LANDESANSTALT FÜR MEDIEN NRW
Der Meinungsfreiheit verpflichtet.

VERIFICATION HANDBOOK – DAS HANDBUCH ZUR ÜBERPRÜFUNG VON DESINFORMATION UND MEDIEN-MANIPULATION

Deutsche Übersetzung im Auftrag der Landesanstalt für Medien NRW

Herausgegeben von Craig Silverman und
übersetzt von Marcus Engert

DataJournalism.com

European Journalism Centre

Craig Newmark Philanthropies

VORWORT

Welche Quelle ist vertrauenswürdig? Sollte ich diesen Beitrag teilen? Wie kann ich überhaupt erkennen, ob diese Information wahr ist? Es fällt Nutzerinnen und Nutzern heute zunehmend schwerer, Antworten auf diese Fragen zu finden – das zeigten nicht zuletzt die COVID-19-Pandemie oder die Geschehnisse rund um die Präsidentschaftswahlen in den Vereinigten Staaten von Amerika.

Regulierung kann dabei Unterstützung bieten – und mit dem Medienstaatsvertrag nun Verstöße gegen die journalistische Sorgfaltspflicht ahnden. Sie ist ein Mittel gegen Pseudojournalismus im Netz, schlechte Recherchen und den Versuch unter dem Deckmantel der vertrauenswürdigen Nachricht, Halbwahrheiten oder Meinungen zu verbreiten.

Doch Regulierung kann die Eigenverantwortung des Nutzers nicht ersetzen: Eingriffe durch Medienanstalten, die vor allem die Meinungsfreiheit schützen wollen und sollen, sind stets ultima ratio. Nutzerinnen und Nutzer müssen selber Kompetenzen entwickeln und lernen, Informationen richtig einzuschätzen, zu bewerten und mit ihnen umzugehen.

Ich freue mich, dass Craig Silverman als Herausgeber der englischsprachigen Originalausgabe des Verification Handbook auf sehr anschauliche Art und Weise Techniken zusammengestellt hat, mit deren Hilfe in der digitalen Welt Desinformation und Medienmanipulation erkannt und untersucht werden können. Die Landesanstalt für Medien NRW dankt Marcus Engert und dem European Journalism Center für die Zusammenarbeit, durch die wir die deutschsprachige Ausgabe anbieten können.

Das „Verification Handbook – Handbuch zur Überprüfung von Desinformation und Medienmanipulation“ wird Journalistinnen und Journalisten sowie Regulierung aber auch interessierten Nutzerinnen und Nutzern eine gute Unterstützung bei der Wahrnehmung ihrer jeweiligen Verantwortung für die Meinungsbildung bieten.

Dr. Tobias Schmid
Direktor der Landesanstalt für Medien NRW

INHALTSVERZEICHNIS

DESINFORMATION UND MEDIENMANIPULATION UNTERSUCHEN

von: Craig Silverman
deutsche Bearbeitung: Marcus Engert

Craig Silverman *ist als Medienredakteur von BuzzFeed News in New York für einen weltweiten Themenbereich zuständig, von Plattformen über Online-Falschinformationen bis hin zu Medienmanipulation. Er ist Herausgeber des „Verification Handbook" und des „Verification Handbook for Investigative Reporting" und Autor des Buches „Lies, Damn Lies, and Viral Content: How News Websites Spread (and Debunk) Online Rumors, Unverified Claims and Misinformation".*[1]

Marcus Engert *ist Investigativ-Journalist und Senior-Reporter im Deutschland-Büro von BuzzFeed News in Berlin.*

Im Dezember 2019 teilte Twitter-Nutzer @NickCiarelli ein Video, das, so seine Beschreibung, angeblich zeigen sollte, wie Unterstützer von Präsidentschaftskandidat Michael Bloomberg eine Tanzsequenz einstudieren. Weil das Ganze nicht sehr fröhlich, nicht sehr enthusiastisch und nicht sehr synchron aussah, bekam es schnell viele Interaktionen und wurde weiter verbreitet, vor allem von Menschen, die es unterhaltsam fanden, sich darüber lustig zu machen. Am Ende haben sich mehr als fünf Millionen Menschen das Video auf Twitter angesehen.

Der Tweet von Nick Ciarelli und ein Ausschnitt aus dem angeblichen Video. Dazu schreibt er an das Team eines Konkurrenten gerichtet: „Achtung, wir haben unseren eigenen Tanz! Aufgenommen bei einer Bloomberg-Wahlkampfveranstaltung in Beverly Hills."

In Ciarellis Twitter-Profil stand, er sei Praktikant im Kampagnenteam von Bloomberg und seine nachfolgenden Tweets sollten dafür eine Reihe von Belegen liefern, darunter auch ein Screenshot einer E-Mail, die angeblich von einem Kampagnenmitarbeiter stammte und in der dieser ein Budget für das Video freigab.

Eine schnelle Google-Suche zeigte: Ciarelli ist ein Komiker, der schon vorher Video-Parodien produziert hat. Die E-Mail des angeblichen Kampagnenmitarbeiters? Gesendet von Brad Evans, ebenfalls Komiker, mit dem Ciarelli schon oft zusammengearbeitet hatte. Beide Informationen waren nur eine einzige kurze Google-Suche entfernt. Trotzdem glaubten in den ersten Minuten und Stunden zahlreiche Menschen, das eigenartige Video sei eine offizielle Bloomberg-Produktion

1 Auf Deutsch würden die Titel ungefähr lauten: „Handbuch für Verifikation", „Handbuch für Verifikation im investigativen Journalismus" und „Lügen, verdammte Lügen und viraler Content: Wie Nachrichtenseiten Online-Gerüchte, unbestätigte Behauptungen und Falschinformationen verbreiten (und entlarven)"

gewesen. Maggie Haberman, bekannte Politikreporterin der New York Times, twitterte, Journalisten, die über frühere Bloomberg-Wahlkämpfe berichtet hätten, hätten nicht unbedingt einen Grund gehabt, das Video direkt als unglaubwürdig zu verwerfen:

Maggie Haberman twitterte: „Die Leute, die die Bloomberg-Parodie produzierten, verstehen nicht, warum Reporter, die schon über frühere Bloomberg-Wahlkämpfe berichtet hatten, es nicht unmittelbar als Parodie erkannt haben."

Wissen kann viele Formen annehmen und auch in dieser neuen, digitalen Welt müssen Journalistinnen und Journalisten sich davor hüten, sich zu sehr auf eine einzige Informationsquelle zu verlassen – auch dann, wenn es ihre eigenen Erkenntnisse aus erster Hand sind. Denn ganz offenbar waren einige der Reporter, die Bloomberg und seinen Wahlkampfstil kannten, der Ansicht, das Video könnte echt sein. Auf der anderen Seite hätten auch Journalisten, die überhaupt nichts über Bloomberg wussten und das Video also nur nach seiner Herkunft beurteilten, die richtige Antwort sofort bei Google finden können – in diesem Fall durch eine simple Suche nach dem Namen des Mannes, der das Video geteilt hat.

Es geht nicht darum, dass es wenig Spaß macht, über Bloomberg zu berichten. Der Punkt ist: Wir können uns zu jedem Zeitpunkt von dem, was wir zu wissen glauben, in die Irre führen lassen. In manchen Fällen kann unser eigener Wissens- und Erfahrungsschatz sogar ein Nachteil sein. Und wir können durch Digitales getäuscht werden, wie zum Beispiel durch in die Höhe schnellende Retweets und Aufrufzahlen (oder die Versuche, diese zu manipulieren).

Wie das Bloomberg-Video zeigt, macht es nicht viel Arbeit, Verwirrung zu stiften, zum Beispiel durch eine Biographie im Twitter-Profil oder einen Screenshot einer angeblichen E-Mail, die irgendetwas glaubwürdiger erscheinen lässt. So etwas trägt dazu bei, dass sich Inhalte viral verbreiten. Und je mehr Retweets und Interaktionen das Ganze dann bekommt, desto mehr Leute lassen sich wiederum durch solche irreführenden Signale davon überzeugen, dass das Video echt sein könnte.

Natürlich gibt es weitaus dramatischere Beispiele als dieses hier. Und anders als Ciarelli enthüllen die Menschen hinter bestimmten Desinformationskampagnen ihre Hintergründe selten. Doch zeigt dieses Beispiel, wie verwirrend und frustrierend es für alle, Journalistinnen und Journalisten eingeschlossen, ist, in einer Informationsumgebung, die leicht mit irreführenden Signalen geflutet werden kann, die zuverlässigen Signale zu finden. Vertrauen ist das Fundament von Gesellschaft. Es ist das Schmiermittel, der Schlüssel für alle Beziehungen. Wenn Sie glauben, je mehr Profile ein Video verbreiten, desto wichtiger ist es, dann wird man Sie überlisten. Wenn Sie glauben, alle Kundenbewertungen zu einem Produkt kommen von echten Kunden, werden Sie Ihr Geld zum Fenster rauswerfen. Wenn Sie meinen, jeder Nachrichtenartikel in Ihrem News-Feed stünde für eine ausgewogene, unvoreingenommene Auswahl dessen, was Sie unbedingt wissen sollten, werden Sie am Ende falsch informiert sein.

Diese Wahrheit zu erkennen ist wichtig für jeden, aber es ist essentiell für Journalistinnen und Journalisten. Wir sind das Ziel von koordinierten und gut ausgestatteten Kampagnen, die unsere Aufmerksamkeit wollen, die uns dazu verleiten wollen, ihre Botschaften zu verstärken und unseren Willen dem Willen von Staaten und anderen mächtigen Kräften zu beugen. Die gute Nachricht aber ist: Das schafft eine Möglichkeit – vielleicht sogar einen Auftrag – zur Aufklärung.

Dieses Handbuch sammelt das Wissen und die Erfahrungen von Top-Journalistinnen und -Journalisten und Wissenschaftlerinnen und Wissenschaftlern, die Anleitungen für die Analyse und Untersuchung von Manipulation, Desinformation und Informationsoperationen geben. Wir bewegen uns in einem komplexen und sich rapide entwickelnden Informations-Ökosystem. Das erfordert, dass wir unseren Zugang dazu genauso schnell anpassen, indem wir unsere Vorannahmen überprüfen, unsere Gegenspieler beobachten, ihre nächsten Schritte vorausahnen und das Beste aus digitalen und traditionellen Recherchemethoden anwenden. Die Schwachstellen in unserer digitalen und datengetriebenen Welt verlangen von uns Journalistinnen und Journalisten, alles genau zu hinterfragen und unsere Fähigkeiten dafür einzusetzen, die Öffentlichkeit zu korrekten, vertrauenswürdigen Informationen zu führen. Sie verlangen von uns auch, darüber nachzudenken, wann wir jenen böswilligen Akteuren und Kampagnen unwissentlich Futter geben, die uns schaden wollen, und wann wir womöglich überstürzt mit dem Finger auf staatliche Akteure zeigen, auch wenn die Beweise das nicht hergeben.

Das Ziel dieses Handbuchs ist es, die Leserinnen und Leser mit jenen Fähigkeiten und Techniken auszurüsten, um das effektiv und verantwortungsvoll tun zu können. Es bietet auch eine Einführung in jene Theorien, Zusammenhänge und Denkweisen, die es ermöglichen, qualitativ wertvolle Arbeit abzuliefern – Arbeit, die die Öffentlichkeit informiert, böswillige Akteure entlarvt und hilft, das Ökosystem zu verbessern, in dem wir unsere Informationen finden. Am wichtigsten aber ist es, zu verstehen, dass alle diese Werkzeuge und Anleitungen nichts bringen, wenn man nicht mit der richtigen Denkweise an die Arbeit geht. Das heißt, zu erkennen, dass absolut alles in einer digitalen Welt gespielt oder manipuliert sein kann und dass die Gruppe derer, die ein Motiv dafür haben, dies zu tun, groß ist.

Das Schöne an dieser digitalen Welt ist, dass es oft, wenn auch nicht immer, eine Spur von Daten gibt, von Interaktionen, Verbindungen und anderen „digitalen Brotkrumen", denen man folgen kann. Und vieles davon ist öffentlich zu finden, wenn man denn weiß, wo und wie man suchen muss.

Im Digitalen recherchieren heißt, nichts so hinzunehmen, wie es auf den ersten Blick scheint. Es heißt, zu erkennen, dass Dinge, die so aussehen, als müsste man sie einfach nur zählen, als seien sie datengetrieben – Likes, Shares, Retweets, Zugriffe, Ansichten, Werbeaufrufe –, einfach manipuliert werden können und es auch oft werden. Es heißt, zu erkennen, dass Journalistinnen und Journalisten ein Hauptziel von Medienmanipulations- und Informationsoperationen sind, sei es, dass sie selbst angegriffen und attackiert werden, oder dass versucht wird, sie zu einem Verbreitungskanal für Falsch- und Desinformation zu machen. Und es heißt, sich selbst und seine Kolleginnen und Kollegen mit den richtigen Einstellungen, Techniken und Werkzeugen auszustatten, um sicherzustellen, dass man verlässliche, akkurate Informationen verteilt – und nicht zum Verstärker von Falschbehauptungen, manipulierten Inhalten oder Troll-Kampagnen wird.

Im Zentrum dieser Einstellung steht das Paradox digitaler Recherchen: Wir dürfen erst einmal nichts glauben und niemandem vertrauen, weil nur das uns eine Arbeitsweise ermöglicht, mit der wir freilegen können, was und wem wir vertrauen können. Nur so können wir Ergebnisse recherchieren, denen die Gesellschaft, für die wir arbeiten, vertrauen kann und will.

Damit einher gehen einige Grundlagen, die in den Kapiteln und Fallbeispielen dieses Handbuchs immer wieder auftauchen werden:

- **Denken Sie wie Ihr Gegenüber.**
  Jedes neue Feature auf einer Plattform, jeder digitale Service kann irgendwie ausgenutzt werden. Es ist entscheidend, dass Sie sich in die Rolle desjenigen begeben, der nach Möglichkeiten zur Manipulation sucht: aus ideologischen, politischen, finanziellen oder anderen Gründen. Wer sich digitale Inhalte und Botschaften anschaut, der sollte sich auch fragen, welche Motive hinter ihrer Erstellung und Verbreitung stehen könnten. Es ist ebenso wichtig, die neuesten Techniken zu kennen, die böswillige Akteure, digitale Vermarkter und andere nutzen, deren Lebensunterhalt davon abhängt, immer neue Wege zu finden, um Aufmerksamkeit zu erhalten und Einnahmen im digitalen Raum zu machen.

- **Konzentrieren Sie sich auf Akteure, Inhalte, Verhaltensweisen und Netzwerke.**
  Das Ziel ist, Akteure, Inhalte und Verhaltensweisen zu analysieren und zu zeigen, wie sie als Netzwerk zusammenwirken. Indem man diese vier Dinge gegenüberstellt, beginnt man zu verstehen, was man da sieht. Wie sich in zahlreichen Kapiteln und Beispielen in diesem Handbuch zeigen wird, besteht ein grundlegender Ansatz dafür darin, mit einem einzelnen Puzzleteil des Ganzen, zum Beispiel einer Website, zu beginnen und von diesem ausgehend größere Zusammenhänge und weitere Verbindungen im Netzwerk zu suchen. Eine Möglichkeit dafür kann sein, den Kurs nachzuvollziehen, auf dem sich Inhalte und Akteure über verschiedene Plattformen und mitunter auch über verschiedene Sprachen verbreiten.

- **Beobachten und Sammeln.**
  Die beste Möglichkeit, Medienmanipulation und Desinformation zu identifizieren, ist, permanent danach Ausschau zu halten. Permanent die Augen offenzuhalten und zu beobachten, was bekannte Akteure, Themen und Gruppen im Netz tun, ist essentiell. Sammeln und organisieren Sie, was Sie finden, ob in einer Tabelle, in Ordnern mit Screenshots oder mit kostenpflichtigen Diensten wie Hunchly.

- **Vorsicht mit Zuschreibungen.**
  Manchmal ist es schlicht unmöglich, zu sagen, wer genau hinter einem bestimmten Konto, einem speziellen Inhalt oder einer größeren Informationsoperation steckt. Ein Grund dafür: Akteure mit unterschiedlichen Zielen können sich ähnlich verhalten, ähnliche Inhalte produzieren und diese ähnlich verteilen. Sogar die Plattformen selbst – die viel besseren Zugang zu den Daten und viel mehr Ressourcen haben – machen oft Fehler bei den Zuschreibungen. Die erfolgreichsten und überzeugendsten Beweise sind in der Regel jene, die digitale Belege mit menschlichen Quellen kombinieren – also ein Mix aus der Online-Recherche und der traditionellen investigativen Arbeit. Leider wird das zunehmend schwer, da sowohl staatliche Akteure als auch andere sich weiterentwickeln und neue Wege finden, ihre Fingerabdrücke zu verwischen. Zuschreibungen sind schwer; hier Fehler zu machen, kann die ganze mühsame Arbeit zunichtemachen, die vorher nötig war, um überhaupt bis dorthin zu kommen.

Ein Wort noch zu den beiden Handbüchern, die diesem Handbuch vorausgegangen sind.[2] Dieses Handbuch baut auf den Grundlagen der ersten Ausgabe des „Verification Handbook" und des „Verification Handbook for Investigative Reporting" auf. Beide bieten grundlegende Techniken und Lösungen zur Beobachtung sozialer Medien, zur Überprüfung von Bildern, Videos und Social-Media-Accounts sowie für die Nutzung von Suchmaschinen zur Identifizierung von Personen, Unternehmen und anderen Einrichtungen. Wenn Sie Schwierigkeiten haben, den Beispielen in diesem Buch zu folgen, ermutigen wir Sie, auch dort nachzulesen.

Und damit: An die Arbeit!

2 Das „Verification Handbook" und das „Verification Handbook for Investigative Reporting" stehen online zur Verfügung, allerdings bislang nicht in Deutsch. Sobald eine Neuauflage der Bücher in Arbeit ist, werden wir uns bemühen, diese dann auch auf Deutsch zu veröffentlichen.

DAS ZEITALTER DER INFORMATIONSSTÖRUNG

von: Claire Wardle
deutsche Bearbeitung: Marcus Engert

Claire Wardle *ist für die strategische Ausrichtung und Forschung von First Draft verantwortlich. First Draft ist eine internationale gemeinnützige Nichtregierungsorganisation, die Journalistinnen und Journalisten, Wissenschaftlerinnen und Wissenschaftler und Technikexpertinnen und Technikexperten unterstützt, die sich mit Vertrauen und Wahrheit im digitalen Zeitalter beschäftigen. Sie war Fellow am Shorenstein Center for Media, Politics and Public Policy an der Kennedy School in Harvard, Forschungsdirektorin am Tow Center for Digital Journalism an der Columbia University und Leiterin des Bereichs Social Media für UNHCR, dem Flüchtlingshilfswerk der Vereinten Nationen.*

Wie wir alle wissen, sind Lügen, Gerüchte und Propaganda keine neuen Konzepte. Menschen hatten immer die Fähigkeit zu betrügen, und es gibt einige beeindruckende Beispiele aus der Geschichte, wie frei erfundene Dinge genutzt wurden, um die Öffentlichkeit in die Irre zu führen, Regierungen zu destabilisieren oder Aktienmärkte explodieren zu lassen.[3] Was jedoch neu ist, ist die Leichtigkeit, mit der heute jedermann unredliche, falsche und irreführende Inhalte erstellen kann, sowie die Geschwindigkeit, in der sich diese über die Welt verbreiten. Uns war immer klar, dass es bei Täuschungen komplex zugehen kann. Es gibt nicht die eine Formel, die für alles passt. Eine Notlüge, die zum Beispiel um des lieben Friedens willen in einem Familienstreit erzählt wird, ist nicht das Gleiche wie ein irreführendes Statement eines Politikers, der damit mehr Wähler gewinnen will. Eine staatlich finanzierte Propagandakampagne ist nicht das Gleiche wie eine Verschwörungstheorie über die Mondlandung. Doch leider wurde in den letzten Jahren alles hier Genannte in einen Topf geworfen und mit Fake News bezeichnet, ein sehr einfacher Begriff, der sich weltweit durchgesetzt hat, meist ohne dass man ihn übersetzen müsste.

Ich sage leider, weil der Begriff völlig ungeeignet ist, um die Komplexität zu beschreiben, mit der wir es zu tun haben. Die meisten Inhalte, die auf irgendeine Art und Weise irreführend sind, sind nicht einmal als Nachrichten getarnt. Es sind Meme, Videos, Bilder oder koordinierte Gruppenaktionen auf Twitter, YouTube, Facebook oder Instagram. Und das meiste davon ist gar kein *Fake;* es ist irreführend oder, was noch häufiger vorkommt, zwar echt, aber aus dem Zusammenhang gerissen. Den größten Effekt haben Desinformationen, die einen Kern Wahrheit in sich tragen: irgendetwas, das wahr ist, aber in einen neuen Rahmen eingebettet (also „geframed") wird, oder das so geteilt wird, als ob es neu wäre, in Wahrheit aber alt ist.

Vielleicht ist das Gefährlichste dabei, dass der Begriff Fake News zu einer Waffe geworden ist, oft genutzt von Politikern und ihren Unterstützern, um Redaktionen weltweit anzugreifen. Meine Frustration über den Begriff hat mich dazu geführt, mit meinem Co-Autor Hossein Derakhshan die Formulierung „Informationsstörung" zu prägen. 2017 haben wir den Bericht „Information Disorder" vorgelegt und darin die Schwierigkeiten untersucht, die die jeweils verschiedenen Vokabeln in diesem Feld mit sich bringen.[4] In diesem Kapitel will ich daher einige der wichtigsten Aspekte erläutern, um dieses Thema zu verstehen und kritisch darüber diskutieren zu können.

3 Zum Beispiel: https://www.zdf.de/dokumentation/die-glorreichen-10/die-dreistesten-fake-news-der-geschichte-102.html
4 Der Bericht wurde für den Europarat erstellt und ist hier verfügbar: https://rm.coe.int/information-disorder-report-version-august-2018/16808c9c77

DIE SIEBEN TYPEN DER INFORMATIONSSTÖRUNG

In 2017 habe ich die nachfolgende Typologie vorgeschlagen, um die Unterschiede zwischen den verschiedenen Typen von Informationsstörung deutlich zu machen:

Schaubild der sieben häufigsten Formen von Informationsstörung nach Wardle.

Quellennachweis: https://de.firstdraftnews.org/fake-news-es-ist-kompliziert/

Satire oder Parodie

Verständlicherweise haben viele Menschen Einwände dagegen geäußert, dass ich Satire in diese Übersicht aufnehme, und auch ich selbst habe gezögert. Leider aber labeln viele Akteure, die Desinformation verbreiten, ihre Inhalte als Satire, um sicherzustellen, dass diese nicht auf Fakten hin überprüft werden und um jeden Schaden, den sie damit anrichten, direkt im Vorfeld zu entschuldigen. Doch in einem Informations-Ökosystem, in dem Kontexte, Hinweise oder gedankliche Abkürzungen (Heuristik) oft weggelassen werden, ist es wahrscheinlicher, dass auch satirische Inhalte Leserinnen und Leser verwirren. Ein Amerikaner wird wissen, dass The Onion eine satirische Website ist. Eine Deutsche wird wissen, dass Der Postillon keine echte Nachrichtenseite ist. Aber wussten Sie, dass es laut Wikipedia weltweit 57 satirische Nachrichtenwebsites gibt? Wenn Sie die nicht alle kennen und etwas von einer solchen Seite irgendwo im Facebook-Feed an Ihnen vorbeirast, ist es gar nicht so schwer, getäuscht zu werden.

Kürzlich hat Facebook bekanntgegeben, dass Satire nicht von den Faktencheckern geprüft werden soll, die das Unternehmen weltweit beauftragt hat, um falsche Informationen zu überprüfen und kenntlich zu machen.[5] Jene, die in diesem Bereich arbeiten, wissen, wie häufig Satire als Deckmantel eingesetzt wird. Im August 2019 hat die amerikanische Website Snopes, die sich auf die Richtigstellung von Falschinformationen spezialisiert hat, in einem Text erklärt, warum sie nun auch bei Satire einen Faktencheck macht. Inhalte, die behaupten, satirisch zu sein, entziehen sich den Faktencheckern, aber im Laufe der Zeit geht dieser ursprüngliche Kontext häufig verloren: Menschen beginnen, die Inhalte zu teilen und von da ab weiter zu teilen, ohne zu erkennen, dass es sich um Satire handelt – und alles in dem Glauben, der Inhalt sei wahr.

Falsche Verknüpfungen

Hier geht es um altmodisches sogenanntes Clickbait: Menschen werden auf eine bestimmte Website gelockt anhand von Versprechungen mit einer oftmals sensationellen Überschrift – nur um dann bei der Lektüre des Artikels zu merken, dass die Überschrift erschreckend wenig mit dem eigentlichen Inhalt zu tun hat. Es mag leicht sein für Nachrichtenmedien, beim

5 https://www.facebook.com/journalismproject/programs/third-party-fact-checking

Problem Desinformation zu denken, dass es nur von böswilligen Akteuren ausginge. Dem möchte ich entgegenhalten, dass es wichtig ist zu erkennen, wie schlechtes journalistisches Handwerk ebenfalls eine Gefahr werden kann.

Irreführende Inhalte
Irreführende Inhalte waren schon immer ein Problem, im Journalismus wie in der Politik. Sei es die Auswahl eines bestimmten Teils für ein Zitat, das Erstellen von Statistiken zur Unterstützung einer bestimmten Behauptung, ohne die Grundlage der Datenerhebung zu zeigen, oder das Zurechtschneiden von Fotos, um einen bestimmten Moment in einem bestimmten Licht dastehen zu lassen – diese Formen der Irreführung sind allesamt nicht neu.

Falsche Zusammenhänge
Aus dieser Kategorie sehen wir die meisten Beispiele: Das ist fast immer der Fall, wenn älteres Material nochmals neu gestreut wird. Oft passiert das bei Eilmeldungen und aktuellen Nachrichtenereignissen, indem alte Bilder neu geteilt werden. Mitunter geschieht es aber auch, dass ältere Artikel nochmals neu verteilt werden, deren Überschriften auch zu den aktuellen Ereignissen passen könnten.

Betrügerische Inhalte
Das ist zum Beispiel dann der Fall, wenn das Logo einer bekannten Marke oder ihr Name neben falschen Inhalten verwendet wird. Diese Taktik ist strategisch, denn sie spielt mit der Bedeutung von Heuristik, also von Daumenregeln, Erfahrungswerten, Bauchgefühl und Urteilsvermögen. Wir beurteilen Inhalte auch, indem wir schauen, ob sie von einer Organisation oder Person kommen, der wir schon vertrauen. Indem man also das Logo einer renommierten Nachrichtenredaktion nimmt und es zu einem Foto oder Video hinzufügt, erhöht man automatisch die Chance, dass Menschen diesen Inhalten vertrauen, ohne sie nochmals zu überprüfen.

Manipulated Content
Das geschieht, wenn echte Inhalte mit manipulierten Inhalten irgendwie zusammengebaut werden. Ein Video von Nancy Pelosi ist ein Beispiel dafür: Die Sprecherin des US-Repräsentantenhauses hielt im Mai 2019 eine Rede und wurde dabei gefilmt. Nur wenige Stunden später tauchte ein Video davon auf, in dem sie bei ihrer Rede betrunken klang. Die Geschwindigkeit war verringert worden und damit klang es so, als ob Pelosi lallte. Solche Methoden sind mächtig, denn sie basieren auf realem Material. Da die Menschen wussten, dass Pelosi diese spezielle Rede an diesem Tag vor diesem Hintergrund gehalten hatte, wurde es wahrscheinlicher, dass sie einem so gefälschten Video glaubten.

Erfundene Inhalte
Inhalte, die in diese Gruppe fallen, sind zu 100 % erfunden. Ein Beispiel wäre, dass ein neues Social-Media-Konto komplett gefälscht ist und bestimmte Inhalte verteilt. In diese Kategorie fallen auch sogenannte Deepfakes, bei denen künstliche Intelligenz benutzt wird, um eine Video- oder Audio-Aufnahme zu manipulieren, so dass es so aussieht, als ob jemand etwas gesagt oder getan hätte, was aber in der Realität nie der Fall war.

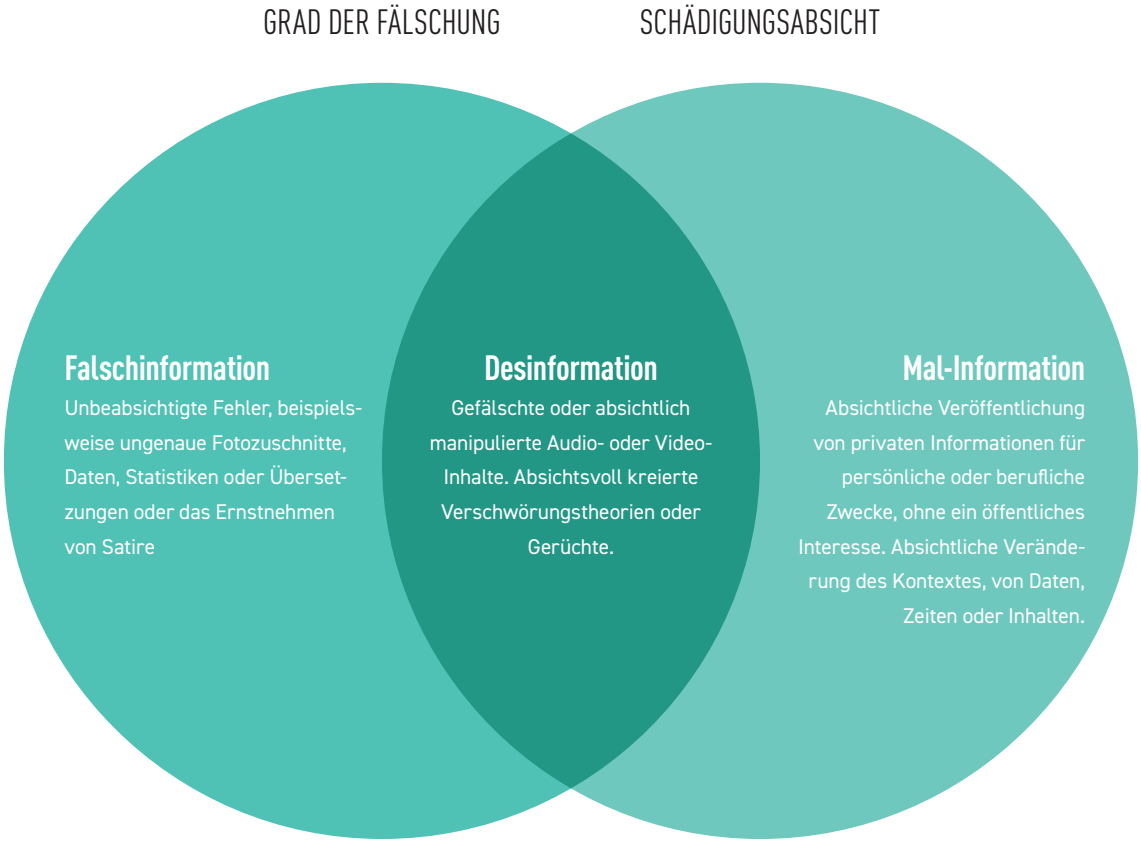
ABSICHT UND MOTIVATION VERSTEHEN

Diese Typisierungen können dabei helfen, die Komplexität einer vergifteten Informationsumgebung zu beschreiben, aber sie sagen nichts über die dahinterstehende Absicht. Das allerdings ist für das Verständnis dieser Phänomene ebenfalls ein entscheidender Teil. Um das zu verdeutlichen, haben Derakhshan und ich das nachfolgende Schaubild erstellt, das den Unterschied zwischen „Mis-Information“ (Falschinformation), „Dis-Information“ (Desinformation) und einer dritten Sorte, die wir „Mal-Information“ nennen, darzustellen.[6]

6 Mit dem Begriff **Desinformation** sind hergestellte Falschinformationen gemeint, die bewusst und unter Verfolgung bestimmter Ziele von Akteuren geplant produziert und verbreitet werden.
Mis-Information bzw. Falschinformation beschreibt hingegen die ungewollte und nicht absichtsvolle Veröffentlichung und Verbreitung von Falschinformationen.
Mal-Information sind nach Wardle jene Informationen, die für sich genommen zwar richtig sein mögen, aber von bestimmten Akteuren so verkürzt, einseitig oder unvollständig verwendet werden, dass sie eine verzerrte Realität zeichnen und das Potential haben, Gesellschaften zu spalten. Das kann zum Beispiel das auszugsweise Veröffentlichen privater Kommunikation, privater Informationen oder Fotos sein, um jemanden öffentlich in ein schlechtes Licht zu rücken.

Mis-Information und Desinformation sind beides Beispiele für unwahre Inhalte, allerdings ist Desinformation von Menschen hergestellt und verbreitet, die Schaden anrichten wollen, sei es finanziellen, rufschädigenden, politischen oder physischen Schaden. Mis-Information ist zwar ebenfalls unwahr, aber die Menschen, die sie teilen, realisieren in diesem Moment nicht, dass das so ist. Das passiert häufig in Momenten, in denen plötzlich etwas geschieht, das Eilmeldungen auslöst. Menschen teilen dann oft Gerüchte oder alte Fotos und realisieren nicht, dass diese keine Verbindung zu den Ereignissen haben.

TYPEN DER INFORMATIONSSTÖRUNG



Diese Begriffe sind wichtig, da die Absicht ein Teil der Antwort auf die Frage ist, wie eine bestimmte Information zu verstehen ist.

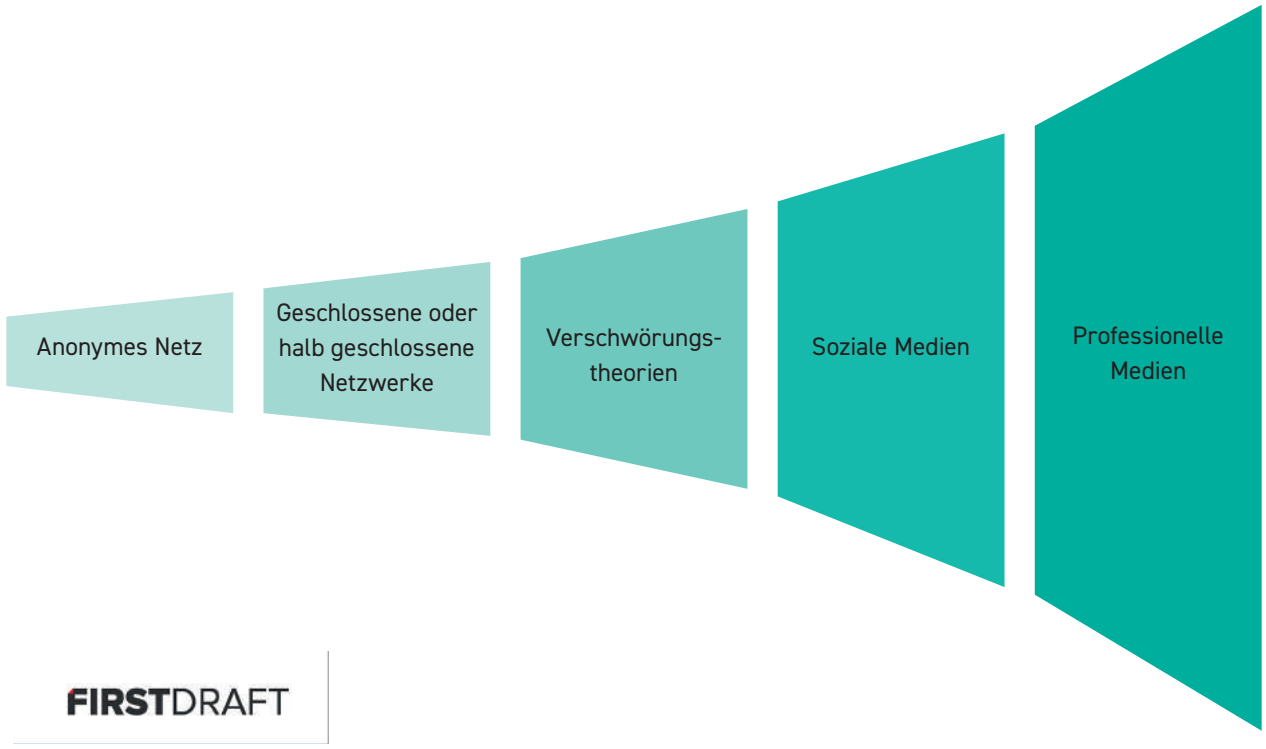
Es gibt drei Hauptmotive, warum falsche und irreführende Inhalte erstellt werden. Das *erste* ist politisch, egal ob für außen- oder innenpolitische Fragen. Es kann sein, dass ein Staat versucht, die Wahlen in einem anderen Land zu beeinflussen. Oder das Motiv kann innenpolitisch motiviert sein, indem sich zum Beispiel eine Kampagne unlauterer Mittel bedient, um den Konkurrenten schlecht dastehen zu lassen. Das *zweite* ist finanzieller Natur. Mit Werbung auf einer Website kann man Geld verdienen. Wer einen sensationellen, falschen Artikel mit einer übertriebenen Überschrift auf seiner Seite hat, kann damit so lange Geld verdienen, wie Menschen damit auf diese Seite gelockt werden. Dafür gibt es Beispiele von allen Seiten des politischen Spektrums.[7] Und *drittens* gibt es soziale und psychologische Motive. Manche Menschen motiviert einfach nur die Vorstellung, Probleme zu verursachen und zu schauen, wie weit sie damit kommen: zu sehen, ob sie Journalisten reinlegen können, eine Veranstaltung auf Facebook anzukündigen, die die Leute zum Protestieren auf die Straßen treibt, Frauen zu tyrannisieren und zu schikanieren. Andere teilen irgendwann Falschinformationen aus keinem anderen Grund als dem, ein bestimmtes Narrativ oder eine bestimmte Rolle zu bestärken. Nach dem Motto: „Es ist mir egal, ob es stimmt, ich will nur meinen Freunden auf Facebook zeigen, wie sehr ich Kandidat XY hasse."

7 https://www.npr.org/sections/alltechconsidered/2016/11/23/503146770/npr-finds-the-head-of-a-covert-fake-news-operation-in-the-suburbs?t=1578235215369&t=1598862620081

DIE TROMPETE DER VERSTÄRKUNG

Um dieses umfassendere Ökosystem wirklich zu verstehen, müssen wir uns anschauen, wie es miteinander verflochten ist. Allzu oft sieht jemand irgendwo irreführende oder falsche Inhalte und glaubt dann, diese seien auch dort entstanden. Leider wissen diejenigen, die Desinformation am besten verstehen, genau, wie sie sich das für ihre eigenen Zwecke zunutze machen können.

Man muss sich klar machen: Wenn Gerüchte, Verschwörungstheorien und falsche Inhalte nicht geteilt würden, würden sie auch keinen Schaden anrichten. Es ist das Teilen, das zerstörerisch ist. Ich habe dieses Schaubild erstellt, das ich die Trompete der Verstärkung nenne, um zu beschreiben, wie koordiniert die Agenten der Desinformation koordiniert vorgehen, um Informationen durch ein Ökosystem zu bewegen.



Häufig werden Inhalte in 4chan oder Discord veröffentlicht.[8]

Diese Orte sind anonym, sie erlauben es Menschen, zu veröffentlichen, ohne dafür Verantwortung zu übernehmen. Oft werden dort auch Anweisungen zur Koordination veröffentlicht, nach dem Prinzip „Wir werden versuchen, diesen Hashtag zu einem Trend zu machen" oder „Benutzt dieses Mem, um heute auf Facebook auf ein bestimmtes Thema zu antworten". Von hier verbreiten sich solche Anweisungen zur Koordination erst als private Nachricht, also noch nicht öffentlich, in großen Nutzergruppen auf Twitter, WhatsApp oder Telegram, so dass Knoten innerhalb dieser Netzwerke die Inhalte an eine größere Gruppe von Menschen verbreiten. Von da ausgehend gelangen die Inhalte häufig geballt auf Seiten wie Gab, Reddit oder YouTube, gelangen von dort auf Mainstream-Plattformen wie Facebook, Instagram oder Twitter – und hier werden sie dann oft von professionellen Medien aufgegriffen, weil sie entweder

8 **4chan** ist ein sogenanntes Imageboard: ein soziales Netzwerk, auf dem die Nutzer in bestimmten Gruppen Bilder und Grafiken, mittlerweile aber auch Videos und Fotos austauschen und darüber kommunizieren. Alle Nutzer posten und kommentieren anonym. Die Seite ist eine der meistbesuchten des Internets. Inhalte, die hier veröffentlicht werden, verbreiten sich anschließend häufig viral. Sie wird oft zum Verstärker und Ausgangspunkt weltweiter Netzphänomene. Häufig fallen Nutzer von hier durch organisierte Aktionen im Netz auf. Auch die Anonymous-Bewegung hat hier ihre Wurzeln. Kritisch wird gesehen, dass von Gruppen auf 4chan ausgehend häufig Inhalte mit Bezug zu Rechtsextremismus, Rassismus oder Frauenhass organisiert verteilt werden.
Discord ist ein Online-Service zum Chatten und für Sprach- oder Videokonferenzen, der vorrangig für Computerspieler gedacht war, inzwischen aber zunehmend auch von anderen Gruppen genutzt wird.

den Ursprung und die Vorgeschichte nicht erkennen und ohne angemessene Überprüfung die Inhalte in ihre Berichterstattung aufnehmen oder weil sie die Inhalte durchleuchtet haben und die Hintergründe entlarven, sie also „debunken".

In beiden Fällen werden die Urheber der Desinformation das als Erfolg verbuchen. Ob schlechte Überschriften, die das Gerücht oder die irreführende Behauptung weiterverbreiten, oder Richtigstellungen, die den falschen Inhalt einbetten in eine recherchierte Story – beides ist im Sinne des ursprünglichen Plans: nämlich den Effekt zu verstärken, das Gerücht zu befeuern.

Bei First Draft haben wir für dieses Dilemma ein Konzept, das wir den Wendepunkt nennen: Für Journalistinnen und Journalisten besteht einerseits die Gefahr, dass sie Gerüchte oder Lügen befeuern und verstärken, wenn sie zu früh darüber berichten. Zu spät berichten kann andererseits bedeuten, dass sich die Lüge bereits durchgesetzt hat und man kaum noch etwas dagegen tun kann. Diesen genauen Wendepunkt zu erkennen ist schwierig. Er unterscheidet sich je nach Standort, Inhalt oder Plattform.

FAZIT

Sprache ist wichtig. All diese Phänomene sind komplex und vielschichtig und die Worte, die wir für sie wählen, machen den Unterschied. Es gibt bereits wissenschaftliche Erkenntnisse darüber, dass immer mehr Menschen die Bezeichnung Fake News mit mangelhaftem journalistischem Handwerk von Redaktionen gleichsetzen.[9] Einfach alles nur Desinformation zu nennen, auch dann, wenn es sich nicht im Wortsinne um falsche Inhalte handelt oder wenn diese unabsichtlich von Menschen geteilt wurden, die dachten, die Inhalte seien korrekt – auch das steht einer genauen Beschreibung dessen, was passiert, im Weg.

Wir leben in einem Zeitalter der Informationsstörung. Es birgt neue Herausforderungen für Journalistinnen und Journalisten, Wissenschaftlerinnen und Wissenschaftler, Expertinnen und Experten. Berichten oder nicht berichten? Wie eine Überschrift formulieren? Wie ein Video oder Bild effektiv richtigstellen? Wie bewerten, ob der Wendepunkt schon erreicht ist? Diese Fragen stellen sich für all jene, die in diesem Feld arbeiten. Es ist kompliziert.

9 Vgl. dazu auch: https://shop.freiheit.org/download/P2@792/214061/FNF_FakeNews_Broschuere_DE_final.pdf, S. 23, sowie https://www.stiftung-nv.de/sites/default/files/fake_news_methodenpapier_deutsch.pdf und https://www.stiftung-nv.de/de/publikation/kurzanalyse-zu-trumps-crime-tweet-deutschland-viel-aufmerksamkeit-wenig-unterstuetzung

DER LEBENSZYKLUS VON MEDIENMANIPULATION

von: Joan Donovan
deutsche Bearbeitung: Marcus Engert

Dr. Joan Donovan *ist Forschungsdirektorin am Harvard Kennedy Shorenstein Center on Media, Politics and Public Policy.*

In einer Zeit, in der eine Handvoll einflussreicher globaler Tech-Plattformen die klassischen Kanäle, über die eine Gesellschaft sich informiert, durcheinandergewirbelt hat, fordern Medienmanipulation und Desinformationskampagnen alle politischen und sozialen Institutionen heraus. Tricksereien und Fälschungen werden von einer bunten Mischung aus politischen Akteuren, Marken, sozialen Bewegungen und lose agierenden Trollen verbreitet, die neue Techniken zur Beeinflussung der öffentlichen Meinung entwickelt und verfeinert haben und damit auf lokaler, nationaler und sogar globaler Ebene einiges an Schaden anrichten. Es gibt weitgehende Einigkeit darüber, dass Medienmanipulation und Desinformation wichtige Probleme sind, mit denen sich unsere Gesellschaft konfrontiert sieht. Doch es bleibt schwer, Desinformation zu definieren, aufzuspüren, zu dokumentieren und richtigzustellen, zumal die Attacken vom Journalismus über das Rechtssystem bis hin zu Plattformen viele Bereiche tangieren. Für die Untersuchung, Aufdeckung und Eindämmung dieser Angriffe ist es darum zentral, zuerst ein Verständnis von Medienmanipulation als strukturiertem Vorgang zu entwickeln.

MEDIENMANIPULATION UND DESINFORMATION DEFINIEREN

Um Medienmanipulation zu definieren, teilen wir den Begriff zunächst in zwei Teile. In ihrer allgemeinsten Form sind Medien ein Produkt von Kommunikation. Beispiele dafür sind Texte, Bilder, Audios und Videos, sowohl auf analogen als auch auf digitalen Trägern. Bei der Analyse von Medien kann jedes Fundstück als aufgezeichneter Beweis eines Ereignisses betrachtet werden. Von entscheidender Bedeutung dabei ist, dass Medien von Individuen zum Zwecke der Kommunikation hergestellt werden. So vermitteln Medien eine bestimmte Bedeutung zwischen Individuen, aber die Interpretation dieser Bedeutung ist immer relativ und eingebettet in einen bestimmten Kontext der Übertragung. Wer sagt, Medien seien manipuliert, der geht über die bloße Feststellung, dass Medien von Individuen zur Übertragung einer bestimmten Bedeutung gestaltet wurden, hinaus. Das Wörterbuch Merriam-Webster beschreibt Manipulation als „etwas mit raffinierten oder unfairen Mitteln so zu verändern, dass es dem eigenen Zweck dient". Der Duden als „undurchschaubares, geschicktes Vorgehen, mit dem sich jemand einen Vorteil verschafft". Auch wenn es mitunter schwierig sein kann zu erkennen, für welchen konkreten Zweck ein bestimmtes Fundstück kreiert wurde, kann es helfen, das Wer, das Was, das Wo und das Wie der Kommunikation zu bestimmen, um festzustellen, ob manipulative Taktiken als Teil des Verteilungsprozesses angewandt wurden.

