Radikalisierer zu wirken, da vor allem Empfehlungen des Algorithmus immer extremere Inhalte präsentieren.[29] Das liegt auch daran, dass besonders viele extreme Inhalte auf YouTube präsent sind, dagegen aber wenig klassische Qualitätsmedien. Auch in Chemnitz war das der Fall, wie die taz berichtete:

> *„Serratos Analysen sagen, dass die meisten YouTube-Videos zu Chemnitz, die in den Tagen nach den Ausschreitungen populär waren, stramm rechts waren. Und dass in der Empfehlungsspalte, die auf YouTube rechts neben dem Video auftaucht, auch dann schnell Verschwörungstheoretisches und Rechtsextremes vorgeschlagen wurde, wenn man eigentlich nur nach Nachrichten zu Chemnitz gesucht hat.“*[30]

Inzwischen haben sich die Suchergebnisse zum Fall Chemnitz verändert und Medien wie *Welt, Bild, Tagesschau* und *MDR* werden anzeigt. Die Plattform arbeite an ihrem Empfehlungsalgorithmus, so die taz weiter: „Seit dem 11. September dieses Jahres hat YouTube auch in Deutschland zwei Funktionen eingeführt, die die Empfehlungen bei Nachrichtenthemen beeinflussen. „Breaking News“ soll dabei helfen, Falschmeldungen und unseriöse Nach-

29 taz (13.10.2018): Verschwörungstheorien auf YouTube. Mit jedem Klick immer weiter nach rechts. URL: http://www.taz.de/!5540029/

30 ebd.

richtenquellen bei aktuellen Großereignissen weniger prominent anzuzeigen, „Top News“ sortieren nach der gleichen Logik Nachrichten auch dann noch, wenn die Aktualität etwas abgeebbt ist.“[31]

Werbeanzeigen

Ob zusätzlich zur Mobilisierung von einigen Facebook-Seiten auch noch Werbeanzeigen geschaltet wurden, ist derzeit leider überhaupt nicht nachvollziehbar, denn Facebooks Werbebibliothek enthält derzeit keine Daten für Deutschland. Doch auch wenn diese freigeschaltet würde, wäre es derzeit unklar, ob Posts, die zu einer Mobilisierung aufrufen würden, unter die von Facebook definierte Kategorie von *Politischer Werbung* fielen. Dann müssten sich die Werbetreibende vorher freischalten lassen, um entsprechende Werbungen zu veröffentlichen.

4. Wie diese Daten zur Verfügung gestellt werden könnten

Datenschnittstellen

Wir fassen zusammen: Um Desinformationskampagnen systematisch zu untersuchen, braucht es eine Vielzahl von Daten. Nur die wenigstens davon stellen die Plattformen bisher über Datenschnittstellen oder Transparenztools für Wissenschaft, Journalismus, NGOs oder die Öffentlichkeit vollumfänglich zur Verfügung. *Abb. 11* gibt einen abschließenden Überblick, wie es um die Datenschnittstellen derzeit bestellt ist. Was unter den jeweiligen Datentypen zu verstehen ist, wurde in *Kap. 2 & Abb. 3* erläutert. Deutlich wird: Zu vielen Daten existieren nur eingeschränkt funktionierende Datenschnittstellen oder gar keine. Den besten Zugang zu Daten bietet im Vergleich noch Twitter, wodurch es allerdings auch in der Wissenschaft einen massiven Überhang von Studien zu Twitterdaten gibt.

31 ebd.