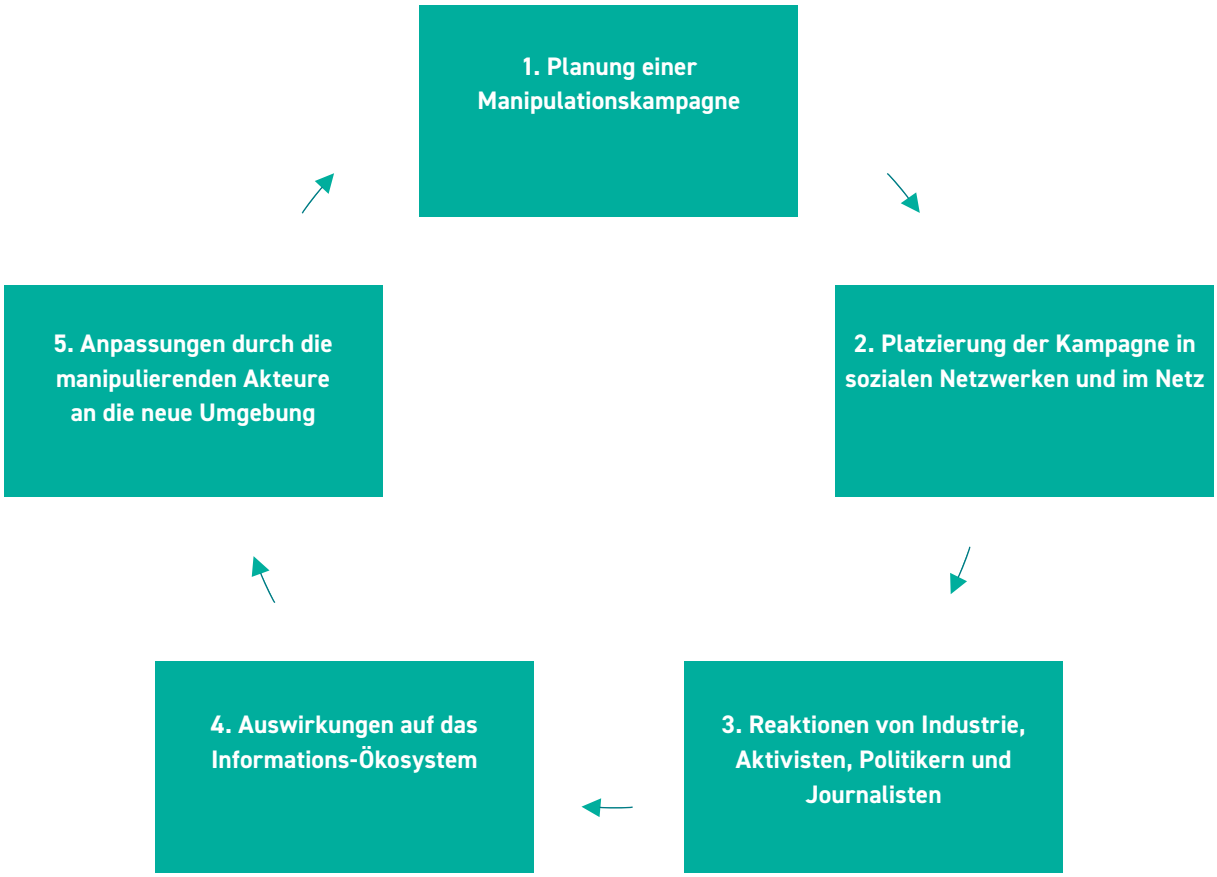
Manipulationstechniken können sein, die eigene Identität, die Quelle des Fundstücks, seine Bearbeitung, seine Bedeutung oder auch seinen Kontext zu verschleiern oder zu verändern. Es kann auch darum gehen, Algorithmen durch technische Mittel, wie zum Beispiel Bots oder Spamming-Tools (das sind automatisierte massenhafte Aussendungen oder Interaktionen, hinter denen keine einzelnen Menschen sitzen, obwohl es den Anschein erwecken soll) zu überlisten.

In diesem Zusammenhang ist Desinformation also eine Unterart von Medienmanipulation und bezieht sich auf das Herstellen und Verbreiten von vorsätzlich falschen Informationen zu politischen Zwecken.

Expertinnen und Experten, Wissenschaftlerinnen und Wissenschaftler, Journalistinnen und Journalisten, Politikerinnen und Politiker – sie alle müssen sich über den besonderen Stellenwert von Desinformation bewusst sein, denn ihre Bekämpfung erfordert die Kooperation dieser Gruppen.

Wir selbst, das Technology and Social Change research team (TaSC) am Harvard Kennedy Shorenstein Center, nutzen zur Darstellung des Lebenszyklus von Medienmanipulationen einen Fallstudienansatz. Dieser methodische Ansatz versucht, das Durcheinander zu ordnen, den Ablauf und den Umfang von Manipulationskampagnen zu analysieren, indem wir verfolgen, wie sich das Fundstück durch Raum und Zeit verbreitet und wann es wo welche Beziehungen knüpft. Als Teil dieser

Arbeit haben wir einen Überblick über den typischen Lebenszyklus von Medienmanipulationskampagnen erstellt, der für Journalistinnen und Journalisten hilfreich sein kann, Medienmanipulation zu identifizieren, zu verfolgen und zu entlarven.



Dieser Lebenszyklus hat fünf Handlungsmomente. Zu jedem können Taktiken und Medienmanipulatoren dokumentiert werden, sowohl quantitativ als auch qualitativ. Wichtig ist, dass die meisten Manipulationskampagnen nicht in dieser Reihenfolge „entdeckt" werden. Daher sollte man während seiner Recherchen nach jedem dieser Momente Ausschau halten und von da ab die Kampagne in ihrem Lebenszyklus sowohl vorwärts beobachten als auch rückwärts rekonstruieren.

FALLBEISPIEL: „BLOW THE WHISTLE" – EIN GEHEIMNIS VERRATEN

Untersuchen wir einmal die Social-Media-Aktivitäten im Zusammenhang mit dem Whistleblower, der sich zu Donald Trumps Aktivitäten in der Ukraine äußerte, um zu sehen, wie eine Medienmanipulationskampagne sich entfaltet und wie ethisches Handeln von Journalisten und Plattformen zu Beginn des Lebenszyklus einer solchen Kampagne dazu beitragen kann, Manipulationsbemühungen zu vereiteln.

Ein Tweet verbreitet den angeblichen Namen und das angebliche Foto des Whistleblowers. Quelle: https://twitter.com/realcandaceo/status/1192817302422601730?lang=de

Planung und Platzierung (Phasen 1 und 2): Im Medien-Ökosystem von Verschwörungstheorien war die Identität des Whistleblowers bereits bekannt und sein Name zirkulierte in Blogs, Twitter, Facebook, YouTube-Videos und Diskussionsforen. Wichtig ist, dass Klarnamen wie Suchwörter oder Hashtags fungieren, also als spezifische Suchphrase, die gesucht und gefunden werden kann. Es gab einen konzertierten Vorstoß mit dem Ziel, seinen Namen und sein Foto so breit wie möglich zu streuen. Dennoch blieb die Verbreitung des Namens weitgehend auf eine Echokammer aus rechten Accounts und Verschwörungstheoretikern beschränkt. Selbst mit koordinierten Bemühungen von reichweitenstarken Influencern aus dem Verschwörungsbereich gelang es nicht, den Namen in den Mainstream zu rücken und aus der eigenen Filterblase herauszukommen. Warum?

Reaktionen von Journalisten, Aktivisten etc. (Phase 3): Gemäßigte, liberale, progressive oder eher linke Medien veröffentlichten den Namen des angeblichen Whistleblowers nicht und verbreiteten auch nicht die Behauptung, er sei enttarnt worden. Obwohl es sich um eine relevante und berichtenswerte Geschichte handelte, haben große Redaktionen es unterlassen, Aufmerksamkeit auf die Verbreitung des Namens in sozialen Netzwerken zu lenken. Jene, die darüber berichteten, betonten zwar meist, wie mit der Verbreitung des Namens der Versuch unternommen werden sollte, die Diskussion zu manipulieren, verzichteten aber selbst auf die Namensnennung. Das geht in weiten Teilen auf ein Berufsethos im Journalismus zurück, nach dem Reporterinnen und Reporter eine besondere Verpflichtung haben, ihre Quellen (also auch Whistleblower) zu schützen.

Auswirkungen auf das Informations-Ökosystem (Phase 4): Während Journalisten seriöser Medien den Namen nicht verbreiteten, wurde der angebliche Name Eric Ciaramella selbst zu einem unverwechselbaren Schlüsselbegriff. Menschen, die speziell danach suchten, konnten eine große Vielzahl von Inhalten finden, die einem von Verschwörungstheorien dominierten Milieu entstammten.

Zusätzlich zu ethisch korrekt handelnden Journalistinnen und Journalisten, die eine Story, obwohl sie viel Aufmerksamkeit hätte bekommen können, nicht veröffentlichten, begannen auch Plattformen damit, Inhalte aktiv zu moderieren, die den angeblichen Namen des Whistleblowers als Schlüsselwort enthielten. YouTube und Facebook löschten Inhalte, die den Namen enthielten, während Twitter den Namen aus den „trending topics" (einer Auflistung von Themen, die gerade große Aufmerksamkeit erhalten) entfernte. Anders die Suche von Google: Google ließ es zu, den Namen zu suchen, und brachte als Suchergebnisse tausende Links zu Verschwörungsblogs.

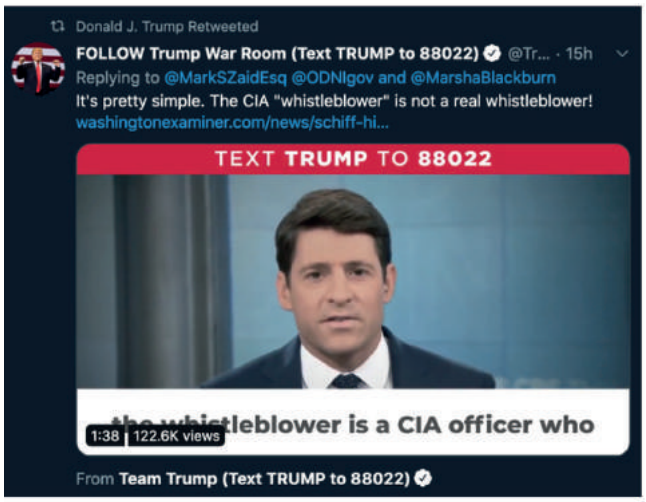
Anpassungen durch die manipulativen Akteure (Phase 5): Durch diese Maßnahmen, die die Verbreitung von Falschinformationen eindämmen sollten, wurden die manipulativen Akteure noch angestachelt und änderten ihre Taktiken. Statt weiterhin auf die Verbreitung des angeblichen Namens zu drängen, begannen sie damit, Bilder eines anderen weißen Mannes (mit Brille und Bart) zu verbreiten, die dem Bild ähnelten, welches sie zunächst gemeinsam mit dem angeblichen Namen verbreitet hatten. Diese neuen Bilder wurden mit einem Narrativ eines geheimen „Staates im Staat" vermischt, und es wurde behauptet, der Whistleblower sei eng befreundet mit hochrangigen und prominenten Demokraten und habe also aus parteilichen Motiven gehandelt. Tatsächlich handelte es sich um Bilder von Alexander Soros, dem Sohn des Milliardärs, Investors und Philanthropen George Soros, der ein Dauerziel von Verschwörungstheorien ist. Als auch dieser Versuch, die Aufmerksamkeit von Redaktionen zu bekommen, scheiterte, verbreitete das Twitter-Profil von US-Prä-

Ein Tweet von Steve King, einem Kongressmitglied, der vier vermeintliche Bilder des angeblichen Whistleblowers verbreitete.
Quelle: https://twitter.com/SteveKingIA/status/1194983687508676609 (zwischenzeitlich gelöscht, Archiv verfügbar unter https://web.archive.org/web/20191114145218if_/https://twitter.com/SteveKingIA/status/1194983687508676609)

sident Donald Trump, @RealDonaldTrump, einen Artikel an seine 68 Millionen Follower, der den angeblichen Namen des Whistleblowers enthielt, zusammen mit folgender Botschaft: „Der CIA Whistleblower ist kein echter Whistleblower!" Der ursprüngliche Tweet kam vom Profil @TrumpWarRoom, dem Account seiner Wahlkampagne, ebenfalls ein von Twitter als echt eingestuftes Profil.

Eine Kaskade von Medienberichterstattungen folgte, inklusive großer Mainstream-Medien, die sich alle bemühten, den angeblichen Namen zu entfernen oder zu schwärzen. Viele Menschen forderten daraufhin in sozialen Netzwerken, der Whistleblower solle in einer Anhörung des Senats über ein Amtsenthebungsverfahren gegen Donald Trump aussagen, wobei sein Name (neben dem anderer wichtiger Zeugen) aufgerufen worden wäre, was dann wiederum die Wahrscheinlichkeit erhöhen würde, dass andere über diesen Namen stolpern. Und so beginnt ein neuer Zyklus von Medienmanipulation.

Die Anfragen nach dem Namen des Hinweisgebers nehmen zu und in Blogs kursieren zahlreiche Vermutungen und Theorien zu seinen möglichen persönlichen und beruflichen Beweggründen, Geheimnisse über Trumps Aktivitäten zu verbreiten. Die Beiträge von Journalisten, die über diese Tweets berichteten, reichten von Sorge über die Einschüchterung von Zeugen über die Gefahr, dass ein solches Vorgehen künftige Informanten abschrecken könnte, bis zu reißerischer Neugierde und Tratscherei über Trumps Motive, den angeblichen Informanten zu entlarven. Insofern ist es zwar lobenswert, wie einige Redaktionen versuchten, Verantwortungsträger zur Verantwortung zu ziehen, dies muss aber scheitern, solange die Plattformen und sozialen Netzwerke sich nicht damit befassen, wie ihre Produkte zu nützlichen politischen Werkzeugen für Medienmanipulation und die Verbreitung von Desinformationen geworden sind.



DEN LEBENSZYKLUS DOKUMENTIEREN

Jene, die großen Schaden mit Medienmanipulation anrichten wollten, wollten erreichen, dass Redaktionen dieses Bemühen noch verstärken, indem sie den Namen und Fotos in sozialen Netzwerken verbreiteten – bei den Plattformen, die es zuließen, dass der Name als Suchbegriff zu einem Trend wurde und leicht such- und auffindbar war.

Es waren Entscheidungen und Maßnahmen von Redaktionen, Journalistinnen und Journalisten, die diesen Versuch, die angebliche Identität des Whistleblowers in das Bewusstsein der Mainstream-Medien zu bekommen, großflächig scheitern ließen, zumindest bis zu dem Moment, an dem eine so wichtige Figur wie Donald Trump das Thema groß machte. Während viele Redaktionen bemüht waren, sich an ihren Berufsethos zu halten, sind soziale Medien zu einer Waffe der ohnehin schon Mächtigen geworden, um die Agenda für Medien zu setzen und gefährliche Verschwörungstheorien zu verstärken.

Im Allgemeinen aber zeigt dieser Fall erhebliche Verbesserungen gegenüber früheren Bemühungen, die Verbreitung von Desinformationen zu stoppen, bei denen Journalistinnen und Journalisten Desinformationskampagnen noch verstärkten, indem sie versuchten, sie zu entlarven, und Plattformbetreiber sich nicht verpflichtet fühlten, dem Publikum richtige Informationen zur Verfügung zu stellen. Dieser Trend ist vielversprechend, aber eines fehlt immer noch: nämlich, dass sich Verantwortungsträger auch verantworten müssen. Sowohl für Journalistinnen und Journalisten als auch für Forscherinnen und Forscher steht viel auf dem Spiel, wenn sie Kampagnen zur Medienmanipulation aufdecken, dokumentieren und entlarven. In einem solchen Moment kann jedes Benennen einer Desinformationskampagne auch Horden von Trollen und unerwünschte Aufmerksamkeit mit sich bringen. Die Auseinandersetzung mit dem Inhalt und dem Kontext von Desinformation erfordert von uns allen, forensisch genau zu dokumentieren, wie Kampagnen beginnen, sich verändern und enden. Und wir müssen erkennen, dass jedes gefühlte Ende einer solchen Kampagne immer auch der Beginn einer neuen sein kann.

1. SOCIAL-MEDIA-PROFILE UNTERSUCHEN

von: Brandy Zadrozny
deutsche Bearbeitung: Marcus Engert

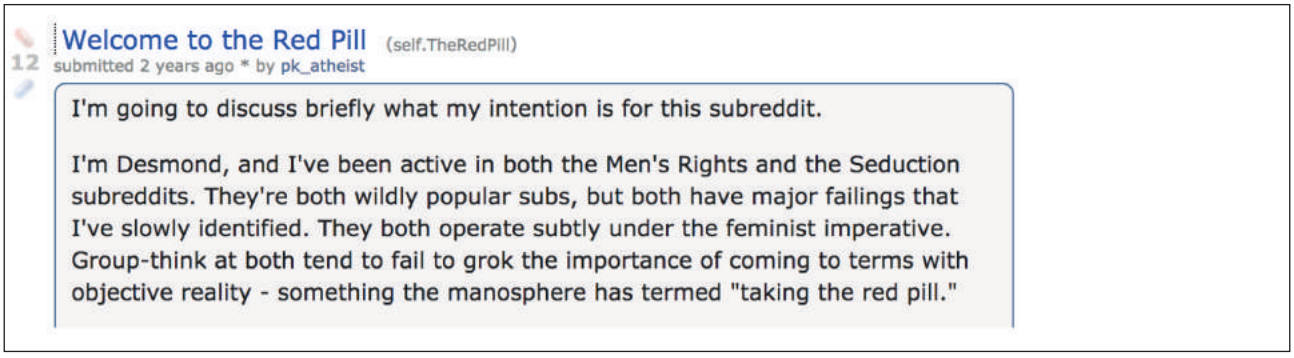
Brandy Zadrozny *ist Investigativ-Reporterin bei NBC News in New York und beschäftigt sich dort hauptsächlich mit Falschbehauptungen, Desinformation und Extremismus im Internet.*

So ziemlich jede Story, über die ich berichte, beinhaltet das Recherchieren in sozialen Netzwerken. Von den Hintergründen eines Profils über Breaking News bis hin zu längeren Recherchen: Soziale Netzwerke sind eine der besten Möglichkeiten, etwas über das echte Leben von jemandem zu erfahren – Freunde, Familie, Berufe, Verbindungen zu Politik oder gesellschaftlichen Gruppen – und auch über die verborgenen Gedanken und weitere Online-Identitäten. Es ist eine unglaubliche Zeit, um Journalistin oder Journalist zu sein; die Menschen leben ihr Leben zunehmend online, und die Werkzeuge, um die sozialen Profile zu finden und zu durchsuchen, sind überall. Inzwischen haben Plattformen wie Facebook auch auf negative Presse über die Verletzung der Privatsphäre und gefährliche Ideologien, die sich über die Plattformen verbreiten, reagiert – indem sie die Tools entfernt, geschlossen und ausgesperrt haben, auf die Journalistinnen und Journalisten, Forscherinnen und Forscher angewiesen sind, um diese Geschichten aufzudecken und die Menschen dahinter zu identifizieren.

Im folgenden Kapitel werde ich darum einige zentrale Ansätze zeigen, um Profile auf sozialen Netzwerken zu untersuchen. Ich zeige Werkzeuge und Tools, mit denen ich derzeit gerade arbeite. Schon bald werden sie von Facebook ausgesperrt worden sein oder ersetzt durch bessere. Reporterinnen und Reporter, die diese Arbeit machen, haben ihre eigenen Abläufe und Kniffe, aber wie bei jedem Berichterstattungsgebiet führen auch hier Hartnäckigkeit und Ausdauer zu den besten Ergebnissen. Man muss sich darauf einstellen, abertausende von Tweets zu lesen, bis zur letzten Seite der Google-Suchergebnisse zu klicken, immer tiefer in einem Social-Media-Profil zu versinken, vom Hölzchen aufs Stöckchen zu kommen, um genau die winzigen biographischen Hinweise zu finden, die uns am Ende die Frage beantworten lassen: „Wer ist das?"

NUTZERNAMEN

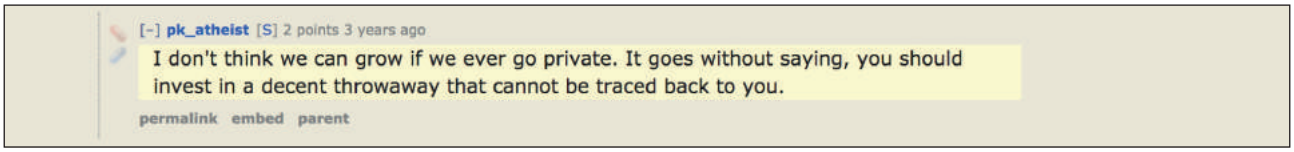
Manchmal ist ein Nutzername alles, was wir haben. Das ist okay, denn meistens ist es auch genau das, womit wir ohnehin beginnen. So war das auch im Falle jenes konservativen Abgeordneten aus New Hampshire, der eine der größten und widerlichsten Männergruppen auf Reddit aufgebaut hatte.[10] Auch diese Recherche nach dem Architekten von „The Red Pill", mittlerweile ist die Gruppe stillgelegt, begann mit einem Nutzernamen: „pk_atheist".



Screenshot aus der Gruppe „The Red Pill" auf Reddit. Im hier gezeigten Begrüßungstext heißt es, der Ersteller der Gruppe, Desmond, sei vorher bereits in anderen „Männerrechte"-Gruppen aktiv gewesen, habe dort aber „erhebliche Mängel" festgestellt. Sie würden „subtil unter dem feministischen Imperativ" arbeiten. In den dortigen Gruppendiskussionen begreife man nicht, wie wichtig es sei, sich mit der „objektiven Realität" auseinanderzusetzen – etwas, dass die „Manosphere" als „Die rote Pille schlucken" bezeichnen würde.

10 Reddit ist ein Social-News-Aggregator: In Gruppen und Communitys zu allen möglichen Themen sammeln und teilen Menschen dort Inhalte. Reddit nennt sich daher auch „Die Startseite des Internets".

Manche Menschen hängen an ihren Benutzernamen und verwenden sie, manchmal mit kleinen Variationen, über verschiedene Plattformen und E-Mail-Anbieter hinweg. Vorsichtigere Menschen, so wie der Abgeordnete aus New Hampshire einer ist, erstellen und löschen Benutzerkonten mit jedem neuen Unterfangen.

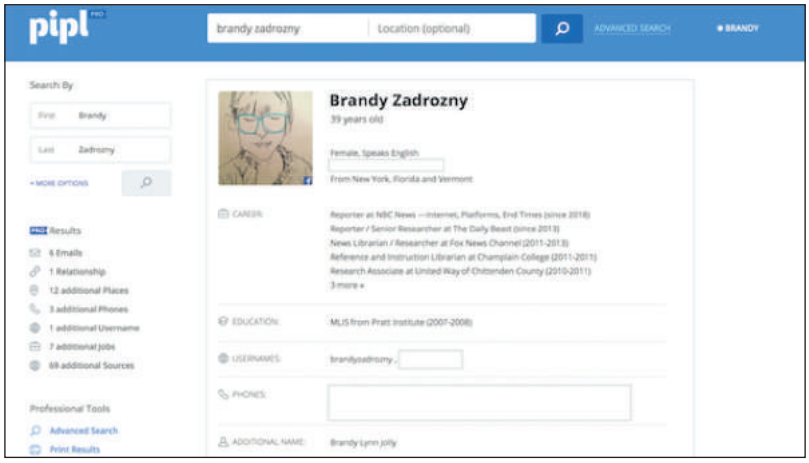


Screenshot einer Mitteilung in der Reddit-Gruppe. Der Autor äußert darin Zweifel, dass die Gruppe wachsen könne, wenn sie auf Privat gestellt, also nicht öffentlich sei. Er sagt, die Mitglieder sollten sich stattdessen Wegwerf-Mailadressen besorgen, die man nicht zu ihnen zurückverfolgen könne.

In jedem Fall gibt es ein paar Websites, bei denen man den Nutzernamen einfach einmal eingeben sollte.

Als Erstes gebe ich Nutzernamen bei Google ein. Menschen – vor allem jüngere, die die größeren sozialen Netzwerke meiden – tendieren dazu, an unerwarteten Stellen eine Spur zu hinterlassen, so zum Beispiel in Kommentarspalten, Rezensionen oder Foren für Bewertungen und Erfahrungen, die wiederum zu anderen Konten führen können. Sinnvoll sind neben einer Google-Suche auch eigenständige Dienste. Die können Geld kosten, und je nach Budget ist der Zugang dazu möglich oder nicht. Nexis ist ein Dienst, den viele große Firmen haben und der großartig ist, um öffentliche Dokumente und Gerichtsakten zu finden, aber leider ist Nexis nicht sonderlich gut darin, nach Nutzernamen oder Mailadressen zu suchen. Dazu kommt, dass der Dienst sich nur für die Suche nach Menschen in den Vereinigten Staaten eignet.[11] Pipl und Skopenow gehören, wie ich finde, zu den besten Diensten, wenn es um Informationen aus der „realen" Welt geht, wie Telefonnummern, Akten über Immobilieneigentum, um E-Mail-Adressen oder Benutzernamen abzugleichen, und beide funktionieren weltweit.[12] Diese kostenpflichtigen Suchmaschinen können oftmals Telefonnummern oder Aufzeichnungen über Grundbesitz liefern, aber sie können auch Facebook- oder LinkedIn-Profile finden, mitunter auch dann noch, wenn der Account dahinter geschlossen wurde.

Sie zeigen auch Verbindungen zu Nutzerkonten, die die Menschen längst vergessen haben: Alte Blogs oder Amazon-Wunschlisten zum Beispiel – eine Goldgrube, um herauszufinden, was jemand liest, kauft oder möchte. Mit solchen Suchen bekommt man immer auch viele Falsch-Treffer, weshalb ich dazu neige, meine Recherchen bei diesen Ergebnissen zu beginnen und sie dann mit weiteren Überprüfungsschritten zu kombinieren.



Screenshot der kostenpflichtigen Personen-Suchmaschine Pipl. Sie listet Alter, Mailadressen, Beziehungen, Lebensmittelpunkte, Ausbildungsweg, Berufe und die dazugehörigen Quellen auf.

11 Anmerkung: Mitunter verfügen größere öffentliche Bibliotheken oder Hochschulbibliotheken über einen Nexis-Zugang. Gerichtsurteile können in Deutschland über den Dienst juris.de gefunden werden, den ebenfalls viele Bibliotheken abonniert haben. Darüber hinaus unterhalten einzelne Bundesländer eigene Entscheidungsdatenbanken. In Nordrhein-Westfalen ist dies zum Beispiel nrwe.de. Eine Übersicht findet sich hier: https://justiz.de/onlinedienste/rechtsprechung/index.php. Informationen über Firmen sowie Jahresabschlüsse und Handelsregisterauszüge können über unternehmensregister.de und handelsregister.de gegen Gebühr heruntergeladen werden. Vor privaten Anbietern, die den gleichen Service bieten, sollte gewarnt werden: Meist erhält man dort die gleichen Informationen, die Gebühren allerdings sind höher.

12 Anbieter wie diese können nur jene Informationen bereitstellen, die jemand zu einem früheren Zeitpunkt einmal herausgegeben hat. In Ländern mit höheren Datenschutzstandards kann das wenig sein. Für Deutschland lohnt sich vor dem Abschluss eines Abos daher in jedem Fall ein kostenloser Testmonat.

Wenn ich einen Nutzernamen oder eine Mailadresse gefunden habe, von der ich glaube, sie gehöre zu meinem Untersuchungsgegenstand, jage ich sie durch ein Tool wie namechk or namecheckr, mit dem man prüfen kann, ob ein Name auf verschiedenen Plattformen noch frei ist.

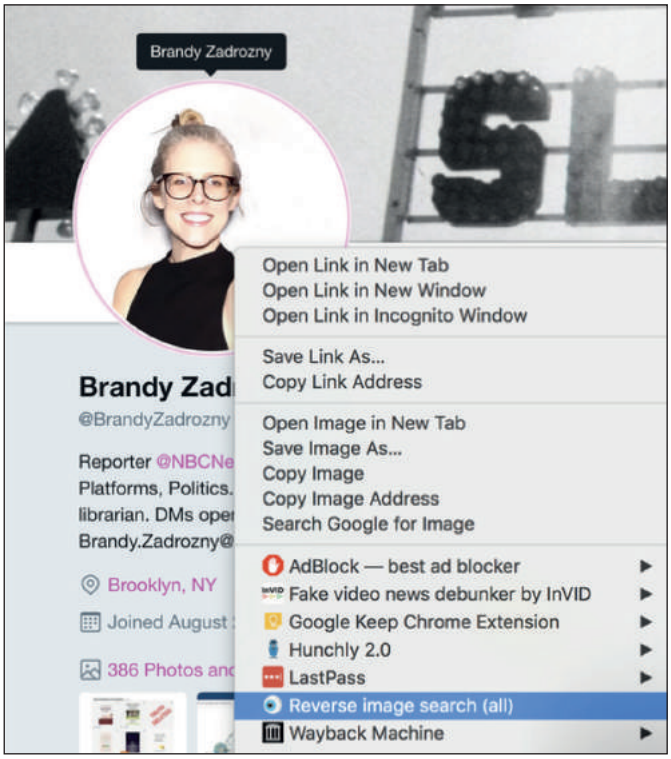
Diese Tools sind eigentlich für das Marketing gedacht, um zu überprüfen, ob ein bestimmter Name, den man sich als Marke schützen lassen möchte, über verschiedene Plattformen hinweg verfügbar ist. Aber sie sind ebenso hilfreich, um zu überprüfen, ob ein Nutzername, zu dem man recherchiert, auf anderen Plattformen existiert. Klar, nur weil ein Name irgendwo existiert, heißt das noch lange nicht, dass alle diese Konten zusammengehören. Aber es ist ein guter Ausgangspunkt, um von hier plattformübergreifend weiterzusuchen.

Nach Eingabe des Namens zeigt die Website namechk.com an, auf welchen Online-Diensten dieser noch verfügbar ist.

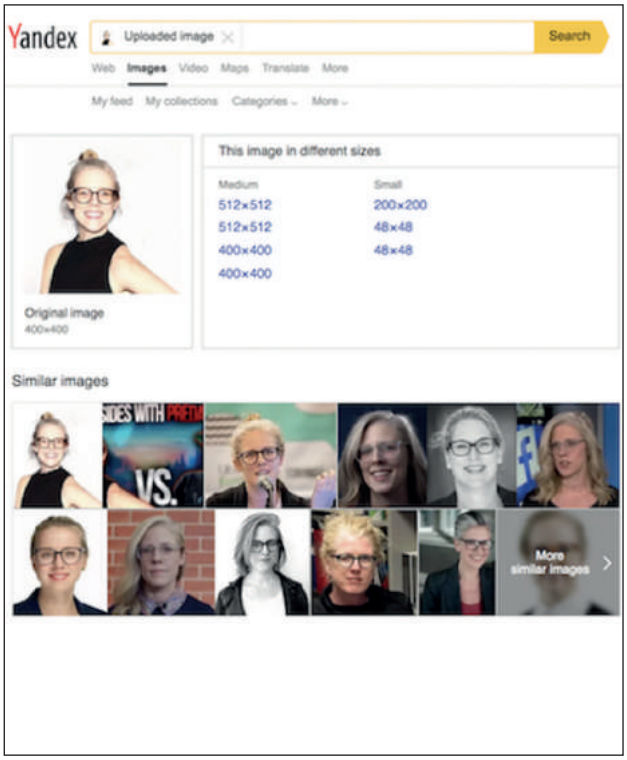
Zur weiteren Überprüfung von Benutzernamen gibt es haveibeenpwned.com und Dehashed.com, die beide Datenlecks und gestohlene Daten nach Benutzerinformationen durchsuchen – und damit auch eine schnelle Möglichkeit sind, um nachzusehen, ob es eine bestimmte Mailadresse wirklich gibt – und so neue Rechercherichtungen zu eröffnen.

FOTOS

Ein Benutzername allein ist meist nicht genug, um schon weitermachen zu können, und nichts überzeugt Menschen so sehr wie ein Bild. Profilbilder sind eine weitere Möglichkeit, eine Person über mehrere Konten hinweg zu identifizieren. Googles Bilder-Rückwärtssuche ist okay, aber oft liefern andere Suchmaschinen bessere Ergebnisse – vor allem die aus Russland stammende Yandex. Ich habe mir die Chrome-Erweiterung von Reveye installiert, so dass ich per Rechtsklick auf ein Bild mehrere Plattformen auf einmal danach durchsuchen kann: inklusive Google, Bing, Yandex and Tineye. Eine andere Erweiterung namens Search by Image hat eine praktische Funktion, mit der man nur einen Ausschnitt des Bildes markieren und damit quasi nach einem Bild in einem Bild suchen kann.

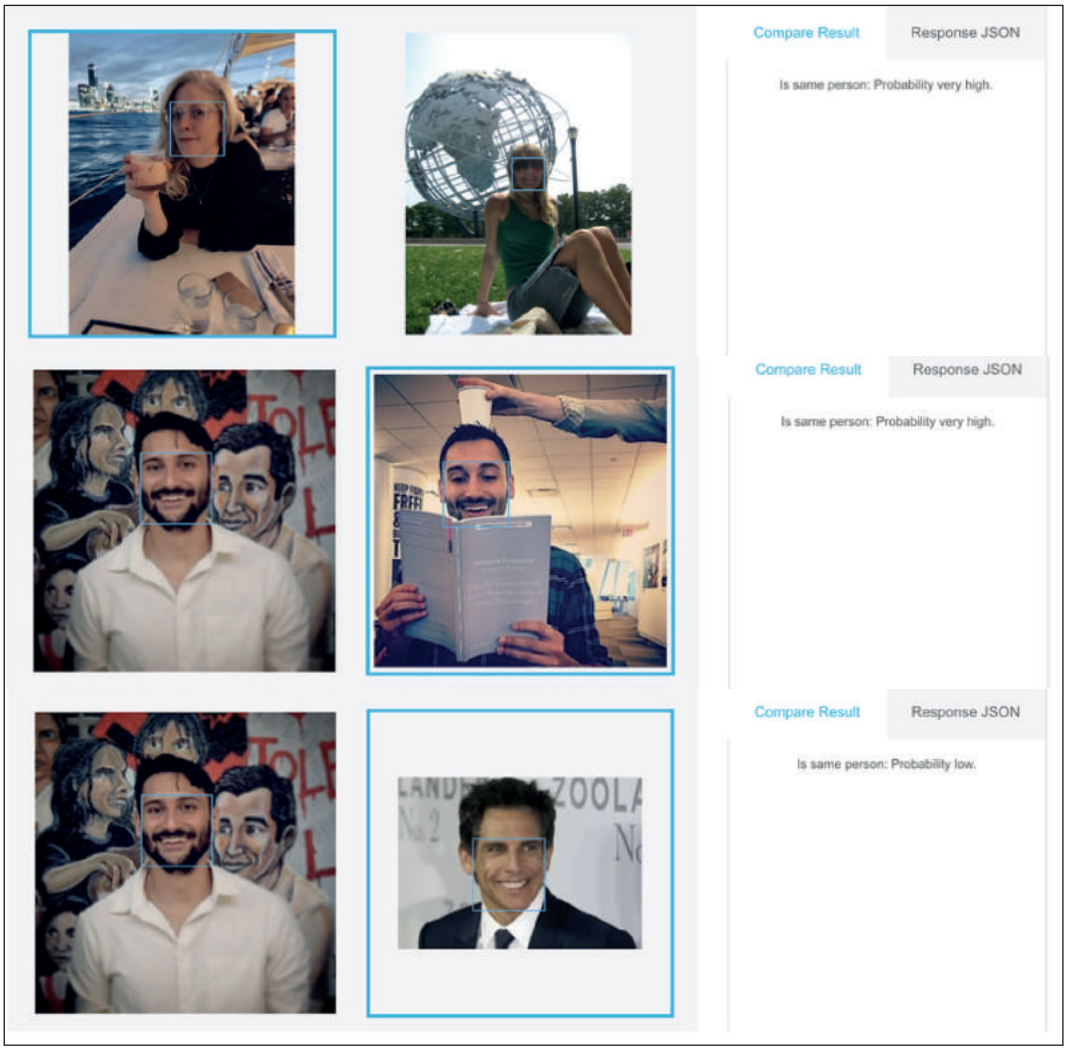


Hat man die Erweiterung von Reveye installiert und klickt mit der rechten Maustaste auf ein Bild, hat man direkt dort die Optionen, das Bild per Rückwärtssuche im Netz zu finden.



Die Website yandex.com liefert oft bessere Suchergebnisse bei der Bilder-Rückwärtssuche als andere Suchmaschinen.

Natürlich gibt es auch Probleme mit der Bilder-Rückwärtssuche. Die oben genannten Dienste sind ziemlich schlecht darin, Bilder auf Twitter zu finden und so gut wie nutzlos für Seiten wie Instagram und Facebook. Also schaue ich mir oft stundenlang verschiedene Bilder von Menschen an. Ich kann gar nicht mehr zählen, wie oft ich auf meinen Monitor gestarrt und meine Kollegen gefragt habe: „Ist das die gleiche Person?" Ich vertraue meinen Augen einfach nicht. Auf Fotos charakteristische Merkmale wie Muttermale oder Gesichtshaare zu identifizieren ist hilfreich; in letzter Zeit habe ich mir angewöhnt, sie auch mit Gesichtserkennungsprogrammen wie Face++ zu überprüfen, die es einem erlauben, zwei Fotos hochzuladen, und dann eine Wahrscheinlichkeit angeben, dass diese beiden zur gleichen Person gehören. In diesen Beispielen war das Programm in der Lage, mich selbst zu identifizieren – auf Fotos, die zehn Jahre auseinander liegen. Es hat auch meinen Kollegen Ben erfolgreich über Profilbilder von Facebook und Twitter hinweg gefunden und dabei korrekt festgestellt, dass er nicht Ben Stiller ist.

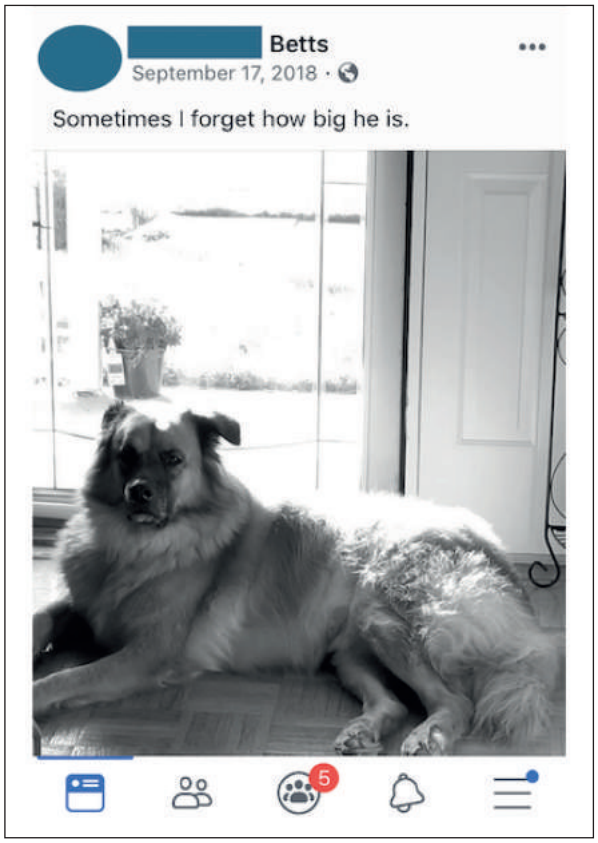


Der Dienst Face++ vergleicht zwei Gesichter, die man zuvor hochgeladen hat. Anschließend wird eine Wahrscheinlichkeit genannt, ob es sich um ein und dieselbe Person handelt oder eher nicht.

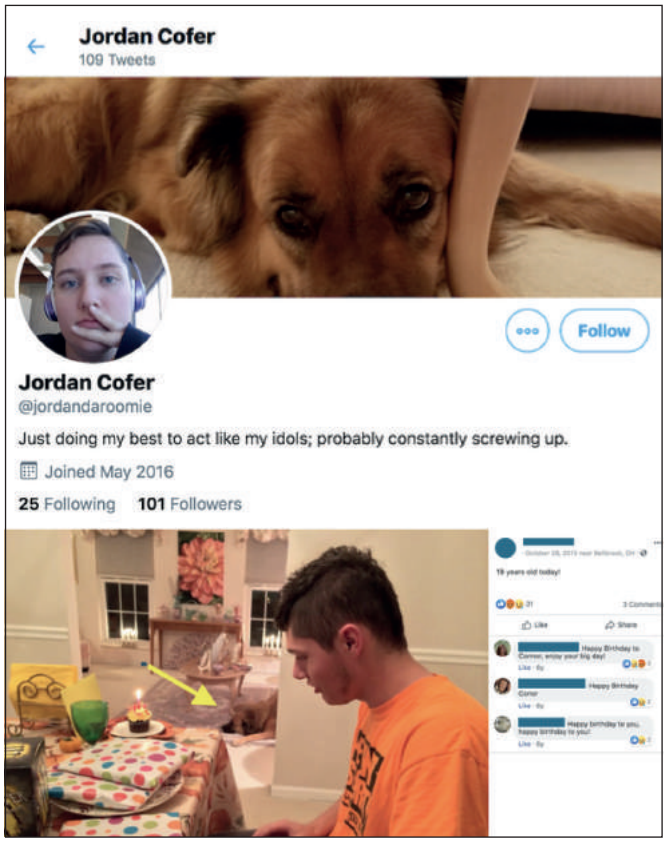
Wenn Sie auf der Suche nach Trollen oder Betrügerinnen und Betrügern sind, dann werden Sie womöglich feststellen, dass diese sich durchaus Mühe geben, ihr Bild zu verbergen oder gefälschte Profilbilder zu benutzen. Hier kann es helfen, das Profilbild zu bearbeiten, zu beschneiden oder zu spiegeln, um den Prozess wieder umzudrehen.

Nicht nur Profilbilder können Wegweiser sein. Auch wenn die Menschen sich zunehmend Gedanken über ihre Privatsphäre und die ihrer Familie machen, neigen sie immer noch dazu, Fotos von Dingen zu teilen, auf die sie stolz sind. Ich konnte Menschen bereits durch die Verbindung zu Bildern von Autos, Häusern oder Tieren identifizieren. In diesem Sinne sind Fotos ein Mittel geworden, um Accounts und die hinter ihnen stehenden Menschen miteinander zu verbinden und so ein Netzwerk um das Ziel herum aufzuzeigen.

Für die Recherche zu Social-Media-Profilen sind das Grundlagen. Wir wollten zum Beispiel die Profile eines Mannes bestätigen, der neun Menschen vor einer Bar in Dayton, Ohio erschossen hatte. Sein Twitter-Konto gab erste Anhaltspunkte bezüglich seiner Ideologie, aber sein Twitter-Benutzername, @iamthespookster, war eigenwillig und hatte keinerlei Ähnlichkeit mit dem bürgerlichen Namen, den die Behörden veröffentlicht hatten. Die Tatsache, dass eines seiner Opfer sein Bruder war, ein transsexueller Mann, dessen Name in den öffentlich bekanntgegebenen Informationen nicht enthalten war und auch nicht anderweitig kursierte, machte es zusätzlich kompliziert, die Schlüsselfiguren zu identifizieren. Aber sowohl durch sein Profil als auch die Profile seiner Familienangehörigen zogen sich Bilder eines Hundes – jenes Tieres, das sein transsexueller Bruder als Bannerbild seines bislang unbekannten Kontos genutzt hatte.



Screenshot aus einem Facebook-Profil, das zu dem Mann gehören soll, der in Dayton neun Menschen erschossen hatte. Der Hund auf dem Foto tauchte auch in den Profilen von Familienangehörigen auf.



Twitter- und Facebook-Profil seines Bruders, den der Amokläufer in Dayton mit acht anderen Menschen erschossen hatte. Darauf ebenfalls zu sehen: besagter Hund.

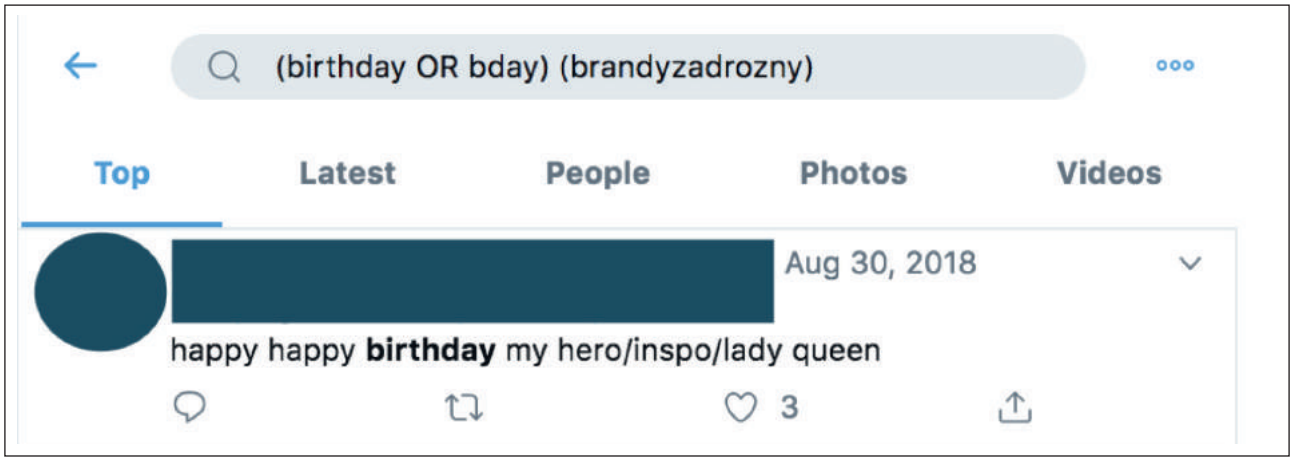
Der Hund war nicht das einzige hilfreiche Detail in diesem Bild hier. Es stammte vom Vater des Amokläufers und half uns dabei, sein persönliches Konto wie auch die Konten von Familienangehörigen zu verifizieren.

Wenn Sie ein Konto bei Facebook oder Twitter haben, dann kann man womöglich ihren Geburtstag herausfinden, auch dann, wenn sie ihn weder veröffentlicht noch etwas dazu geschrieben haben. Bei Eilmeldungen und Breaking-News-Situationen ist das Alter oft eines der ersten Puzzleteile, das von Behörden bekanntgegeben wird – und es ist eine zuverlässige Möglichkeit, ein Social-Media-Konto zu überprüfen: Man muss sich lediglich alle Beiträge ansehen bis zu dem Tag oder

Ein Tweet der Sängerin und Schauspielerin Dolly Parton, mit dem sie Mick Jagger von den Rolling Stones zum Geburtstag gratuliert. Wer Jaggers Geburtstag anderswo nicht finden konnte, hat hier einen ziemlich glaubhaften Hinweis, da beide Profile auch „verifiziert" sind, also in Bezug auf den Inhaber von Twitter überprüft und bestätigt wurden (erkennbar am blauen Haken).

Ebenfalls ein Tweet von Schauspielerin Dolly Parton. Hier gratuliert sie dem Sänger und Musiker Willie Nelson und lässt uns damit wissen, wann dieser Geburtstag hat.

Monat, um den es geht, und dort nachschauen, ob sich Glückwünsche von anderen finden. Auch wenn das eigene Profil nichts dazu enthält, schreiben Mütter, Väter und Verwandte oft etwas am Geburtstag ihrer Kinder. Das ist in Twitter nicht anders, und wer freut sich nicht über Geburtstagswünsche?



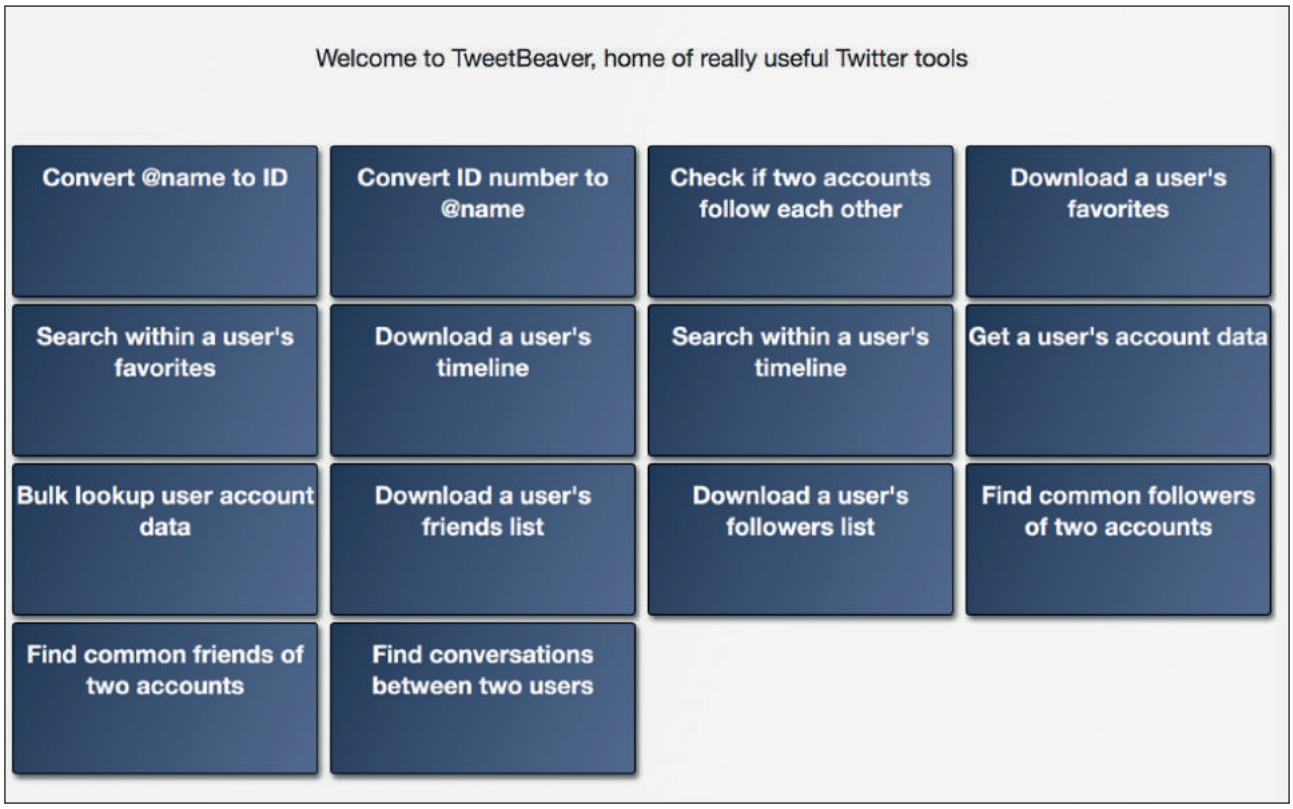
Geburtstage sind nur ein Beispiel. Hochzeiten, Beerdigungen, Feiertage, Jahrestage, Jubiläen, Abschlussfeiern – so ziemlich jeder wichtige Lebensabschnitt wird auch auf sozialen Netzwerken gefeiert. Das öffnet zahllose Möglichkeiten, ein Konto zu finden und zu durchsuchen. Man kann nach diesen Schlagwörtern in Kombination mit anderen Suchfiltern auch über spezielle Seiten suchen, die die Facebook-Suche anzapfen. Diese sind nicht mehr ganz so mächtig wie früher, weil Facebook sich irgendwann mehr um Privatsphäre zu sorgen begann, aber es gibt sie noch. Einer meiner Favoriten ist whopostedwhat.com.[13]

BEZIEHUNGEN

Man kann die Verbindungen und Beziehungen untersuchen, die Menschen in sozialen Netzwerken pflegen. Wir können sogar ziemlich viel über das Leben einer Person und ihre Neigungen erzählen, wenn wir uns anschauen, mit wem diese Person online interagiert. Als ich begonnen habe, Twitter zu nutzen, habe ich meinen Mann und meinen besten Freund überredet, sich dort auch anzumelden, so dass sie mir folgen konnten. Wenn ich mir heute beruflich Profile anschaue, muss ich oft daran denken. Auch die Plattformen wollen nicht, dass man sich dort allein fühlt, und so legen deren Algorithmen los, sobald man erstmals einen Account eröffnet. Beeinflusst von der Kontaktliste im Telefon, dem Auftauchen in den Kontaktlisten der Telefone anderer, die schon ein Profil haben, dem Standort und anderen Faktoren beginnen die Plattformen, Vorschläge für Profile zu unterbreiten, denen man folgen könnte. Und weil das so ist, ist es stets sehr erhellend, sich die allerersten Follower und Freunde eines Profils anzuschauen. TweetBeaver ist ein gutes Tool, um die Verbindungen zwischen großen Accounts zu untersuchen und um sich Sachen herunterzuladen, zum Beispiel die kompletten Timelines und Gefällt-mir-Markierungen von kleineren Profilen.[14] Für größere Datensätze ist man auf den Zugriff über eine Entwicklerschnittstelle angewiesen.

13 Viele erweiterte Funktionen von Facebook sind nur für US-Nutzer freigeschaltet. Es empfiehlt sich für Nutzer aus Deutschland, die alle Möglichkeiten der Suche ausschöpfen wollen, in den Einstellungen des eigenen Facebook-Profils den Standort auf „US-amerikanisch“ zu setzen.

14 Der Dienst bietet noch deutlich mehr Möglichkeiten: gemeinsame Freunde finden, die individuelle ID-Nummer hinter einem Profil finden, Unterhaltungen zwischen zwei Profilen finden oder die Profildaten einer großen Anzahl von Profilen auf einmal durchsuchen.



TweetBeaver eröffnet eine Vielzahl von Möglichkeiten: die ID-Nummer eines Profils herausfinden oder die Profilnummer hinter einer ID-Nummer prüfen, bis sich zwei Accounts gegenseitig folgen, die Favoritenliste eines Profils herunterladen, innerhalb der Favoriten suchen, die komplette Veröffentlichungshistorie eines Profils herunterladen, in der Historie nach etwas suchen, die Account-Daten eines Profils oder auch gleich einer Reihe von Profilen anschauen, Listen von Freunden oder Followern herunterladen, gemeinsame Follower oder Freunde zweier Profile finden oder schauen, welche Unterhaltung zwei Profile geführt haben.

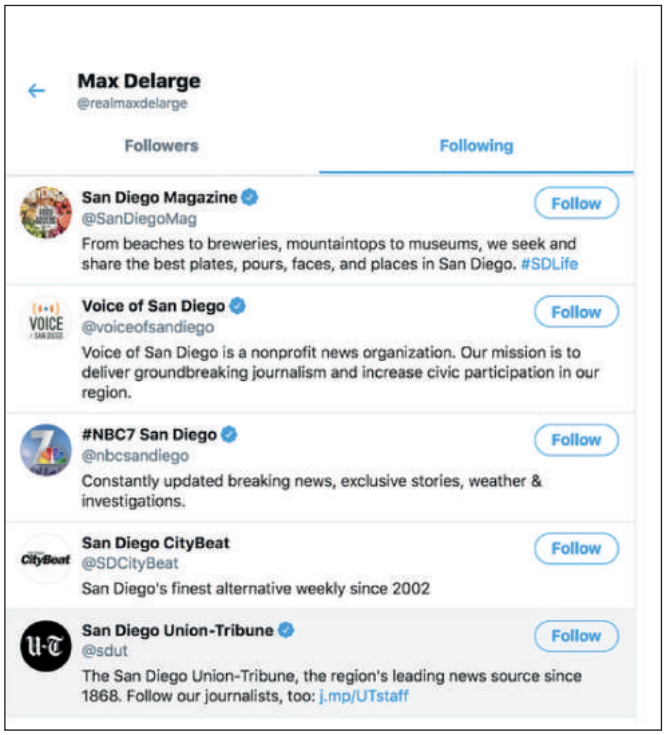
Schauen wir uns einen der beliebtesten rechten anonymen Twitter-Accounts an, angefeuert durch die Tatsache, dass Donald Trump dessen Inhalte bereits zweimal weiterverteilt hat: The Columbia Bugle.

Screenshot des Twitter-Profils von The Columbia Bugle.
Quelle: https://twitter.com/columbiabugle

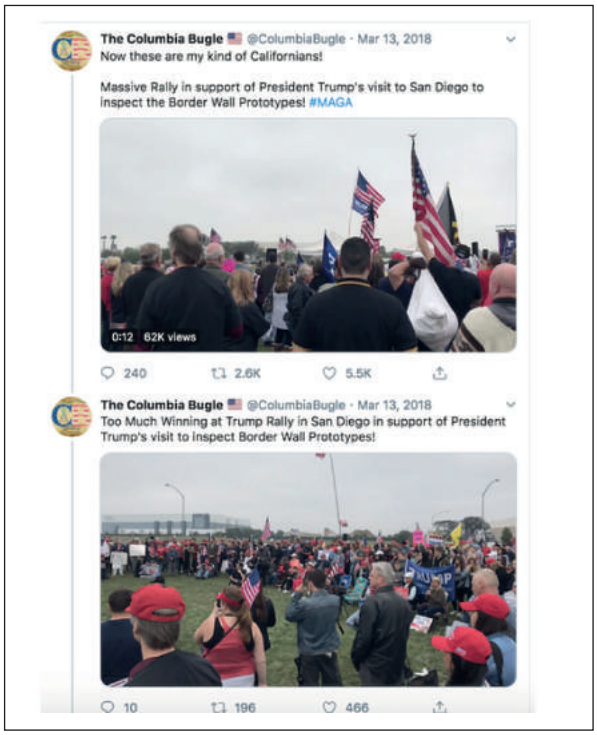
Die Selbstbeschreibung lautet: „Ehrliche und konservative politische Kommentare zu ‚America first'. Zweimal von @realDonaldTrump geteilt!", gefolgt von Hashtags mit der Aufforderung, die Grenzmauer zu bauen und „sie alle" zu deportieren.

TweetBeaver zeigt, dass das Profil von The Columbia Bugle und die Autorin dieses Kapitels zwölf gemeinsame „Freunde" auf Twitter haben.

Die Profile, denen Max Delarge als Erstes folgte, ein Profil, das behauptet, der Herausgeber von The Columbia Bugle zu sein, sind San Diego-spezifische Nachrichtenseiten und San Diego-spezifische Sport-Accounts. Da viele der Tweets von The Columbia Bugle Videos von Wahlkampfveranstaltungen für Donald Trump in San Diego sowie Veranstaltungen an der University of California in San Diego sind, können wir einigermaßen sicher sein, dass die Person hinter dem Bericht in der Nähe von San Diego lebt.

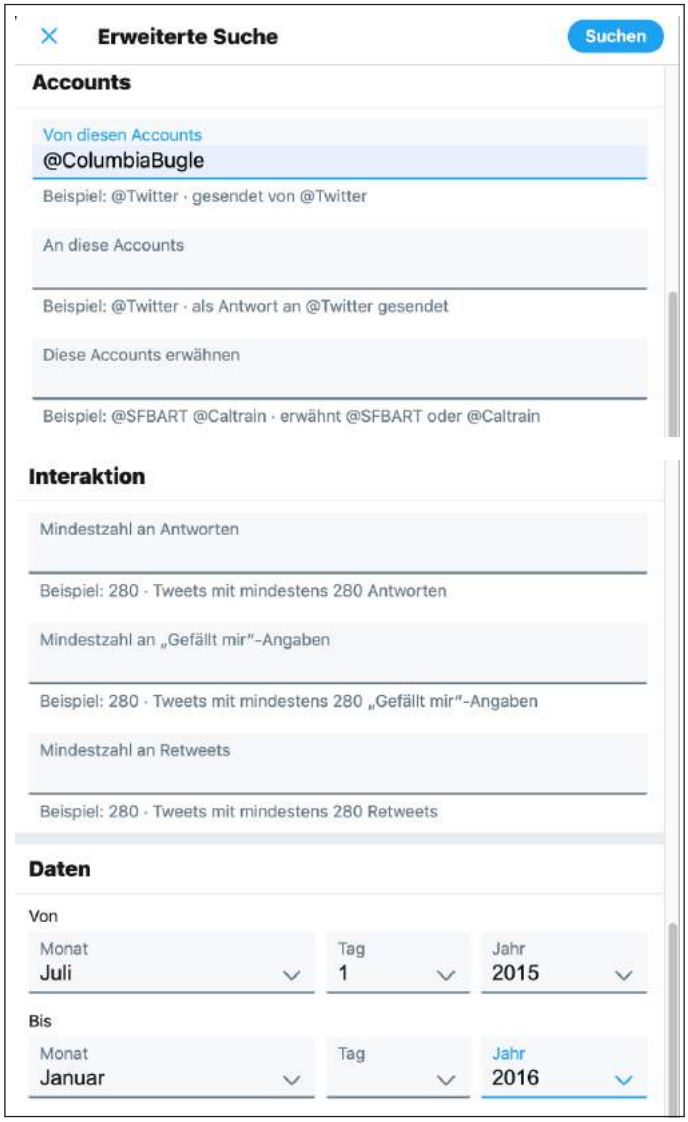


Ein Klick auf die Übersicht der Profile, denen Max Delarge folgt, listet viele Nachrichtenseiten aus San Diego auf.
Quelle: https://twitter.com/realMaxDelarge/following



Das Profil von The Columbia Bugle verbreitet Tweets, die sich auf Veranstaltungen aus San Diego beziehen.

Bei einer neuen Recherche beginne ich gern am Beginn der Twitter-Historie eines Profils und arbeite mich dann auf dem Zeitstrahl nach vorn. Das geht per Hand, mit einer Chrome-Erweiterung, die automatisch scrollt, oder mit Hilfe der erweiterten Suche von Twitter, indem man dort den Zeitrahmen der anzuzeigenden Tweets auf die ersten paar Monate eines Accounts beschränkt.

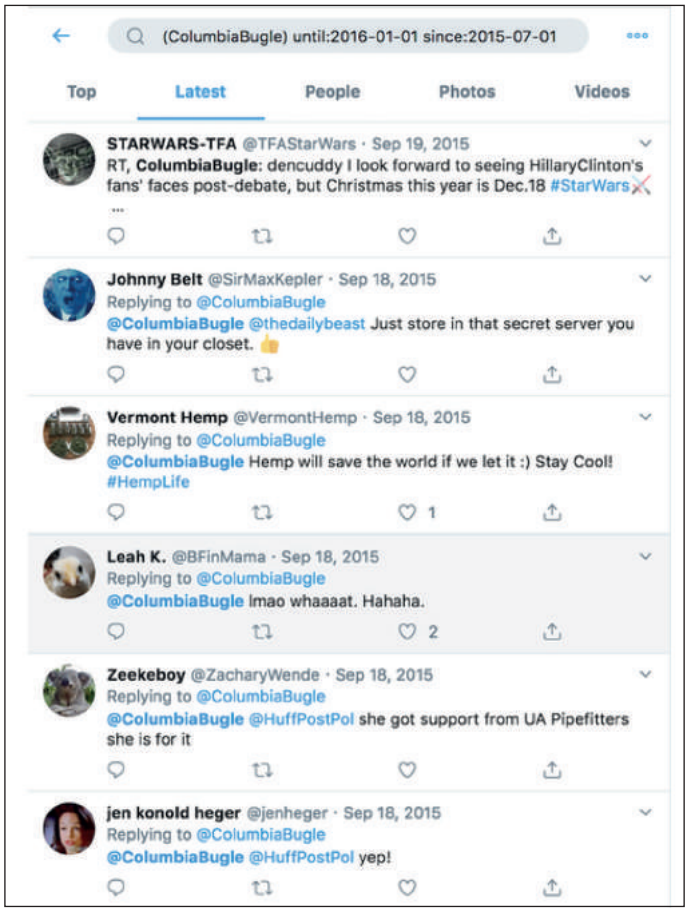


Die einfache Twitter-Suche bietet nicht viele Möglichkeiten, die erweiterte Twitter-Suche hingegen ist ein wertvolles Werkzeug mit vielen Optionen. Sie ist über https://twitter.com/search-advanced zu finden.

Interessanterweise zeigen die ersten sechs Monate dieses Accounts null Tweets.

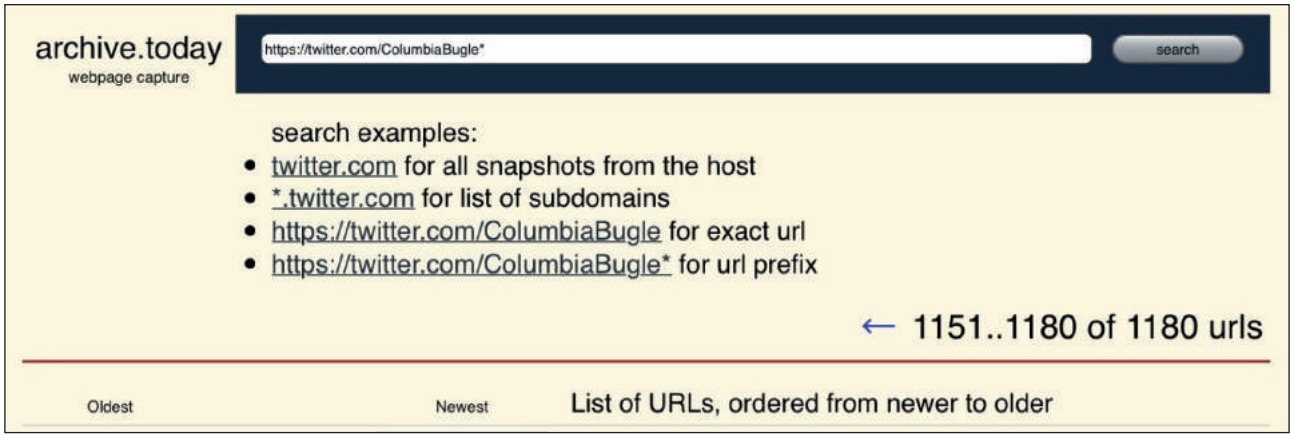
Anzeige des Suchergebnisses der erweiterten Twitter-Suche aus der vorhergehenden Grafik. Die Suchoperatoren im Suchfeld können frei benutzt werden, auch ohne über die erweiterte Suche zu gehen. Eine Übersicht dazu gibt es hier: https://www.mr-gadget.de/allgemein/2012-11-12/tutorial-suchen-in-twitter-such-operatoren-und-syntax

Das könnte darauf hindeuten, dass die Person hinter The Columbia Bugle ältere Tweets gelöscht hat. Um herauszufinden, warum das so sein könnte, müssen wir die Suche verfeinern. Statt nach Tweets zu suchen, die von dem Profil kommen, suchen wir nach Tweets, die das Profil selbst erwähnen.



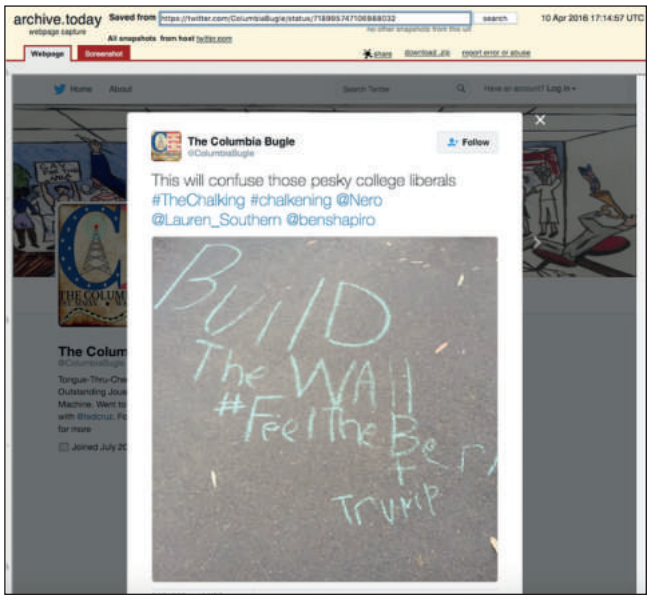
Diese Unterhaltungen bestätigen unseren Verdacht, dass The Columbia Bugle seine Tweets aus dem ersten Jahr gelöscht hat, sagen uns aber noch nichts darüber, warum das geschehen ist, und auch die ersten Profile, die mit dem Konto interagierten, geben dazu keine hilfreichen Tipps.

Um kürzlich gelöschte Tweets zu finden, kann man den Cache von Google durchsuchen; ältere gelöschte Tweets können manchmal über die „Wayback Machine" des Internet Archive gefunden werden.[15] Eine weitere Archivierungsseite ist archive.is, und über diese waren zahlreiche gelöschte Tweets zu finden, wonach The Columbia Bugle an einer Veranstaltung teilnahm, bei der College-Studenten Pro-Trump-Botschaften verbreiteten. Um alle Tweets zu sehen, die jemand von einem bestimmten Konto archiviert hat – so wie ich es getan habe, um diesen Tweet zu finden –, kann man nach einer vollständigen URL suchen und hinter den Namen des Kontos ein Sternchen anfügen:



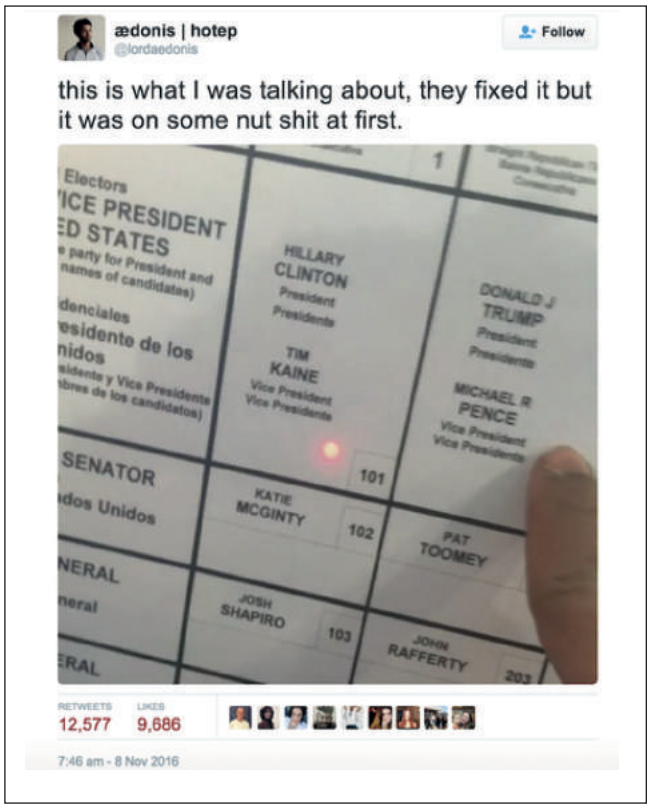
Bei archive.is lassen sich ältere Inhalte, die nicht mehr im Internet verfügbar sind, wiederfinden – wenn diese vorher dort archiviert wurden.

15 Es handelt sich bei archive.org und archive.is um kostenfreie Archivierungsdienste, mit denen Abbilder von Websites gespeichert werden können. Ist beabsichtigt, einen im Internet verfügbaren Inhalt zu sichern und auch nach einer eventuellen Löschung noch vorzufinden, empfiehlt es sich, diesen Link manuell hier zu sichern.



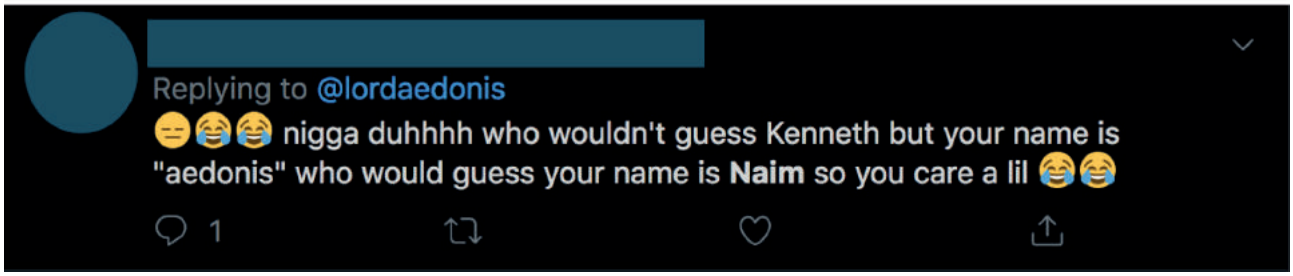
So fand archive.is unter anderem auch diesen mittlerweile nicht mehr online verfügbaren Tweet wieder.

Nur selten gelingt es, erfolgreich das wirkliche Leben von allen Online-Aktivitäten getrennt zu halten. Ein Beispiel, um das zu illustrieren: Gemeinsam mit einem Kollegen bei NBC News recherchierte ich die Geschichte hinter dem viralsten – und irreführendsten – Gerücht über einen Wahlbetrug im Jahr 2016, und das mit ein klein wenig Unterstützung aus der Nachbarschaft des rechtsextremen Trolls, der sie verbreitet hatte.



Hinter diesem Tweet verbarg sich eines der viralsten Gerüchte über einen angeblichen Wahlbetrug im Jahr 2016. Der Twitter-Nutzer hatte zuvor behauptet, die Wahlmaschine habe seine Stimme weiterhin für Hillary Clinton gezählt, obwohl er mehrfach auf Donald Trump gedrückt hätte. Tatsächlich war der Nutzer der Anleitung zum Ändern einer schon eingegebenen Stimme nicht gefolgt.

Obwohl der ursprüngliche Tweet von einem Mann stammte, der seinen Followern nur als @lordaedonis bekannt ist, hatten Menschen aus seiner echten Nachbarschaft auf frühere Tweets mit seinem Klarnamen geantwortet, so dass bei uns das Bild eines aufmerksamkeitshungrigen Unternehmers entstand, dessen Tweet dann von einem aus dem Kreml unterstützten Twitter-Profil geteilt, vom kommenden US-Präsidenten beworben und schließlich von Millionen gesehen wurde.

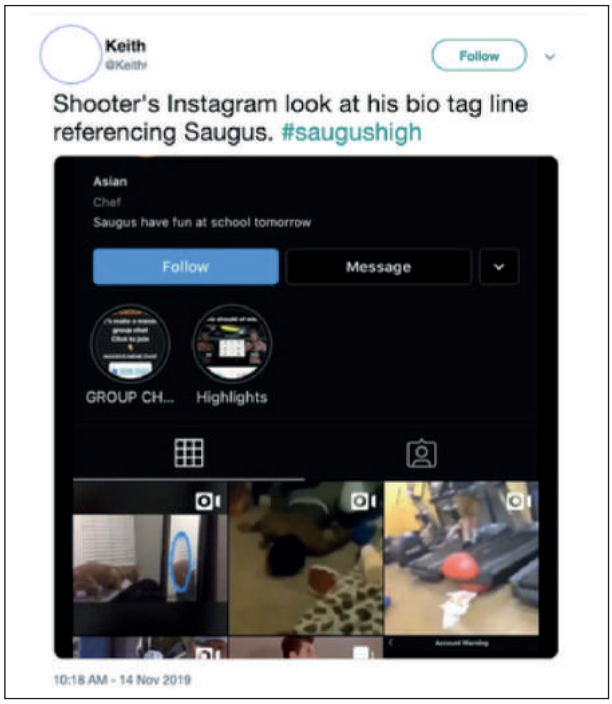


Ein Tweet eines Nutzers, der @lordaedonis unter seinem bürgerlichen Namen zu kennen scheint. Er schreibt sinngemäß, man würde bei der Betrachtung des Profils nicht ahnen, dass sein Name Naim sei.

Am liebsten mache ich Geschichten, die die wahre Identität hinter einflussreichen, anonymen Social-Media-Profilen offenlegen. Diese geheimen Accounts verlassen sich weniger auf den Algorithmus, und jemand hat sich mehr Mühe gegeben, damit der Öffentlichkeit zu entfliehen. Sie erlauben es demjenigen, mit einem privaten Profil weiterhin mit Freunden und Familie in Kontakt zu bleiben und mit einem öffentlichen Profil jene Ideen und Ansichten zu kommunizieren, die man sich aus persönlichen oder politischen Gründen nicht laut zu sagen traut. Die Journalistin Ashley Feinberg kann man mit Fug und Recht die Mutter für diese Art von Geschichten nennen. Geschichten, die die geheimen zwei Accounts von öffentlichen Personen wie James Comey oder Mitt Romney entlarven. Ihr Geheimrezept war, kleinere Konten von Familienangehörigen zu finden, denen Comey und Romney natürlich folgten, und dann durch diese Profile zu stöbern, um zu schauen, ob einem dort ein unbekannter Account ins Auge sticht, der vielleicht erstmal unecht wirkt, aber dessen Beiträge und dessen Netzwerk aus Freunden und Followern zu denen des bestätigten „Promi"-Profils passten.

VORSICHT BEI FAKE-ACCOUNTS

Jede Plattform hat ihren eigenen Charakter, ihre eigenen Suchmöglichkeiten, und jede Plattform macht je nach Situation mehr oder weniger Sinn. Aber auf jeder gilt: Vertrauen ist gut, Kontrolle ist besser. Ganze Gruppen erfreuen sich daran, Journalistinnen und Journalisten zu täuschen. Vor allem in Situationen mit Eilmeldungen entstehen falsche Accounts, viele mit zweifelhaften oder bedrohlichen Beiträgen, die nichts anderes tun sollen, als Reporterinnen und Reporter anzuziehen. Dieser Instagram-Account zum Beispiel benutzt den Namen eines Amokläufers, wurde aber nach der Schießerei in der Saugus High School in Kalifornien erstellt. Er erhielt über Screenshots viel Aufmerksamkeit auf Twitter; BuzzFeed News enthüllte später, dass er gar nicht dem Schützen gehörte.



Ein Tweet verbreitet das Gerücht, dass das gezeigte Instagram-Profil zum Amokläufer der Saugus High School gehören soll. Er verweist auf den dortigen Profiltext, in dem es heißt: „Saugus, viel Spaß in der Schule morgen."

Ein Profil über die Person selbst, die Familie, Freunde, Strafverfolgungsbehörden und soziale Medien zu verifizieren, kann einen davor bewahren, getäuscht zu werden. Und zu guter Letzt, und das ist vielleicht der wichtigste Hinweis: Es gibt nicht die eine richtige Reihenfolge, in der solche Schritte abgearbeitet werden müssen. Wie oft falle ich in ein Fass ohne Boden und habe viel mehr Tabs offen, als mir lieb ist. Hier hilft nur, sich selbst ein System zu überlegen und jeden einzelnen Schritt aufzuzeichnen: Ob als Google-Dokument oder über bezahlte Dienste wie Hunchly, die das eigene Stöbern protokollieren – das ist der Schlüssel, um Verbindungen zwischen Menschen und den Doppelleben, die sie online leben, zu erkennen und aus den Schlussfolgerungen dann Geschichten zu machen.

1a. Fallbeispiel: Wie wir über eine Gruppe von Facebook-Konten den Versuch entdeckten, Propaganda auf den Philippinen zu verbreiten

von: Vernise Tantuco und Gemma Bagayaua-Mendoza
deutsche Bearbeitung: Marcus Engert

*Seit gut 20 Jahren ist **Gemma Bagayaua-Mendoza** Journalistin. Bei Rappler, einer philippinischen Nachrichtenwebsite, ist sie für Recherche und Strategie zuständig und leitet dort die Faktenprüfung und Recherchen zu Desinformation und Falschinformation im Netz.*
***Vernise Tantuco** ist Mitglied des Rechercheteams von Rappler, wo sie an Faktenchecks und der Untersuchung von Desinformationsnetzwerken auf den Philippinen arbeitet.*

Im Herbst 2016 schickte John Victorina, ein Investment-Analyst, Rappler eine Liste von, wie er sagte, 26 verdächtigen philippinischen Facebook-Profilen. Wir begannen, die Konten zu untersuchen und zu beobachten und stellten schnell fest, dass die persönlichen Details in den Profilen falsch waren. Im Laufe wochenlanger Recherchen führten uns diese 26 Konten schließlich zu einem viel umfangreicheren Netzwerk von Seiten, Gruppen und Konten.

Diese Accounts sowie einige Seiten und Gruppen, mit denen sie verbunden waren, wurden schließlich von Facebook gelöscht. Sie haben uns auch inspiriert, Sharktank zu bauen, ein Tool, um zu beobachten, wie sich Informationen durch Facebook verbreiten. Diese Arbeit wurde zur Grundlage für eine ganze Serie von investigativen Geschichten darüber, wie Propaganda- und Informationsoperationen auf Facebook die Demokratie in den Philippinen beeinflussen. Die Serie drehte sich um die Recherche zu den Aktivitäten dieser 26 Konten und sie war der Auftakt unserer anhaltenden Berichterstattung darüber, wie Facebook auf den Philippinen zum Werkzeug wurde: für die Verbreitung von politischer Desinformation, für die Belästigung von Menschen und letztlich mit dem Ziel, die Demokratie zu untergraben. Dieses Fallbeispiel soll zeigen, wie wir diese 26 Konten analysierten und damit das dahinterstehende größere Netzwerk aufdeckten.

IDENTITÄTEN VERIFIZIEREN, „HANDPUPPEN" ENTLARVEN

Der erste Schritt bei der Untersuchung der Profile bestand für uns darin, herauszufinden, ob diese Verbindungen zu echten Menschen hatten. Dieser Teil beinhaltete ganz altmodisch eine Faktenprüfung und begann mit der Erstellung einer Tabelle, in der wir Details zu den einzelnen Accounts festhielten, zum Beispiel die persönlichen Informationen, die im Profil angegeben waren, die Seiten, die mit „Gefällt mir" markiert wurden und andere Informationen. Die Facebook-Nutzerin Mutya Bautista beispielsweise hatte angegeben, als Software Analyst bei ABS-CBN zu arbeiten, der größten Fernsehsenderkette der Philippinen. Rappler hatte bei ABS-CBN nachgefragt, die dem widersprachen und uns wissen ließen, dass sie **nicht** dort arbeitet.

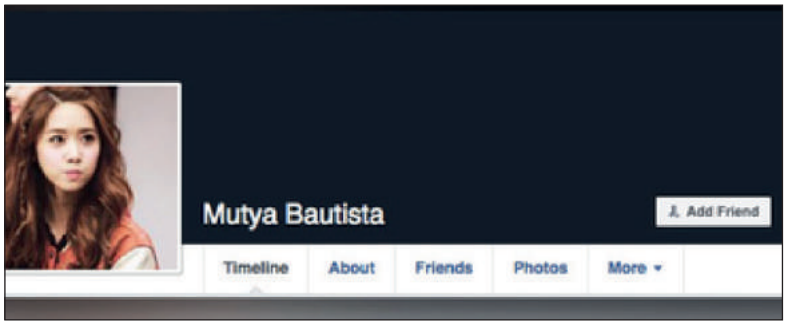
Mit Tools zur Bilder-Rückwärtssuche fanden wir heraus, dass viele der 26 Konten für ihre Profilbilder Fotos von bekannten Personen benutzten. Bautista beispielsweise nutzte ein Foto von Im Yoona von der koreanischen Pop-Band „Girl's Generation". Ein anderes Profil unter dem Namen Lily Lopez nutzte ein Bild der koreanischen Schauspielerin Kim Sa-rang.

Personal Information	
Facebook ID	https://www.facebook.com/profile.php?id=10
Profile Name	Mutya Bautista
Occupation	Software Analyst
Current Company	ABS-CBN Corporation
Former Occupation 1	
Former Occupation 2	
Former Occupation 3	
Former Occupation 4	
Former Occupation 5	
Studied	Computer Engineering
Studied at	University of the Philippines
Went to	
Lives in	
Married to	
From	
Account Set-up Date	October 19, 2015

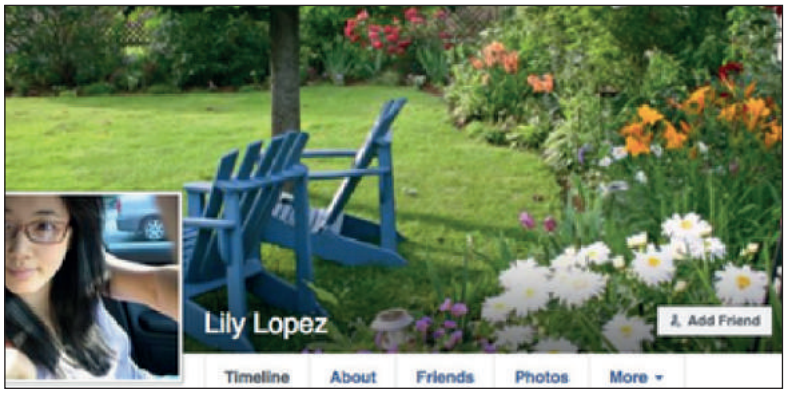
Liked Pages	Liked Pages Facebook ID
Okay Dito	https://www.facebook.com/vidtimestories/
The Philippine Pride	https://www.facebook.com/sirangplaka2/

Photos	Source of Photo
Profile Photo	Numerous sources. Im Yoona of SNSD
Cover Photo	

Rappler begann die Recherche mit dem Erfassen von Informationen zu den fraglichen Facebook-Konten in einer Tabelle.



Auch wenn der Name Mutya Bautista ist, das Bild gehört zur koreanischen Musikerin Im Yoona.

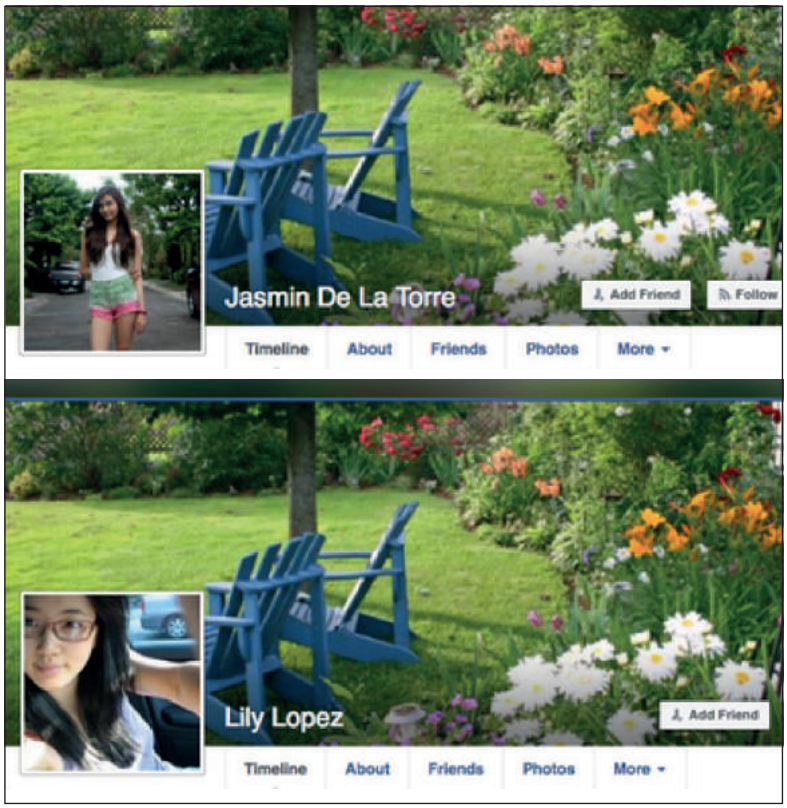


Auch das Profilbild dieses Accounts entstammt einer anderen Quelle.

Wieder ein anderes Konto mit dem Namen Luvimin Cancio hatte ein Profilbild von softcorecams.com, einer Porno-Website. Wir konnten diese Seite als Quelle des Bildes über die Bilder-Suchmaschine TinEye nachweisen.

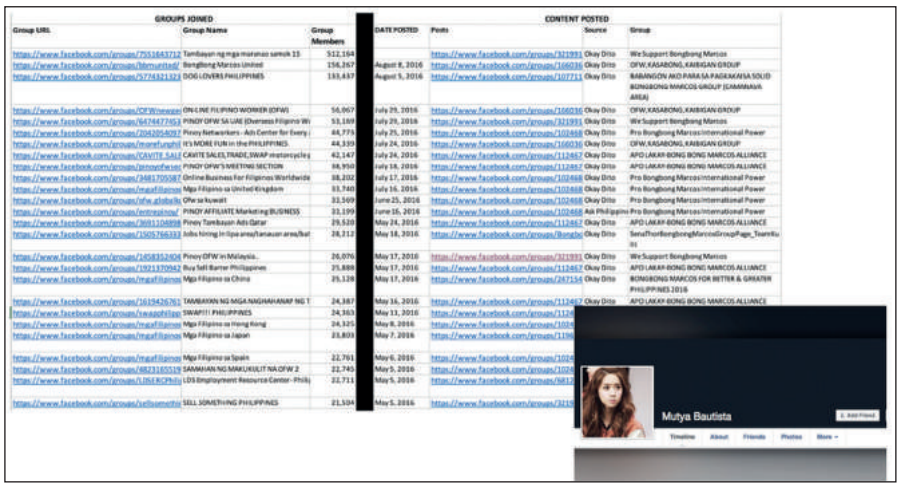
Auch dieses Facebook-Profil arbeitet mit fremden Bildern.

Außerdem nutzten manche der Profile ähnliche oder gleiche Hintergrundbilder. In diesem Beispiel hat das Profil von Jasmin De La Torre das gleiche Foto verwendet wie die angebliche Lily Lopez.



Zwei unterschiedliche Menschen, angeblich, doch beide mit dem gleichen Hintergrundbild im Profil.

Noch etwas Merkwürdiges fiel uns zu den 26 Profilen auf: Die Nutzer waren Mitglieder in mehr Gruppen, als sie Freunde hatten. Das war ungewöhnlich, denn auf den Philippinen haben die meisten Menschen Freunde und Familie im Ausland. Facebook ist hier oft der Kommunikationskanal, über den sie miteinander in Verbindung bleiben. Darum haben die Menschen üblicherweise ziemlich viele Freunde, sind aber in nicht so vielen Gruppen. Die Freundesliste von Bautista, die zu diesem Zeitpunkt öffentlich war, zeigte, dass sie nur 17 Freunde hatte. Tatsächlich hatte sogar jedes der 26 Profile, das wir uns anschauten, weniger als 50 Freunde, als wir sie 2016 entdeckten. Bautista allerdings war in mehr als 100 Gruppen, darunter welche für den Wahlkampf des damaligen Vizepräsidentschaftskandidaten Ferdinand Marcos Jr., für Filipinos im Ausland, für private Kleinanzeigen, allesamt mit zehntausenden bis hunderttausenden Mitgliedern. Insgesamt hatten diese Gruppen 2,3 Millionen Mitglieder auf Facebook. Die Grafik zeigt eine Liste der größten Gruppen sowie eine Liste von Beiträgen, die Bautista in diesen Gruppen erstellt hat.



Erfassen von Informationen zu den fraglichen Facebook-Konten in einer Tabelle.

Indem wir alle diese Beobachtungen und damit verbundenen Daten miteinander kombinierten, kamen wir zu dem Schluss, dass es sich bei den Profilen um Handpuppen handeln muss: erfundene Identitäten, kreiert, um bestimmte Positionen und Standpunkte öffentlich zu bestärken.

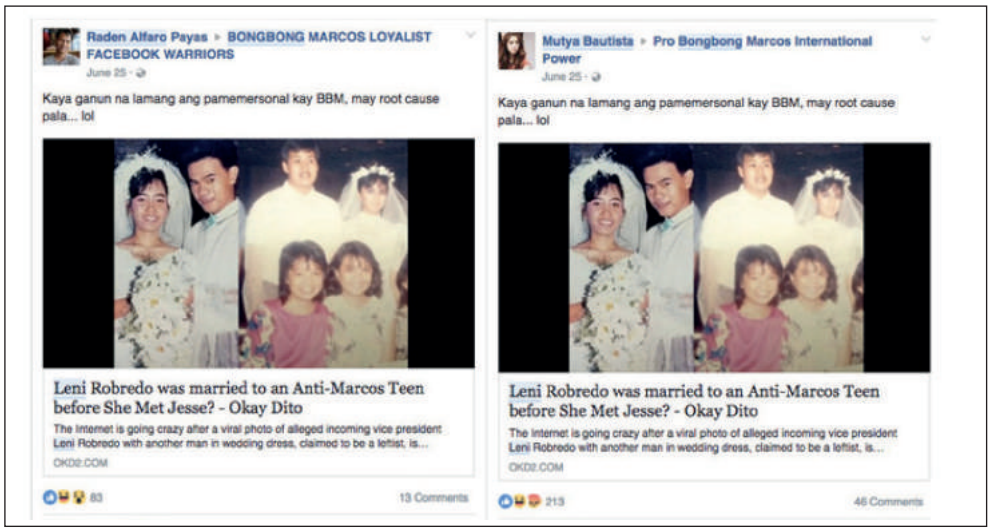
PRO-MARCOS-NETZWERK

Anhand der Datumsstempel der ersten Profilbilder und der ersten Beiträge konnten wir erkennen, dass die 26 Profile im ersten Quartal 2015 erstellt wurden und damit **vor** den Wahlen im Mai 2016. Wir stellten außerdem fest, dass sie kontinuierlich Inhalte verbreiteten, die den weithin dokumentierten Missbrauch des Kriegsrechts im Marco-Regime der 1970er-Jahre leugneten.

Die Profile attackierten auch Konkurrenten des Sohnes des ehemaligen Diktators, Vizepräsidentschaftskandidat Ferdinand „Bongbong" Marcos Jr.

Im folgenden Beispiel teilte die Nutzerin Mutya Bautista ein zwischenzeitlich als Falschmeldung entlarvtes Gerücht, wonach Bongbongs Konkurrentin — die damals frisch gekürte Vizepräsidentin Leni Robredo – vor der Ehe mit ihrem zweiten Ehemann, dem verstorbenen Innenminister Jesse Robredo, mit einem Aktivisten verheiratet gewesen sein soll. Bautista teilte die Meldung mit der Überschrift „Leni Robredo war mit einem Anti-Marcos-Teenager verheiratet, bevor sie Jesse traf?" in der Gruppe „Pro Bongbong Marcos International Power", und zwar mit dem Kommentar *„Kaya ganun na lamang ang pamemersonal kay [Bongbong Marcos], may root cause pala"* („Es [Bongbong Marcos] ist also eine persönliche Sache, und sie hat einen tieferen Grund").