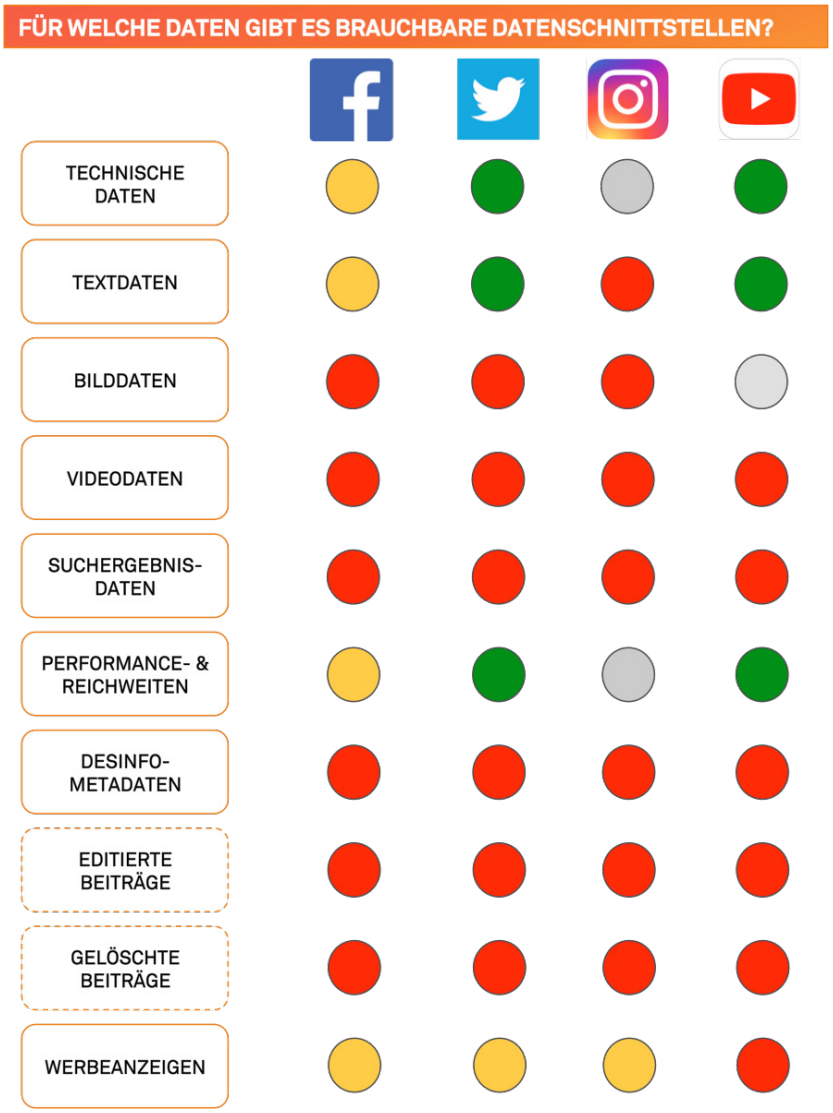

Abbildung 11: Für welche Datentypen gibt es brauchbare Datenschnittstellen? Rot = keine Schnittstelle vorhanden, gelb = Schnittstelle nur eingeschränkt oder nicht vollständig (Beispiel: Werbeanzeigenregister – bisher nur für die USA bei Google bzw. für USA/GB/BRA/IND bei Facebook). Grün = Schnittstelle vorhanden. Grau = trifft nicht zu (bei YouTube gibt es keine Bilder) oder: nicht einschätzbar. Die Beurteilung beruht auf der Einschätzung von Anbietern von Social Media Intelligence, Datenjournalisten und dem SNV-Team. Alle abgetragenen Daten beziehen sich nur auf öffentliche Daten. Inwiefern es sinnvolle Schnittstellen für *gelöschte* oder *editierte* Beiträge geben kann, muss im Einzelfall für die jeweiligen Netzwerke diskutiert werden.