Ein weiterer verdächtiger Account mit dem Namen Raden Alfaro Payas teilte den gleichen Beitrag in der gleichen Gruppe mit dem gleichen Kommentar – Wort für Wort, identisch bis zur Zeichensetzung – am gleichen Tag.



Angeblich zwei verschiedene Menschen, die den gleichen Beitrag mit komplett identischem Kommentar dazu teilten.

Solche falschen Profile werden oft genutzt, um Gruppen mit Links zu fluten, und manchmal kann man sie dabei ertappen, wie sie die exakt gleichen Texte benutzen, wenn sie das tun. Damals war es noch möglich, Facebooks „Graph Search" zu benutzen, um sich die öffentlichen Beiträge einer Gruppe genauer anzuschauen. Leider hat Facebook viele Funktionen von Graph Search 2019 abgestellt, diese Funktion inbegriffen. In der Folge ist es jetzt leider nötig geworden, direkt in die Gruppen zu gehen und dort zu suchen, um einen Eindruck davon zu bekommen, welche Nutzer dort was teilen.

VERBUNDENE WEBSITES

Indem wir analysierten, welche Inhalte die Profile teilten, waren wir in der Lage zu erkennen, dass alle 26 „Handpuppen"-Profile die gleichen Websites bewarben: Okay Dito (OKD2.com), Ask Philippines (askphilippines.com), why0why.com und andere.

OKD2.com hat ziemlich viele Falschmeldungen und anderes Propagandamaterial geteilt, und das zugunsten der Marcos-Familie und von Präsident Rodrigo Duterte. Heute tarnt sie sich als angebliche Kleinanzeigen-Website. Aber im September 2016 fanden wir heraus, dass Inhalte der Seite mehr als 11.900 Mal auf Facebook geteilt worden waren – auch dank der „Handpuppen". Über diese Websites fand Rappler schließlich den mutmaßlichen Kopf hinter den 26 „Handpuppen"-Profilen: jemanden namens Raden Alfaro Payas.

DEN PUPPENSPIELERN AUF DER SPUR

Wie bei vielen anderen Seiten, die Rappler beobachtet, sind auch die Registrierungsdaten der Domain OKD2.com auf privat gestellt. Auch findet sich auf der Seite selbst nichts zu Autoren oder einem Eigentümer, keine Kontaktinformationen, nur ein Kontaktformular. Über den (kostenpflichtigen) Dienst domaintools.com konnten wir sehen, dass OKD2.com im Juli 2015 auf einen Raden Payas aus Tanauan City, Batangas, registriert worden war.

Wir fanden außerdem heraus, dass OKD2.com die gleiche Google AdSense-ID benutzt wie andere Websites, die die 26 Profile ebenfalls teilten.[16] AdSense-ID findet man heraus, indem man sich den Quellcode einer Seite anzeigen lässt und dort dann nach „ca-pub" sucht – die dahinterstehende Zahlenkombination ist die gesuchte ID. Jedem Google AdSense-Account wird eine eigene Nummer gegeben, die mit „ca-pub" beginnt, und jede Seite, die zu einem solchen Account gehört, hat diese ID in ihrem Quellcode.

Passend zu den Registrierungsdaten sahen wir, dass eines der 26 Konten Raden Alfaro Payas (Unofficial) hieß. Wir fanden außerdem ein anderes Profil mit dem Nutzernamen realradenpayas, das mit einigen der „Handpuppen" interagierte. Zum Beispiel kommentierte dieses Profil einen Beitrag von Luvimin Cancio, der auf einen Text verwies, in dem unter dem Kriegsrecht begangene Gräueltaten unter Marcos geleugnet wurden. Das „echte" Payas-Konto behauptete, während der Jahre des Kriegsrechts im Gymnasium gewesen zu sein und „nie gehört" zu haben, dass jemand getötet oder gefoltert worden sein soll.

Facebook-Kommentar von einem der 26 Profile. Er schreibt unter anderem: „Ich war ein Jugendlicher, als das Kriegsrecht 1972 eingeführt wurde. (...) Es haben darunter nur jene gelitten, die in der Untergrundbewegung waren, aber nicht jene Bürger, die sich an die Gesetze gehalten haben."

16 Eine Adsense-ID ist ein individueller Code, den Google Werbetreibenden zuteilt. Wer über Google Werbung ausspielt und dafür entweder zahlt oder damit Einnahmen macht, mit dem rechnet Google über diesen individuellen Code ab. Ein Code kann auf mehreren Websites genutzt werden.

STARTSCHUSS FÜR DEN SHARKTANK

Diese 26 gefälschten Accounts und ihre Reichweite inspirierten uns bei Rappler, eine Datenbank namens Sharktank zu bauen und das Datensammeln auf öffentlichen Facebook-Gruppen und -Seiten zu automatisieren. Zum August 2019 beobachteten wir so rund 40.000 Seiten mit Millionen Followern.

Was als eine Untersuchung einer kleinen Gruppe von auffälligen Profilen begann, wurde zu einer fortdauernden Beobachtung eines ganzen Netzwerks tausender echter und unechter Profile, Gruppen und Seiten, die Desinformation und Propaganda verteilen, Politik verzerren und so die Demokratie einer ganzen Nation schwächen.

1b. Fallbeispiel: Wie wir herausfanden, dass die größte Black Lives Matter-Seite auf Facebook ein Fake war

von: Donie O'Sullivan
deutsche Bearbeitung: Marcus Engert

Donie O'Sullivan *ist Reporter bei CNN und recherchiert an der Schnittmenge von Technologie und Politik. Er ist Teil des CNN Business Teams und arbeitet eng mit der Investigativ-Einheit zusammen, um Online-Desinformationskampagnen zu beobachten und zu identifizieren, die die amerikanischen Wahlen im Visier haben.*

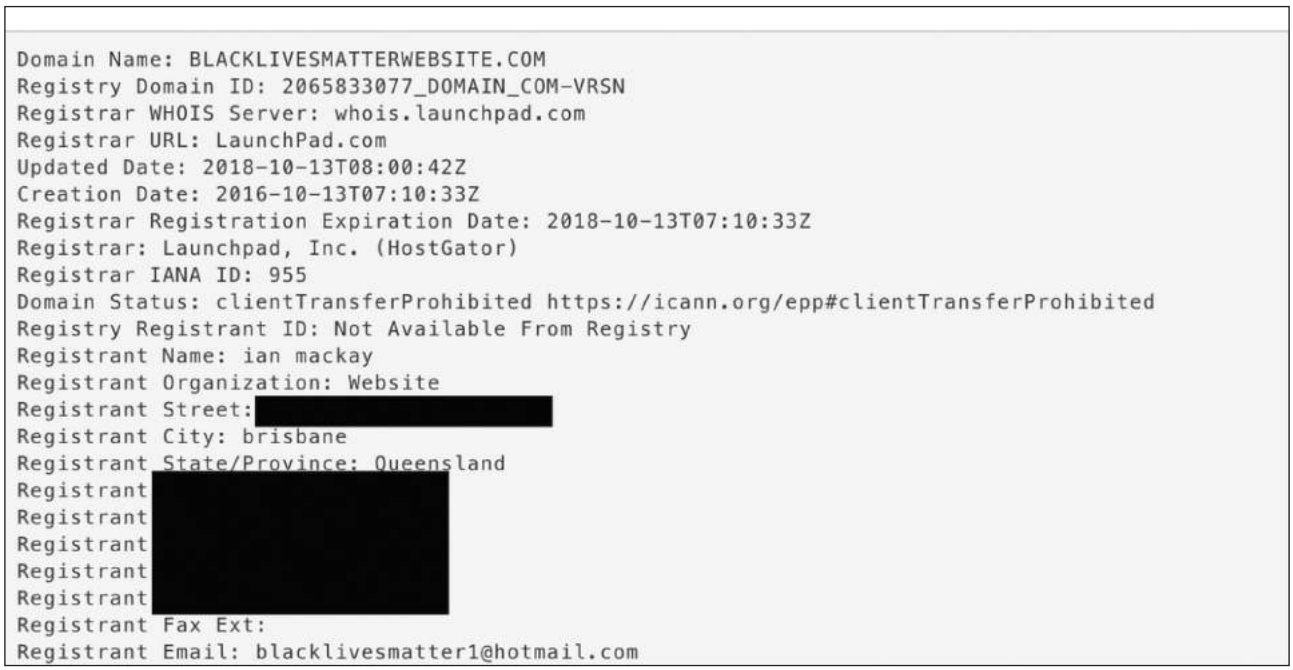
Im Sommer und Herbst 2017, als die Welt von Russlands umfangreichen Bemühungen erfuhr, die amerikanischen Wählerinnen und Wähler über soziale Medien zu beeinflussen, wurde langsam klar, dass Afroamerikaner und die Black Lives Matter-Bewegung zu den Hauptzielen der Kampagne aus dem Kreml gehörte, deren Ziel es war, Spaltung zu sähen. Meine Kollegen bei CNN und ich hatten Monate damit verbracht, darüber zu berichten, dass Russland hinter einigen der wichtigsten Black Lives Matter (BLM)-Profilen auf Facebook steckte. Wenn ich mit BLM-Aktivisten sprach, wurde ich manchmal gefragt: „Wissen Sie, wer hinter der größten Black Lives Matter-Seite auf Facebook steckt?"

Verrückterweise wusste niemand die Antwort – selbst die prominentesten BLM-Aktivisten nicht. Manche hatten nachvollziehbarerweise angenommen, die Seite könnte aus Russland betrieben werden. Aber unsere Recherchen hatten herausgefunden, dass es kein Russe und auch kein Amerikaner war – es war ein weißer Mann aus Australien. Die Seite selbst, mit dem einfachen Titel „Black Lives Matter", sah glaubhaft aus. Im April 2018 folgten ihr fast 700.000 Menschen. Sie teilte durchgehend Geschichten über Polizeigewalt und Ungerechtigkeit, sie machte Online-Spendensammlungen, sie hatte sogar einen Online-Shop, in dem man BLM-Produkte kaufen konnte.

Startseite der damals größten Facebook-Seite zur Black Lives Matter-Bewegung.
Rund 700.000 Menschen verfolgten die Seite.

Es ist gar nicht ungewöhnlich, dass Seiten von solcher Größe anonym gemacht werden. Manche Aktivisten wollen ihre Namen nicht auf eine große Seite schreiben und damit riskieren, die Aufmerksamkeit von Trollen oder von Strafverfolgungsbehörden zu bekommen, die versuchen, große Proteste zu beenden. Außerhalb der USA war es für digitalen Aktivismus und soziale Bewegungen mitunter entscheidend, dass Aktivisten hinter ihren Seiten anonym bleiben konnten. (Es war auch genau das, was Russland dann ausnutzte, und was nun den Verdacht gegenüber dieser BLM-Seite nährte.)

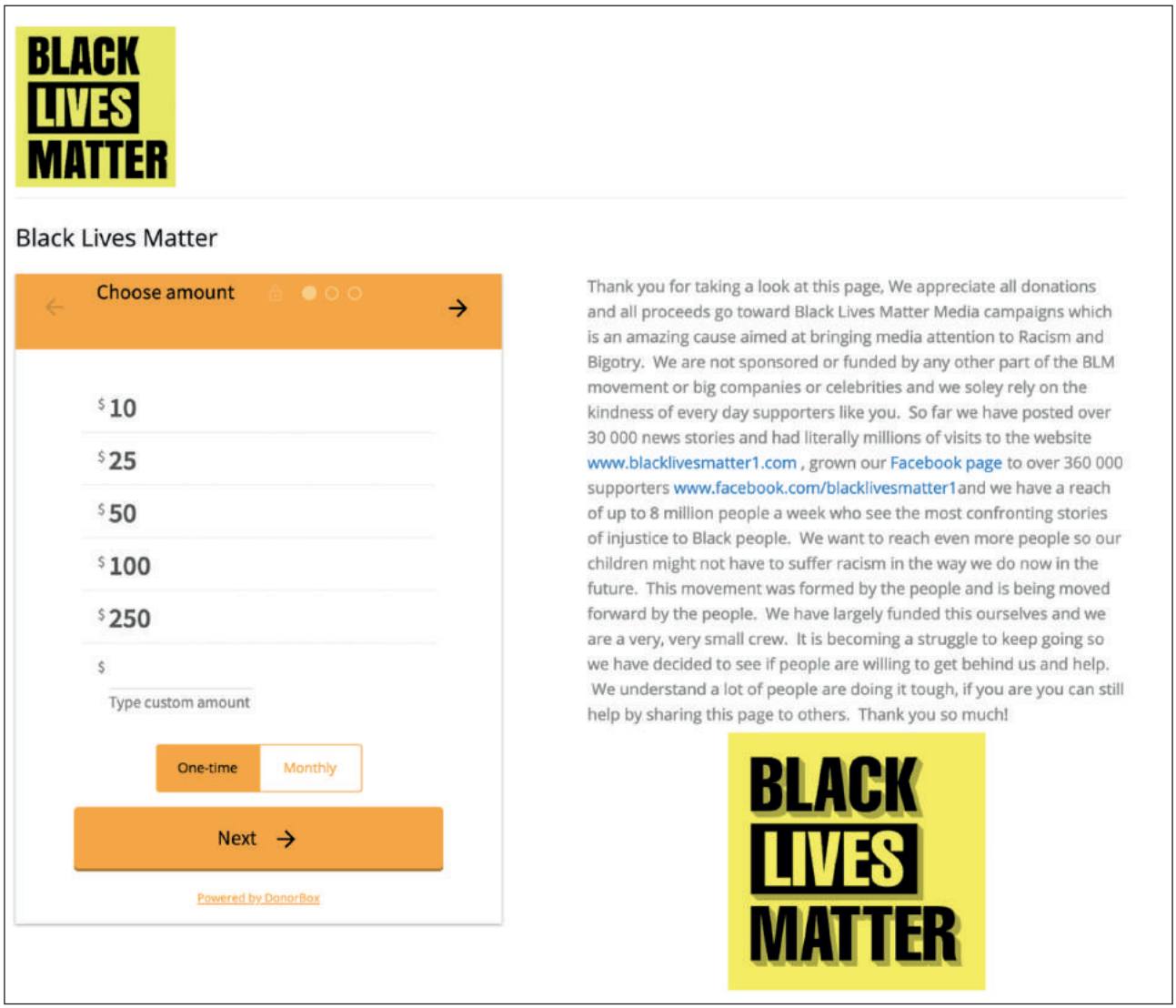
Etwa um diese Zeit herum, als ich begann, dieser mysteriösen Seite meine Aufmerksamkeit zu schenken, meldete sich Jeremy Massler, ein freiberuflicher Rechercheur und unglaublicher Online-Schnüffler, mit einem Tipp. Massler hatte sich die Domain-Registrierungsdaten der Websites angesehen, die die große BLM-Seite immer wieder verlinkte. Obwohl alle diese Seiten anonym registriert worden waren, fand er heraus, dass eine von ihnen für eine kurze Zeit in 2016 einer Person namens Ian MacKay in Perth, Australia, gehörte – einem weißen Mann.



```
Domain Name: BLACKLIVESMATTERWEBSITE.COM
Registry Domain ID: 2065833077_DOMAIN_COM-VRSN
Registrar WHOIS Server: whois.launchpad.com
Registrar URL: LaunchPad.com
Updated Date: 2018-10-13T08:00:42Z
Creation Date: 2016-10-13T07:10:33Z
Registrar Registration Expiration Date: 2018-10-13T07:10:33Z
Registrar: Launchpad, Inc. (HostGator)
Registrar IANA ID: 955
Domain Status: clientTransferProhibited https://icann.org/epp#clientTransferProhibited
Registry Registrant ID: Not Available From Registry
Registrant Name: ian mackay
Registrant Organization: Website
Registrant Street: [redacted]
Registrant City: brisbane
Registrant State/Province: Queensland
Registrant [redacted]
Registrant
Registrant
Registrant
Registrant
Registrant Fax Ext:
Registrant Email: blacklivesmatter1@hotmail.com
```

Über den kostenpflichtigen Online-Dienst domaintools.com lassen sich die Registrierungsdaten von Websites herausfinden, aktuelle wie auch solche, die schon Jahre zurückliegen.

Massler kontaktierte also den Mann namens MacKay, der ihm sagte, er kaufe und verkaufe Websites als ein Hobby und habe nichts zu tun mit der Facebook-Seite. Es war die gleiche Antwort, die MacKay, ein Gewerkschaftsfunktionär mittleren Alters, mir gab, als ich ihn ein paar Monate später per Telefon erreichte. Zu diesem Zeitpunkt aber hatten wir schon herausgefunden, dass MacKay Dutzende Websitenamen registriert hatte, viele davon mit einer Verbindung zur Black Lives Matter-Bewegung. Trotz meiner Skepsis bezüglich der Seite und bezüglich der Tatsache, dass mehrere Aktivisten sie verdächtig fanden, fand ich MacKays Erklärung nicht komplett abwegig. Internetadressen können wertvoll sein, und Menschen kaufen und verkaufen sie ständig. Der Punkt, dass er auch Seiten kaufte und verkaufte, die keine Verbindung zum BLM-Aktivismus aufwiesen, machte ihn da nur noch glaubwürdiger. Dann aber geschah etwas Eigenartiges. Ein paar Minuten, nachdem ich mit MacKay gesprochen hatte, ging die Seite offline. Sie war nicht gelöscht worden von Facebook, aber wer auch immer sie verwaltete, hatte sie temporär deaktiviert. Das wirkte verdächtig, also begannen Massler und ich, tiefer zu graben. Die Seite, die ein paar Wochen nach meinem Anruf wieder online kam, hatte in der Vergangenheit Spendensammlungen beworben, die angeblich für die BLM-Bewegung gedacht waren. In einem Fall hatte sie behauptet, Gelder für Aktivisten in Memphis, Tennessee, zu sammeln. Aber als ich mit den Aktivisten dort sprach, wusste niemand etwas davon oder darüber, wo das Geld hingeflossen sein konnte. Andere Aktivisten erzählten uns sogar, sie hätten sich schon bei Facebook gemeldet, weil sie glaubten, die Seite könne Betrug sein – doch Facebook habe nichts unternommen.



Die Facebook-Seite war auch sehr erfolgreich im Spendensammeln. Eine solche Aktion ist hier zu sehen. Im Text daneben steht, man wolle die Aufmerksamkeit auf die Belange der BLM-Bewegung richten, bekomme aber Sponsoring oder öffentliche Gelder und sei nur ein kleines Team. Weil die Seite von Menschen für Menschen gemacht sei, bitte man um Unterstützung.

Als ich die verschiedenen Zahlungsdienstleister und Spendenplattformen kontaktierte, die die Seite benutzt hatten, begannen diese, die Kampagnen zu entfernen und erklärten, sie hätten ihre Regularien verletzt. Mit Verweis auf den Datenschutz und die Privatsphäre ihrer Nutzer gab mir keine der Seiten zitierfähige Informationen darüber, wohin das Geld geflossen war. Eine Herausforderung, die wir oft erleben. Aus Gründen der Privatsphäre geben die Plattformen fast nie die Namen oder Kontaktdaten hinter einem Benutzerkonto an die Presse. Später erfuhr ich von einer Quelle, die mit einigen der abgewickelten Zahlungen vertraut war, dass mindestens eines der Konten zu einem australischen Bankkonto und einer australischen IP-Adresse gehörte. Eine andere Quelle berichtete mir, dass um die 100.000 Dollar gesammelt worden sein sollen.

Quellen innerhalb der wichtigen Internetfirmen zu erschließen, die bereit sind, mehr zu erzählen als das, was man von der Firma on the record bekommt, sind zunehmend wichtig geworden. Viele Geschichten können mit öffentlich verfügbaren Informationen allein nicht mehr aufgedeckt werden, da Betrüger und böswillige Akteure ebenfalls immer raffinierter geworden sind.

Ich habe Facebook mit diesen Informationen konfrontiert und Mitarbeitern gesagt, ich hätte Beweise, dass die Seite Verbindungen nach Australien habe, dass Zahlungsdienstleister die Kampagnen gelöscht hätten, nachdem sie sie untersucht hatten, und dass wir wüssten, dass manches von dem Geld nach Australien geflossen sei. Ein Facebook-Sprecher erklärte, die Untersuchungen der Plattform hätten „nichts ergeben, was unsere Gemeinschaftsregeln verletzt".

Erst ganz kurz, bevor wir die Story veröffentlichten — und erst, nachdem ich meine Bedenken über Facebooks Untersuchung und die Antwort des Sprechers mit einem höherrangigen Facebook-Mitarbeiter geteilt hatte –, handelte Facebook und entfernte die Seite. Die Gewerkschaft, bei der MacKay arbeitete, führte nach der CNN-Veröffentlichung eine eigene Untersuchung durch. Noch am Ende der gleichen Woche war MacKay entlassen, ebenso ein zweiter Mitarbeiter, der ebenfalls in den Betrug verwickelt gewesen sein soll.

Was an dieser Geschichte so bemerkenswert ist, ist die Vielzahl von Techniken, die Massler und ich angewandt hatten, um unser Ziel zu erreichen. Wir haben uns auf Archiv-Websites wie die Wayback Machine gestützt, die es uns ermöglichte, zu sehen, wie die Seite und andere Seiten, auf die sie verlinkte, aussahen, bevor wir sie entdeckten. Das war besonders hilfreich, denn nachdem Massler MacKay zum ersten Mal kontaktiert hatte, begannen die Leute hinter der Seite damit, ihre Spuren zu verwischen.

Um die Seiten zu untersuchen, die MacKay alle registriert hatte, und um seine privaten Kontaktdaten zu finden, benutzten wir auch Dienste, die die Registrierungsdaten von Websites speichern, wie zum Beispiel domaintools.com. Massler hatte außerdem ausgiebig mit Facebooks Graph Search gearbeitet (was leider nicht mehr verfügbar ist), um jene Fake-Profile auf Facebook zu untersuchen, mit denen in anderen Facebook-Gruppen für die Seite geworben wurde. Die Auswertung von öffentlichen oder halb öffentlichen Informationen und die Benutzung von Recherchewerkzeugen wie jenen, mit denen wir die Registrierungsdaten verglichen haben, sind zentrale Instrumente – aber sie sind nicht die einzigen. Der einfache Griff zum Telefon, um MacKay anzurufen, und Quellen, die Informationen preisgeben, die andernfalls nicht öffentlich geworden wären und die man sich vorher mit der Zeit erschlossen hat – beides ist ziemlich klassisches journalistisches Handwerk, und beides war hier elementar bei der Aufdeckung dieses Betrugs.

2. DEN PATIENTEN NULL FINDEN

von: Henk van Ess
deutsche Bearbeitung: Marcus Engert

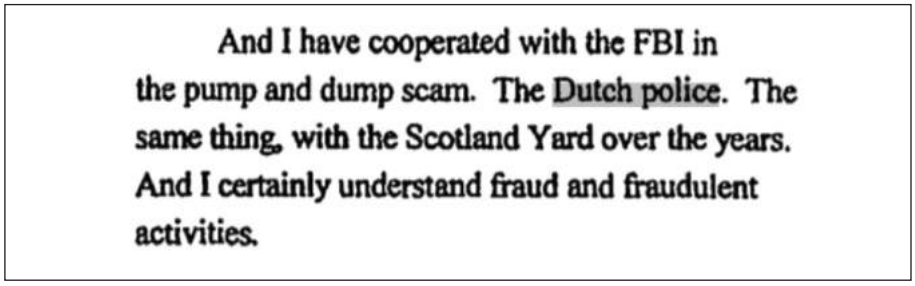
Henk van Ess *ist Gutachter im International Fact-Checking Network des renommierten Poynter Instituts in St. Petersburg, Florida. Er ist besessen davon, Geschichten in Daten zu finden. Van Ess bildet weltweit Medienprofis in der Recherche im Netz und in sozialen Netzwerken aus. Zu seinen Kunden gehören NBC News, BuzzFeed News, ITV, Global Witness, SRF, Axel Springer sowie zahlreiche Nichtregierungsorganisationen und Universitäten. Die vom ihm betriebenen Websites whopostedwhat.com und graph.tips zählen zu den meistgenutzten Werkzeugen, um soziale Netzwerke zu durchsuchen. Auf Twitter heißt er @henkvaness.*

Für Jahrzehnte galt der Flugbegleiter Gaëtan Dugas als „Patient Null": Als der erste Mann, der AIDS in die Vereinigten Staaten gebracht haben soll. Verstärkt durch Bücher, Filme und zahllose Artikel wurde er zur Personifizierung einer Epidemie, die allein in Nordamerika 700.000 Menschen getötet hat. Jedoch zu Unrecht. Bill Darrow, ein Beamter der amerikanischen Seuchenschutzbehörde, hatte Dugas befragt und als „Patient 0 außerhalb Kaliforniens" in die Unterlagen aufgenommen. Das wurde schon bald als Nummer null missverstanden und setzte eine Kettenreaktion von Falschinformationen in Gang, die erst vor kurzem endete. Wenn man nicht weiß, wie man seine Suche anlegen muss, kann es einer Journalistin/einem Journalisten passieren, dass man sich auf den falschen Menschen, den falschen Patienten Null fokussiert. Dieses Kapitel soll dabei helfen, Primärquellen online zu finden, sich von ungenauen Ergebnissen zu befreien und tiefer zu graben.

1. RISIKEN BEI DER KONTAKTAUFNAHME MIT PRIMÄRQUELLEN UND WIE MAN SIE VERMEIDEN KANN

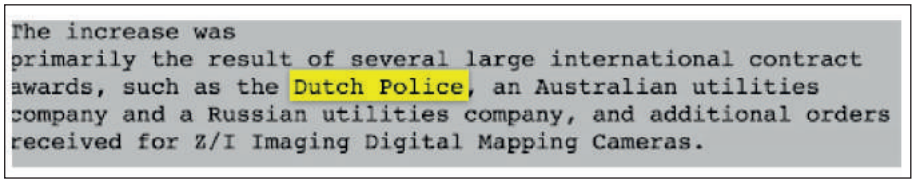
Journalistinnen und Journalisten lieben Primärquellen, die sie online finden können. Das können Zeitungsartikel, wissenschaftliche Studien, Pressemitteilungen, soziale Netzwerke und auch jeder andere denkbare „Patient Null" sein. Eine einfache Suche nach einem Schlagwort auf einer offiziellen Seite kann uns glauben machen, wir würden dort gezeigt bekommen, was vorhanden ist. Oft stimmt das nicht. Ein Beispiel: Besuchen wir einmal die Seite der amerikanischen Börsen- und Wertpapieraufsicht SEC, auf der man finanzielle Informationen sowohl über US-Bürger als auch über ausländische Unternehmer finden kann. Nehmen wir an, wir wollten wissen, wann zum ersten Mal die „holländische Polizei" in deren Unterlagen auftaucht. Die eingebaute Suche auf der Seite sec.gov hilft uns dabei:

Wir bekommen nur einen Suchtreffer – ein Dokument aus 2016. Also kommt die holländische Polizei in den Akten der US-Börsenaufsicht nur einmal vor, richtig?


And I have cooperated with the FBI in the pump and dump scam. The Dutch police. The same thing, with the Scotland Yard over the years. And I certainly understand fraud and fraudulent activities.

Auszug aus einem Dokument der US-Börsenaufsicht. Ins Deutsche übersetzt steht dort: „Und ich habe während des Aktienbetrugs mit dem FBI kooperiert. Mit der niederländischen Polizei. Ebenso mit Scotland Yard all die Jahre. Und mit Sicherheit verstehe ich etwas von Betrug und betrügerischen Aktivitäten."

Falsch. Die erste Erwähnung der holländischen Polizei auf sec.gov war 2004, also zwölf Jahre früher, und zwar in einer ursprünglich geheimen, mittlerweile öffentlichen verschlüsselten E-Mail:


The increase was primarily the result of several large international contract awards, such as the Dutch Police, an Australian utilities company and a Russian utilities company, and additional orders received for Z/I Imaging Digital Mapping Cameras.

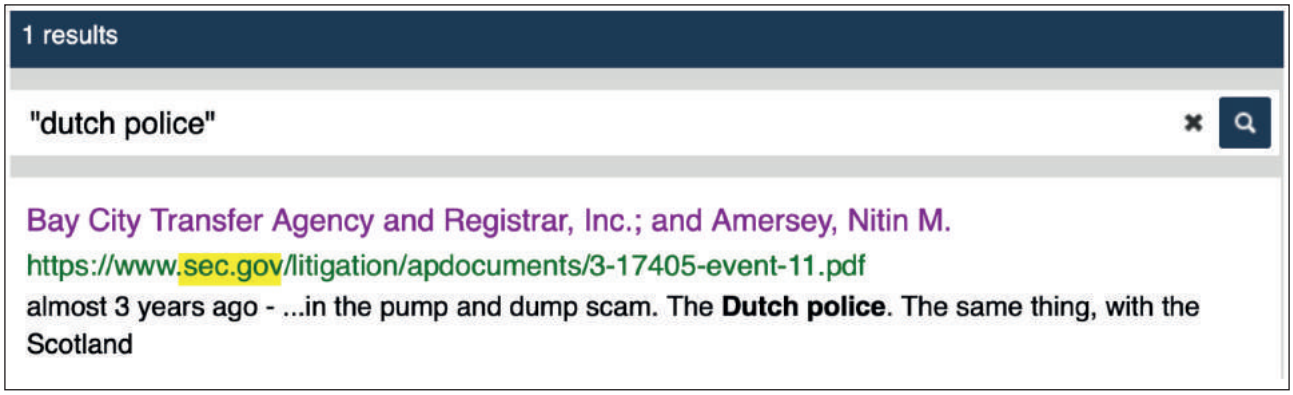
Ein Textausschnitt, der ebenfalls das gesuchte Stichwort enthält, jedoch nicht in den Suchergebnissen auftauchte.

Das hier wird nicht in den Suchergebnissen auf sec.gov angezeigt, und das, obwohl es genau von dieser Website stammt. Was also ist der Unterschied?

Wir sehen: Man sollte Suchmaschinen von Primärquellen grundsätzlich misstrauen. Sie geben manchmal ein falsches Bild davon ab, was auf einer Website an Inhalten vorhanden ist. Der richtige Weg, um sie zu durchsuchen, ist ein eigenständiger Primärquellen-Check.

PRIMÄRQUELLEN-CHECK

Schritt 1: der richtige Link
Nachdem wir auf der Seite sec.gov die Suche benutzt haben, haben wir folgenden Link auf dem Schirm:



Nach Ausführen der Suche werden uns Suchergebnisse präsentiert sowie der jeweils zugehörige Link (= Schrift in grün).

Ja, nur ein Ergebnis, das ist enttäuschend, aber damit können wir weiterarbeiten. Als Erstes ignorieren wir alles, was vor dem Namen der Website steht, also „https://www". Dann suchen wir den ersten Schrägstrich (/), in unserem Fall kommt dieser vor dem Wort „/litigation". Was davor steht, das ist der Teil, den wir brauchen: sec.gov

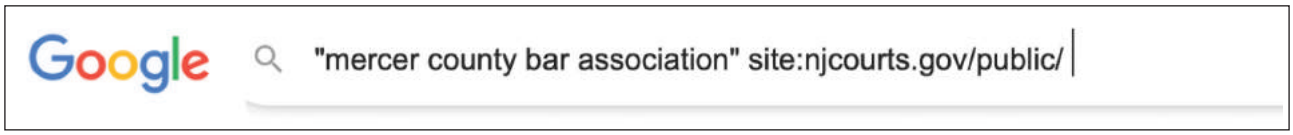
Schritt 2: Wir benutzen den Suchbefehl „site:“
Gehen Sie zu einer beliebigen Suchmaschine. Geben Sie zuerst den Suchbegriff ein (in unserem Beispiel war das „niederländische Polizei“) und geben Sie anschließend den Suchbefehl „site:“ ein, gefolgt vom Namen der Website, wie wir ihn in Schritt 1 extrahiert haben. Wichtig: Das muss hintereinander weg geschehen, also ohne Leerzeichen. Das ist nun die Formel, mit der wir überprüfen können, ob uns die Suchmaschine einer Primärquelle wirklich alles zeigt, was es dort zu finden gibt:

Der Teil in Anführungszeichen ist unser Stichwort, nach dem wir suchen. Der Befehl „site:“ sagt der Suchmaschine, in diesem Beispiel ist es Google, dass sie nur auf der nachfolgend genannten Seite suchen soll. Diese Seite muss dem Befehl direkt folgen, ohne Leerzeichen, ohne „http“ und ohne „www“ davor.

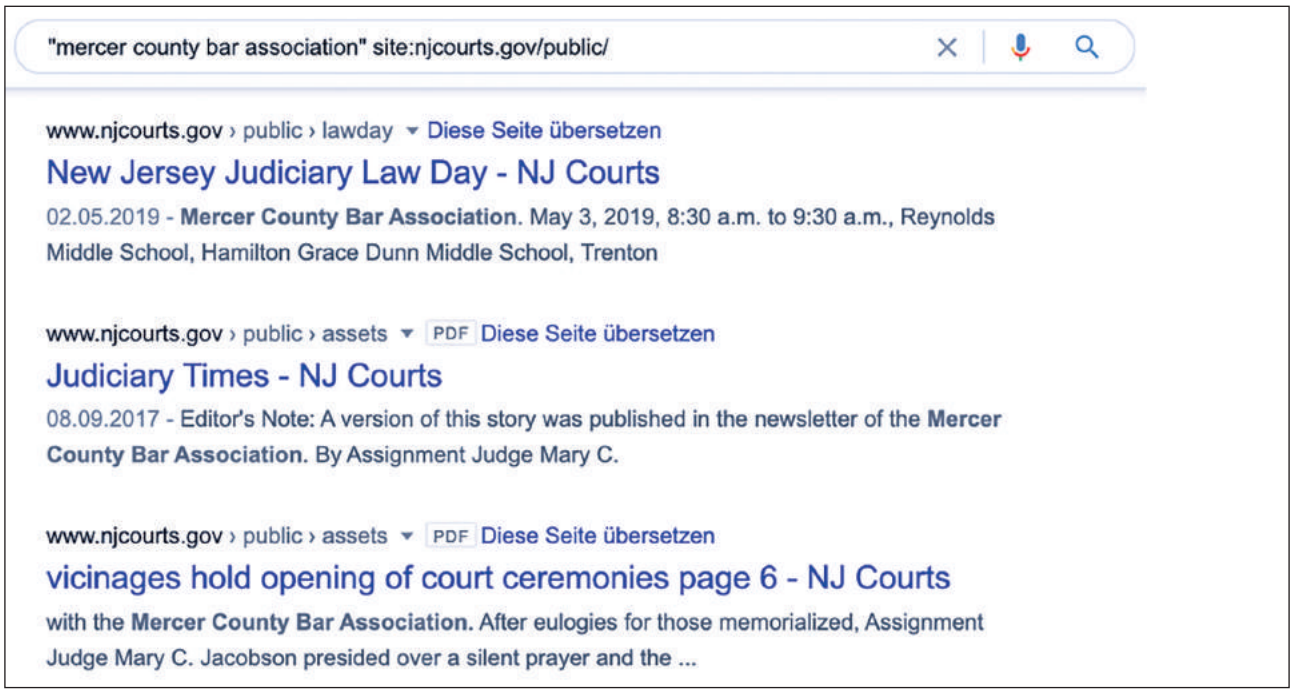
DIE RICHTIGEN VERZEICHNISSE BENUTZEN

Wir können diese Grundformel jetzt an unsere Bedürfnisse anpassen. Schauen wir uns dafür zum Beispiel einmal den Pressebereich auf der Website der Gerichte in New Jersey an. Sagen wir, wir wollen wissen, ob die Anwaltskammer des Landkreises Mercer County dort einmal eine Veranstaltung unterstützt hat. Auf der Seite unserer Primärquelle finden wir nichts: In keiner Pressemitteilung steht die Anwaltskammer von Mercer County in der Überschrift. Aber beim Blick auf den Link dieser Seite fällt uns etwas auf:

Die Pressemitteilungen befinden sich alle im gleichen Bereich auf der Website. Das Verzeichnis heißt „/public“. Das sollten wir in unsere Google-Suche einbauen:

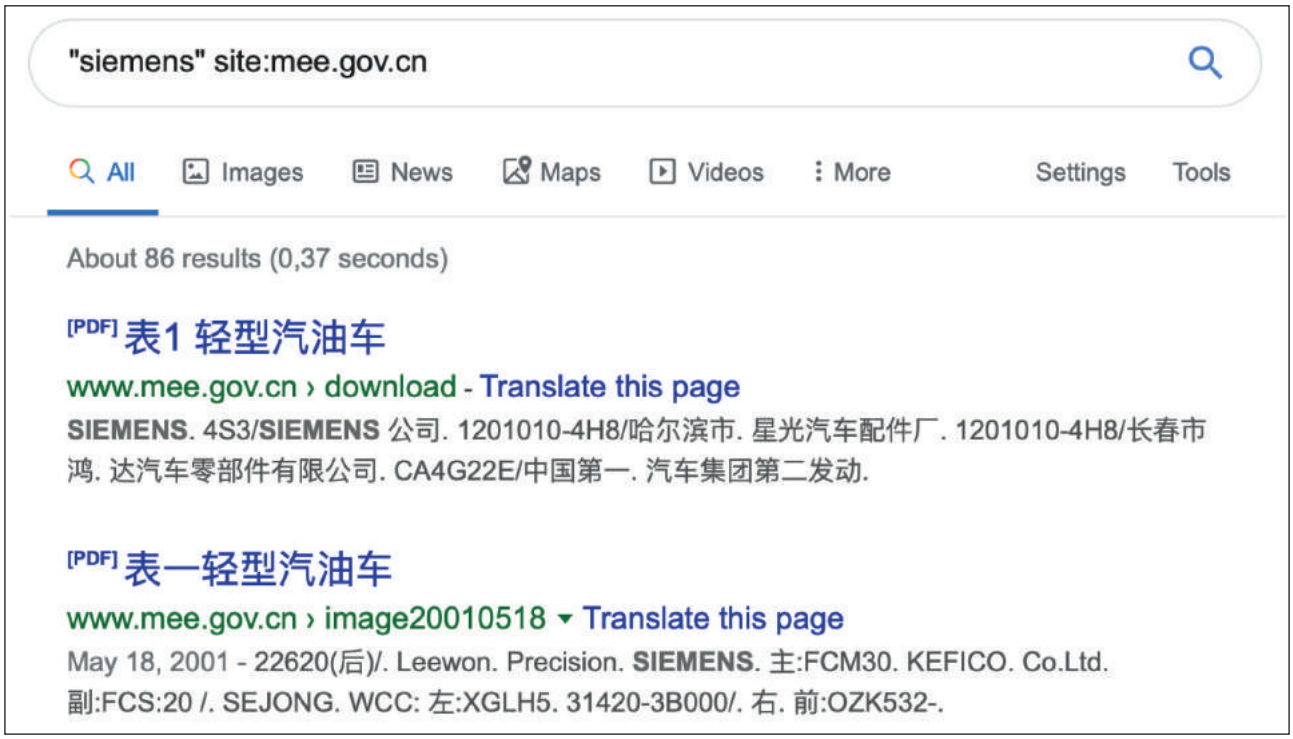


Und siehe da – es gibt solche Veranstaltungen:



VERZEICHNISSE KANN MAN MANCHMAL AUCH ERRATEN

China hat ein Ministerium für Ökologie und Umwelt. Gibt es dort Dokumente über die deutsche Firma Siemens? Mit der nachfolgenden Formel kann man diese in den Suchergebnissen sehen:



Wir wollten ja eigentlich nur die Ergebnisse sehen, die in Englisch verfasst wurden – vielleicht wurde das Wort „english“ also im Link benutzt? Probieren wir es aus:

2. DER SPUR DER DOKUMENTE FOLGEN

Manchmal steckt die Information, die wir brauchen, nicht auf einer Website, sondern tatsächlich in einem Dokument, das irgendwo auf der Website liegt. Schauen wir uns also an, wie man der Spur von Dokumenten über Google folgt.

Der Wirtschaftswissenschaftler Ross McKitrick gilt als eines der bekanntesten Gesichter der Bewegung, die den menschengemachten Klimawandel leugnet.

Ross McKitrick ist außerordentlicher Professor im Fachbereich Wirtschaft der Universität von Guelph, Ontario. Schon 2014 hielt er einen Vortrag für eine Gruppe von Klimaskeptikerinnen und Klimaskeptikern. Wir wollen mal sehen, ob wir die Einladung für dieses Treffen finden können. Was wir wissen: Dass es am 13. Mai 2014 stattfand, dass es das elfte jährlich stattfindende Bankett war und dass es von den „Freunden der Wissenschaft (FOS)" organisiert wurde. Wenn wir Google nach diesen Begriffen durchsuchen, finden wir nichts:

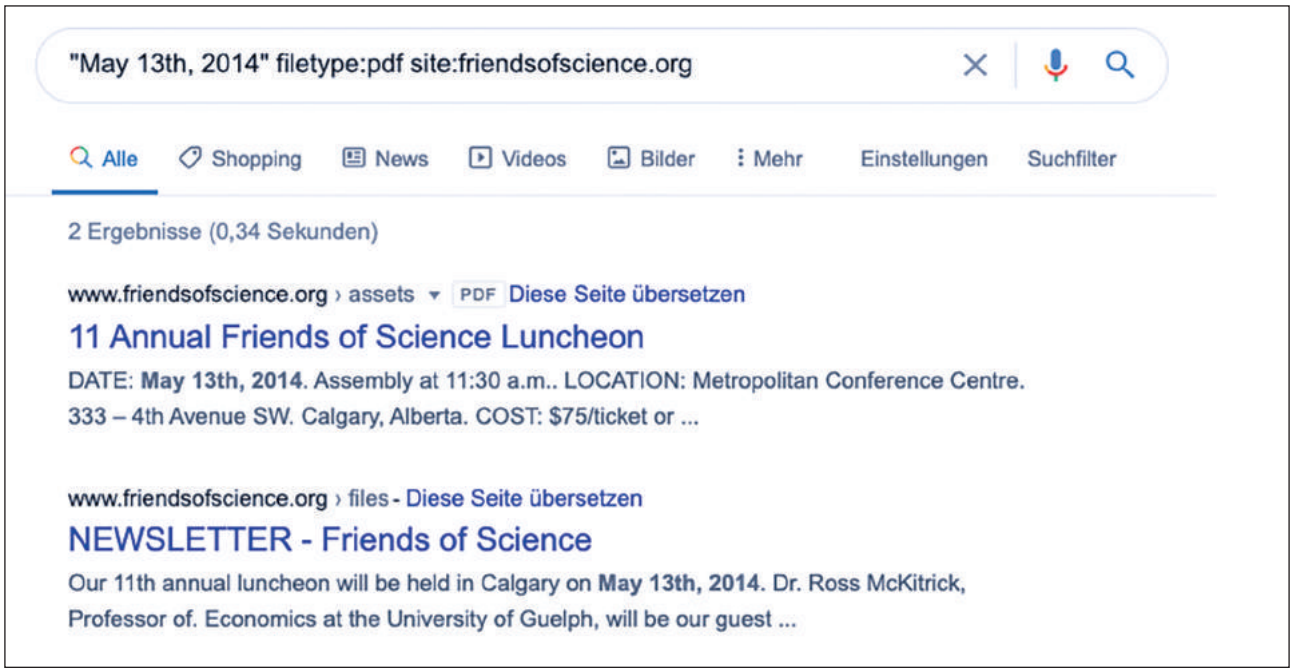
> Keine Ergebnisse für **"Friends of Science 11th Annual Luncheon 2014" "invitation"** gefunden

Warum das? Ganz einfach: Weil das Wort Einladung nicht in so vielen Einladungen vorkommt. Das gleiche gilt für Interview. Auch dieses Wort enthalten die meisten Interviews nicht. Sogar die meisten Karten kommen ohne das explizit dahingeschriebene Wort Karte aus. Also? Aufhören zu raten und durchatmen.

Schritt 1: nen Dateityp festlegen
Suchen wir nach einem gemeinsamen Nenner aller Online-Einladungen, für die wir uns interessieren. Den Dateityp zum Beispiel. Oftmals ist es das PDF. Wenn wir Google mit dem Befehl „filetype:pdf" füttern, bekommen wir nur PDFs in den Suchergebnissen.

Schritt 2: neutral bleiben
Wir wissen den genauen Wortlaut der Einladung nicht. Aber was wir wissen, ist, dass das YouTube-Video von einer Veranstaltung am 13. Mai 2014 stammt. Es ist naheliegend, dass das Datum in der Einladung erwähnt wird. Denken Sie daran, nach verschiedenen Datumsformaten zu suchen: So kann in den USA der 13. Mai als „May 13" oder „May 13th, 2014", aber auch im Format „05/13/2014" angegeben sein.

Schritt 3: Wer ist involviert?
Wir kennen die Veranstalter von „Friends of Science", und wir wissen, dass ihre Website friendsofscience.org lautet. Kombinieren wir nun das alles zu einer Suchanfrage, lautet die Formel für Google wie folgt:



Und da haben wir sie, im ersten Treffer – die Die Einladung zu unserer gesuchten Veranstaltung:

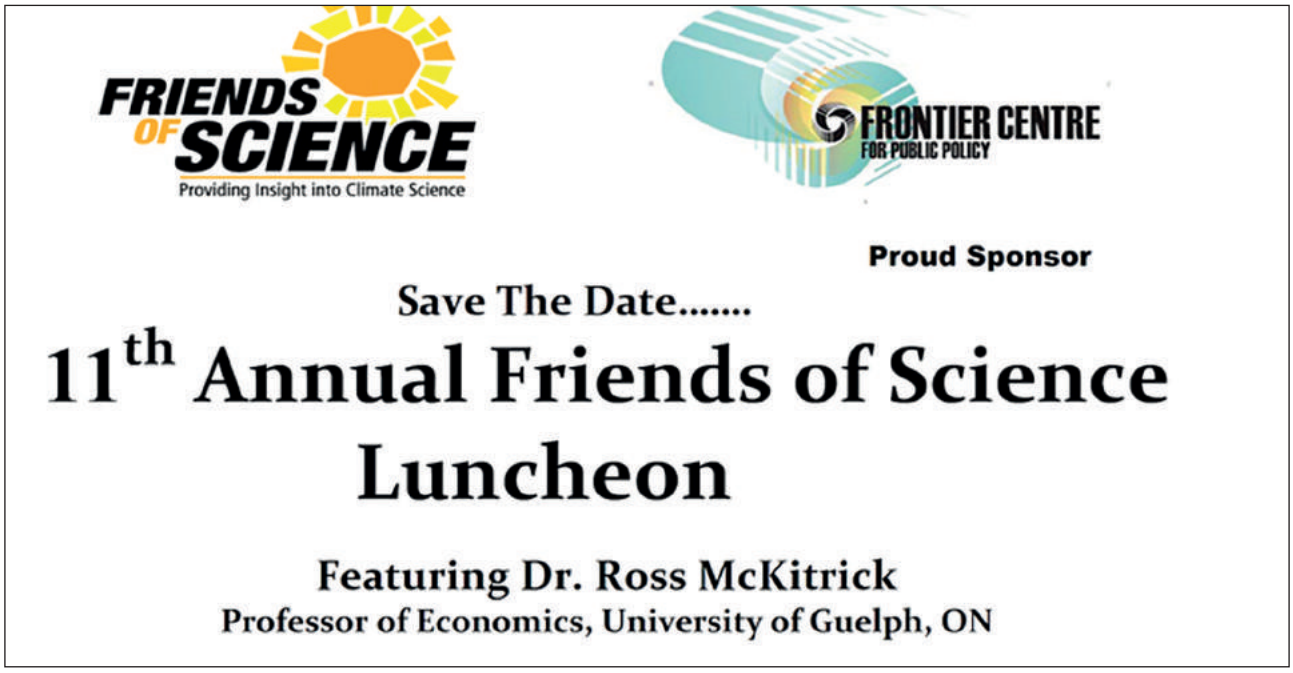


Abbildung aus der gesuchten Einladung.