Werden Daten von den Plattformen doch bereitgestellt, sollten vergangene Daten nicht erst durch langwierige und komplizierte Antragsverfahren an nur eine Handvoll Wissenschaftler:innen übergeben werden, wie es Facebook derzeit mit *Social Science One*[32] betreibt. Auch, weil die sich bewerbenden Institutionen sehr dezidiert Auskunft geben müssen, welche Daten sie genau

32 Webseite von https://socialscience.one/

zu welchen Forschungsfragen auswerten wollen. Ein so umständliches Verfahren widerspricht der Dringlichkeit des Erkenntnisgewinns und der Logik induktiver Verfahren und ist ebenso hinderlich für alle Institutionen, die ein sinnvolles Echtzeit-Monitoring durchführen müssen, wie z.B. Fact-Checking-Organisationen oder zivilgesellschaftliche Organisationen, die Desinformation oder Hate Speech beobachten.[33]

Möglich gemacht werden sollte demnach beides: ein Echtzeit-Monitoring von sozialen Netzwerken – wie es Fact-Checking-Organisationen oder NGOs benötigen, aber auch der Zugriff auf vergangene Daten durch die Wissenschaft, die im Nachhinein dann versucht, zu rekonstruieren, wie beispielsweise die Debatte zum UN-Migrationspakt in den sozialen Netzwerken entstanden ist. Dabei können zuweilen auch gelöschte Daten wichtig sein, wie es im konkreten Fall von Chemnitz ebenfalls deutlich wurde.

Transparenztools

Im Idealfall geht die Transparenz weit über die Datenschnittstellen hinaus, sodass es auch für die Öffentlichkeit möglich ist ohne Social-Media-Intelligence-Anbieter oder ausgewiesene Programmierkenntnisse, die öffentlichen Debatten in sozialen Netzwerken systematisch nachzuverfolgen. Dafür bräuchte es ein Transparenztool, dass beispielsweise (politische) Themen durchsuchbar macht und aufbereitet – und diese Daten aber auch im Hintergrund als Schnittstelle für eigene Auswertungen bereit hält.

Was utopisch klingt, ist mit den **Wahlwerberegistern**[34] von Facebook und Google bereits einigermaßen funktionierende Realität und kann als Blaupause für andere Transparenztool herangezogen werden, auch gibt es eine Datenschnittstelle.[35] Allerdings gibt es auch hier Einschränkungen in der

33 Facebook stellt journalistischen Institutionen und Fact-Checking-Organisationen auf Antrag das Tool Crowdtangle zur Verfügung. Laut Aussagen uns bekannten Organisationen, die mit dem Tool arbeiten, bestehen hier aber die gleichen Probleme durch die Einschränkung der Datenschnittstellen.

34 Wahlwerberegister. Facebook nennt dieses Tool „Werbebibliothek“, diese ist (Stand: 22.01.2019) bisher für die USA, Großbritannien und Brasilien verfügbar: https://www.facebook.com/ads/archive/?active_status=all&ad_type=political_and_issue_ads&country=ALL, Google nennt seine Übersicht „Politische Werbung bei Google“, diese enthält bisher (Stand 22.01.2019) nur die USA: https://transparencyreport.google.com/political-ads/overview

35 Für das Wahlwerberegister bei Facebook („Werbebibliothek“) existiert eine solche Schnittstelle: https://www.facebook.com/ads/archive/access

Funktion und einen langen "Leidensweg" bis zur Verwirklichung, der sich so nicht wiederholen sollte:

Die Wahlwerberegister sind die transparenzpolitische Antwort von Google und Facebook auf einen nun schon über fast drei Jahre andauernden Prozess. Dahinter stehen zwei Fragen: 1) Welchen Einfluss hatten auf den Plattformen geschaltete Werbeanzeigen (früher: als sogenannte *Dark Ads*) auf die US-Wahlen oder den Brexit[36] und 2) inwiefern haben welche (ausländische) Akteure sich in Wahlen und Referenden eingemischt, indem u.a. auch gezielte Desinformationskampagnen geschaltet wurden (z.B. durch die Alt-Right-Bewegung beim Referendum zu Schwangerschaftsabbrüchen in Irland[37]). Fragen, die sich in Bezug auf Desinformationskampagnen auf den Social-Media-Plattformen ebenfalls stellen – auch ganz ohne, dass sie als "Werbeanzeigen" ausgespielt werden.