Die „Friends of Science", beheimatet in Calgary, werden häufig als Leugner des Klimawandels bezeichnet und zum Teil aus der Öl- und Gas-Industrie finanziert. Wie sollte also eine Suchanfrage aussehen, wenn wir über dieses Netzwerk von Unterstützern und Geldgebern noch mehr herausfinden wollen?

Schritt 1: das Ziel eingrenzen
„Friends of Science" allein erzielt zu viele Suchergebnisse, daher nehmen wir „Calgary" mit in die Suche auf.

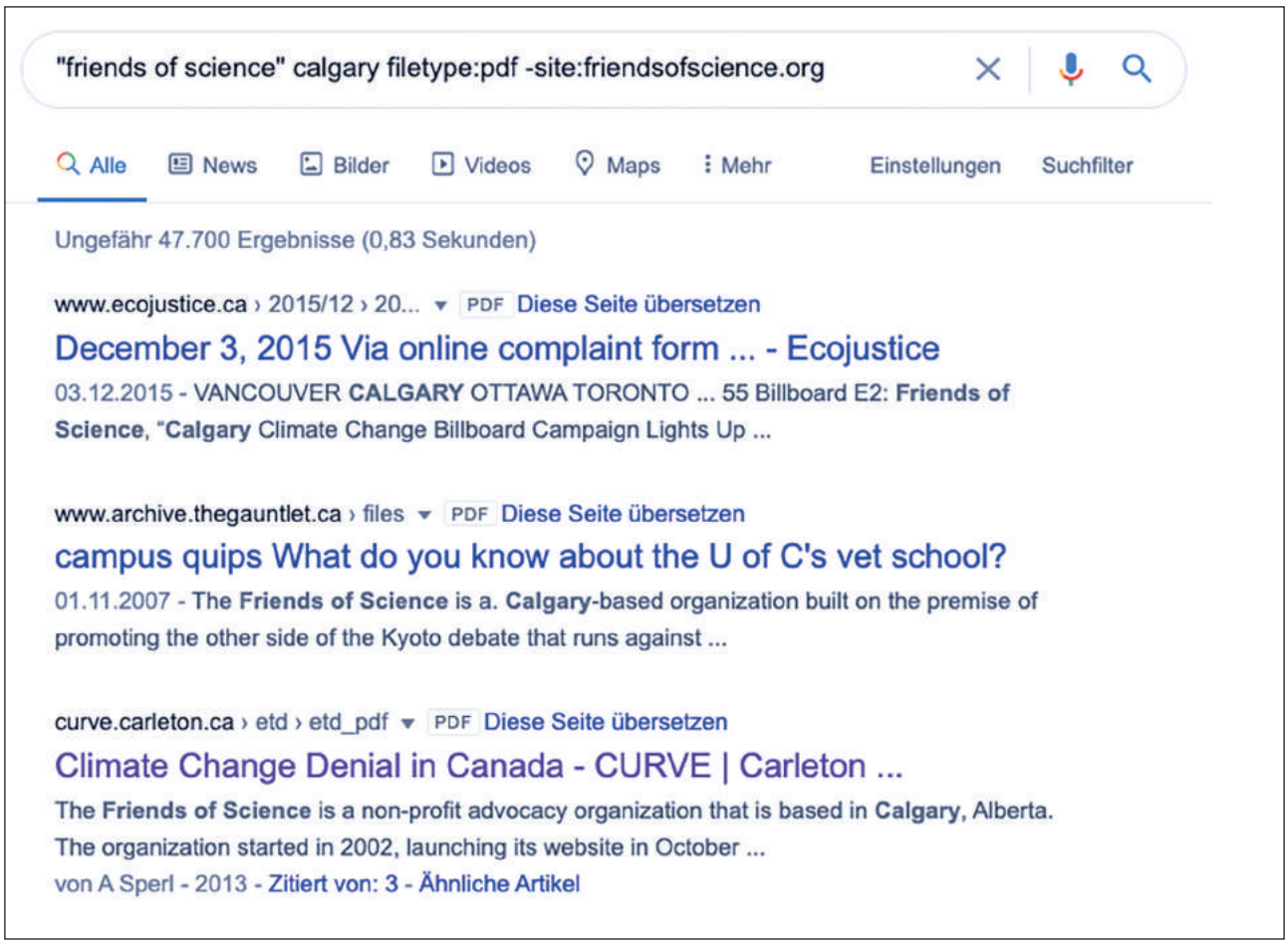
Schritt 2: mit „filetype" den Dateityp festlegen
Beschränken wir uns auf das Format, in welchem die allermeisten offiziellen Dokumente veröffentlicht werden: mit „filetype:pdf".

Schritt 3: die eigene Website unseres Ziels ausklammern
Wir schließen die Website unseres Ziels, friendsofscience.org, aus den Suchergebnissen aus, denn dort werden wir nicht finden, was wir suchen. Der Befehl dafür lautet: „-site:friendsofscience.org". So finden wir Informationen von Dritten.

Und damit lautet unsere komplette Suchanfrage:

```
"friends of science" calgary filetype:pdf -site:friendsofscience.org
```

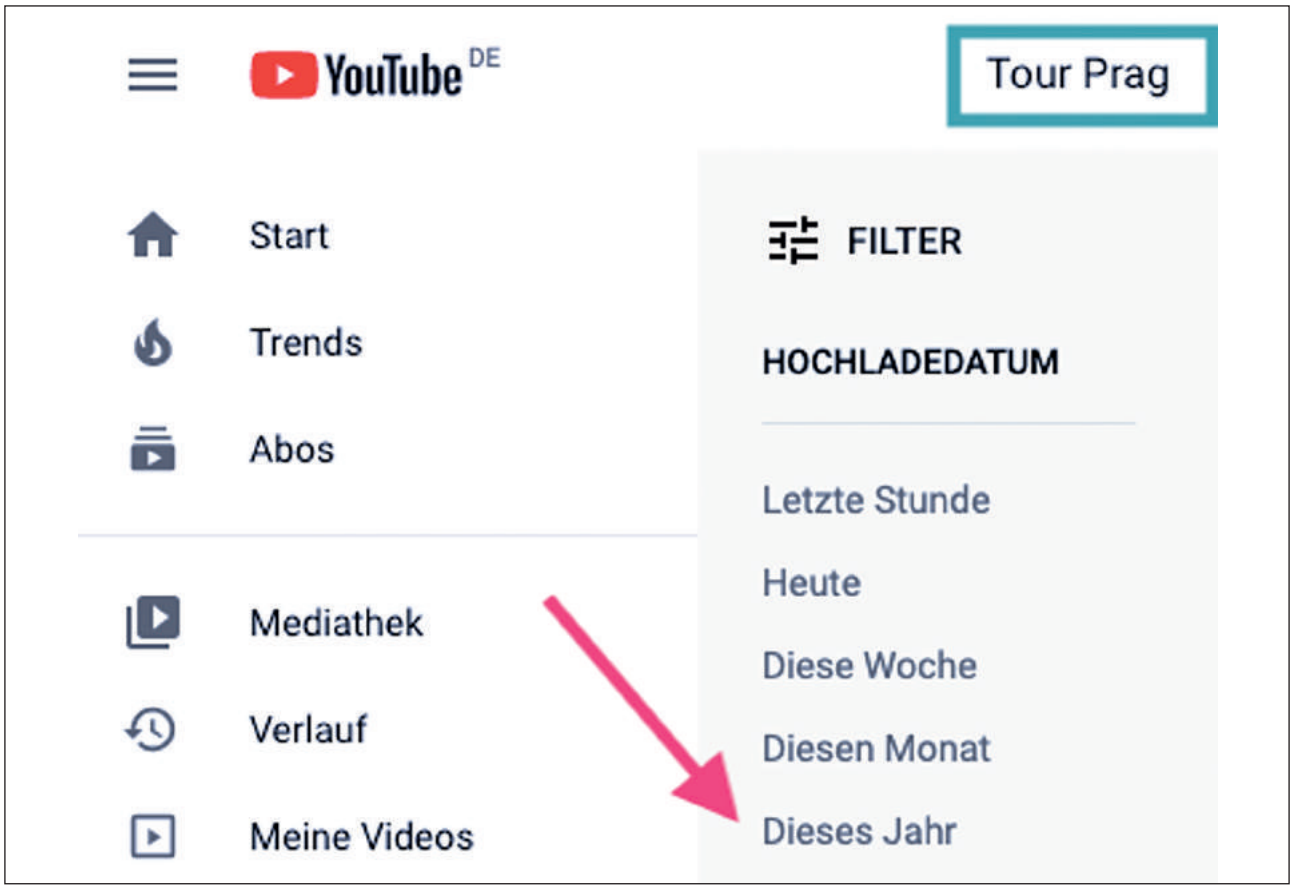
Weil wir nach offiziellen Dokumenten gesucht haben, die aber nicht auf der Website unseres Ziels liegen, finden wir nun beides – Verbündete und Kritiker der Organisation:



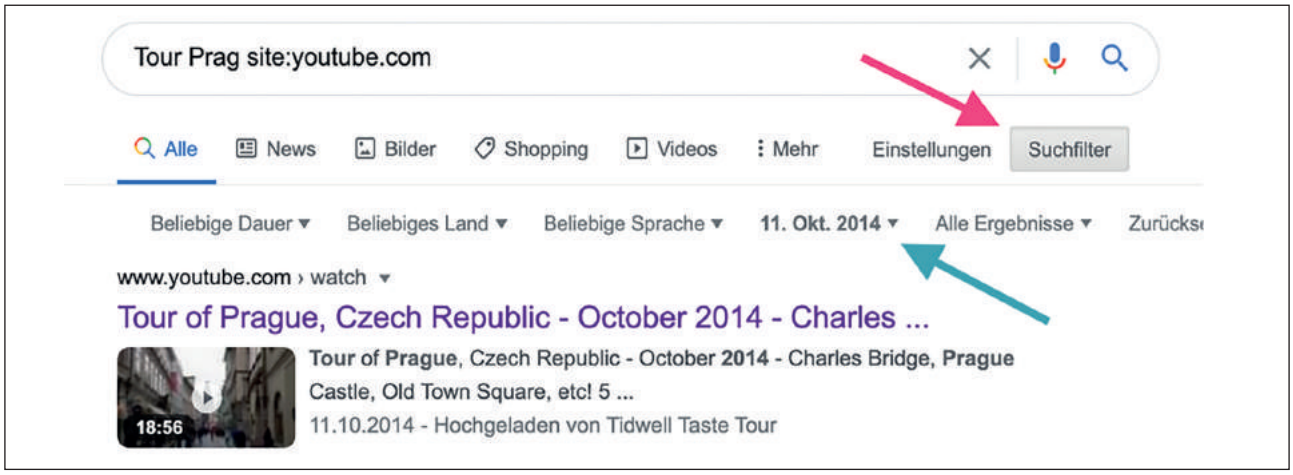
3. SUCHERGEBNISSE IN SOZIALEN NETZWERKEN AUF PRIMÄRQUELLEN EINGRENZEN

YOUTUBE

Das Such-Werkzeug von YouTube hat ein Problem: Es gibt uns keine Möglichkeit, nur nach Videos zu suchen, die älter als ein Jahr sind. Wenn wir zum Beispiel ein Video von einer Tour durch Prag vom 11. Oktober 2014 finden wollen, rennen wir gegen diese Wand:



Also müssen wir einen Umweg gehen und über Google suchen, nicht über YouTube. Dazu fügen wir zu unserem Befehl von oben („site:") noch ein bestimmtes Zeitfenster hinzu. Das tun wir, indem wir zunächst unsere Suche durchführen, danach auf die Schaltfläche „Suchfilter" klicken und dort unter „Beliebige Zeit" ein bestimmtes Zeitfenster einstellen. Nun erhalten wir die Ergebnisse, die wir brauchen:

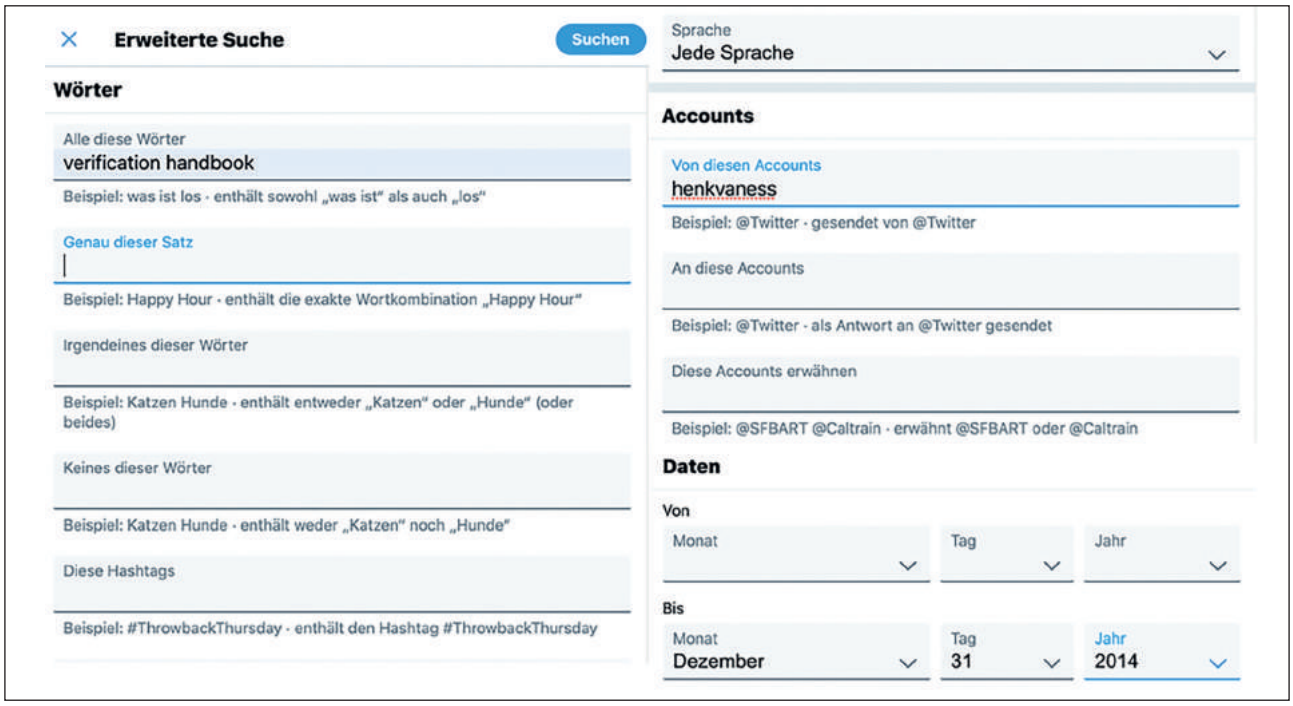


TWITTER

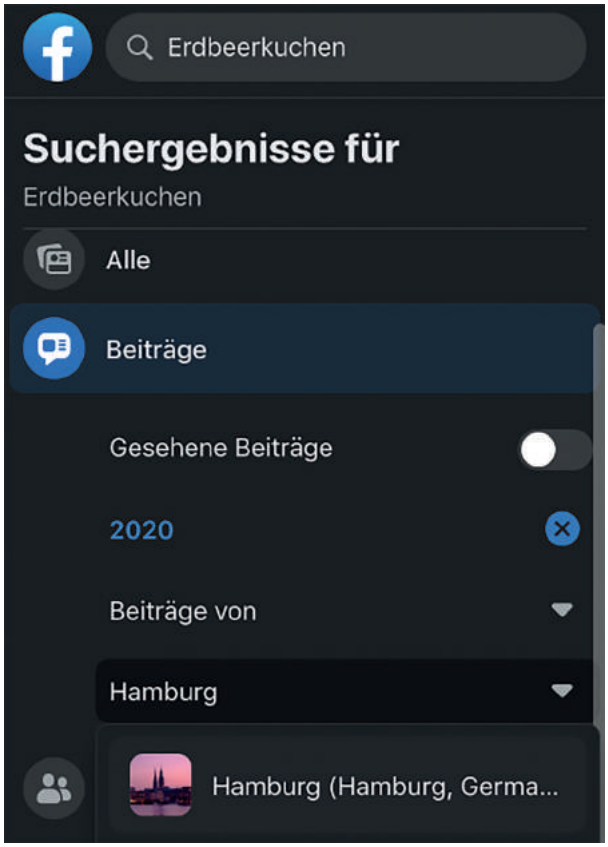
Unabhängig vom Nutzen des Suchbefehls „site:" werden Sie enttäuscht sein, wenn Sie die Google-Suche benutzen, um Twitter zu durchsuchen. Wenn wir zum Beispiel herausfinden wollen, wann ich das erste Mal etwas über dieses Handbuch getwittert habe, könnten wir diese Suchabfrage nutzen:

```
"verification handbook" site:twitter.com/henkvaness
```

Wirklich zum Ziel führt das nicht. Allgemeine Suchmaschinen wie Google haben oft ihre Probleme damit, aus den Billionen Beiträgen auf Twitter, Facebook oder Instagram die qualitativ besten Ergebnisse herauszufinden. Für Twitter lautet die Antwort: Twitters erweiterte Suche – und dort Suchbegriffe, Zeitraum und Benutzernamen miteinander kombinieren, so wie hier:



Die anzeigten Suchergebnisse sortiert Twitter stets danach, was es selbst am wichtigsten findet. Das ist selten das, wonach wir eigentlich suchen. Darum: Nicht vergessen, die Suchergebnisse mit einem Klick auf „Neueste“ nach Datum zu sortieren.

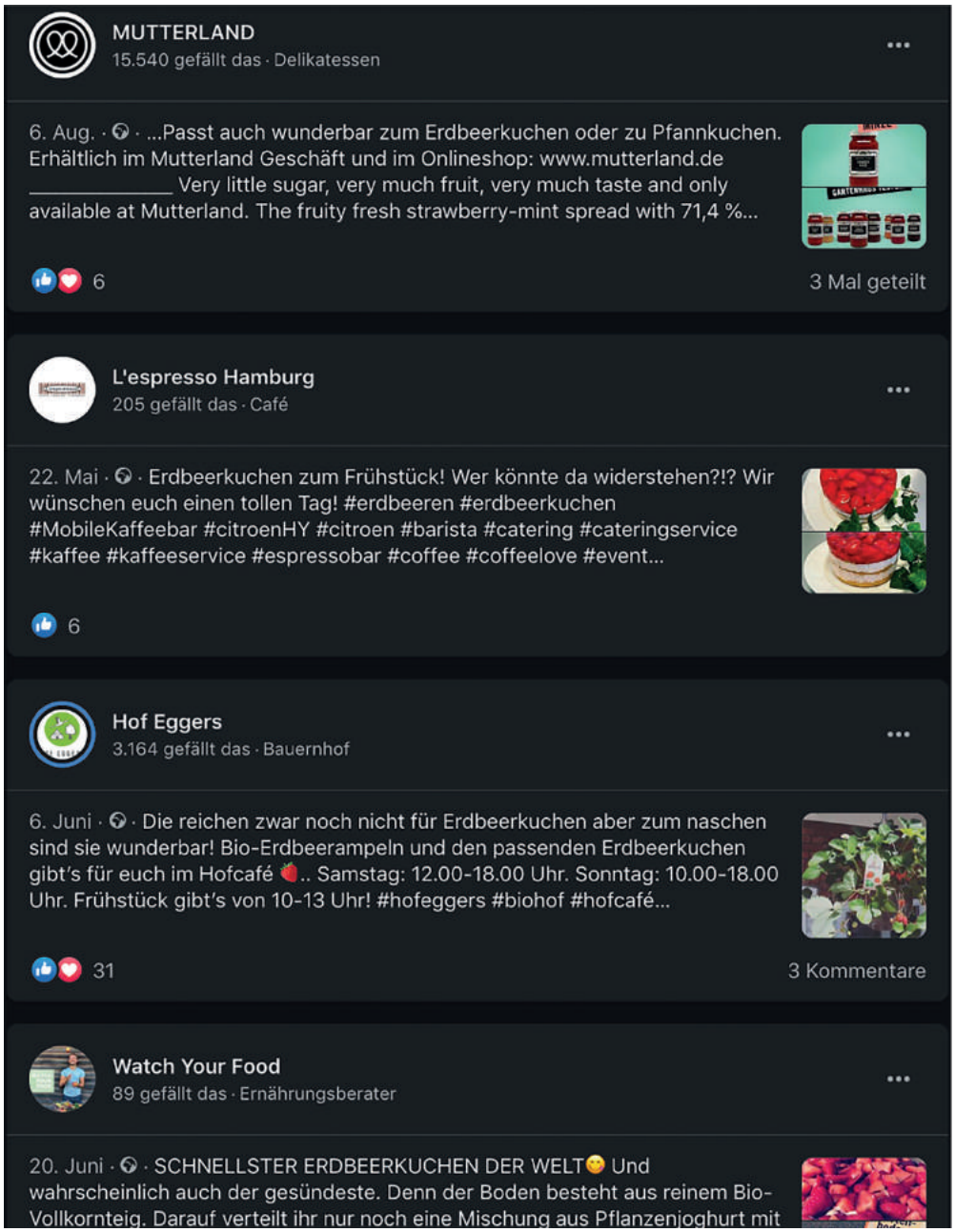


FACEBOOK

Nicht viel besser läuft es mit dem Suchbefehl „site:“ für Facebook. Immerhin können wir die Facebook-eigene Suchmaschine ein wenig an unsere Bedürfnisse anpassen. Sagen wir, wir wollen jüngere Beiträge aus 2020 sehen, die sich mit Erdbeerkuchen beschäftigen und von Menschen aus Hamburg stammen. Das sind die Schritte dafür:

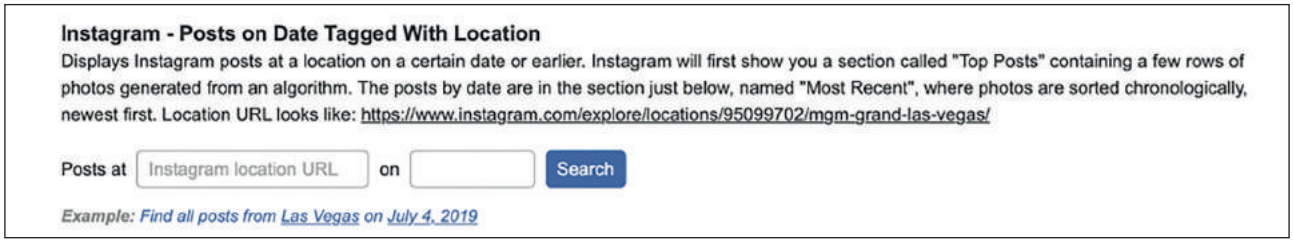
Schritt 1: Den Suchbegriff eingeben
Schritt 2: Auf „Beiträge“ klicken
Schritt 3: Einen Ort wählen
Schritt 4: Das Datum eingrenzen

Und schon haben wir, was wir wollten:



INSTAGRAM

Um Instagram nach Beiträgen von einem bestimmten Ort oder speziellen Datum zu durchsuchen, können Sie eine Website von mir benutzen: whopostedwhat.com. Einfach dort die entsprechenden Felder ausfüllen:



3. BOTS, CYBORGS UND UNECHTE AKTIVITÄTEN ENTDECKEN

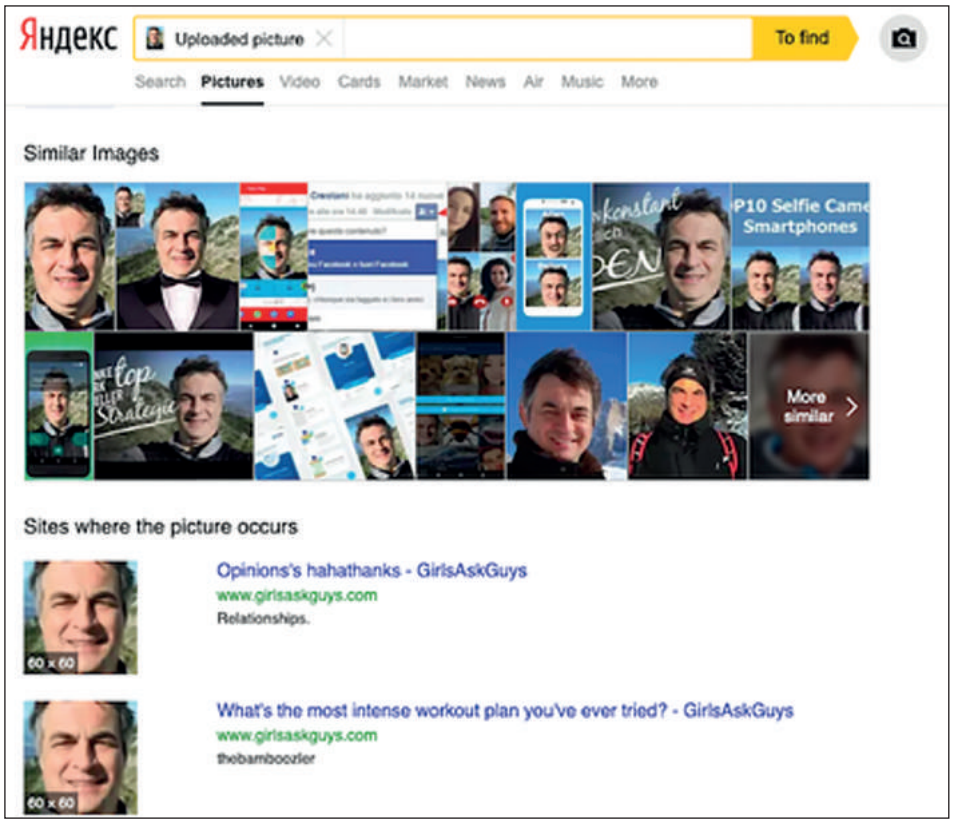
von: Johanna Wild, Charlotte Godart
deutsche Bearbeitung: Marcus Engert

Charlotte Godart *ist Rechercheurin und Trainerin bei Bellingcat. Davor hat sie am Zentrum für Menschenrechte der Universität Berkeley (Kalifornien) im Investigations Lab Studierenden beigebracht, wie man bei globalen Konflikten anhand öffentlicher Quellen Recherchen für Menschenrechtsorganisationen durchführt.*
Johanna Wild *ist eine auf öffentliche Quellen spezialisierte Rechercheurin bei Bellingcat. Dort hat sie sich auf die Entwicklung von Techniken und Werkzeugen für digitale Recherchen spezialisiert. Sie kommt aus dem Journalismus und hat in der Vergangenheit mit Journalistinnen und Journalisten in (ehemaligen) Konfliktgebieten gearbeitet. In Ostafrika hat sie für Voice of America Journalistinnen und Journalisten bei der Produktion von Sendungen unterstützt.*

Ende August untersuchte Benjamin Strick, einer unserer Kollegen bei Bellingcat und ein Rechercheur für „Africa Eye" von der BBC, einige Tweets, die die Hashtags #WestPapua und #FreeWestPapua verbreiteten. Dabei fielen ihm Konten mit ungewöhnlichem Verhalten auf. Zu einem Zeitpunkt, als der Konflikt in Westpapua gerade begann, internationale Aufmerksamkeit zu bekommen, verbreiteten diese Konten Nachrichten, die die indonesische Regierung unterstützten. Eine Unabhängigkeitsbewegung hatte begonnen, auch auf den Straßen für die Befreiung von indonesischer Kontrolle zu kämpfen, was zu Gewalt zwischen der indonesischen Polizei und den Demonstranten führte. Die Konten, die Strick sah, wiesen zahlreiche auffällige Ähnlichkeiten auf. Bald schon erkannte er darin die ersten Anzeichen für ein koordiniertes und nicht authentisches Verhalten. Doch zunächst waren es nur Kleinigkeiten.

Zum einen hatten viele Konten gestohlene Profilbilder. Dieses Profil zum Beispiel behauptete, jemand namens Marco zu sein.

Mit der Bilder-Rückwärtssuche von Yandex fand Strick heraus, dass dieses Bild schon früher auf anderen Websites und unter anderen Namen genutzt worden war. Keines der Profile, die das Foto nutzten, führte zu einer Person namens Marco. Das zeigte: Diese Konten agierten mindestens irreführend, was ihre eigene Identität betraf.



Die russische Suchmaschine Yandex liefert bei Rückwärtssuchen von Bildern oft bessere Ergebnisse als andere Suchmaschinen.

Neben dem Hinwegtäuschen über die eigene Identität bemerkte Strick außerdem, dass die Konten oft ähnliche, mitunter sogar gleiche Inhalte teilten und sich oft gegenseitig retweeteten. Noch auffälliger war, dass einige von ihnen sich bei den Uhrzeiten ihrer Tweets präzise an ein Muster hielten, was die Zeiten ihrer Beiträge betraf. Zum Beispiel twitterten die Profile @bellanow1 und @kevinma40204275 auffällig oft in Minute 7 und in Minute 32 einer jeden Stunde.

26/8/19	17:07:37	bellanow1	26/8/19	23:07:20	kevinma40204275
26/8/19	5:27:06	bellanow1	26/8/19	21:32:52	kevinma40204275
26/8/19	5:27:06	bellanow1	26/8/19	20:32:52	kevinma40204275
26/8/19	5:27:05	bellanow1	26/8/19	18:32:51	kevinma40204275
26/8/19	5:27:04	bellanow1	26/8/19	15:07:22	kevinma40204275
26/8/19	5:27:04	bellanow1	26/8/19	12:32:54	kevinma40204275
26/8/19	3:32:55	bellanow1	26/8/19	9:32:54	kevinma40204275
26/8/19	0:32:56	bellanow1	26/8/19	5:32:54	kevinma40204275
26/8/19	0:07:33	bellanow1	26/8/19	5:07:36	kevinma40204275
25/8/19	23:32:54	bellanow1	26/8/19	3:32:54	kevinma40204275
25/8/19	22:32:53	bellanow1	26/8/19	0:32:54	kevinma40204275
25/8/19	22:07:06	bellanow1	25/8/19	23:32:52	kevinma40204275
25/8/19	20:32:53	bellanow1	25/8/19	23:07:16	kevinma40204275
25/8/19	10:07:19	bellanow1	25/8/19	19:32:53	kevinma40204275
25/8/19	9:32:56	bellanow1	25/8/19	15:07:24	kevinma40204275
25/8/19	9:07:27	bellanow1	25/8/19	10:32:55	kevinma40204275
25/8/19	8:32:56	bellanow1	25/8/19	7:32:55	kevinma40204275
25/8/19	7:07:23	bellanow1	25/8/19	6:32:54	kevinma40204275
25/8/19	6:32:56	bellanow1	25/8/19	6:08:01	kevinma40204275
24/8/19	13:07:57	bellanow1	25/8/19	3:07:21	kevinma40204275
24/8/19	10:07:19	bellanow1	25/8/19	0:07:26	kevinma40204275
24/8/19	7:32:56	bellanow1	24/8/19	20:32:51	kevinma40204275
24/8/19	7:07:20	bellanow1	24/8/19	20:07:08	kevinma40204275
24/8/19	5:32:56	bellanow1	24/8/19	19:32:51	kevinma40204275
24/8/19	4:32:56	bellanow1	24/8/19	15:07:24	kevinma40204275
24/8/19	0:07:31	bellanow1	24/8/19	13:32:55	kevinma40204275
			24/8/19	10:07:17	kevinma40204275
			24/8/19	7:32:54	kevinma40204275
			24/8/19	7:07:18	kevinma40204275
			24/8/19	5:32:54	kevinma40204275
			24/8/19	1:32:54	kevinma40204275

Dass ein Mensch ein solches Twitter-Verhalten an den Tag legt, ist unwahrscheinlich. Die zeitliche Synchronisierung, über mehrere Profile hinweg, kombiniert mit den irreführenden Bildern, legte nahe, dass die Konten nicht zu echten Menschen gehörten, und dass sie automatisiert agierten. Nachdem Strick auffällige Muster unter den Konten suchte, schloss er letztendlich, dass diese Konten Teil eines pro-indonesischen Bot-Netzwerks auf Twitter waren, welches einseitige und irreführende Informationen über den Konflikt in Westpapua verteilte. (Im Fallbeispiel 11 b erfahren Sie mehr darüber.)

WAS IST EIN BOT? DIE ANTWORT IST KOMPLIZIERTER, ALS SIE GLAUBEN

Der Fall aus Westpapua ist weit davon entfernt, die einzige Informationsoperation zu sein, bei der sogenannte Social Bots verwendet wurden. Andere Operationen wurden breiter öffentlich gemacht und auch heftiger kritisiert, obwohl sie im Kern alle ähnlich funktionieren.

Ein Bot ist eine Software-Anwendung, die Aufgaben automatisch ausführt, die ihr vorher von Menschen gestellt wurden. Ob ein Bot etwas Gutes oder etwas Schlechtes tut, hängt komplett von den Absichten seines „Besitzers" ab. Die Bots, um die sich die öffentliche Debatte meist dreht, sind Social Bots, die in sozialen Netzwerken wie Facebook, Twitter oder LinkedIn agieren. Auf diesen Plattformen können sie genutzt werden, um bestimmte ideologische Botschaften zu verbreiten. Oftmals geschieht dies mit dem Ziel, es so aussehen zu lassen, als ob es für ein Thema, eine Person, eine Veröffentlichung oder einen Hashtag eine überwältigende Menge an Unterstützung gäbe.

Üblicherweise fallen Social Bots in eine von drei Hauptkategorien: der vorgeplante Bot, der beobachtende Bot und der verstärkende Bot. Es ist wichtig, sich vorab zu überlegen, für welchen Typ man sich interessiert, denn jeder verfolgt einen anderen Zweck. Und mit jedem Zweck gehen eine andere Sprache und ein anderes Kommunikationsmuster einher. Im Zusammenhang mit Desinformation interessieren uns verstärkende Bots am meisten.

Verstärkende Bots haben genau den Zweck, den ihr Name nahelegt: Inhalte verstärken und verbreiten, mit dem Ziel, die öffentliche Meinung im Netz zu beeinflussen. Sie können auch genutzt werden, um den Eindruck zu erwecken, Individuen oder Organisationen hätten größere Unterstützung als es tatsächlich der Fall ist. Ihre Macht kommt also mit den Zahlen. Ein Netzwerk von Verstärkungs-Bots kann versuchen, Hashtags zu beeinflussen, Links oder visuelle Inhalte zu verbreiten, sich zu massenhafter Werbung zusammenzuschließen oder eine Person online zu belästigen, sie zu diskreditieren oder sie kontrovers und unter Druck stehend erscheinen zu lassen. Indem sie in großer Zahl zusammenarbeiten, erwecken verstärkende Bots den Eindruck, sie seien glaubwürdiger, was ihnen wiederum dabei hilft, die öffentliche Meinung zu formen. Verstärkende Bots, die Falschinformationen verbreiten, tun das meist durch Hashtag-Kampagnen oder indem sie Nachrichten in Form von Links, Videos, Memes, Fotos oder mit anderen Inhaltstypen verbreiten. Hashtag-Kampagnen basieren dann auf Bots, die fortwährend den oder die gleichen Hashtag(s) koordiniert verbreiten. Ihr Ziel ist, Twitters Algorithmus damit auszutricksen, so dass dieser Hashtag in der Liste der „trending topics" gezeigt wird: einer Auflistung jener Themen, die auf der Plattform gerade großes Interesse erfahren. Ein Beispiel ist der Hashtag „#Hillarysick", der weiträumig von Bots verbreitet wurde, nachdem Hillary Clinton im September 2016, kurz vor der US-Präsidentschaftswahl, gestolpert war. (Wichtig zu erwähnen ist auch, dass Hashtag-Kampagnen nicht immer auf Bots angewiesen sind, sondern ohne sie sogar effektiver sind. Hier gibt es eine Recherche über pakistanische Hashtag-Farmen, in denen echte Menschen sitzen. Bots zu kaufen oder zu erstellen ist verhältnismäßig einfach.) Zahlreiche Websites verkaufen Ihnen ihre eigene Bot-Armee für ein paar 100 Euro, vielleicht sogar für weniger. Aber ein fortschrittliches Bot-Netzwerk, das sich menschenähnlich verhält, ist viel schwerer zu bauen und zu pflegen.

WIE MAN BOTS ERKENNT

Entwickler und Forscher haben viele Werkzeuge entwickelt, um festzustellen, ob ein Konto automatisiert sein könnte. Diese Werkzeuge können hilfreich bei der Analyse sein, allerdings ist ihre Bewertung mitnichten endgültig und sollte in keinem Fall die alleinige Grundlage für eine Berichterstattung darstellen.

Eines der bekanntesten Werkzeuge hierfür ist Botometer, entwickelt von Wissenschaftlern an der Indiana University in Bloomington, Indiana. Basierend auf verschiedenen Kriterien berechnet die Seite einen Wert, der angibt, mit welcher Wahrscheinlichkeit ein Twitter-Profil und dessen Follower Bots sind:

Für Reddit hat Jason Skowronski eine Seite zur Überwachung in Echtzeit gebaut. Einmal für ein bestimmtes Unterforum von Reddit eingerichtet, versucht diese, zu bewerten, ob die Kommentare von Bots, Trollen oder Menschen stammen.

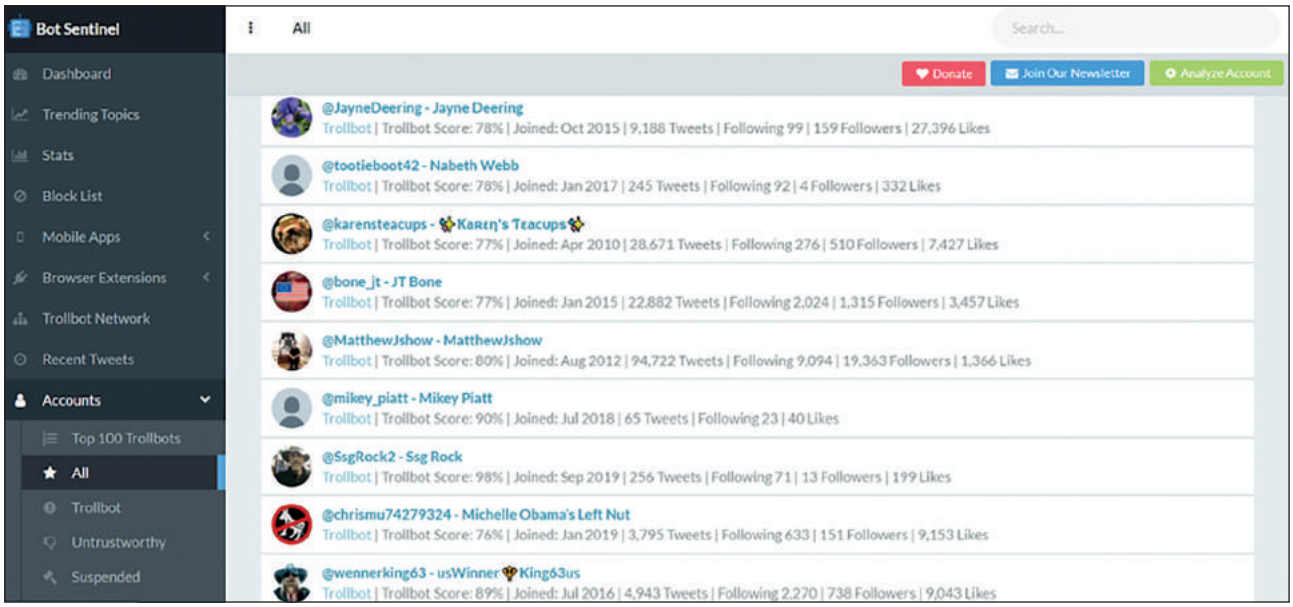
Das „Reddit Bot and Troll Dashboard" beobachtet Unterforen des Social-News-Aggregators Reddit, auf dem sich Nutzer in Gruppen und Communitys zu allen möglichen Themen sammeln und Inhalte dazu teilen.

Mit ein paar Ausnahmen sind die meisten öffentlich verfügbaren Bot-Erkennungswerkzeuge für Twitter entwickelt worden. Das liegt daran, dass die meisten Plattformen ihre Programmierschnittstellen so eingegrenzt haben, dass sie für die Öffentlichkeit keine Möglichkeit mehr bieten, einen entsprechenden Datensatz zu bilden und zu analysieren.

Wie schon früher angemerkt, sind Bot-Erkennungswerkzeuge gut geeignet als Ausgangspunkt, aber sie sollten nicht der einzige Beleg einer Recherche sein. Ein Grund für den doch schwankenden Grad ihrer Genauigkeit ist, dass es schlicht keine ganz genaue Liste von Erkennungsmerkmalen gibt, die Bots zu 100 % erkennbar machen. Außerdem gibt es wenig Konsens darüber, wie genau man ein Konto als Bot klassifiziert. Forscher am Computational Propaganda Project des Oxford Internet Institute klassifizieren Profile, die mehr als 50 Mal am Tag etwas veröffentlichen, als stark automatisiert. Das Digital Forensic Lab des Atlantic Council bezeichnet 72 Tweets am Tag als verdächtig (einen alle zehn Minuten für zwölf Stunden am Stück) und 144 Tweets täglich als sehr verdächtig.

Es kann eine Herausforderung sein, zu sagen, ob eine Desinformationskampagne von Social Bots gemacht wird oder aber von Menschen, die motiviert sind oder dafür bezahlt werden, eine große Zahl von Tweets zu einem bestimmten Thema abzusetzen. Die BBC zum Beispiel fand heraus, dass viele Profile, die im November 2019 ähnliche Facebook-Nachrichten mit unterstützenden Inhalten für Boris Johnson veröffentlicht hatten, von Menschen stammten, die nur vorgaben, Social Bots zu sein.

Möglicherweise stoßen Sie auch auf Cyborgs: Profile in sozialen Medien, die zum Teil automatisiert arbeiten und zum Teil von Menschen gemacht werden, was sich durch eine Mischung aus natürlichem und auffällig unnatürlichem Verhalten zeigt. Journalisten sollten in einem solchen Fall die Profile nicht schlicht als Bots bezeichnen, da falsche Anschuldigungen die eigene Glaubwürdigkeit untergraben. Eine Möglichkeit, mit diesen verschiedenen Typen von Bots, Cyborgs und hyperaktiven menschlichen Bots umzugehen, ist, die Recherche auf sämtliches unauthentisches Verhalten zu fokussieren anstatt zu versuchen, ausschließlich verdächtige Profile zu finden. Bot Sentinel bietet eine öffentlich verfügbare Datenbank von (US-amerikanischen) Twitter-Konten, die verdächtiges Verhalten an den Tag legen. Die Betreiber haben dafür entschieden, jene Konten zu erfassen, die wiederholt gegen die Regularien von Twitter verstoßen haben, anstatt allein nach Social Bots zu suchen.



SCHRITTE ZUR UNTERSUCHUNG UNAUTHENTISCHEN VERHALTENS

Im Wesentlichen schlagen wir den folgenden Zugang zur Identifizierung unauthentischen und potentiell automatisierten Verhaltens in sozialen Netzwerken vor:

1. Die Konten von Hand nach verdächtigem Verhalten durchsuchen.
2. Dies mit der Benutzung von Tools und technischen Netzwerk-Analyse-Diensten kombinieren.
3. Im so gefundenen Netzwerk nach Inhalten und Aktivitäten von anderen Konten suchen, mit denen das ursprüngliche Konto interagiert. Das anschließend mit klassischen Recherchetechniken kombinieren: wie zum Beispiel dem Versuch, die Profile oder Menschen dahinter zu kontaktieren.
4. Sich Rat von Experten einholen, die sich auf Bots und unauthentisches Verhalten in sozialen Netzwerken spezialisiert haben.

Zu lernen, wie man verdächtige Profile von Hand untersucht, ist wichtig, um die typischen Warnsignale zu verstehen, sowohl auf Twitter als auch auf anderen sozialen Netzwerken. Jeder Social Bot braucht eine Art Identität. Die Ersteller von Bots wollen sie damit so überzeugend wie möglich wirken lassen, aber das Aufsetzen und das Betreuen von glaubwürdig erscheinenden Profilen brauchen Zeit, vor allem dann, wenn das Ziel das Betreiben eines großen Bot-Netzwerks ist. Je mehr Konten jemand hat, desto zeitaufwendiger ist es, diese in einer Art zu betreuen, die sie immer echt wirken lässt. Das ist der Punkt, an dem sie auffliegen. In vielen Fällen investieren ihre Ersteller auch das absolute Minimum, um ein Konto zu erstellen, und Rechercheure können genau das ausnutzen. Hier ein paar Punkte, nach denen man Ausschau halten sollte:

Kein echtes Profilbild
Ein gestohlenes Profilbild oder gar kein Profilbild – beides kann ein Hinweis auf ein unechtes Profil sein. Da die Ersteller von Bots viele Profile auf einmal eröffnen wollen, müssen sie auch eine Sammlung von Bildern verfügbar haben – und kopieren diese oftmals von anderen Seiten. Das allerdings sorgt natürlich für Ungereimtheiten. Zum Beispiel wird ein Profil mit einem als weiblich gelesenen Namen und einem als männlich verstandenen Bild nahelegen, dass etwas nicht stimmt. Um so etwas zu umgehen, nutzen viele Bot-Ersteller mittlerweile Cartoons oder Tiere als Profilbilder, was selbst allerdings wiederum zu einem Muster werden kann, dem man im Rahmen der Recherche folgen kann.

Automatisch generierte Benutzernamen
Als Nächstes richten wir unseren Fokus auf die Namen und Benutzernamen. Jedes Twitter-Konto hat einen eindeutigen Nutzernamen. Der Name, den man sich wünscht, ist natürlich meist bereits vergeben. Für die meisten Menschen ist das unbequem, wer aber 50, 500, 5.000 Accounts in kürzester Zeit anlegen will, für den wird das zu einer wirklichen Herausforderung.