Abbildung 12: Politische Werbung bei Google – Transparenzbericht mit Benutzeroberfläche. Bisher funktionstüchtig für die USA. URL: https://transparencyreport.google.com/political-ads/overview?top_advertisers=q:obama&lu=top_advertisers

36 Correctiv (06.08.2018): „Dark Ads" bei Facebook: Wie mit Falschnachrichten Stimmung für den Brexit gemacht wurde. https://correctiv.org/faktencheck/hintergrund/2018/08/06/dark-ads-bei-facebook-wie-mit-falschnachrichten-stimmung-fuer-den-brexit-gemacht-wurde/

37 Netzpolitik (25.05.2018): Irland: Mit Dark Ads gegen Abtreibung. https://netzpolitik.org/2018/irland-mit-dark-ads-gegen-abtreibung/

Abbildung 13: Werbebibliothek für Facebook und Instagram. Bisher funktionstüchtig für USA, Großbritannien und Brasilien. URL: https://www.facebook.com/ads/archive/?active_status=all&ad_type=political_and_issue_ads&country=ALL

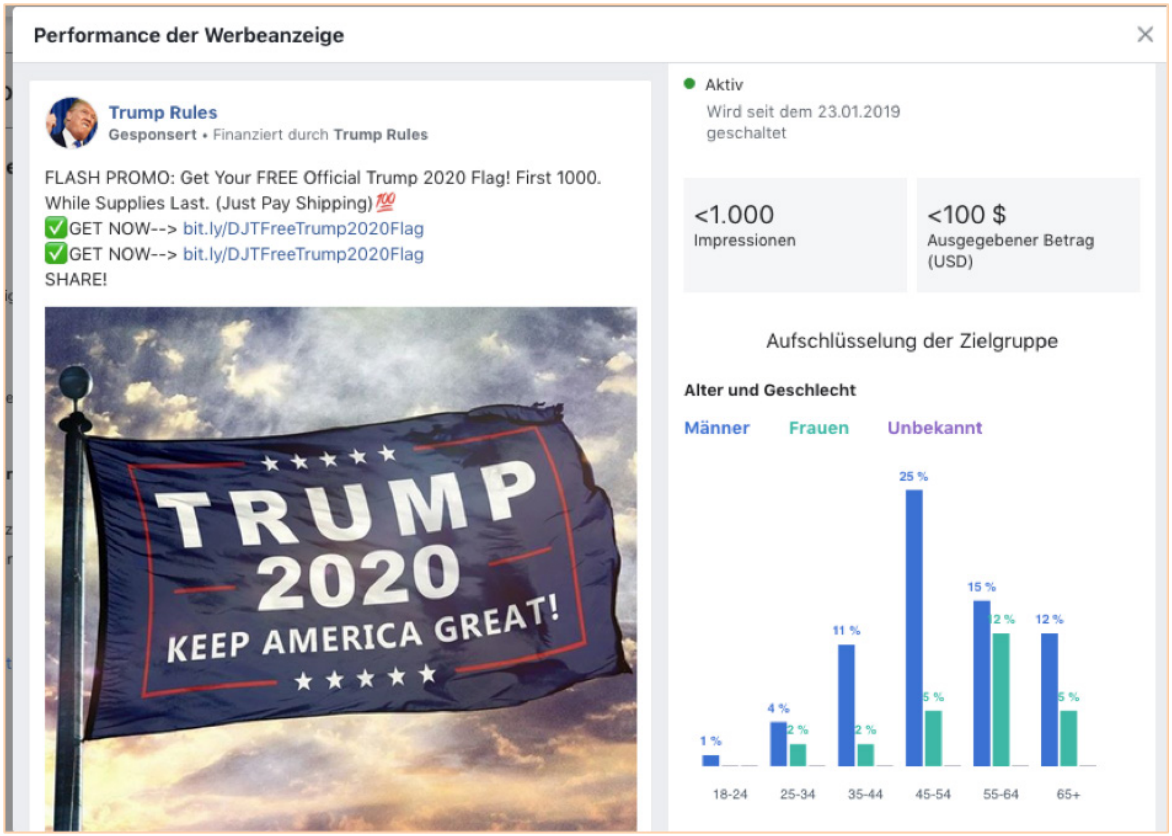

Abbildung 14: Beispiel für Transparenz in der Werbebibliothek bei Facebook. Angezeigt werden auch Informationen über den Financier des Posts, wie viel Geld investiert wurde, in welchem Zeitraum der Post online ist/war und die Zielgruppen.

Die Wahlwerberegister ermöglichen nun – unabhängig von der Datenschnittstelle – auch der interessierten Öffentlichkeit die Möglichkeit der „Selbstbeobachtung". Diese Möglichkeit ist sicherlich der richtige Weg hin zu transparenten Kommunikationsprozessen in den digitalen Öffentlichkeiten, bergen aber dennoch einige Schwachstellen, die nicht unerwähnt bleiben sollen, weil sie auch die Komplexität zwischen "gut gemacht" und "gut gemeint" aufzeigen. So existieren diese Transparenztools bei Google bisher nur für die USA, beziehungsweise bei Facebook nur für die USA, Brasilien, Indien und Großbritannien. Allerdings steht die Europawahl vor der Tür: Für die 27-EU-Staaten warten alle Beteiligten noch immer auf die Freischaltung der Wahlwerberegister, die von beiden Konzernen angekündigt wurde.