Bot-Ersteller fahren daher beim Finden von Nutzernamen oft eine gewisse Strategie. Daher werden häufig Skripte – also ein programmierter Code, der einem Computer einen festen Ablauf an zu erledigenden Aufgaben vorgibt – für die automatische Generierung von Nutzernamen benutzt, die ungefähr so ausschauen:

Nutzername, gefolgt von einer vierstelligen Zahl	12 beliebige Zeichen in Reihe, die bestehen können aus (a–z, A–Z und 0–9)	Ein beliebiger Vorname, gefolgt von einer beliebigen 8-stelligen Nummer, was nahelegt, dass ein von Twitter generierter Standard-Benutzername genutzt wurde
superman_1230 superman_2313 superman_9832 superman_3934 superman_4920	vP1tfl1ZoPG1 dNi29j2utANQ YQBrodhbPC84 TUq3R6GBWYyA XI87NreGshx8	Neil03211977 Sarah92839820 Claire02938593 John09340293 Stephen83749284

Wenn Ihnen bestimmte Twitter-Konten auffallen, deren Namen einem ähnlichen Strickmuster folgen und aus der gleichen Anzahl von Buchstaben und Zahlen bestehen, können Sie händisch weitere Profile recherchieren, indem Sie unter den Followern des verdächtigen Profils nach anderen Konten suchen, auf die das erkannte Muster im Namen ebenfalls zutrifft – und so möglicherweise auf ein Netzwerk stoßen.

In diesem Beispiel hier haben die drei Profile, obwohl augenscheinlich sehr unterschiedlich, etwas gemeinsam: Sie wurden alle im September 2019 erstellt und hinter jedem Vornamen kommen acht Zahlen. In Kombination mit anderen Signalen kann so etwas ein Zeichen dafür sein, dass hinter allen die gleiche Person steht.

WAS DAS KONTO TUT, PASST NICHT ZUM ANGEBLICHEN ALTER

Man sollte zusätzlich misstrauisch werden, wenn ein neues Konto schon eine größere Zahl an Followern aufweist oder wenn es in sehr kurzer Zeit bereits eine große Zahl von Tweets abgesetzt hat. Andersherum gilt: Hat ein älterer Account sehr wenige Follower, obwohl er sehr aktiv ist, ist ebenfalls Vorsicht geboten.

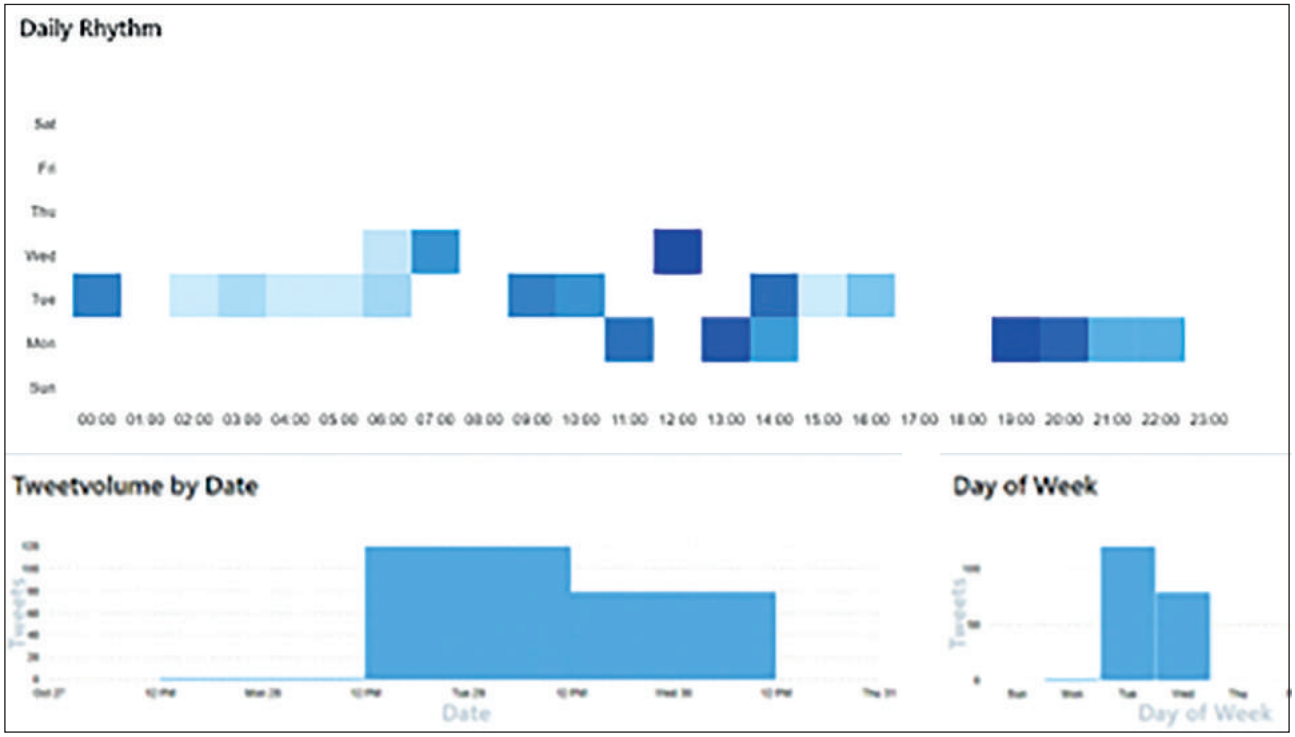
Wurde ein solches Konto identifiziert, dann ist der nächste Schritt, sich dessen Aktivitäten genauer anzuschauen. Eine erste Frage dafür könnte sein, wie viele Tweets durchschnittlich pro Tag abgesetzt werden. Dazu wird die Anzahl der Tweets (diese werden auf der Profilübersichtsseite oben links angezeigt) durch die Zahl der Tage geteilt, die das Konto bereits be-

Dieses Profil, obwohl erst seit September 2019 online, hat bereits weit über 10.000 Tweets abgesetzt.

steht. Hat ein Profil beispielsweise mit Stand 11. November 2019 insgesamt 3.489 Tweets veröffentlicht und wurde am 15. August 2019 erstellt, errechnen sich aus 3.489 Tweets geteilt durch 89 Tage im Schnitt 39,2 Tweets pro Tag. Daran kann sich dann eine Bewertung anschließen, ob diese Zahl gemessen am Auftreten des Profils als zu hoch, unglaubwürdig oder vertretbar erscheint.

VERDÄCHTIGE MUSTER IM VERHALTEN

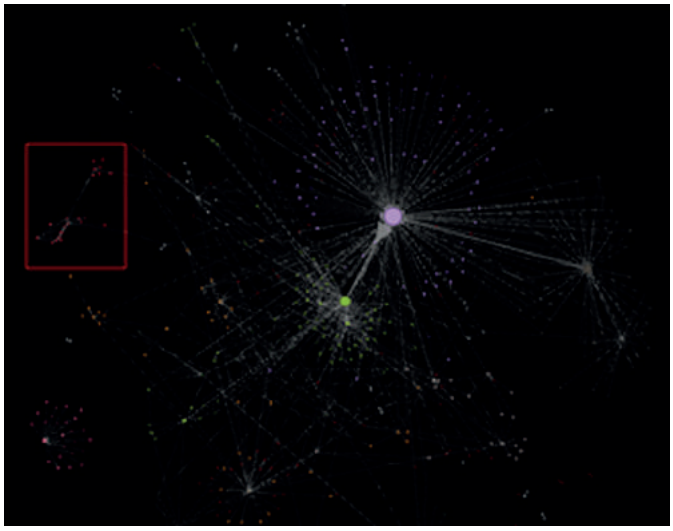
Eine andere Kategorie, die sich untersuchen lässt, ist der Rhythmus, in welchem Tweets verfasst werden. Bei Menschen dürften sich Vorlieben für bestimmte Tageszeiten und Wochentage beobachten lassen. Dass ein Mensch hingegen über eine längere Zeit hinweg durchgängig nur montags, dienstags und mittwochs Beiträge verfasst und an allen anderen Wochentagen nicht, ist ziemlich unwahrscheinlich. Um solche Muster zu erkennen, gibt es das Account Analysis Tool von Luca Hammer:



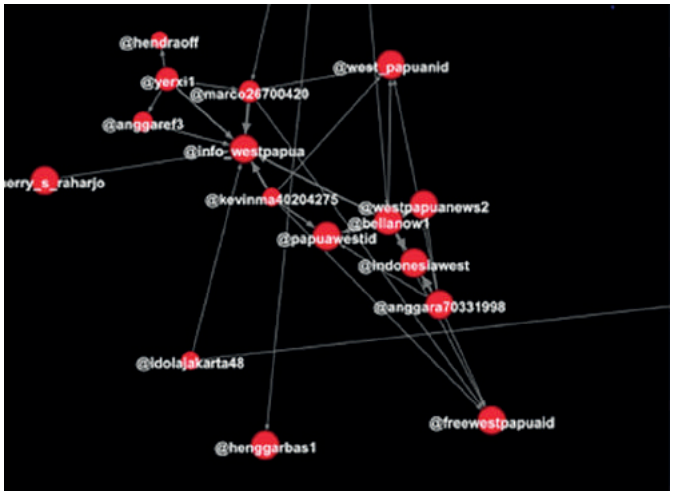
Das vom deutschen Social-Media-Analysten Luca Hammer entwickelte Account Analysis Tool visualisiert die Veröffentlichungsrhythmen und -häufigkeiten eines Twitter-Profils.

SICHTBARMACHUNG ALS TEIL DER RECHERCHE

Um die Aktivitäten eines Netzwerks im Ganzen sichtbar zu machen, empfiehlt sich die Benutzung einer Visualisierungsplattform wie Gephi. Bellingcat-Rechercheur Benjamin Strick nutzte dieses Tool, um die Verbindungen zwischen den Konten aufzuzeigen, die zu dem von ihm gefundenen pro-indonesischen Bot-Netzwerk gehörten. Bei der Betrachtung des entstandenen Schaubildes mit zahlreichen Verbindungen zwischen einer großen Zahl von Konten fiel ihm auf, dass eine Struktur am linken Rand des Bildes (hervorgehoben in Rot) herausstach.



Indem er den betreffenden Bildausschnitt vergrößerte, konnte er erkennen, welche Twitter-Konten zu diesem Teil des Netzwerks gehörten.



Jeder rote Punkt steht für ein Twitter-Konto, und jede Linie zeigt eine Verbindung zwischen zwei Profilen an. Normalerweise sind kleinere Profile um einen größeren Kreis herum angeordnet. Das bedeutet, dass sie alle mit einem einflussreichen Profil in Verbindung stehen. Diese Konten hier interagierten allerdings auf andere Weise miteinander. Das brachte Strick dazu, sich das Verhalten dieser Konten einmal genauer anzuschauen.

DIE ZUKUNFT VON SOCIAL BOTS: KÖNNEN WIR SIE ÜBERLISTEN?

Die Technologie hinter Social Bots hat in den vergangenen Jahren große Fortschritte gemacht. Das hat es diesen kleinen Software-Anwendungen ermöglicht, menschliches Verhalten besser imitieren zu können. Wir sind an einem Punkt angekommen, an dem Menschen erwarten, dass künstliche Nutzer sich an anspruchsvollen Online-Unterhaltungen beteiligen, ohne dass das menschliche Gegenüber realisiert, dass es gerade eine Unterhaltung mit einem Bot führt. Allerdings gibt es,

Stand heute, keinerlei Nachweis dafür, dass solch hochentwickelte, selbstlernende Social Bots bereits existieren oder eingesetzt werden. Noch sieht es so aus, dass die Unterstützung für viele Desinformationskampagnen von deutlich weniger komplexen Verstärkungs-Bots kommt.

„Ich glaube nicht, dass es viele fortschrittliche Social Bots gibt, die in der Lage sind, echte Unterhaltungen mit Menschen zu führen und diese von bestimmten politischen Positionen zu überzeugen", so Dr. Ole Pütz, der als Wissenschaftler an der Universität Bielefeld im Projekt „Unbiased Bots that Build Bridges" forscht. Seiner Ansicht nach ist der beste Weg für die Öffentlichkeit, um unechtes Verhalten in sozialen Netzwerken zu erkennen, auf all die Faktoren zu achten, die ein Konto auffällig machen. Als Beispiel nennt er: „Dieses Konto benutzt ein Skript, um Nachrichten zu retweeten, jenes folgt anderen automatisiert, und das Konto verwendet nie Sprachmuster, die Menschen normalerweise nutzen würden." Noch bleibt also eine kombinierte methodische Analyse von Verhalten, Inhalten, Interaktionen und Mustern der beste Weg, um unechtes Verhalten zu identifizieren. Im nachfolgenden Kapitel geben wir einen tieferen Einblick und weitere technische Erläuterungen zu Methoden, die wir benutzt haben, um ein verdächtiges Netzwerk von Twitter-Konten mit Verbindung zu den Protesten in Hongkong zu analysieren.

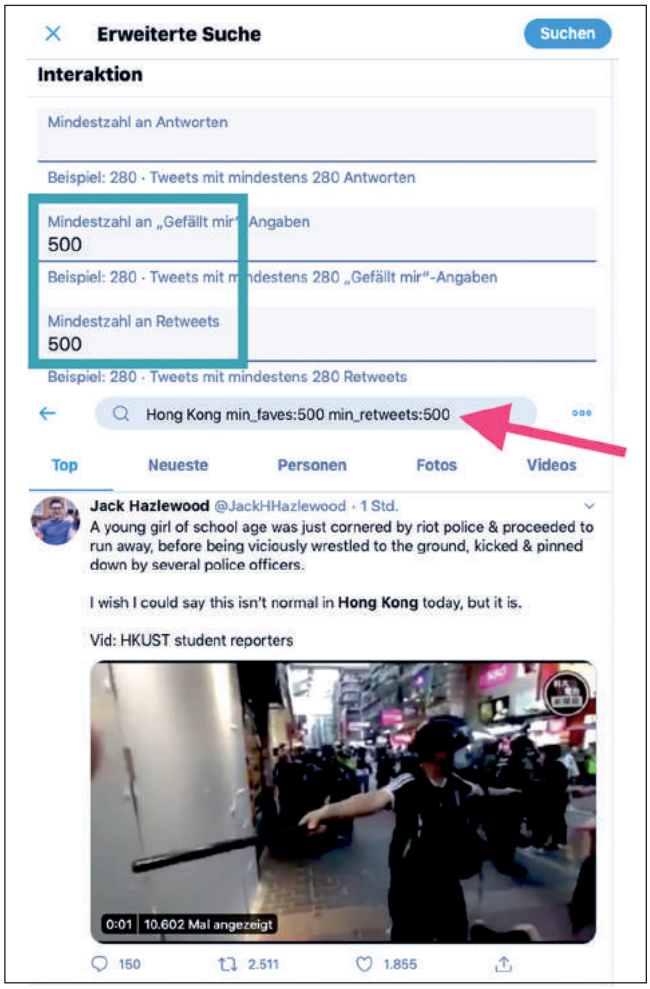
3 a. Fallbeispiel: Wie wir Beweise für automatisierte Aktivitäten auf Twitter während der Proteste in Hongkong fanden

von: Charlotte Godart, Johanna Wild
deutsche Bearbeitung: Marcus Engert

Charlotte Godart *ist Rechercheurin und Trainerin bei Bellingcat. Davor hat sie am Zentrum für Menschenrechte der Universität Berkeley (Kalifornien) im Investigations Lab Studierenden beigebracht, wie man bei globalen Konflikten anhand öffentlicher Quellen Recherchen für Menschenrechtsorganisationen durchführt.*
Johanna Wild *ist eine auf öffentliche Quellen spezialisierte Rechercheurin bei Bellingcat. Dort hat sie sich auf die Entwicklung von Techniken und Werkzeugen für digitale Recherchen spezialisiert. Sie kommt aus dem Journalismus und hat in der Vergangenheit mit Journalistinnen und Journalisten in (ehemaligen) Konfliktgebieten gearbeitet. In Ostafrika hat sie für Voice of America Journalistinnen und Journalisten bei der Produktion von Sendungen unterstützt.*

Im August 2019 gab Twitter bekannt, tausende Konten gelöscht zu haben. Diese hätten als Teil „einer koordinierten, staatlich geführten Operation" dabei geholfen, Falschinformationen über die Protestbewegung in Hongkong zu verbreiten. Schon bald gaben auch Facebook und YouTube ähnliche Statements heraus. Auch sie hatten Konten gelöscht, die durch koordiniertes Verhalten aufgefallen waren. Anders als Facebook und YouTube veröffentlichte Twitter eine Liste mit den gelöschten Profilen – und bot uns dadurch eine Möglichkeit, sie genauer zu untersuchen. Mit einem Teilnehmer eines Bellingcat-Workshops entschied sich unser Team, in den nicht gelöschten Beiträgen auf Twitter über die Proteste in Hongkong nach Zeichen für koordiniertes unechtes Verhalten zu suchen.

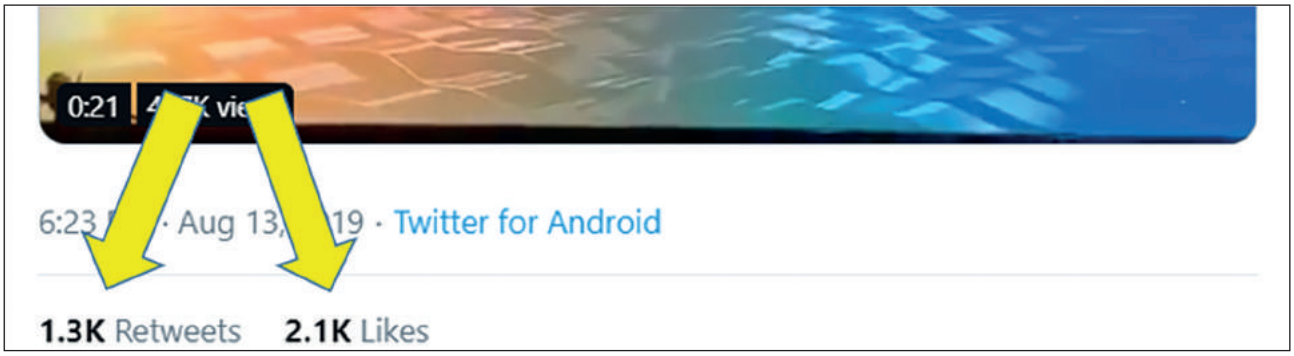
Auffällige Aktivitäten finden
Wir begannen unsere Suche bei den Hashtags, die zu den Protesten genutzt wurden. Eine einfache Stichwortsuche nach „Hongkong Riots" brachte eine große Zahl von Tweets hervor, darunter viele Hashtags. Wir wollten uns auf chinafreundliche Profile und Inhalte konzentrieren, da so auch jene Profile agierten, die von Twitter bereits gelöscht wurden, nachdem sie durch unechtes und künstlich gesteuertes Verhalten aufgefallen waren. Wir nutzten daher Suchformeln wie *„Schande über Hongkong" -Polizei -Regierung* – mit dem Ziel, Tweets wie „Schande über die Polizei von Hongkong" oder „Die Regierung von Hongkong ist eine Schande" herauszufiltern und Tweets wie „Schande über die Protestbewegung in Hongkong" angezeigt zu bekommen. Andere gängige Vokabeln waren „Hongkong Mob" oder „Hongkong Kakerlaken", da diese unter den chinafreundlichen Accounts gängige Beschreibungen für die Protestbewegung waren. Nachdem wir solche und andere Suchen durchgeführt hatten, untersuchten wir von den jüngeren Tweets über Hongkong jene, die durch eine große Zahl an Interaktionen (Likes und Retweets) auffielen. Diese zu finden ist leicht möglich: über Twitters „erweiterte Suche", die einen Filter für eine Mindestanzahl an „Gefällt mir"-Angaben oder Likes bereitstellt.



Beides geht auch händisch, indem man „min_retweets:500" oder „min_faves:500" in das Suchfeld eingibt. Auf diese Weise zeigt die Suche nur Tweets mit mindestens 500 Retweets oder 500 „Gefällt mir"-Markierungen an. Wir haben uns danach die Profile angeschaut, die mit diesen Tweets interagiert haben. So auch diesen Tweet von Hu Xijin, Chefredakteur der englischen und chinesischen Ausgabe der Global Times, die unter der Schirmherrschaft der Kommunistischen Partei Chinas herausgegeben wird:

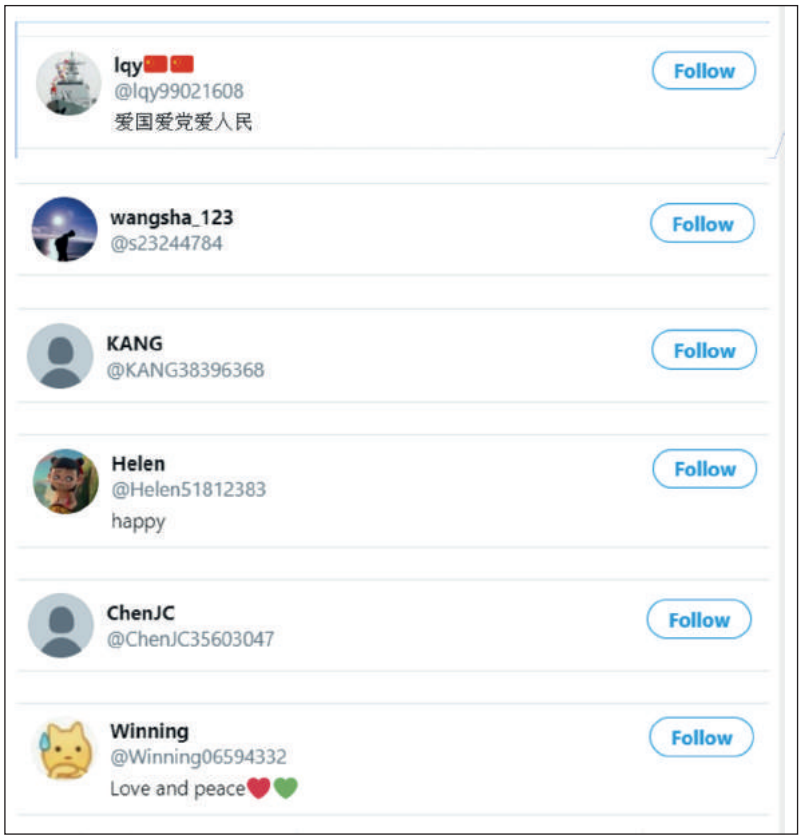
Der Tweet lautet:
„Fu Guohao, ein Reporter der GT (Global Times), wird von Demonstranten am Flughafen von Hongkong festgehalten. Ich bestätige, dass der in diesem Video gezeigte Mann der Reporter selbst ist. Er hat keine andere Aufgabe als Berichterstattung. Ich bitte die Demonstranten nachdrücklich, ihn freizulassen. Ich bitte außerdem westliche Reporter um Hilfe."

Wir klickten anschließend auf die Anzahl der Retweets und die Anzahl der „Gefällt mir"-Markierungen. Tut man dies, listet Twitter jene Profile auf, die entsprechend interagiert haben.



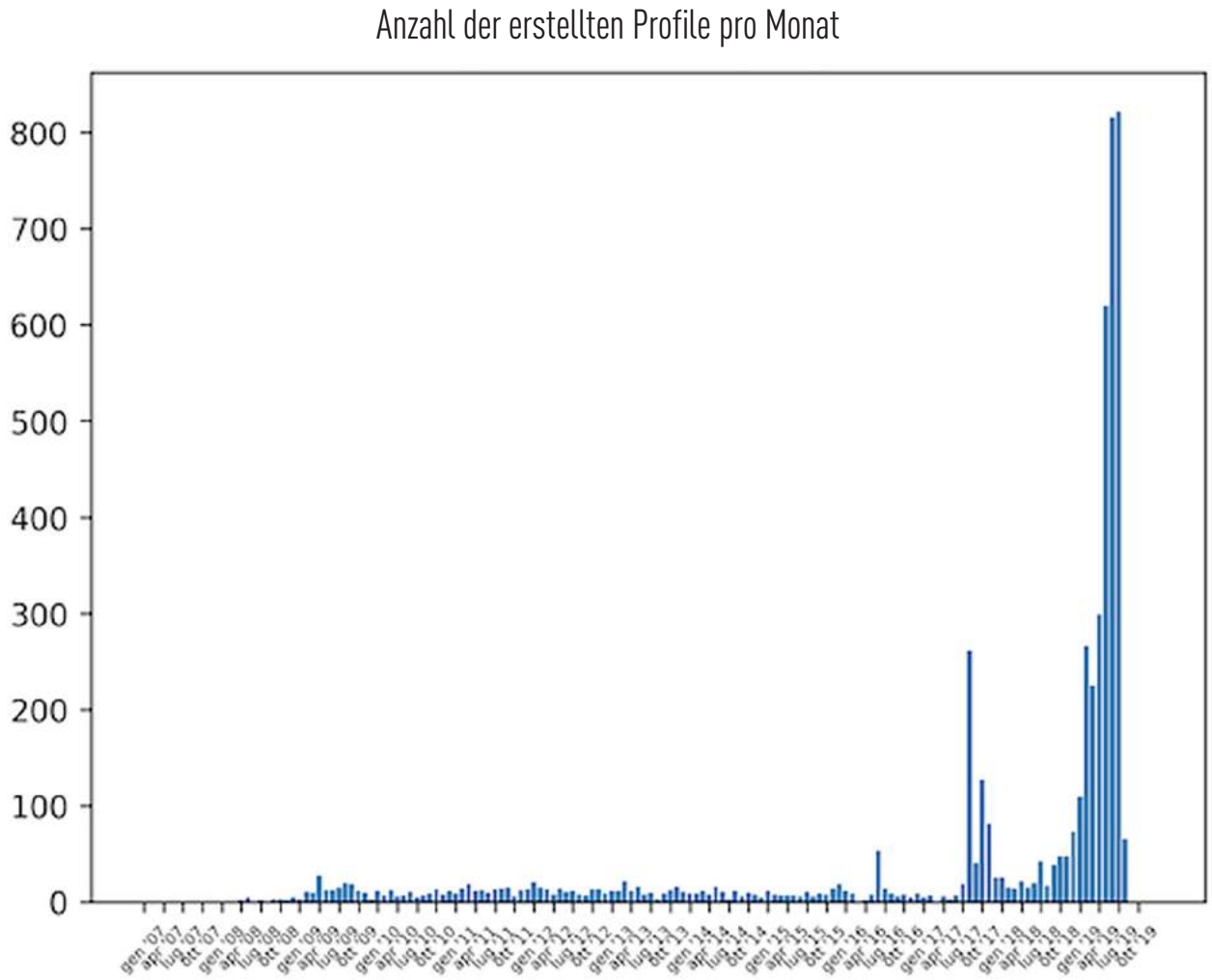
Unter einem Tweet wird die Gesamtzahl der Interaktionen angezeigt. Das können „Retweets", also Weiterverbreitungen, sein oder auch Markierungen mit „Gefällt mir". Mit Klick auf diese Gesamtzahlen bekommt man alle Nutzer angezeigt, die entsprechend interagiert haben.

Unsere Hypothese war, dass unechte Pro-China-Konten Beiträge von prominenten Mitarbeitern chinesischer Staatsmedien verstärken würden. In diesem Fall fielen uns sofort viele Konten auf, die achtstellige Zahlen in den Nutzernamen hatten. Das zeigt, dass die Ersteller der Profile die von Twitter automatisch vorgeschlagenen Benutzernamen akzeptierten, als sie die Profile registrierten, was natürliche Nutzer selten tun – und es legte so weitere Recherchen zu ihrem Verhalten und ihren Charakteristika nahe.



Als wir die Accounts weiter untersuchten, stellten wir fest, dass sie nur wenige Konten als Follower hatten und selbst nur wenigen Konten folgten, keine persönlichen Angaben im Profil hatten, häufig die Tweets von anderen weiterverteilten, ohne eigene Tweets abzusetzen sowie beinahe ausschließlich Beiträge unterstützten, die sich kritisch zu den Protesten äußerten. Wir stellten außerdem fest, dass die Profile alle ungefähr zur gleichen Zeit erstellt worden waren, nämlich im August 2019. Da Twitter ja bereits eine Liste der deaktivierten chinafreundlichen Profile veröffentlicht hatte, konnten wir prüfen, ob sich Ähnlichkeiten finden lassen. Mit der Hilfe von Luigi Gubello, einem Programmierer, der sich für offene Software (Open Source) engagiert, konnten wir ein einfaches Skript in der Programmiersprache Python nutzen (der Code dazu

ist auf GitHub veröffentlicht, mehr Informationen dazu finden sich hier: https://www.gubello.me/blog/), um Muster in den Daten zu finden. Die Grafik unten zeigt, dass die gelöschten Profile ebenfalls alle frisch erstellt worden waren, was zur Charakteristik der Gruppe von Profilen passte, die wir untersuchten.



Den Prozess automatisieren

Da wir jetzt ein Set von Tweets vorliegen hatten, die verdächtige Merkmale und auffälliges Verhalten gezeigt haben, wollten wir eine tiefergehende Analyse durchführen. Dafür waren wir auf Automation angewiesen. Ein Teilnehmer eines Workshops bei Bellingcat hatte Erfahrung mit Software-Entwicklung und schrieb uns einen kleinen JavaScript-Code – regulärer Ausdruck (\w+\d{8}) –, um zwei Funktionen automatisiert durchzuführen: Die Nutzernamen jener Konten herausfiltern, die mit einem bestimmten Tweet interagiert hatten, und aus den Ergebnissen jene herausfiltern, die einem bestimmten Muster entsprachen. Das Muster, nach dem wir filterten, war ein Name, gefolgt von einer achtstelligen Zahlenreihe. Indem wir dieses Skript in die Entwicklerkonsole des Browsers Chrome luden (welcher Entwicklerwerkzeuge direkt im Browser anbietet), lief das Skript im Hintergrund, wann immer wir für einen bestimmten Tweet auf die Liste der „Likes" oder „Retweets" klickten. Anschließend erhielten wir eine Aufstellung, in der jene Namen hervorgehoben waren, die unserem Suchmuster entsprachen. Wie das genau aussah, kann man hier sehen: https://regex101.com/r/zmNya7/1. Fortan konnten wir das Skript benutzen, um jene Konten zu finden, die mit anderen prominenten chinafreundlichen Tweets interagierten.

In der Hochphase der Proteste in Hongkong teilte die chinesisch-amerikanische Schauspielerin Liu Yifei einen Beitrag im sozialen Netzwerk Weibo, welcher die Polizei unterstützte. Das führte dazu, dass einige Menschen in sozialen Medien dazu aufriefen, ihren neuen Film „Mulan" zu boykottieren. Gleichzeitig fiel uns aber auch auf, dass viele Twitter-Profile die Schauspielerin und ihren Film mit dem Hashtag #SupportMulan unterstützten (hier gibt es einen Bericht von der NZZ darüber: https://nzzas.nzz.ch/kultur/disney-dilemma-mulan-unterstuetzt-honkongs-polizei-ld.1505400?reduced=true).[17] Wir entschieden uns, das Skript zu nutzen, um jene User zu finden, die mit diesen unterstützenden Tweets interagierten.

17 Link geändert, Original-Link: https://edition.cnn.com/2019/08/22/entertainment/china-hong-kong-disney-mulan-intl-hnk-trnd/index.html

Der Text des Tweets lautet sinngemäß: „Es ist ihre Pflicht, für ihr Heimatland zu kämpfen. Sie ist eine Heldin für ihre Nation."

In diesem Tweet wird gefordert, beide Seiten anzuhören und sich erst dann ein Bild zu machen. Die Demonstrierenden würden den Film benutzen, um Unwahrheiten zu verbreiten. Man solle aufhören, Gerüchte zu verbreiten.

Wir sammelten die Namen jener Konten, die ins Muster passten, und überprüften dann, wann sie erstellt wurden. Heraus kam, dass die meisten der Accounts am 16. August eröffnet worden waren.

https://twitter.com/monicaG62882882	created: 16 August, 20.07h
https://twitter.com/Min85741833	created: 16 August, 05.29h
https://twitter.com/cherry71737735	created: 16 August, 19.22h
https://twitter.com/Catheri57246362	created: 16 August, 06.13h
https://twitter.com/crystal09837022	created: 16 August, 04.16h
https://twitter.com/Suqing26464572	created: 16 August, 06.30h
https://twitter.com/Yates52905656	created: 16 August, 22.16h
https://twitter.com/hu02261927/	created: 16 August, 04.53h
https://twitter.com/xinjin66947005	created: 16 August, 19.18h
https://twitter.com/Ta99869608	created, 16 August, 21.15h

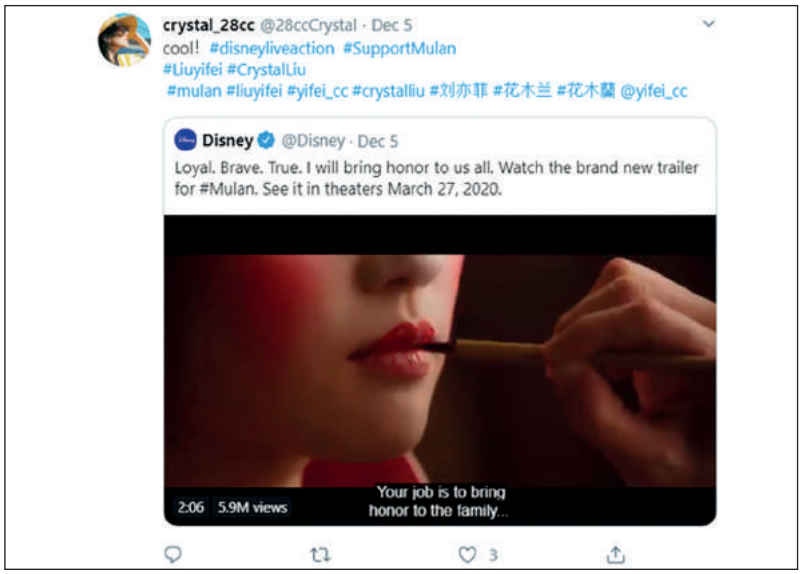
Wir lasen das genaue Datum und die Uhrzeit, zu der das Konto eröffnet wurde, aus, indem wir mit dem Mauszeiger über die Datumsangabe fuhren, ohne daraufzuklicken:

Mit den Konten vor uns begannen wir nun damit, die Inhalte, die die Profile verteilten, von Hand zu analysieren. Schnell wurde klar, dass die Profile aus unserer Liste sämtlichst chinafreundlich und in Opposition zur Protestbewegung twitterten.

Im Tweettext steht: „Die Demonstranten haben unser normales Leben beeinträchtigt, die Polizei von Hongkong schützt die Öffentlichkeit. Und ihr Fremden, die ihr die Wahrheit nicht kennt, ihr werdet von den Medien gehirngewaschen" – dazu steht über den gezeigten Bildern die Frage, ob das friedlicher Protest sei.

Viele der Konten aus unserer Liste blieben nach dem 17. oder 18. August inaktiv, was einmal mehr auf eine koordinierte Vorgehensweise hindeutete. Warum das so war, wissen wir nicht genau. Es wäre möglich, dass Twitter von den Erstellern der Konten zusätzliche Verifikationsschritte forderte und diese nicht in der Lage waren, diese Sicherheitsanforderungen zu erfüllen. Eine andere Möglichkeit wäre, dass man keine Aufmerksamkeit erregen wollte, nachdem Twitter die Löschung einer großen Zahl von chinafreundlichen Konten bekanntgegeben hatte. Wie dem auch sei: Einige Monate später bemerkten wir, dass zahlreiche der Konten wieder aktiv wurden. Diesmal verbreiteten sie positive Botschaften über Yifei und ihren Film „Mulan".

Wir konnten auch andere Konten finden, deren Profilnamen anderen Mustern folgten oder die an anderen Daten erstellt wurden und die Beiträge zur Unterstützung von Yifei verbreiteten. Auf diese Konten stießen wir über eine Suche nach den Hashtags #SupportMulan oder #liuyifei.



Hier wird ein Werbe-Tweet von Disney für den Film genutzt, um darauf politische Botschaften in Form von Hashtags aufzusetzen.

Im Tweettext steht: „Die wirklichen Schläger sind die Demonstranten, nicht die Polizei. Wir unterstützen die beste Künstlerin in Mulan und die Polizei von Hongkong."

4. WIE MAN BEI EILMELDUNGEN UND PLÖTZLICHEN EREIGNISSEN FÄLSCHUNGEN UND KAMPAGNEN AUFDECKT

von: Jane Lytvynenko
deutsche Bearbeitung: Marcus Engert

Jane Lytvynenko *ist Senior-Reporterin bei BuzzFeed News und auf Desinformation, Cybersicherheit und Online-Recherchen spezialisiert. Sie hat Medienmanipulationskampagnen enttarnt sowie Kryptowährungsbetrüger und Menschen, die aus finanziellen Motiven Falschinformationen verbreiten. Ihre Faktenchecks in Krisensituationen erreichen regelmäßig eine große Leserschaft. Jane Lytvynenko stammt aus Kiew in der Ukraine und lebt in Toronto, Kanada.*

Wenn etwas Unvorhersehbares geschieht, dann kann es Stunden oder sogar Tage dauern, bevor Journalisten und Politiker die Situation vollends erfassen können. Mitunter beginnen Übeltäter schon damit, Zwietracht und Misstrauen zu säen und vielleicht sogar ein Geschäft mit der Aufmerksamkeit eines besorgten Nachrichtenpublikums zu machen, während sich die ersten Informationen und Puzzleteile gerade erst über soziale Netzwerke und andere Plattformen verbreiten. Selbst ein wohlwollendes Publikum kann ohne Absicht irreführende Informationen verbreiten. Die Mischung aus hochkochenden Emotionen und nur langsam eintrudelnden Informationen in den ersten Minuten und Stunden eines Ereignisses macht es nötig, dass sich Journalistinnen und Journalisten dafür rüsten, Eilmeldungen effektiv zu sichten, zu verifizieren und – wenn nötig – Falsches richtigzustellen. Es braucht nur wenige Minuten, um einen Tweet, ein Bild, ein Social-Media-Konto oder einen Artikel zu fälschen – und die richtigen Informationen haben es oftmals schwer, diesen Vorsprung dann später wieder einzuholen.

Das Erfolgsgeheimnis, um Fälschungen in solch hektischen Situationen erfolgreich zu beobachten und richtigzustellen ist: Das Fundament dafür schon vorher zu legen. Dazu gehört ein solides Wissen über Verifizierung (wie in der ersten Ausgabe dieses Handbuchs dargelegt), ein Verständnis für die richtige Beobachtung von sozialen Netzwerken und Plattformen sowie eintrainierte Abläufe – und: Sicherheitsvorkehrungen für Situationen, in denen man selbst oder die Kollegen Ziel von Angriffen werden. Wer in diesem Feld arbeitet, sollte Online-Sicherheit niemals auf die leichte Schulter nehmen.

Wenn eine Eilmeldung entsteht, ist der erste Schritt, zu erkennen, wo sich die davon betroffenen Gruppen befinden. Während der Schießerei an einer High School in Parkland, Florida, im Jahr 2018 durchkämmten Reporter die Karte des Video-Kurznachrichtendienstes Snapchat nach Videos, auf denen zu sehen war, was sich in den Klassenräumen abspielte. Während des Hurricanes Irma 2017 war es Facebook, wo die davon Betroffenen Informationen suchten. Zu verstehen, wie jedes soziale Netzwerk funktioniert und wann es sich mit bestimmten Situationen und Gruppen überschneidet, ist essentiell. In diesem Kapitel soll daher der Fokus auf Werkzeugen und Hilfsmitteln liegen, die man in solchen Situationen nutzen kann. Nicht jedes Werkzeug wird zu jeder Situation passen. Darum ist die Frage, wer wie betroffen ist, stets die erste, um zu entscheiden, worauf man den Fokus legt.

DREI DINGE, NACH DENEN MAN AUSSCHAU HALTEN SOLLTE

Nachdem inzwischen auch die Plattformen und Medien daran arbeiten, Desinformation zu bekämpfen, haben die Übeltäter ihre Anstrengungen und Taktiken verbessert, um nicht aufzufliegen. Doch einige Auffälligkeiten und Muster finden sich nach wie vor in ihren Inhalt und ihrem Verhalten:

1. Bearbeitete, verfälschte oder aus dem Zusammenhang gerissene Bilder
Das berühmte Bild eines Hais auf einer überfluteten Autobahn kursiert seit Jahren und täuscht die Menschen. (In der ersten Ausgabe dieses Handbuchs findet sich auch eine Beispielrecherche dazu.) Wenn sich Fotos und Videos von Sachverhalten, die längst als falsch entlarvt wurden, wieder einmal verbreiten, dann nennen wir das gern Zombie-Schwindel. Dennoch ist es wichtig, sie weiterhin zu beobachten: Weil Bilder sich in sozialen Netzwerken viel schneller verteilen als Texte es tun, kann es aufschlussreich sein, sich auf diese zu konzentrieren.

Zwei verschiedene Profile, eines im August 2019 erstellt, eines im Dezember 2017, benutzen das gleiche Profilbild. Darüber hinaus verbreiten sie die gleichen Tweets weiter.

Es scheint, dass diese Konten ihre Strategie geändert hatten: weg von der Kritik an der Protestbewegung in Hongkong und hin zu Werbung für eine Schauspielerin und ihren Film – möglicherweise um zu vermeiden, von Twitter gesperrt zu werden.

Diese Fallstudie zeigt, wie händische und automatisierte Analyseverfahren kombiniert werden können, um in kurzer Zeit ein Netzwerk verdächtiger Twitter-Konten zu finden. Sie zeigt weiterhin, dass es sich lohnt, auch dann noch nach solchen Konten zu suchen, wenn die Plattform bereits bekanntgegeben hat, sie gelöscht zu haben. Nur so waren wir in diesem Fall in der Lage, mit einfachen Suchanfragen und dem genauen Anschauen der Profile eine größere Gruppe von Konten zu identifizieren, die deutliche Anzeichen dafür aufwiesen, dass sie koordiniert und unauthentisch agieren.

Ein Nutzer verteilt ein YouTube-Video, das schon erheblich älter ist, mit dem Hinweis, darauf sei der Attentäter von El Paso zu sehen.

So haben beispielsweise während der Schießerei in einem Supermarkt in El Paso 2019 Rechtsextreme versucht, ein älteres YouTube-Video umzudeuten, das keinerlei Verbindung zum Verdächtigen hatte.

2. Falsche Opfer oder Täter

Als es zu einer Schießerei am Hauptquartier von YouTube kam, wurden soziale Netzwerke mit falschen Behauptungen über den Verdächtigen geradezu überschwemmt. Während der US-Zwischenwahlen verbreitete der US-Präsident falsche Gerüchte über Stimmzettel, die angeblich von illegalen Einwanderern abgegeben wurden. Falsche Aussagen über angebliche Täter tauchen eigentlich bei fast allen großen Ereignissen auf.

Während der Schießerei an der High School in Parkland, Florida, 2018 verbreitete ein gefälschtes Konto namens Bill O'Reilly einen falschen Namen des Verdächtigen. Der Tweettext lautete: „EILMELDUNG: Zweiter Parkland-Schütze festgenommen. Namen in den Polizeiberichten lauten Nicholas Cruz und Sam Hyde. Gefahr vor Ort könnte noch nicht vorüber sein, da Berichte von zwei oder mehr Bomben. Florida High School."

3. Belästigung und Schikanierung

Auch wenn keine Falschinformationen im eigentlichen Sinne kursieren, versuchen Übeltäter regelmäßig, Menschen durch Belästigung zum Schweigen zu bringen. Das kann auch ein Anzeichen dafür sein, dass eine Gruppe von Menschen beginnt, diesem Ereignis Aufmerksamkeit zu schenken und sich vielleicht später mit anderen Taktiken darauf konzentrieren wird. Die Schikane kann vielfältig sein: Es kann passieren, dass eine Gruppe von Profilen koordiniert zusammenarbeitet, um den Eindruck eines Ansturms zu erwecken. Es kann sein, dass ein bestimmter Beitrag in einer Abstimmung nach unten oder oben gedrückt wird. Oder dass ein Nutzer mit Kommentaren geradezu überflutet wird.

Der Tweet lautet „Viele verdächtige Profile, die heute Abend das ‚Kamala Harris ist nicht Schwarz' - Narrativ verbreiten. Es ist quasi überall und es zeigt alle Zeichen einer koordinierten/künstlichen Operation."

So verbreiteten nach einer Debatte unter Kandidaten der Demokraten im Wahlkampf 2019 anonyme Profile in Massen die immer gleiche Botschaft: Sie zweifelten öffentlich an, dass Kamala Harris (die Vizepräsidentschaftskandidatin der Demokraten) schwarz sei.

EMPFEHLUNGEN FÜR ARCHIVIERUNG UND VERÖFFENTLICHUNG

Bevor Sie beginnen, solche Schwindeleien und Fälschungen zu untersuchen, legen Sie sich einen Ordner oder ein Ablagesystem für Ihre Funde an und beginnen Sie damit, diese vom ersten Moment an in einer Tabelle zu erfassen. Was auch immer Sie finden: Machen Sie auf der Stelle einen Screenshot von jedem Beitrag und jedem kleinen Schnipsel und archivieren Sie die Seite. (Die Browser-Erweiterung von archive.org ist eine kostenfreie, schnelle und effektive Methode dafür.) Achten Sie darauf, in Ihrer Tabelle sowohl den originalen Link als auch den Link der archivierten Seite zu erfassen. Das erlaubt es Ihnen, in Ruhe nach Mustern und Auffälligkeiten unter all Ihren Funden zu suchen, wenn die Aufregung vorbei ist. Um zu vermeiden, dass man selbst Seiten verbreitet und erfolgreicher macht, die mit Des- oder Falschinformation arbeiten, sollte man in Artikeln oder sozialen Netzwerken zu der archivierten Seite verlinken und nicht zum Originalbeitrag. Darüber hinaus sollte man auf den Bildern, die man verbreitet, ein deutlich lesbares „Fake", „Falsch", „Fälschung" oder Ähnliches anbringen, um sicherzustellen, dass sie nicht in einen anderen Kontext gerückt und isoliert weiterverbreitet werden können.

Wenn Sie einen Artikel schreiben, fokussieren Sie sich in der Überschrift und im Text darauf, das zu sagen, was stimmt, anstatt Falsches zu wiederholen. Studien haben gezeigt, dass auch die auszugsweise Wiederholung von Unwahrheiten dazu führen kann, dass Menschen ebenjene Unwahrheit behalten, und nicht die Richtigstellung (nach dem Prinzip: Sie können nicht nicht daran denken, dass es fliegende Elefanten gibt). Ihre Aufgabe besteht darin, die Wiederholung von Unwahrheiten so weit wie möglich zu minimieren und die Menschen auf korrekte Informationen hinzuweisen.

SCHLAGWORTE UND ORTE FINDEN

Wenn die Ereignisse beginnen, sich zu überschlagen, erstellen Sie eine Liste mit relevanten Schlagworten und Orten. Was die Orte betrifft, sollten Sie nicht nur die Stadt, den Staat und das Land einbeziehen, sondern auch lokale Bezeichnungen wie Slogans oder Spitznamen einer Stadt sowie benachbarte Stadtteile. Während Wahlen sollte ebenfalls der Landkreis oder relevante Wahlbezirk mit in die Suchliste. Diese Informationen werden benötigt, um Beiträge, die mit einer Ortsmarke versehen wurden, zu beobachten und nach Ortserwähnungen in den Beiträgen und Antworten zu suchen. Darüber hinaus sollte man die Profile der relevanten örtlichen Behörden und offiziellen Stellen suchen und verfolgen. Das umfasst neben Polizei und Feuerwehr auch Katastrophenschutz, Politiker oder Lokalredaktionen.

Als Nächstes identifizieren Sie die Schlüsselworte. Das können Schlagworte wie Opfer, Verdächtiger, Schütze, Schießerei, Flut, Feuer oder auch der von Behörden bestätigte Name jeder involvierten Person sein, aber auch Phrasen wie „wird gesucht". Denken Sie daran, was Menschen in solchen Situationen neben den üblichen Suchanfragen benutzen. Wenn Sie ein glaubwürdiges Konto finden, das angibt, inmitten der konkreten Situation zu sein, die Sie beobachten, notieren Sie sich den Kontonamen und studieren Sie das gesamte Profil. Schauen Sie sich auch Freunde und Follower an, um zu prüfen, ob Sie dort noch weitere Menschen finden, auf die das zutrifft. Bedenken Sie außerdem, dass in stressigen Situationen Menschen die Namen von Orten oder Personen oft falsch schreiben. Beispielsweise wurde während des massiven Buschfeuers in Kincade in Kalifornien der Hashtag #kinkaidfire genutzt, vermutlich wegen der Rechtschreibkorrektur auf Mobilgeräten. Nehmen Sie daher häufige oder naheliegende falsche Varianten in Ihre Suche auf und versuchen Sie, andere Varianten davon zu finden, indem Sie die ersten Buchstaben in die Suchmaske eingeben und prüfen, was vom Gerät an Vorschlägen zur automatischen Vervollständigung angezeigt wird.

In sich überschlagenden Nachrichtensituationen kann man als Journalistin oder Journalist auch öffentlich dazu aufrufen, die Recherche zu unterstützen. In diesem Fall schreibt die Autorin dieses Kapitels: „Ich sammle unter diesem Tweet Falschmeldungen und Irreführendes über die möglicherweise stattfindende Schießerei am Hauptquartier von YouTube. Wer etwas sieht, kann sich per Direktnachricht (DM = direct message) oder E-Mail bei mir melden."

Diese Momente sind auch der richtige Zeitpunkt, um mit Quellen in Kontakt zu treten – alle, die man vor Ort kennt, die Teil der betroffenen Gemeinschaft sind, die Ziel von Belästigung, Angriffen oder gezielt über sie verbreiteten Falschinformationen werden könnten – und sie zu fragen, was sie online beobachtet haben. Außerdem kann man das eigene Publikum wissen lassen, dass man auf der Suche nach Falschinformation und anderen problematischen Inhalten ist, die mit dem Ereignis in Zusammenhang stehen. Dazu empfiehlt es sich auch, sich mit Kolleginnen und Kollegen abzustimmen, die die offiziellen Profile Ihrer Institution in sozialen Netzwerken betreuen, und sie zu bitten, die Information zu verbreiten und zu beobachten, ob Antworten eingehen.

DIE WICHTIGSTEN WERKZEUGE FÜR BILDER

1. Bilder suchen

Die Bilder-Rückwärtssuche ist ein unverzichtbares Werkzeug. In Google nach einem Bild zu suchen ist einfach: Im Browser Chrome per Rechts-Klick auf das Bild klicken und anschließend „Mit Google nach Bild suchen" anklicken. Es ist allerdings immer eine gute Idee, diese Suche auf mehreren Diensten durchzuführen. Wer die Browser-Erweiterung von InVID installiert hat, kann das für Bilder und Videos direkt mit zwei Klicks über verschiedene Dienste hinweg tun – ebenfalls per Rechts-Klick.

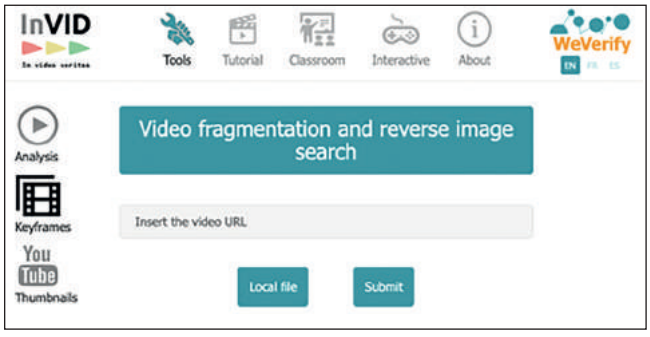
Nicht jede Suchmaschine ist für jeden Inhalt gleich gut geeignet. Das zeigt diese Übersicht von Domain-Tools, die die Stärken und Schwächen verschiedener Dienste auflistet:

	Identifizierte Elemente	Gesichter	Strukturen/ Gebäude	Orte	Logos	Andere Größen	Gespiegelt oder gedreht
Google	1	neutral	sehr gut	sehr gut	sehr gut	gut	neutral
Yandex	2+	sehr gut	sehr gut	sehr gut	gut	gut	gut
Bing	3+	gut	gut	gut	gut	neutral	sehr gut
TinEye	1	neutral	neutral	neutral	sehr gut	sehr gut	gut

Google eignet sich sehr gut für Gebäude, Orte und digitale Inhalte und Logos. Die russische Suchmaschine Yandex bringt sehr gute Ergebnisse für Gesichter sowie für Gebäude und Orte. Die Suchmaschine Bing von Microsoft erkennt am besten, ob ein Bild gedreht oder gespiegelt wurde. TinEye erreicht die besten Ergebnisse mit digitalen Inhalten, Logos und bei der Suche nach anderen Größen oder Auflösungen desselben Bildes.

InVID

InVID ist eine kostenfreie Browser-Erweiterung und die beste Plattform, um Videos zu analysieren und zu verifizieren. Sie erlaubt es Nutzerinnen und Nutzern, einen Link zu einem Video einzufügen oder ein Video hochzuladen, welches dann anschließend in einzelne Standbilder zerlegt wird. Mit diesen Bildern lässt sich dann wiederum eine Rückwärtssuche durchführen, um zu prüfen, wo im Netz das Video bereits aufgetaucht ist.



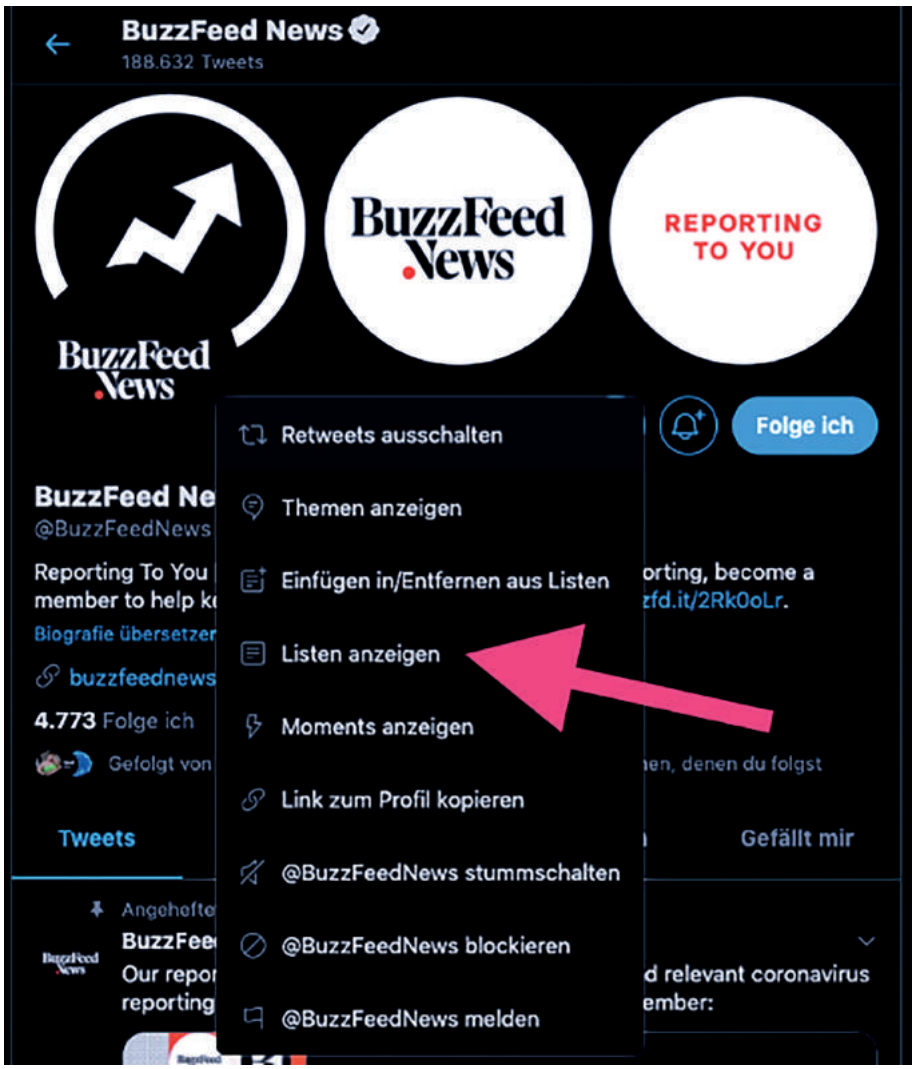
2. Twitter per TweetDeck durchsuchen

Der beste Weg, um Twitter zu durchsuchen, ist TweetDeck. Das Programm stellt verschiedene Suchen und Listen in einer Spalte dar. So lassen sich mehrere Suchen nebeneinander auf den Bildschirm bringen. Auch das Finden von Listen mit Konten, die zu einem bestimmten Thema kommunizieren, kann hilfreich sein. Twitter-Listen lassen sich über Google finden und zwar mit Hilfe einer einfachen Formel:

Einfach site:twitter.com/*/lists in das Suchfeld eintragen und anschließend das Schlagwort oder Ereignis, zu dem man Listen sucht, in Anführungsstrichen. Um nach Twitter-Listen mit Quellen für Eilmeldungen zu suchen, wäre die Suchformel folglich: site:twitter.com/*/lists "Eilmeldungen".

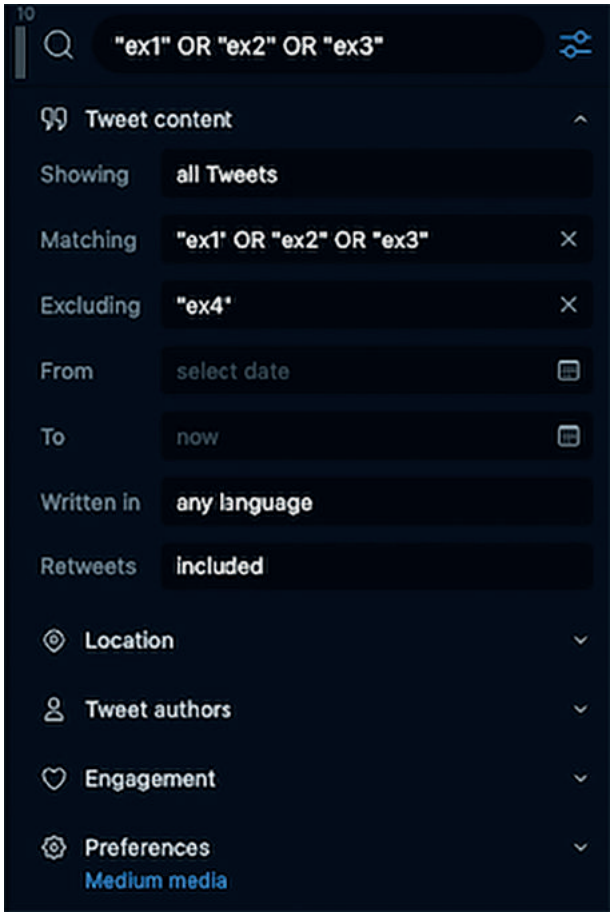
Mit dieser Formel können Listen gefunden werden, die andere Nutzer zuvor angelegt haben und die das Wort „Eilmeldungen" im Titel oder in der Beschreibung enthalten. Wurde eine solche Liste gefunden und für relevant befunden, kann man sie entweder verfolgen – oder man entscheidet sich, sie als Vorlage für eine eigene Liste zu nutzen. Dazu muss man die Liste zunächst duplizieren/kopieren, um sie danach in TweetDeck hinzufügen zu können. Für das Duplizieren der Liste kann diese Anwendung benutzt werden: http://projects.noahliebman.net/listcopy/connect.php

Darüber lassen sich unbegrenzt viele Listen duplizieren. Das Duplizieren ist in der Regel besser, als einer Liste zu folgen: Denn auf der duplizierten Liste kann man anschließend selbst Konten hinzufügen oder sie daraus entfernen.

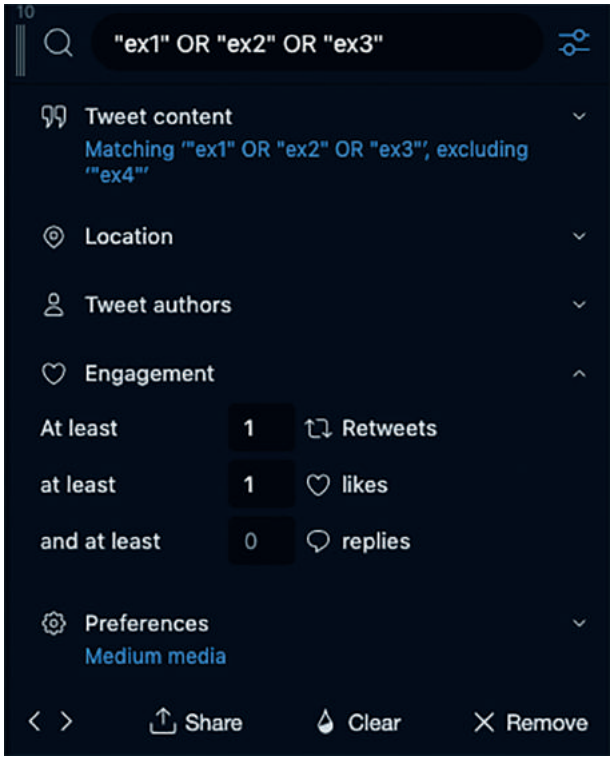


Hat man in Twitter ein Profil geöffnet, kann man sich mit einem Klick auf das Symbol mit den drei Punkten ein Menü anzeigen lassen. Darin befindet sich auch die Option „Listen anzeigen."

Neben dem Finden und Hinzufügen von Listen zu TweetDeck ist es ratsam, sich dort selbst neue Spalten anzulegen und diese mit bestimmten Suchfiltern auszustatten. Das ermöglicht es, schnell nach bestimmten Schlüsselbegriffen, Bildern oder Videos zu suchen. Ein hilfreiches Beispiel dafür ist, als Sucheinstellung mehrere Schreibweisen zu hinterlegen und diese mit einem „oder"-Befehl zu verbinden, indem man den Befehl OR dazwischensetzt. Die Suche lautet dann zum Beispiel „Kincade" OR „Kinkade". Andersherum lassen sich so auch bestimmte Begriffe aus einer Suche ausschließen. Da heute die meisten Nutzer ihre Beiträge nicht mehr mit dem Ort versehen, von wo aus sie twittern, kann man dieses Feld freilassen, um mehr Ergebnisse zu erhalten.



In TweetDeck lassen sich Suchen verfeinern. Dazu legt man zunächst eine neue Spalte an und definiert in der oberen Suchmaske, wonach man suchen möchte. Ein „OR" zwischen zwei Begriffen zeigt Ergebnisse an, die den einen ODER den anderen Begriff enthalten. Ein „AND" lässt nur Ergebnisse erscheinen, in denen beide Begriffe vorkommen.



Mit einem Klick auf das kleine Schieberegler-Symbol oben rechts öffnen sich weitere Einstellungen: Dort lassen sich zum Beispiel bestimmte Begriffe ausschließen, bestimmte Sprachen vorgeben, Retweets ein- oder ausschließen, Orte oder bestimmte Nutzer voreinstellen und Zeitrahmen festlegen.

Sollen die Ergebnisse noch weiter eingegrenzt werden, kann man den Zeitrahmen im Feld „Von" auf einen Tag oder zwei Tage vor dem Ereignis festlegen. Ein solcher Puffer sollte eingebaut werden, um keine Beiträge aufgrund von unterschiedlichen Zeitzonen zu verpassen. Wenn auch danach die Suche noch zu viele Ergebnisse hervorbringt, kann man eine Mindestanzahl von Interaktionen festlegen und sich damit nur jene Tweets anzeigen lassen, die ein gewisses Echo entfaltet haben. Man kann auch Schlagworte wie *„Eil" oder „Eilmeldung" oder „Breaking" oder „Breaking News"* usw. in eine Suche eintragen und fortan in einer separaten Spalte beobachten. So können Sie beispielsweise Orte in einer Spalte und andere Stichworte in einer anderen beobachten. Ich nutze in der Regel noch eine dritte Spalte, in der ich die möglichen Namen von Verdächtigen und Opfern inklusive möglicher falscher Schreibweisen verfolge. Wenn Sie dann eine große Anzahl von Tweets irgendwo registrieren und diese unübersichtlich wird, sollten Sie eine neue Spalte anlegen, dort die Zahl der Schlagworte verringern und in den Einstellungen dieser Spalte die Ergebnisse auf Tweets mit Fotos oder Videos beschränken. Das wird Ihnen dabei helfen, virales oder im Entstehen begriffenes Bildmaterial früh zu entdecken.

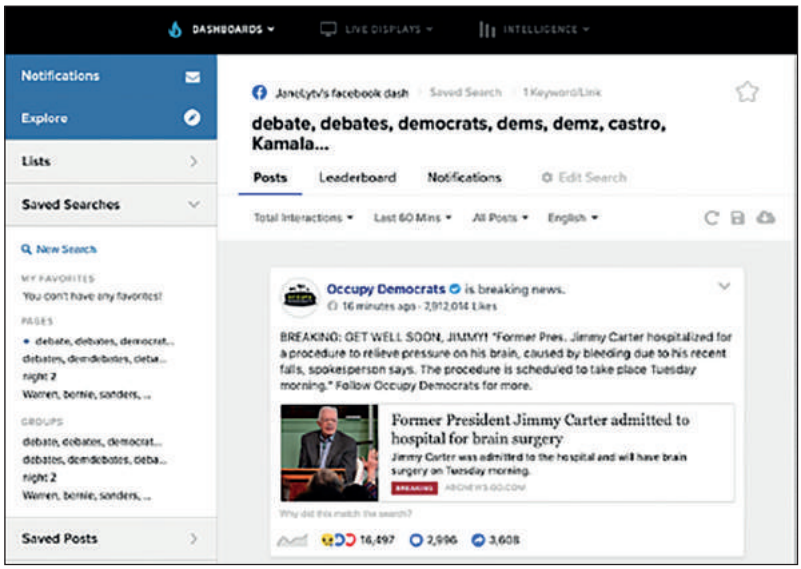
3. CrowdTangle

CrowdTangle ist eine Anwendung, die im Netz oder als Browser-Erweiterung arbeitet, für Redaktionen kostenlos ist und von Facebook angeboten wird. (Wir empfehlen, mit CrowdTangle Kontakt aufzunehmen, wenn noch kein Zugang besteht. Darüber hinaus werden online Kurse angeboten, um den Umgang mit CrowdTangle zu erlernen.) CrowdTangle ist ein mächtiges Werkzeug, mit dem man mehrere Suchen definieren und damit Inhalte über Facebook, Instagram und Reddit hinweg beobachten kann. Es bietet neben klassischer Stichwortsuche auch zahlreiche Filter, darunter Eingrenzungen in Bezug auf den Veröffentlichungstermin, die Sprache oder die Interaktionen.

CrowdTangle ist besonders gut geeignet, um auf Facebook zu prüfen, über welche Seiten oder Gruppen sich ein bestimmter Link oder irreführender Inhalt besonders weit verbreitet hat. Sobald Sie einen Zugang haben, besuchen Sie apps.crowdtangle.com und klicken Sie auf „Create New Dashboard", um eine neue Suche zu definieren. Ein Dashboard ist eine eigene Übersichtsseite, die jede Nutzerin und jeder Nutzer sich nach individuellen Bedürfnissen selbst konfigurieren kann. Wenn Sie noch keinen Zugang zu CrowdTangle haben, können Sie dennoch die Browser-Erweiterung nutzen – diese ist frei für jedermann und bietet Zugang zu diesem vielseitigen Recherche-Instrument.

CrowdTangle: Nach Facebook-Beiträgen suchen

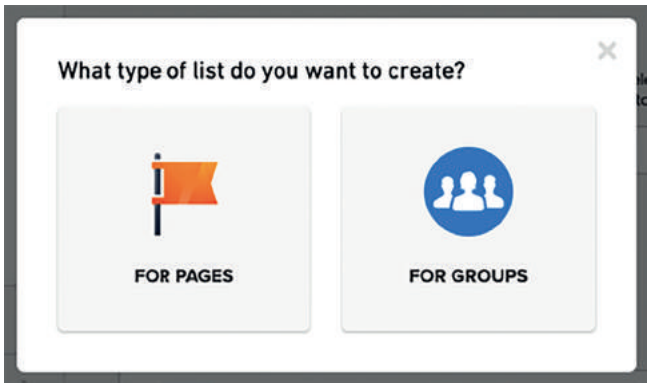
Öffnen Sie die Startseite von CrowdTangle und anschließend eines Ihrer Dashboards. Klicken Sie auf „Saved Searches" (gespeicherte Suchen) in der linken Leiste. Für Facebook werden dort zwei Optionen angeboten: Seiten oder Gruppen durchsuchen. Ich empfehle, beides zu tun. Tragen Sie so viele Suchbegriffe in die Suche ein, wie Sie für sinnvoll halten, und trennen Sie diese mit einem Komma voneinander ab. Anschließend können Sie einstellen, wie Sie die Ergebnisse sortiert bekommen: die jüngsten zuerst, die beliebtesten zuerst oder jene zuerst, die überdurchschnittlich erfolgreich waren (was auf einem Vergleich der Interaktionen mit einer vergleichbaren durchschnittlichen Seite beruht). Ich entscheide mich je nach konkreter Situation für die Option, die mir hilfreich erscheint, je nachdem, ob ich eher virale oder eher neue Inhalte suche. Außerdem können Suchergebnisse auf einen bestimmten Zeitraum oder Inhaltstyp eingegrenzt werden. Erst kürzlich hat CrowdTangle eine Suchoption hinzugefügt, mit der man Beiträge über den Ort der Seite finden kann, über die sie veröffentlicht wurden. Indem man auf „english" klickt und so die Spracheinstellung öffnet und anschließend auf „Country" für eine Übersicht der Länder lassen sich Suchergebnisse beispielsweise auf Beiträge beschränken, die auf Deutsch geschrieben und aus Österreich abgesendet wurden – oder zum Beispiel auch aus dem Iran, Russland, Saudi-Arabien, den Philippinen oder Indien. Achten Sie besonders auf Beiträge mit Bildern und Videos, da diese sich in der Regel weiter verbreiten. Sobald eine Ihrer Suchen für Sie relevante Ergebnisse hervorbringt, stellen Sie sicher, dass Sie die Sucheinstellungen abspeichern, so dass Sie diese auch später noch aufrufen können.



Das Tool CrowdTangle lässt Nutzer individuelle Suchen über die Plattformen Facebook, Instagram oder Reddit hinweg durchführen. Unter „Saved Searches" lassen sich einmal definierte Suchen abspeichern und später wieder öffnen. Die Ergebnisse lassen sich nach Anzahl der Interaktionen („Total interactions"), Alter („Last 60 Minutes"), Inhaltstyp (zum Beispiel Bilder oder Videos – im Beispiel hier ist mit „All Posts" eingestellt, dass alle Ergebnisse angezeigt werden) oder Sprache eingrenzen. Mit CrowdTangle lässt sich so nachverfolgen, wer in welchem Netzwerk welche Inhalte wie effektiv verbreitet hat oder wo diese ihren Ursprung nahmen.

CrowdTangle: Listen

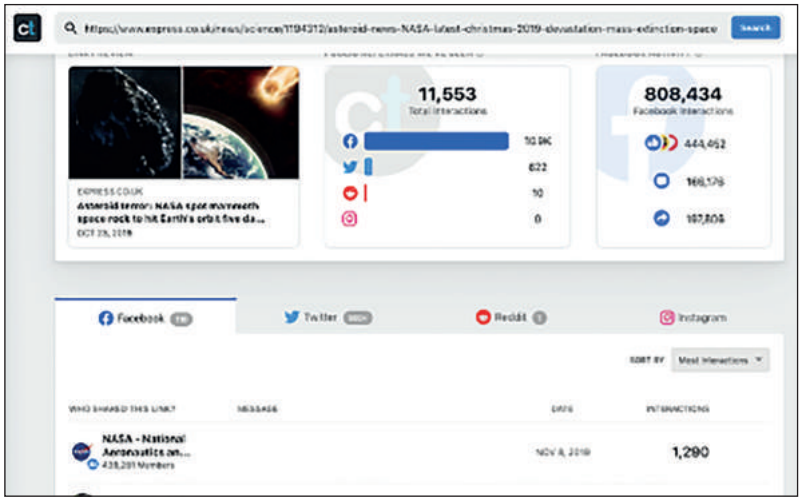
Genau wie bei TweetDeck lassen sich auch bei CrowdTangle Listen und Interessengruppen anlegen. Dazu klickt man im Menü auf der linken Seite auf „Lists" und anschließend auf „Create List". So können mehrere Seiten oder Gruppen beobachtet werden, die zu den gesuchten Stichworten passen oder deren Link man selbst dort hinzugefügt hat. Mit einem Klick auf die Option „Explore" lassen sich außerdem Listen anzeigen, die bereits von CrowdTangle erstellt wurden und öffentlich verfügbar sind. Genau wie bei Twitter auch ist das Anlegen und Beobachten von Listen ein hilfreicher und effektiver Weg, um die Gespräche im Netz zu gewissen Themen zu verfolgen. Sie bieten außerdem den Vorteil, dass man mit den Profilen der entsprechenden Nutzer nicht interagieren muss und folglich nicht ihre Aufmerksamkeit erregt.



Nach dem Anlegen einer neuen Liste für Facebook fragt CrowdTangle, ob es sich um eine Liste für Seiten (Pages) oder für Gruppen handeln soll.

CrowdTangle: Link-Suche

Eine weitere wichtige Funktion von CrowdTangle ist die Suche nach Links. Öffnen Sie dazu https://apps.crowdtangle.com/search/ und kopieren Sie den betreffenden Link oder die Schlagworte für das, wonach Sie suchen, in das Suchfeld. CrowdTangle wird Ihnen danach zeigen, wer den Link am erfolgreichsten geteilt hat und wer den größten Einfluss auf seine Verbreitung über Facebook, Instagram, Reddit oder Twitter nahm (wobei die Ergebnisse für Twitter auf die zurückliegenden sieben Tage beschränkt sind). Das ermöglicht es einzuschätzen, wie sich Inhalte verbreiten, ob es Gruppen oder Individuen gibt, die man sich näher ansehen sollte, und ob sich etwas bereits weit genug verbreitet hat, um einen „Debunking"-Artikel zu veröffentlichen, in dem falsche Behauptungen richtiggestellt werden. Ob dieser Zeitpunkt schon gekommen ist, dafür gibt es nicht die eine richtige Antwort. Es hängt vom konkreten Fall ab. Einige Leitfragen können sein: Hat sich die Falschmeldung schon über das Netzwerk ursprünglicher Autoren und Erstverbreiter hinaus verbreitet? Wurde es von Menschen mit einem gewissen Renommee oder einer größeren Reichweite verbreitet? Hat es Interaktionen in nennenswerter Größenordnung bekommen? Beides, die kostenlose Browser-Erweiterung wie auch die Online-Suchseite von CrowdTangle, gibt hier Einblicke und ist auch ohne CrowdTangle-Profil nutzbar.



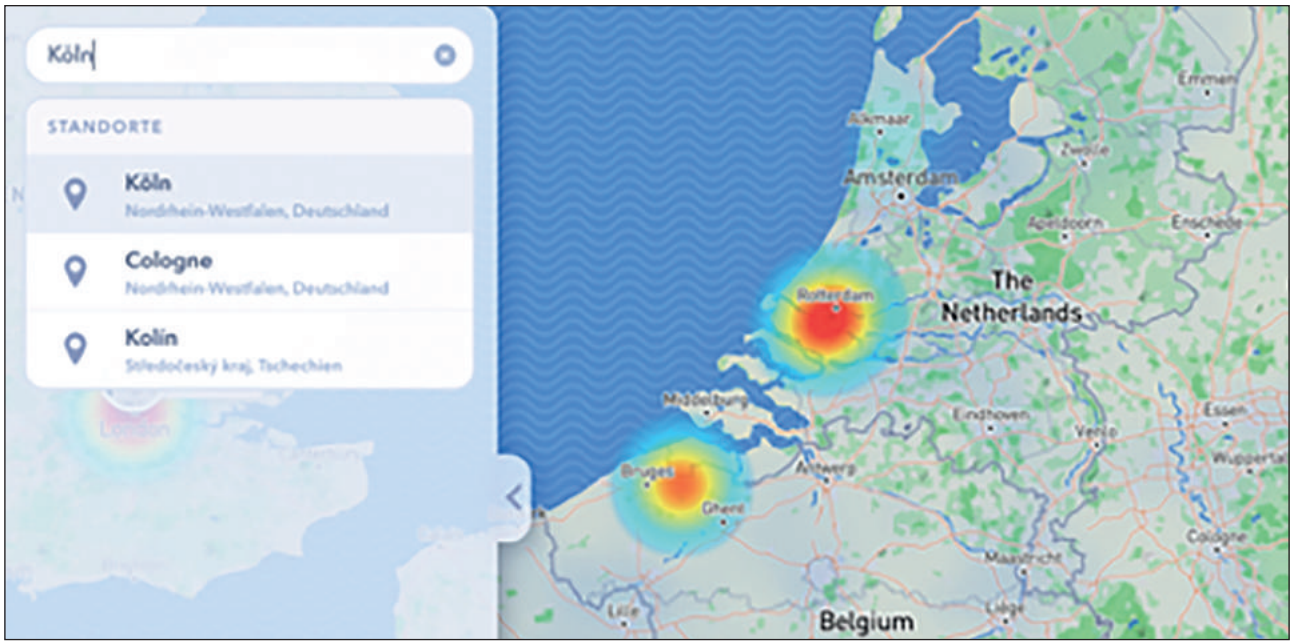
CrowdTangle zeigt an, wie stark sich ein Link über soziale Netzwerke verbreitet hat, wie viele Interaktionen er erhalten hat und welches die einflussreichsten Profile für die Verbreitung waren. In diesem Beispiel hat der Link 11.553 Interaktionen erhalten, davon 10.900 über Facebook, und am einflussreichsten für seine Verbreitung war die Seite der NASA.

4. Instagram.com

Instagram ist sehr hilfreich, um Hashtags und Beiträge mit einer Ortsmarkierung zu beobachten. Suchen Sie nach Örtlichkeiten in der Nähe des Ereignisses, das Sie interessiert, und lassen Sie sich die Einträge von Nutzern anzeigen, die diese Orte auf ihren Beiträgen markiert haben. Denken Sie daran, dass das auch benachbarte Stadtviertel oder Regionen sein können. Wenn Sie über diesen Weg auf ein Profil gestoßen sind, von dem Sie meinen, es sei für Ihre Suche interessant, schauen Sie sich das gesamte Profil an. Vergessen Sie dabei nicht, nicht nur die Beiträge des Nutzers, sondern auch seine Storys anzuschauen – diese kurzen Beiträge werden auf der Plattform nur 24 Stunden lang angezeigt und sind deutlich erfolgreicher als die klassischen Posts. Lesen Sie sich alle Kommentare durch, mitunter finden sich dort Beiträge weiterer Augenzeugen. Achten Sie darauf, ob Sie neue Hashtags finden, die Sie in Ihren Suchen noch nicht aufgenommen haben. Wenn Sie eine Story eines Nutzers archivieren wollen, können Sie das mit der Seite storysaver.net tun.

5. Snap Map

Auf Snapchat kommt es eher selten vor, dass Desinformation verbreitet wird. Die von Snapchat bereitgestellte Karte allerdings ist ziemlich hilfreich, um Schwindeleien gegenzuprüfen und zu entlarven. Sollte es beispielsweise in einer Metropole eine große Explosion gegeben haben, aber auf der Karte von Snapchat findet sich nichts dergleichen, wäre das ein möglicher Grund für Skepsis. Man erreicht sie über den Link map.snapchat.com und gibt dort anschließend die Region ein, in der man suchen möchte.



Über map.snapchat.com lässt sich eine Landkarte aufrufen, auf der in Farben dargestellt wird, wo gerade wie viele Nutzer etwas veröffentlicht haben. Blau und Grün stehen für wenige Beiträge, Gelb für mehrere und Rot für sehr viele.

Daraufhin erscheint eine Art Temperaturkarte, auf der eine große Anzahl von Beiträgen durch gelbe und rote Bereiche angezeigt wird – je intensiver eine Gegend eingefärbt ist, desto mehr Beiträge kommen von dort. Um einen gefundenen und hilfreichen Beitrag zu speichern, drücken Sie auf die drei Punkte oben rechts in der Ecke und wählen Sie „Snap teilen" aus. Sie können anschließend den angezeigten Link kopieren und abspeichern, um ihn später wieder aufzufinden. Zur Sicherheit sollten Sie auch Screenshots davon machen.

Indem man auf eine beliebige Region in der Karte drückt, öffnen sich hintereinander weg alle Beiträge, die von dort veröffentlicht wurden. Mit einem Klick auf die drei Punkte oben rechts in der Ecke kann der Link zu einem speziellen Beitrag eingesehen und gesichert werden.

Die Summe der einzelnen Teile

Es ist wichtig, jedes einzelne Werkzeug und jeden angebotenen Dienst in Ruhe kennenzulernen, bevor es ernst wird und sich die Eilmeldungen überschlagen. Desinformation ist so angelegt, dass sie mit den Emotionen spielt und Profit aus den Lücken in der Berichterstattung schlägt. Das sollte man sich bewusst machen, wenn man Inhalte im Netz durchstöbert.

Diese Arbeit ist auch eine Teamarbeit: Man kann über akkurate Inhalte stolpern, die anderen Kolleginnen und Kollegen helfen können. Daher sollte man sich notieren, was man weiß und was bestätigt wurde – so dass es leichter fällt, Falsches und Unbestätigtes schneller zu erkennen. Scheuen Sie im Gegenzug nicht davor zurück, Kolleginnen und Kollegen zu fragen, die vor Ort sind. Und nachdem sich der Nebel gelichtet hat, arbeiten Sie sich in Ruhe durch alle Bilder, Links und Screenshots, die Sie abgespeichert haben. Während es im Moment des Ereignisses eher darum gehen sollte, Falsches von Richtigem zu unterscheiden und das der Öffentlichkeit zugänglich zu machen, geht es in der Nachbereitung darum, Muster und Abläufe zu erkennen. Gab es Angriffe auf Menschen wegen ihrer Herkunft oder ihres Geschlechts? Haben es Lügen von kleinen, unbekannten Profilen in den Mainstream geschafft? Welche Plattform hat versagt, welche hat effektiv agiert? Eine solche Nachbereitung kann Ihren Leserinnen und Lesern helfen, den Zweck und die Methoden von Desinformations-kampagnen zu realisieren. Und Sie kann Ihnen, Ihren Kolleginnen und Kollegen und anderen Gruppen dabei helfen, beim nächsten Mal schon vorher zu wissen, worauf man achten sollte, wenn sich das nächste Mal alles überschlägt.

5. BILDER VERIFIZIEREN UND BEFRAGEN

von: Hannah Guy, Farida Vis, Simon Faulkner
deutsche Bearbeitung: Marcus Engert

Farida Vis *ist Direktorin des Visual Social Media Lab und Professorin für digitale Medien an der Manchester Metropolitan University. Ihre akademische und datenjournalistische Arbeit konzentriert sich auf die Verbreitung von Falschinformationen im Netz. Sie war Mitglied des Global Agenda Council on Social Media (2013–2016) sowie des Global Future Council for Information and Entertainment (2016–2019) des Weltwirtschaftsforums und ist Direktorin bei Open Data Manchester.*
Simon Faulkner *ist Dozent für Kunstgeschichte und visuelle Kultur an der Manchester Metropolitan University. Seine Forschung dreht sich um den politischen Nutzen und die politische Bedeutung von Bildern, mit einem Fokus auf Aktivismus und Protestbewegungen. Er ist außerdem Co-Direktor des Visual Social Media Lab und arbeitet an der Entwicklung von Methoden zur Analyse der Verbreitung von Bildern in sozialen Medien.*
Hannah Guy *ist Doktorandin an der Manchester Metropolitan University und forscht zur Rolle von Bildern bei der Verbreitung von Falschinformationen in sozialen Netzwerken. Sie ist Mitglied im Visual Social Media Lab. Ihre aktuellen Projekte beschäftigen sich mit den auf Twitter verbreiteten Bildern während der Black Lives Matter-Bewegung und mit visueller Medienkompetenz zur Bekämpfung von Fehlinformationen an kanadischen Schulen.*

Kommunikation in sozialen Medien ist geradezu überwältigend visuell. Fotos und Videos wirken überzeugend, sie sind einnehmend, man kann sie einfacher denn je erstellen, und sie können starke emotionale Reaktionen hervorrufen. In der Folge wurden sie zu mächtigen Verbreitungswegen für Des- und Falschinformation. Bis jetzt war die Diskussion über Bilder im Kontext von Des- und Falschinformation entweder auf Verifizierungstechniken fokussiert oder aber, verstärkt in jüngerer Zeit, auf sogenannte Deepfakes (Videos, bei denen künstliche Intelligenz benutzt wird, um eine Video- oder Audio-Aufnahme so zu manipulieren, dass es aussieht, als ob jemand etwas gesagt oder getan hätte, was aber in der Realität nie der Fall war). Bevor wir uns im nächsten Kapitel mit Deepfakes beschäftigen, ist es wichtig, grundsätzlich die Verwendung von irreführenden Bildern und Videos zu verstehen, besonders wenn sie aus dem Zusammenhang gerissen genutzt werden. Angesichts der weitverbreiteten Nutzung von Grafiken und Bildern in Versuchen, die öffentliche Meinung zu beeinflussen und zu manipulieren, müssen Journalistinnen und Journalisten über ein grundlegendes Wissen zur Verifikation von Bildern verfügen und über die Fähigkeit, Bilder kritisch zu befragen, um zu verstehen, von wo und wie sie verbreitet wurden. Dieses Kapitel beschäftigt sich mit dem zweiten Teil – und benutzt dafür ein Gerüst, das wir am Visual Social Media Lab entwickelt haben.

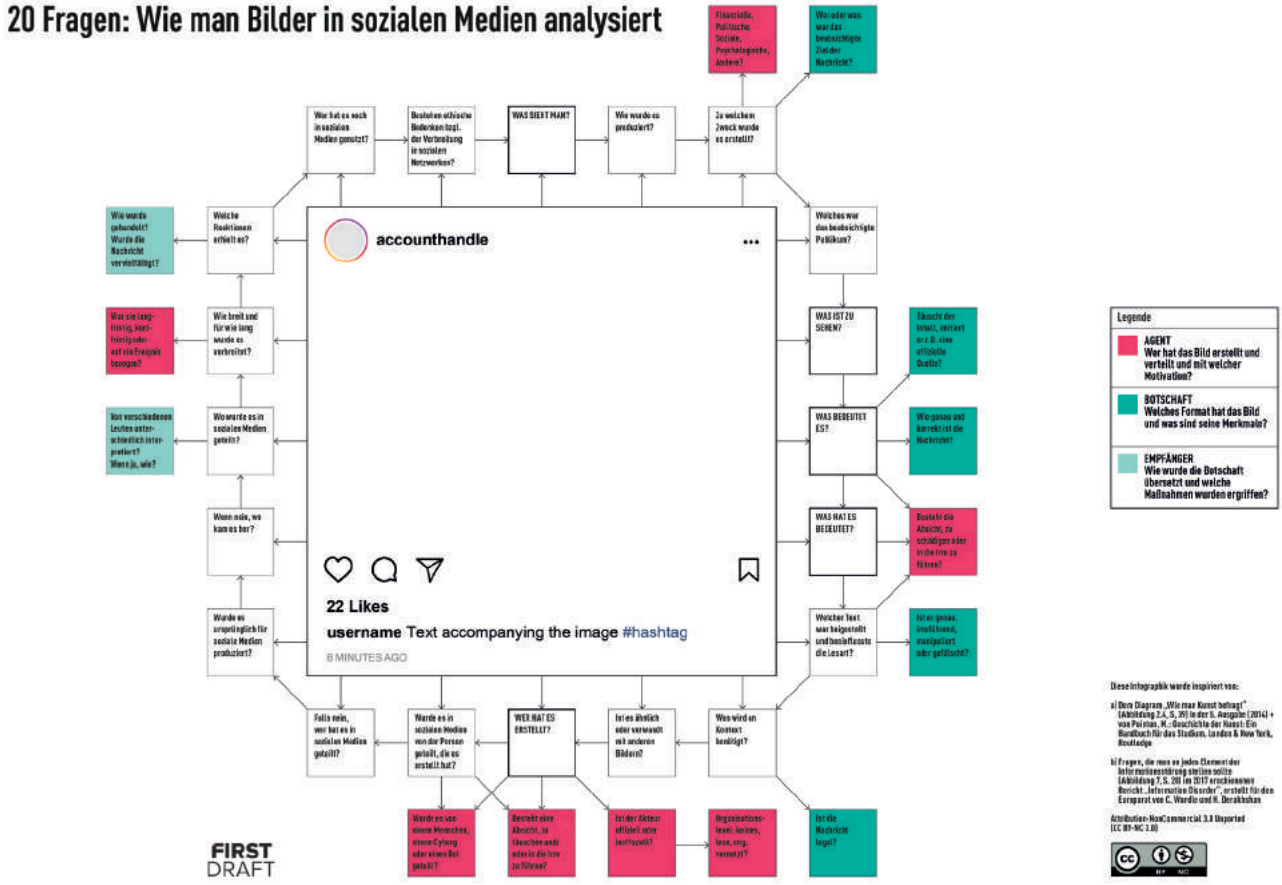
AUFBAUEN AUF VERIFIKATION

Im Visual Social Media Lab konzentrieren wir uns darauf, die Rollen zu verstehen, die Bilder im Netz in unserer Gesellschaft spielen. Obwohl das hauptsächlich Standbilder sind, umfasst das auch eine Reihe anderer Arten von Bildern: Fotos, Collagen, Meme, Grafiken oder Screenshots, um nur einige zu nennen. Die Bekämpfung von visueller Fehl- und Desinformation erfordert eigene Strategien. Bisher konzentrierte sich die Überprüfung von Bildern durch Journalisten darauf, zu prüfen, ob ein Bild das ist, wonach es aussieht. In der ersten Ausgabe dieses Handbuchs hat Trushar Barot vier zentrale Grundprinzipien für die Bildverifizierung skizziert, die nach wie vor von großem Wert sind. Auch der Visual Verification Guide von First Draft ist ein hilfreicher Leitfaden, der ähnliche Prinzipien bei der Bewertung von Bildern und Videos verfolgt und mit fünf Kernfragen arbeitet:

1. Sehen wir die Originalversion?
2. Wissen wir, wer das Foto gemacht hat?
3. Wissen wir, wo das Foto gemacht wurde?
4. Wissen wir, wann das Foto gemacht wurde?
5. Und wissen wir, warum das Foto gemacht wurde?

Die besten Hilfsmittel, um Fotos und Videos zu untersuchen, sind InVID sowie die Bilder-Rückwärtssuchen von Yandex, TinEye, Google und Forensically. Sie alle helfen bei der Suche nach dem Ursprung eines Bildes. Das ist natürlich von entscheidender Bedeutung. Doch bei der Vielzahl von Strategien und Techniken zur Fehl- und Desinformation und auch zur Medienmanipulation ist es ebenso wichtig zu fragen, wie und von wem Bilder verwendet und geteilt werden und welche Rolle Journalistinnen und Journalisten bei der möglichen weiteren Verbreitung problematischer Bilder spielen. Um über die klassischen Arten der Bilder-Verifikation hinauszukommen, haben wir Methoden aus der Kunstgeschichte mit Fragen kombiniert, die speziell für Inhalte mit Des- und Falschinformationen entwickelt wurden. Unser so gemeinsam mit First Draft und mit Journalistinnen und Journalisten entwickeltes Grundgerüst „20 Questions for Interrogating Social Media Images" (20 Fragen in Abbildung unten: Wie man Bilder in sozialen Medien analysiert) ist ein Werkzeug, das bei der Untersuchung von Bildern genutzt werden kann.

BILDER IN SOZIALEN MEDIEN UNTERSUCHEN



Das Gerüst aus 20 Fragen plus 14 weiterführenden Fragen, wie es vom Visual Social Media Lab gemeinsam mit First Draft entwickelt wurde.

Wie der Titel schon ahnen lässt, besteht unser Gerüst aus 20 Fragen, die man über jedes Bild in sozialen Medien stellen kann, sowie 14 zusätzlichen Fragen, die darauf zielen, tiefer in die verschiedenen Aspekte von Des- und Falschinformation vorzudringen. Die Fragen haben keine vorgegebene Reihenfolge, aber es macht Sinn, diese fünf zuerst zu beantworten:

1. Was sieht man?
2. Was ist zu sehen?
3. Wer hat es erstellt?
4. Was hat es bedeutet?
5. Was bedeutet es?

Die Fragen 1 bis 3 ähneln den klassischen Abläufen bei der Verifikation eines Bildes und haben das Ziel zu erfassen, welche Sorte Bild vorliegt (Fotografie, Video-Aufnahme etc.), was es darstellt und wer der Urheber ist. Die Fragen 4 und 5 allerdings leiten uns in eine andere Richtung. Sie führen Überlegungen zur Bedeutung ein, die sich auf das beziehen, was das Bild zeigt, aber auch auf alle Bedeutungen, die durch die Verwendung des Bildes erzeugt werden, auch durch seine Fehlinterpretation. Zusammen betrachtet weisen die Fragen 4 und 5 damit auch auf die sich verändernde Natur der Bedeutung von Bildern hin sowie auf die Art und Weise, wie Bildern beim Wiederverwenden neue Bedeutungen zugeschrieben werden können. Dabei geht es nicht allein darum zu fragen, welchen neuen Kontext Bilder für eine Bedeutung bekommen sollen und wie dadurch das, was sie zeigen, als falsch identifiziert wird. Es geht auch darum zu hinterfragen, was die Folgen solcher Fehlidentifikationen sind. Mit diesem Ansatz geht es also nicht mehr nur um Verifizierung, sondern auch um die Analyse der Bedeutungen von Bildern, wie sie in Disziplinen wie der Kunstgeschichte oder der Theorie der Fotografie angewandt wird.