Eine weitere Schwachstelle ist die derzeitige Regelung dazu, welche Anzeigen in die Datenbanken aufgenommen werden. So hat bespielsweise *ProPublica* mit einem eigenen Tool Werbeanzeigen der *National Rifle Association (NRA)* finden können, die nicht in Facebooks Werbebibliothek für Amerika erscheinen[38] – obwohl sie es als "politische Werbung" oder "Themen nationalen Interesses" tun müssten. Problem dahinter: Damit Werbeanzeigen in Facebooks Werbebibliothek angezeigt werden, müssen sich Werbetreibende selbst kategorisieren, ob sie "politische Werbung" oder "Themen von nationalem Interesse" bewerben. Tun sie dies nicht, tauchen ihre Anzeigen auch nicht auf. Anstatt das Facebook mit ProPublica zusammenzuarbeitet, um ein effizienteres Schwachstellenmanagement zu betreiben, wurde das genutzte Tool durch eine Änderung im Code von Facebook unbrauchbar gemacht – auch das zeigt Facebooks Umgang mit Akteur:innen die versuchen, an Daten zu kommen und diese auszuwerten. Es braucht aber auch eine kritische Begleitung durch Wissenschaftler:innen und NGOs von außen, die wiederum die bereitgestellten Transparenztools der Plattformen ihrerseits überprüfen.

Politik oder Plattformen: Einer muss sich bewegen

Ob der Zugang zu öffentlichkeitsrelevanten und öffentlichen Daten freiwillig durch die sozialen Netzwerke erfolgt oder Regulierer hier weisend eingreifen müssen, soll hier nicht Thema sein. Mit Hinblick auf die kommenden Wahlen – von der Europawahl im Mai dieses Jahres bis zu den drei Landtagswahlen in Deutschland im Herbst – besteht allerdings dringender Handlungsbedarf. Die Befürchtungen der Wahlmanipulation durch gezielte Propaganda und Falschinformationen kommen schließlich aus der Erfahrung der US-Wahlen und dem Brexit-Volksentscheid. Gerade die undurchsichtige Gemengelage

38 Foreign Policy (01.06.2018): How Ireland Beat Dark Ads
https://foreignpolicy.com/2018/06/01/abortion-referendum-how-ireland-resisted-bad-behaviour-online/

von 27 EU-Staaten mit eigenen, spezifischen Mediensystemen, geopolitischen Einflüssen in- und außerhalb der EU, verschiedener Mediennutzung und Freiheitsgrade innerhalb der Mediensysteme machen es dringend notwendig, die Vorgänge in den Plattformen transparent und öffentlich zu machen.

Natürlich tragen die sozialen Netzwerke nicht „die Schuld“ an der Wahl Trumps oder am Brexit – diese Gleichung ist unzulässig, weil sie unterkomplex und verkürzt ist, aber doch immer wieder Teil der öffentlichen Debatte. Die Disruptionen und Transformationen des Mediensystems, werden aber ihren Anteil zu einer komplexen Gleichung von politischen, gesellschaftlichen und ökonomischen Faktoren tragen, die wir nach wie vor noch zu wenig verstehen.

Es scheint unabdingbar, dass wir als Gesellschaft ein tiefes und dezidiertes Verständnis von der Wirkungsweise sozialer Netzwerke erlangen. Dafür brauchen aber Wissenschaft, Journalismus, NGOs und Zivilgesellschaft, als Beobachter dieser neu entstehenden digitalen Teilöffentlichkeiten, Zugang zu Daten. Und diese systematische Beobachtung ist nur möglich, wenn funktionierende Datenschnittstellen bereitgestellt und Transparenztools für die Öffentlichkeit geschaffen werden.

Das Anliegen des Autors dahinter ist es, mehr Forschung, mehr kritische Beobachtung und mehr Verständnis für die digitalen Öffentlichkeiten der Netzwerke zu ermöglichen, die weder vom Wohlwollen der sozialen Netzwerke abhängig sind, noch hohe technische oder kommerzielle Hürden setzen und die sensible und private Daten schützen. Auch gute Medienpolitik ist auf diese Daten dringend angewiesen. Im Datenzeitalter leben die Digitalkonzerne davon, unsere Daten massenweise zu sammeln und für eigene Geschäftszwecke zu verwenden. Es wird Zeit, diese Daten mit der Gesellschaft zu teilen.

Über die Stiftung Neue Verantwortung

Think Tank für die Gesellschaft im technologischen Wandel

Neue Technologien verändern Gesellschaft. Dafür brauchen wir rechtzeitig politische Antworten. Die Stiftung Neue Verantwortung ist eine unabhängige Denkfabrik, in der konkrete Ideen für die aktuellen Herausforderungen des technologischen Wandels entstehen. Um Politik mit Vorschlägen zu unterstützen, führen unsere Expertinnen und Experten Wissen aus Wirtschaft, Wissenschaft, Verwaltung und Zivilgesellschaft zusammen und prüfen Ideen radikal.

Über den Autor

Alexander Sängerlaub leitet das Projekt „Desinformation in der digitalen Öffentlichkeit“. Im Mittelpunkt steht die Frage, wie sich der digitale Strukturwandel der Öffentlichkeit auf die Güte der Kommunikation auswirkt. Welche Akteure profitieren von der Demokratisierung des Diskurses und welchen neuen Herausforderungen durch Soziale Netzwerke müssen sich liberale Demokratien stellen?

So erreichen Sie den Autor

Alexander Sängerlaub
Projektleiter Desinformation in der digitalen Öffentlichkeit
asaengerlaub@stiftung-nv.de
+49 (0)30 81 45 03 78 86

Impressum

Stiftung Neue Verantwortung e. V.

Beisheim Center
Berliner Freiheit 2
10785 Berlin

T: +49 (0) 30 81 45 03 78 80
F: +49 (0) 30 81 45 03 78 97

www.stiftung-nv.de
info@stiftung-nv.de

Design:
Make Studio
www.make-studio.net

Layout:
Johanna Famulok

Free Download:

www.stiftung-nv.de