Als wir dieses Gerüst mit Journalistinnen und Journalisten entwickelt und erprobt haben, hörten wir oft, dass sie noch nie so intensiv über ein Bild nachgedacht hätten. Viele sagten, das Gerüst habe ihnen dabei geholfen zu erkennen, dass Bilder komplexe Formen der Kommunikation sind und dass eine klare Methodik nötig ist, um sie auf ihre Bedeutung hin zu befragen.

In den allermeisten Fällen wird es nicht nötig sein, alle 20 Fragen zu beantworten, um ein umfassendes Verständnis dessen zu erhalten, was mit einem Bild geschieht. Die Fragen sind also kein Pflichtkatalog, sondern eher eine Art Sicherheitsnetz. In unserer eigenen Arbeit fanden wir sie dann besonders hilfreich, wenn wir uns mit komplexen Nachrichtenbildern und -videos befassten, die schnell große Aufmerksamkeit erhalten haben. Um zu illustrieren, wie das in der Praxis aussieht, beschreiben wir im Folgenden drei Fallbeispiele mit komplexen Situationen aus Großbritannien und den USA.

Fallbeispiel 1: Belastungsgrenze, Juni 2016

Auf dem Bild steht: „Belastungsgrenze. Die EU hat uns alle im Stich gelassen. Wir müssen uns von der EU lösen und die Kontrolle über unsere Grenzen zurückgewinnen."

Was sieht man?
Das Bild „Belastungsgrenze" war ein Werbeposter der britischen Unabhängigkeitspartei UKIP während der Kampagne zum EU-Referendum 2016. Es nutzte ein Foto, das der Fotojournalist Jeff Mitchell im Oktober 2015 während der großen Flüchtlingsbewegungen nach Europa gemacht hatte.

Was ist zu sehen?
Eine Schlange von syrischen und afghanischen Flüchtlingen, die im Grenzgebiet zwischen Kroatien und Slowenien von der slowenischen Polizei in das Flüchtlingslager Brežice eskortiert werden. Das Poster nutzte nur einen Bildausschnitt und

fügte in großen Buchstaben das Wort „BELASTUNGSGRENZE“ (BREAKING POINT) hinzu. Darunter steht: „Die EU hat uns alle im Stich gelassen“ sowie „Wir müssen uns von der EU lösen und die Kontrolle über unsere Grenzen zurückgewinnen“. Durch den gewählten Bildausschnitt hat es den Anschein, als ob sich die Flüchtlinge in großer Masse auf den Betrachter zubewegen, was eine starke visuelle Wirkung hat.

Wer hat es erstellt?
Die in Edinburgh ansässige Werbefirma Family Advertising Ltd., die von der UKIP für deren Brexit-Kampagne beauftragt worden war.

Was hat es bedeutet?
Die UKIP versuchte nicht, den Inhalt falsch darzustellen, fügte aber durch die Slogans zusätzliche Bedeutungsschichten hinzu. Diese Manipulation dockte an bestehende einwanderungsfeindliche und rassistische Stimmungen an und konzentrierte sich auf der Grundlage unbegründeter Behauptungen und Unterstellungen zur EU-Grenzpolitik auf die Erzeugung größerer Ängste vor Einwanderung und Flüchtlingen.

Was bedeutet es?
Im November 2019, im Vorfeld der Parlamentswahlen in Großbritannien, nutzte die Kampagne Leave.EU (die für den Austritt aus der EU kämpfte) ebenfalls den gleichen Bildausschnitt des Fotos in einer Grafik, die auf Twitter verbreitet wurde, und stellte damit einen klaren Bezug zu dem Poster der UKIP aus 2016 her.

Welche anderen Fragen könnten hilfreich sein?

Ist der Akteur offiziell oder inoffiziell?
Die Schlüsselfigur für die Herstellung und Verbreitung des Bildes, UKIP, ist eine offizielle politische Partei und damit keiner der Akteure, die normalerweise mit Falsch- oder Desinformation in Verbindung gebracht werden.

Ist das Bild ähnlich oder verwandt mit anderen Bildern?
Manche stellten einen Bezug zu Nazipropaganda her; es gibt sowohl eine Referenz zu früheren migrantenfeindlichen Bildern als auch zu einer längeren Geschichte britischer politischer Plakate mit Schlangen von Menschen, darunter eines, das von der UKIP im Mai 2016 verwendet wurde und sich auf die Einwanderung aus der EU konzentrierte.

Drei zentrale Punkte:

- Offizielle politische Parteien und Politiker können Akteure bei der Verbreitung von Falschinformationen sein.
- Falschinformationen müssen nicht zwingend gefälschte Bilder oder die falsche Identifikation des Inhalts sein. Mitunter können Bilder genutzt werden, um Botschaften zu transportieren, die eine größere Situation falsch oder verzerrt darstellen.
- Manche Falschinformationen brauchen mehr als nur Verifikation. Es gibt einen Bedarf daran, kritisch zu untersuchen, wie reale Bilder für Manipulation genutzt werden und was ein so genutztes Bild bedeutet.

Fallbeispiel 2: Foto von der Westminster Bridge, März 2017

Der Tweettext lautet: „Muslima kümmert sich nicht um Terrorangriff, läuft lässig an einem sterbenden Mann vorbei und schaut dabei auf ihr Telefon"

Ein Tweet eines Twitter-Kontos, das anscheinend von einem weißen Texaner betrieben wird und der in den Medien große Beachtung fand. Wie sich später herausstellte, wurde das Konto von der russischen Internet Research Agency (IRA) betrieben und zur Verbreitung von Falsch- und Desinformationen genutzt. (Die IRA gilt als „Trollarmee" Russlands und soll sich schon seit Jahren in Wahlen im Ausland einmischen.) In dem Tweet wurde ein Foto geteilt, das nach dem Terroranschlag auf die Westminster Bridge in London am 22. März 2017 entstanden war.

Was ist zu sehen?
Eine muslimische Frau geht an einer Gruppe von Menschen vorbei, die um eine am Boden liegende Person herumstehen, welche bei dem Terroranschlag verletzt wurde. Der Text hat eine islamfeindliche Konnotation: Er behauptet, dass die Frau die verletzte Person absichtlich ignoriert habe, und nutzt einen offenkundig gegen den Islam gerichteten Hashtag.

Wer hat es erstellt?
Mitarbeiter der Internet Research Agency (IRA), die das Twitter-Konto @SouthLoneStar betreuen, obschon zum Zeitpunkt des Tweets nicht bekannt war, dass das Konto zur IRA gehört. Das Foto selbst wurde vom Pressefotografen Jamie Lorriman gemacht.

Was hat es bedeutet?
Im März 2017 schien der Tweet von einem rechten Twitter-Nutzer aus Texas zu kommen, der das Foto so interpretierte, dass die muslimische Frau sich nicht um die verletzte Person kümmern würde. Es sollte den Anschein erwecken, als ob dieses konkrete Beispiel für eine größere, allgemeinere Wahrheit über Muslime stehe.

Was bedeutet es?
Der Tweet ist bis heute ein Beleg dafür, dass die russische IRA nach Terrorattacken gezielt islamophobe Desinformation verbreitet.

Welche anderen Fragen könnten hilfreich sein?

Welche Reaktionen gab es darauf?
Der Tweet hat viel Aufmerksamkeit von großen Medien erhalten und es so in den Mainstream geschafft. Dutzende britische Zeitungen berichteten darüber, mitunter mehrmals. Auch wenn die meisten Artikel den Ersteller des Tweets dafür

verurteilten, machten sie den Tweet damit auch einem breiteren Publikum außerhalb des sozialen Netzwerks zugänglich. Nachdem sich das Foto verbreitet hatte, äußerte sich die Frau auf dem Foto. Sie sei über die damaligen Anschläge verzweifelt und am Boden zerstört gewesen und sehe sich nicht nur mit den Folgen eines Terroranschlags konfrontiert, sondern auch damit, dass soziale Medien mit ihrem Bild zugepflastert wurden „von jenen, die nur meine Kleidung sehen und ihre Schlussfolgerungen auf Hass und Fremdenfeindlichkeit stützen."

Ist es ähnlich mit anderen Bildern?
Das Bild ist eines von sieben Bildern der Frau, jedoch das am weitesten verbreitete. Andere Bilder zeigen deutlich, dass die Frau bestürzt war, was jedoch nur wenige Redaktionen aufgriffen.

Wie weit und für wie lange verbreitete sich das Bild?
Die zusätzliche Aufmerksamkeit durch große Redaktionen sorgte dafür, dass sich der Tweet weit verbreitete. Allerdings nahm diese Verbreitung nach ein paar Tagen kontinuierlich ab. Im November 2017, als bekannt wurde, dass das Konto @SouthLoneStar von der russischen Internet Research Agency betrieben wurde, nahm die Verbreitung nochmals zu. Diese zweite Welle der Verbreitung blieb jedoch deutlich beschränkter als die erste im März.

Drei zentrale Punkte:

- Visuelle Falschinformation ist nicht immer komplett falsch und kann Elemente in sich tragen, die wahr sind. Die Fotografie war real, aber ihr Kontext wurde manipuliert und verfälscht. Man nutzte den Umstand aus, dass die Menschen nicht wissen konnten, was diese Frau in dem Moment wirklich dachte und tat.
- Journalistinnen und Journalisten sollten also sorgfältig abwägen, ob sie zusätzliche Aufmerksamkeit auf solch emotional aufgeladene, kontroverse, potentiell schädliche Desinformation lenken, indem sie darüber berichten, auch wenn das mit guten Absichten geschieht.
- Es könnte mehr Aufmerksamkeit darauf gelegt werden, auf Falschinformationen beruhende Berichterstattung zu korrigieren und sicherzustellen, dass ein zutreffendes Bild und eine zutreffende Interpretation der Geschehnisse verbreitet werden. Die begrenzte Verbreitung im November bedeutet auch, dass einige Leserinnen und Leser nicht mitbekommen haben, dass es sich bei dem Tweet um aus Russland gesteuerte Desinformation handelte.

Fallbeispiel 3: die Konfrontation am Lincoln Memorial, Januar 2019

Die Szene einer Konfrontation zwischen jungen Trump-Unterstützern und Vertretern der indigenen Bevölkerung wurde zunächst auf Instagram verteilt, dann bei Twitter verkürzt wiedergegeben und verbreitete sich dort, auch unter Beteiligung großer Redaktionen, rasend schnell.

Was sieht man?
Ein Video einer Schülergruppe der Katholischen Covington High School, die an einem Marsch von Abtreibungsgegnern teilnehmen, sowie einen Angehörigen der indigenen Bevölkerung, Nathan Phillips, der mit anderen Angehörigen der Indigenen an einem eigenen Marsch teilnimmt.

Was ist zu sehen?
Eine Konfrontation zwischen einem der Schüler und Phillips. Beide Demonstrationszüge sind auf einem Platz zusammengetroffen. Eine größere Gruppe Schüler mit „Make America Great Again"-Basecaps – dem Slogan der Kampagne von Donald Trump – steht Phillips gegenüber. Das erzeugt den Eindruck eines Einzelkämpfers für die Indigenen Amerikas, der sich einem Aufgebot junger Neu-Rechter entgegenstellt.

Wer hat es erstellt?
Das zugehörige Video war zunächst von einem Teilnehmer des Marsches der Indigenen auf Instagram veröffentlicht worden. Dort erhielt es um die 200.000 Aufrufe. Stunden später wurde das Video nochmals hochgeladen, diesmal jedoch auf Twitter. Hier erhielt es in kürzester Zeit 2,5 Millionen Aufrufe, bevor es vom Inhaber des Profils, der es hochgeladen hatte, selbst wieder gelöscht wurde. Das Video wurde über verschiedene soziale Netzwerke hinweg weiterverbreitet, erhielt auch die Aufmerksamkeit von Redaktionen und gelang so in den Mainstream. Innerhalb von 24 Stunden wurden zahlreiche Artikel zu dem Video veröffentlicht.

Was hat es bedeutet?
Die ursprünglich dem Video beigestellte Erzählung stellte die Szene als eine direkte Konfrontation zwischen Phillips und den Schülern dar, wobei den Schülern die Absicht einer vorsätzlichen Verhöhnung von Phillips angedichtet wurde.

Was bedeutet es?
Eine deutlich längere Version des Videos tauchte einige Tage später auf und zeigte ein komplexeres Bild. So befand sich ebenfalls eine Gruppe der religiösen Splittergruppe „Schwarze Hebräer" vor Ort, die Passanten verspotteten, darunter auch die Schüler und die Indigenen. Das führte zu einer erhitzten Auseinandersetzung zwischen allen drei Gruppen, wobei Phillips wohl zu schlichten versuchte.

Welche anderen Fragen könnten hilfreich sein?

Was sollte man über den Kontext wissen?
Ohne das längere Video und ohne das Wissen, dass die „Schwarzen Hebräer" auch vor Ort waren und den Konflikt aktiv anheizten, geht jeder Kontext verloren. Es stimmt, dass es Aufnahmen gibt, in denen die Schüler sich rassistisch äußern. Jedoch ist die Entstehungsgeschichte dieser Situation komplexer als das simple Bild von rechten Teenagern, die einen älteren Indigenen beschimpfen.

Wo wurde es in sozialen Netzwerken geteilt?
Als das Video auf Instagram von einem Teilnehmer der Demonstration der Indigenen geteilt wurde, erhielt es zwar Aufmerksamkeit, aber nur in begrenztem Rahmen. Es wurde anschließend auf Twitter und YouTube von anderen Nutzern neu hochgeladen, was ihm große Aufmerksamkeit brachte und dazu führte, dass es von Redaktionen aufgegriffen wurde. Die Aufmerksamkeit entstand durch das Neuhochladen und nicht durch das ursprüngliche Video auf Instagram.

Drei zentrale Punkte:

- Wenn sich derart emotionsgeladenes Bildmaterial so schnell verteilt, dann ist es einfach, den Kontext aus dem Blick zu verlieren. So können sich leicht irreführende, reaktionäre Narrative durchsetzen.

- Rückblickend haben einige Journalisten argumentiert, die ersten Artikel trügen eine Mitschuld daran, dass sich die Debatte so schnell verbreiten konnte. Das legt nahe, dass auch Mainstream-Medien unbeabsichtigt bei der Verbreitung von Falschinformationen mitwirken können.

- Aufgrund der Geschwindigkeit, mit der sich das Video verbreitete, übernahmen viele Redaktionen die falschen Erzählungen und recherchierten nicht mehr eigenständig. Viele Nachrichtenwebsites mussten ihre Artikel später umändern oder korrigieren, einige wurden juristisch belangt.

SCHLUSSFOLGERUNGEN

Vieles, was in sozialen Netzwerken geteilt wird, ist visuell. Journalistinnen und Journalisten müssen deshalb die Fähigkeit haben, Bilder kritisch zu hinterfragen und zu bewerten, um deren Inhalte und Absichten aufzudecken. Gerade die Schnelligkeit, mit der sich visuelle Fehlinformationen verbreiten können, macht deutlich, dass Journalisten mit Vorsicht vorgehen und sicherstellen müssen, dass bildbezogene Geschichten vor der Veröffentlichung umfassend untersucht werden. Unsere 20 Fragen zur Untersuchung von Bildern in sozialen Medien sind ein zusätzliches Hilfsmittel, das sie verwenden können, insbesondere wenn die Geschichte in erster Linie auf etwas Visuelles ausgerichtet ist. Nicht jede Frage unseres Gerüsts ist für jedes Bild relevant, aber die fünf Grundfragen sind ein hilfreicher Ausgangspunkt und bauen auf grundlegenden Verifikationsfähigkeiten auf, mit dem Ziel, eine genauere und tiefergehende Berichterstattung zu entwickeln.

ANHANG

Hier die vollständige Liste der 20 Fragen, inklusive 14 Ergänzungsfragen, die sich speziell auf Des- und Falschinformation beziehen. Wie oben bereits ausgeführt, gibt es fünf Fragen (fett gedruckt), von denen wir denken, es macht Sinn, sie zuerst zu bearbeiten.

Die Erweiterungsfragen für Falsch- und Desinformation beziehen sich auf eine dieser drei Kategorien:

AGENT: Wer hat das Bild erstellt und verteilt und mit welcher Motivation?
BOTSCHAFT: Welches Format hat das Bild und was sind seine Merkmale?
EMPFÄNGER: Wie wurde die Botschaft übersetzt und welche Maßnahmen wurden ergriffen?

1. **Was sieht man?**

2. Wie wurde es produziert?

3. Zu welchem Zweck wurde es erstellt?
   a. A: finanzieller, politischer, sozialer, psychologischer, anderer?
   b. B: Wer oder was war das beabsichtigte Ziel der Nachricht?

4. Welches war das beabsichtigte Publikum?

5. **Was ist zu sehen?**

6. **Was bedeutet es?**
   a. B: Täuscht der Inhalt, imitiert er zum Beispiel eine offizielle Quelle?
   b. B: Wie genau und korrekt ist die Nachricht?

7. Was hat es bedeutet?
   a. A: Besteht die Absicht, zu schädigen oder in die Irre zu führen?

8. Welcher Text war beigestellt und beeinflusste die Lesart?
   a. B: Ist er genau, irreführend, manipuliert oder gefälscht?

9. Was wird an Kontext benötigt?
   a. B: Ist die Nachricht legal?

10. Ist es ähnlich oder verwandt mit anderen Bildern?

11. **Wer hat es erstellt?**
    a. A: Ist der Akteur offiziell oder inoffiziell?
    b. A: Organisationslevel: keines, lose, eng, vernetzt?

12. Wurde es in sozialen Medien von der Person geteilt, die es erstellt hat?
    a. A: Wurde es von einem Menschen, einem Cyborg oder einem Bot geteilt?
    b. A: Besteht eine Absicht, zu täuschen und/oder in die Irre zu führen?

13. Falls nein, wer hat es in sozialen Medien geteilt?

14. Wurde es ursprünglich für soziale Medien produziert?

15. Wenn nein, wo kam es her?

16. Wo wurde es in sozialen Medien geteilt?
    a. E: von verschiedenen Leuten unterschiedlich interpretiert? Wenn ja, wie?

17. Wie breit und für wie lang wurde es verbreitet?
    a. B: langfristig, kurzfristig oder auf ein Ereignis bezogen?

18. Welche Reaktionen erhielt es?
    a. E: Wie wurde gehandelt? Wurde die Nachricht vervielfältigt?

19. Wer hat es noch in sozialen Medien genutzt?

20. Bestehen ethische Bedenken bezüglich der Verbreitung in sozialen Netzwerken?

Unser Fragengerüst wurde inspiriert von:
1. Dem Diagramm „Interrogating the work of Art“ von Marcia Pointon
2. Dem Fragenkatalog „Questions to ask about each element of an example of information disorder“ von Claire Wardle und Hossein Derakhshan

6. NACHDENKEN ÜBER DEEPFAKES UND NEUE MANIPULATIONSTECHNIKEN

von: Sam Gregory
deutsche Bearbeitung: Marcus Engert

***Sam Gregory** ist Programmdirektor von WITNESS (www.witness.org) in New York. Die Plattform möchte Menschen dabei helfen, mit Wissen über Videos und Analysetechniken Menschenrechte zu verteidigen. Der mehrfach ausgezeichnete Technologe und Vordenker gilt als Experte für neue Formen von Des- und Falschinformation mittels künstlicher Intelligenz. Als solcher leitet er Projekte, die sich mit neuen Möglichkeiten und Gefahren für den Journalismus befassen. Er ist außerdem Co-Vorsitzender der Partnership on AI, einer Organisation, die Wissenschaftler, Unternehmen und zivilgesellschaftliche Organisationen zusammenbringen will, um die Auswirkungen von künstlicher Intelligenz (Artificial Intelligence) besser zu verstehen; er ist dort Mitglied der Expertengruppe für künstliche Intelligenz und Medien.*

Im Sommer 2018 veröffentlichte Professor Siwei Lyu, ein führender Wissenschaftler für Deepfakes an der Universität Albany, einen Aufsatz. Darin schrieb er, dass Personen in Deepfake-Videos nicht in derselben Frequenz blinzeln würden wie reale Menschen. Die Behauptung wurde schnell von Redaktionen aufgegriffen, Fast Company, New Scientist, Gizmodo, CBS News und andere berichteten darüber und brachten viele Menschen dazu, zu glauben, es gäbe endlich einen sicheren Weg, um Deepfakes zu erkennen. Nur wenige Wochen nach der Veröffentlichung erhielt der Wissenschaftler Videos mit Deepfake-Personen, die exakt wie Menschen blinzeln. Heute gilt dieser Ansatz weder als hilfreich noch als genau. Zum Zeitpunkt der Veröffentlichung des Papiers war Blinzeln noch die Achillesferse jenes Algorithmus, der Deepfakes berechnete. Nur wenige Monate später nicht mehr.

Das illustriert einen der zentralen Punkte bei der Suche und Verifizierung von Deepfakes: Technische Nachweisverfahren sind nützlich, jedoch sind sie es nur so lange, bis sie von den Produktionsverfahren eingeholt werden. Ein perfektes System zur Erkennung von Deepfakes und synthetischen Medien wird es nie geben. Wie also sollten Journalistinnen und Journalisten Deepfakes und andere Formen synthetischer Medien verifizieren?

Der erste Schritt ist es, zu verstehen, dass dieses Feld ein Katz-und-Maus-Spiel ist. Man muss aufmerksam verfolgen, wie sich die Technologie fortentwickelt. Zweitens müssen Journalistinnen und Journalisten lernen, fundamentale Techniken und Werkzeuge der Verifikation für ihre Recherchen einzusetzen, egal ob reales Material absichtsvoll manipuliert oder synthetisches Material im Computer errechnet wurde. Dafür können sowohl all die Verfahren und Methoden genutzt werden, die in der ersten Ausgabe des Verification Handbooks erklärt wurden, als auch jene, die im Visual Verification Guide von First Draft aufgezeigt sind. Schlussendlich müssen Journalistinnen und Journalisten auch verstehen, dass wir in einem Stadium angekommen sind, in dem auch Inhalte zunehmend fälschlicherweise als Deepfakes bezeichnet werden, die es gar nicht sind. Das wiederum bedeutet: Fähigkeiten, mit denen man ein Bild oder Video auf seine Echtheit hin überprüfen kann, sind inzwischen genauso wichtig wie Fähigkeiten, mit denen man Manipulationen an echten Bildern nachweisen kann. In diesem Kapitel soll es um einige Verfahren gehen, mit denen man das tun kann. Doch zunächst ist es wichtig, ein grundlegendes Verständnis von Deepfakes und synthetischen Medien zu bekommen.

WAS SIND DEEPFAKES UND SYNTHETISCHE MEDIEN?

Hinter Deepfakes stehen neue Formen audiovisueller Manipulation, mit denen realistische Simulationen von Gesichtern, Stimmen oder Bewegungen erstellt werden können. Sie versetzen Menschen in die Lage, es so aussehen zu lassen, als ob jemand etwas gesagt oder getan hätte, was jedoch tatsächlich nie geschehen ist.

Es wird immer leichter, Deepfakes zu erstellen. Und um sie zu errechnen, braucht es immer weniger Rohmaterial als Grundlage. Darüber hinaus werden sie zunehmend kommerzialisiert. In jüngerer Zeit haben sie mitunter furchtbare Folgen für Frauen, da Deepfakes auch dafür genutzt werden, ohne jeweilige Zustimmung sexualisierte Bilder und Videos mit dem

Gesicht einer Person zu erstellen. Es gibt zudem Bedenken, dass Deepfakes einen größeren Einfluss auf die Nachrichtenproduktion, auf Verifikationsprozesse und auf unsere Gesellschaft im Ganzen bekommen könnten. Deepfakes sind nur eine Gattung innerhalb einer ganzen Familie von Techniken, die mittels künstlicher Intelligenz synthetische Medien generieren können. Diese Werkzeuge und Techniken erlauben die Erstellung von realistischen Simulationen von Menschen davon, wie sie etwas sagen oder tun, was sie im realen Leben nie gesagt oder getan haben. Sie erlauben sogar die glaubwürdige Erschaffung von Menschen oder Objekten, die nie existierten, oder auch von Veranstaltungen, die nie stattfanden. Die technischen Möglichkeiten für Manipulationen durch synthetische Medien umfassen aktuell:

- Objekte aus einem Video entfernen oder sie hinzufügen.
- Hintergründe und Umgebungen in einem Video verändern. Man kann beispielsweise ein Video aus dem Sommer so aussehen lassen, als ob es im Winter aufgenommen worden wäre.
- Die Lippen, den Gesichtsausdruck und die Körperbewegungen eines bestimmten Individuums in einer realistischen Video-Simulation nachahmen oder kontrollieren. Obwohl die Diskussion über Deepfakes sich generell auf Gesichter fokussiert, können ähnliche Techniken bereits heute auf den ganzen Körper oder auch nur auf einzelne Teile des Gesichts angewandt werden.
- Eine realistische Simulation der Stimme einer Person erstellen.
- Eine existierende Stimme modifizieren, ihr zum Beispiel ein anderes Geschlecht geben oder sie mit der Stimme einer anderen Person bestimmte Dinge sagen lassen.
- Ein realistisches, aber vollkommen erfundenes Foto einer Person erstellen. (Die dahinterliegende Technologie kann auch für weniger problematische Bereiche genutzt werden, wie zum Beispiel Werbung, Unterhaltung oder Forschungszwecke.)
- Ein realistisches Gesicht, das von einer Person stammt, auf eine andere Person übertragen. (Das ist das, was ursprünglich einmal Deepfake hieß.) Wichtig ist, dass die so entstehenden Audios, Videos und Bilder eben nicht auf den ersten Blick als gefälscht zu erkennen sind.

Diese Techniken verlassen sich primär, aber nicht allein auf eine Form von künstlicher Intelligenz, die als „deep learning" bekannt ist. Dahinter steht in diesem Fall in aller Regel ein sogenanntes Generative Adversarial Network (kurz: GAN).

> Ein Generative Adversarial Network (GAN) – zu Deutsch etwa erzeugendes konkurrierendes Netzwerk – ist ein künstliches neuronales Netzwerk. Es gehört zu den generativen Modellen, es erzeugt also etwas. Bei einem GAN laufen stets zwei neuronale Netze parallel. Sie konkurrieren miteinander, um ihre Vorhersagen genauer zu machen. Dieser Machine-Learning-Prozess findet selbstständig statt, er muss nicht überwacht werden.

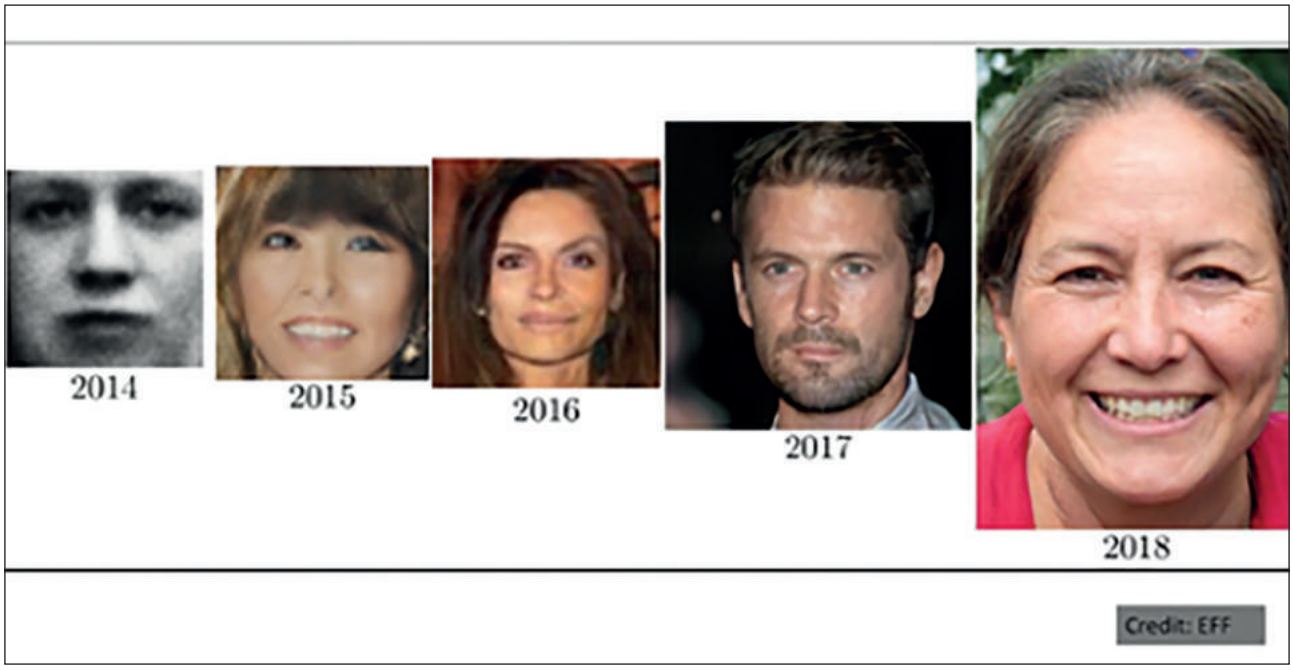
Um einen synthetischen Medieninhalt zu erzeugen, braucht es zunächst Bilder oder Video-Material der Person oder des Gegenstandes, der künstlich erzeugt werden soll. Anschließend kümmert sich ein GAN selbstständig darum, daraus die Simulation zu berechnen, egal ob von einer echten Person oder zum Beispiel durch den Austausch eines Gesichts, und zwar so:

Ein Netzwerk erzeugt Nachbildungen der Quellbilder. Das zweite Netzwerk arbeitet daran, diese als Fälschungen zu entlarven. Jene Informationen, die zur Aufdeckung führten, werden an das erste Netzwerk zurückgespielt, damit es diese beim nächsten Versuch nicht wiederholt.

Ende 2019 erforderten viele dieser Techniken – insbesondere die Erstellung von Deepfakes – noch erhebliche Rechenleistung, ein Verständnis dafür, wie man das eigene Modell nachjustieren und abstimmen muss, und oft auch erhebliche Nachbearbeitungen in der Bild- und Film-Postproduktion, um das Endergebnis zu verbessern. Doch selbst mit diesen noch bestehenden Einschränkungen gelingt es synthetischen Medien schon heute, Menschen auszutricksen. So konnten Untersuchungen im Rahmen des FaceForensics++-Projekts nachweisen, dass Menschen künstlich erzeugte Lippenbewegungen, die auf eine gänzlich neue Audiospur hin errechnet wurden, nicht zuverlässig erkennen können. Das bedeutet: Menschen sind nicht von Natur aus in der Lage, Manipulationen durch synthetische Medien zu erkennen.

An der Stelle sollte zudem in Betracht gezogen werden, dass sich auch die Audiosynthese schneller als erwartet fortentwickelt und kommerziell verfügbar wird. Mit der Text-zu-Sprache-Schnittstelle von Google lässt sich beispielsweise ein Stück Text mit einem Klick in eine realistisch klingende menschliche Stimme als Audio umwandeln. Jüngste Forschungen haben sich daher auch auf die Möglichkeit konzentriert, ein Video-/Audio-Interview zu bearbeiten, indem man am daraus errechneten Text Änderungen vornimmt.

Dazu kommt: Sowohl technische als auch kommerzielle Trends deuten darauf hin, dass es auch weiterhin einfacher und kostengünstiger werden wird, überzeugende synthetische Medien herzustellen. Das Bild unten zeigt zum Beispiel, wie schnell sich die Technologie der Gesichtsgenerierung weiterentwickelt hat.



Während die ersten beiden Bilder links (aus den Jahren 2014 und 2015) noch verschwommen und nicht natürlich aussehen und das dritte Bild in der Mitte (aus 2016) etwas unsymmetrisch ist, sind die beiden Bilder rechts (aus 2017 und 2018) kaum von echten Fotografien lebender Menschen zu unterscheiden.

Diese Netzwerke funktionieren nach einem Katz-und-Maus-Charakter, und so verbessern sie sich im Laufe der Zeit: je öfter sie Rückmeldungen über erfolgreiche Fälschungen oder erfolgreich entdeckte Fälschungen bekommen, desto schneller. Deswegen ist Vorsicht geboten, was die Wirksamkeit der Methoden betrifft, die wir heute haben, um sie aufzudecken.

Deepfakes und synthetische Medien sind – bis jetzt – noch nicht sehr weit verbreitet. Meist sind es sexualisierte Bilder, die ohne Zustimmung der betroffenen Personen erstellt wurden. Eine Untersuchung von DeepTrace (mittlerweile als Sensity bekannt) kam im September 2019 zu dem Ergebnis, dass mehr als 95 % aller Deepfakes diesem Typus zuzurechnen sind, und entweder Prominente, Pornodarsteller oder auch gewöhnliche Menschen involviert sind. Erschwerend kommt hinzu, dass Menschen begonnen haben, reale Inhalte, die sie kritisieren wollen, pauschal als Deepfakes zu bezeichnen, ganz ähnlich dem Begriff Fake News.

In Workshops, die wir bei WITNESS durchgeführt haben, haben wir potentielle Bedrohungsfaktoren mit einer Reihe von Teilnehmenden erörtert – mit zivilgesellschaftlichen Akteuren, Journalisten, Faktencheckern, Wissenschaftlern aus den Bereichen Des- und Falschinformation und Experten für die Auswertung öffentlich verfügbarer Informationen (Open Source Intelligence, kurz: OSINT). Sie identifizierten Bereiche, in denen neue Formen von Manipulation schon bestehende Bedrohungen verschärfen, variieren oder verstärken können oder neue Bedrohungsszenarien ermöglichen. Dazu gehörten Bedrohungen für Journalisten, Faktenchecker und OSINT-Rechercheure sowie die Gefahr von Angriffen auf deren Arbeitsabläufe und Recherchemethoden. Sie betonten auch die Herausforderungen einer Inflation von Ablehnung zuverlässiger Inhalte mit „Das ist doch Deepfake", analog zu „Das sind doch Fake News". Sie betonten, Deepfakes im Kontext von Faktenchecks und Verifikation zu sehen und nicht als etwas Isoliertes. Sie nahmen zudem an, dass sich Deepfakes und synthetische Medien in bestehende Verschwörungstheorien und Desinformationskampagnen integrieren werden und auf Taktiken und Methoden zurückgreifen, die sich in diesen Bereichen entwickelt haben.

Dies sind einige der spezifischen Gefahren, die in den Gesprächen hervorgehoben wurden:

- **Journalisten und Aktivisten werden erleben, dass ihr Ruf und ihre Glaubwürdigkeit attackiert werden,** anknüpfend an bereits existierende Formen von Online-Belästigung und Gewalt. Es wurden bereits erste Angriffe mit modifizierten Videos gegen Frauen registriert, so zum Beispiel der Fall der prominenten indischen Journalistin Rana Ayyub.

- **Bekanntere Personen werden sich mit sexualisierten Bildern und Videos konfrontiert sehen, die ohne ihre Einwilligung erstellt wurden, wie auch mit geschlechtsbezogener Gewalt oder der Verwendung von Doppelgängern.** Lokalpolitikerinnen und -politiker könnten hierbei besonders angreifbar sein, da es von ihnen eine große Menge frei verfügbaren Bildmaterials gibt und sie gleichzeitig geringer ausgeprägte institutionelle Schutzstrukturen haben, als das für Politiker auf nationaler Ebene der Fall wäre. Dennoch sind sie oft Akteure in Nachrichtensituationen, die sich von der lokalen Ebene zu nationaler Bedeutung auswachsen können.
- **Aneignung bekannter Marken** durch gefälschte Video-Bearbeitung oder andere Methoden, so dass eine Nachrichten-, Regierungs-, Unternehmens- oder Nichtregierungsorganisations-Marke vorsätzlich mit einem unechten Inhalt in Verbindung gebracht wird.
- **Versuche, manipulierte nutzergenerierte Inhalte in den Nachrichtenzyklus zu bekommen,** kombiniert mit anderen Techniken wie Source-Hacking (dem Versuch, Journalistinnen und Journalisten dazu zu bekommen, jemanden oder etwas Manipulatives als Quelle zu übernehmen) oder dem gezielten Zuspielen manipulierter Inhalte an Journalisten in Schlüsselmomenten. Typischerweise besteht das Ziel dann darin, die Journalisten dazu zu bringen, diesen Inhalt weiterzuverbreiten.
- **Das Ausnutzen von Schwachstellen des Nachrichtensammel-/Berichterstattungsprozesses.** Ein Beispiel dafür ist die Anwendung der häufigsten Kameraeinstellung in Nachrichtensendungen. Wie das Reuters UGC-Team feststellte, lässt sich nämlich die Tatsache, dass Interviewte häufig frontal in die Kamera blicken und dabei auf Fragen eines Moderators antworten, für Deepfakes leicht ausnutzen. Das gilt auch für die Tatsache, dass Material aus Kriegs- und Krisengebieten oft nur schwer zu erhalten und zu verifizieren ist.
- In dem Maße, in dem Deepfakes üblicher und in der Erstellung einfacher werden, werden sie **auch in der Anzahl zunehmen.** Das könnte Faktenchecker und Verifizierungsfachleute überlasten oder ablenken.
- **Auf Organisationen, die Nachrichten sammeln und verifizieren, wird Druck ausgeübt werden, um zu beweisen, dass etwas wahr ist, oder auch, dass etwas nicht gefälscht ist.** Menschen, die von Medien mit Sachverhalten konfrontiert werden, werden die Möglichkeit nutzen, etwas plausibel zu leugnen, indem sie schlicht erklären, der Inhalt sei „deepfaked".

WO BEGINNEN BEI DER VERIFIZIERUNG VON DEEPFAKES

Angesichts der neuen Deepfake-Technologien werden wir akzeptieren müssen, dass das Fehlen von Beweisen dafür, dass etwas manipuliert wurde, kein Nachweis dafür ist, dass es tatsächlich nicht manipuliert wurde. Journalisten und Rechercheure müssen eine Mentalität der gesunden Skepsis gegenüber Fotos, Videos und Audios entwickeln. Sie müssen davon ausgehen, dass diese Medienformen in dem Maße häufiger in Frage gestellt werden, als das Wissen und die Furcht vor Deepfakes zunehmen. Auch darum ist es wichtig, sich mit den Werkzeugen der sogenannten Medienforensik vertraut zu machen.

Vor diesem Hintergrund sollten Ansätze zur Analyse und Verifizierung von Deepfakes und synthetischer Medienmanipulation Folgendes beinhalten:

1. Den Inhalt selbst auf jene Fehler hin überprüfen, die als Verzerrungen oder Defekte in der Folge von synthetischer Medienmanipulation entstehen.
2. Bestehende technische Ansätze zur Video-Überprüfung und Forensik anwenden.
3. Neue, auf künstlicher Intelligenz basierende Ansätze und neue forensische Ansätze nutzen, wenn verfügbar.

ÜBERPRÜFUNG AUF VERRÄTERISCHE PANNEN ODER VERZERRUNGEN

Dieser Ansatz ist für die Identifizierung von Deepfakes und anderen Modifikationen synthetischer Medien der unzuverlässigste, insbesondere angesichts der schnellen Fortentwicklung technischer Lösungen. Dennoch können auch durch manuelle Überprüfung in schlecht gemachten Deepfakes oder synthetischen Inhalten Hinweise und Fehler enthalten sein. Zu den Dingen, auf die man achten sollte, gehören:

- Mögliche Verzerrungen an der Stirn und am Haaransatz oder dann, wenn ein Gesicht sich von einem Bereich zu einem anderen bewegt.
- Fehlende Details in der Darstellung der Zähne.
- Ungewöhnlich glatte Haut.

- Fehlendes Blinzeln.
- Statisches Sprechen, ohne wirkliche Bewegung des Kopfes oder eine natürliche Bandbreite an Ausdrucksweisen.
- Bildfehler, wenn eine Person ihr Gesicht zur Seite dreht.

Im Moment gilt für einige dieser Fehler, dass die Wahrscheinlichkeit, sie zu entdecken, größer ist, wenn für einen Video-Ausschnitt eine Bild-für-Bild-Analyse durchgeführt wird. Es kann daher helfen, eine Reihe von Frames – also die einzelnen Standbilder, aus denen ein Video besteht – zu extrahieren und diese anschließend einzeln zu analysieren. Für frontal-seitliche Bewegungen hingegen gilt das nicht – diese können in einer Video-Sequenz am ehesten gefunden werden. Man sollte beide Ansätze ausprobieren.

VORHANDENE ANSÄTZE ZUR VIDEO-VERIFIZIERUNG ANWENDEN

Wie mit anderen Formen der Medienmanipulation und wie auch bei Irreführung durch Videos, die bearbeitet oder, in falsche Kontexte gesetzt, neu hochgeladen wurden, sollte sich der Zugang auf bereits bewährte Verifikationsprozesse stützen. Die Verifikationspraktiken für öffentlich verfügbare Informationen (OSINT) bleiben wichtig, daher sind auch die Kapitel und Anwendungsbeispiele im ersten „Verification Handbook" zu Bildern und Videos weiterhin gute Ausgangspunkte für eigene Recherchen. Da die meisten Deepfakes oder Modifikationen heute noch nicht voll synthetisch sind und stattdessen Quellmaterial benötigen, das anschließend verändert wird, besteht ein Ansatz zur Untersuchung darin, ein Video in seine Standbilder zu zerlegen und für diese einzelnen Bilder anschließend eine Bilder-Rückwärtssuche durchzuführen. Sie können im Video auch nach eindeutigen Merkmalen zum Beispiel einer Landschaft oder eines abgebildeten Ortsausschnitts suchen und prüfen, ob diese sich mit Bildern des Orts auf Google Street View in Einklang bringen lassen. Letztlich kann auch Erfahrung helfen, um zu klären, wie sich Inhalte online verteilen, wer sie teilt und wie sie geteilt werden, wie sich allgemein Informationen finden lassen, um festzustellen, ob man einem Bild oder Video trauen kann. Die Grundlagen der Bestimmung von Quelle, Datum, Zeit und Motivation hinter einem Inhalt sind essentiell, um zu beurteilen, ob eine reale Person oder ein realer Vorgang gezeigt wird. (Einen grundlegenden Zugang dazu kann dieser Workshop von First Draft bieten.) Und wie immer ist es auch hier eine gute Idee, die Person oder die Personen zu kontaktieren, die angeblich zu sehen sind, um zu prüfen, ob sie konkrete Informationen geben können, die die Authentizität bestätigen oder widerlegen können. Neue Werkzeuge werden fortlaufend von Regierungen, Wissenschaftlern, Plattformen oder journalistischen Innovationslaboren entwickelt, um synthetische Medien zu entdecken und auch, um die Verbreitung dieser medienforensischen Werkzeuge zu erhöhen. Sie sollten als Gelegenheit gesehen werden, die eigenen Verifizierungsfähigkeiten zu verbessern. Dienste wie InVID und Forensically helfen bei beidem, bei der Suche nach dem Ursprung eines Bildes wie auch nach begrenzter forensischer Analyse.

Kostenlose Werkzeuge in diesem Feld sind zum Beispiel diese:

- FotoForensics: ein Foto-Forensik-Dienst, der eine Fehlerlevel-Analyse durchführen kann, um zu prüfen, wo dem Bild möglicherweise Elemente hinzugefügt wurden.
- Forensically: ein Paket von Werkzeugen, um geklonte Bilder zu finden, Fehlerlevel-Analysen durchzuführen, die Metadaten eines Bildes auszulesen und für viele andere Untersuchungsmethoden.
- InVID: eine Browser-Erweiterung, die es möglich macht, Videos in einzelne Standbilder zu zerlegen, über verschiedene Suchmaschinen hinweg Bilder-Rückwärtssuchen durchzuführen, Bereiche zur Untersuchung zu vergrößern und zu verbessern sowie forensische Filter auf Standbilder anzuwenden.
- Reveal Image Verification Assistant: ein Angebot mit einer ganzen Reihe von Algorithmen zur Erkennung von Bildmanipulationen wie Metadatenanalyse, GPS-Geolokalisierung, Extraktion von EXIF-Miniaturbildern und integrierter Bilder-Rückwärtssuche über Google.
- Ghiro: ein digitales Forensik-Tool (Open Source), dessen Quellcode und Funktionsweise also öffentlich einsehbar sind.

Alle diese Angebote dienen der Verifikation von Bildern, nicht von Videos. Das ist leider ein Bereich, der in forensischer Hinsicht noch schwach ausgeprägt ist. Für Videos ist es deshalb noch nötig, einzelne Standbilder aus dem Video zur Analyse heranzuziehen, bei deren Generierung InVID helfen kann. Am effektivsten arbeiten alle diese Dienste mit hochauflösenden Bildern aus nicht komprimierten Videos – wenn Sie also ein Video haben, das für den Versand über Mail oder Messenger-Apps in der Dateigröße verkleinert wurde, versuchen Sie, das originale unveränderte Video zu bekommen. Je mehr ein Video komprimiert wurde, je öfter es über verschiedene Plattformen weiterverbreitet wurde, je häufiger es heruntergeladen und anderswo neu hochgeladen wurde, desto unzuverlässiger werden die Analysewerkzeuge.

Um einen Überblick zu haben, wo noch Lücken und Probleme bei der Analyse von Deepfakes bestehen, und um über die aktuellen Bemühungen, diese zu schließen, auf dem Laufenden zu bleiben, sollte man beobachten, was Wissenschaftler dazu veröffentlichen. Eine der auf diesem Gebiet führenden Forschungseinrichtungen arbeitet an der Universität von Neapel und stellt die Codes für ihre Werkzeuge online bereit – darunter unter anderem Anwendungen, um die individuellen digitalen Fingerabdrücke verschiedener Kameramodelle zu erkennen (Noiseprint), um Nahtstellen zwischen zusammengefügten Bildern ohne vorherige Hinweise darauf zu finden (Splicebuster) und um in ein Video hineinmontierte Bewegungen aufzuspüren.

Mit der Evolution synthetischer Medien werden sich auch händische und automatische Forensikmethoden entwickeln und in existierende Verifikationswerkzeuge integriert werden, auf die Journalistinnen und Journalisten, Faktenchecker und möglicherweise auch die Plattformen selbst zugreifen werden. Es ist wichtig, dass sie alle bezüglich neuer Werkzeuge auf dem Laufenden bleiben, ohne sich gleichzeitig zu sehr von einem einzelnen abhängig zu machen.

KÜNSTLICHE INTELLIGENZ UND NEU AUFKOMMENDE ANSÄTZE FÜR MEDIENFORENSIK

Momentan (Stand: Mitte 2020) gibt es noch keine getesteten und kommerziell verfügbaren Erkennungswerkzeuge, die auf erzeugenden konkurrierenden Netzwerken (Generative Adversarial Network – GAN) basieren. Wir sollten davon ausgehen, dass die ersten für Journalisten zeitnah auf den Markt kommen werden; ob als Programme, Programmerweiterungen oder Services auf Plattformen. (Eine aktuelle Übersicht über den Status quo der Medienforensik und der Werkzeuge dafür bietet Luisa Verdolivas Aufsatz „Media Forensics and Deepfakes: An overview".) Alle diese Werkzeuge werden auf umfangreiche Trainingsdatensätze angewiesen sein, also synthetische Medien, die von erzeugenden konkurrierenden Netzwerken „erfunden" wurden, um darauf aufbauend sich selbst in die Lage zu versetzen, andere Beispiele zu erkennen, die mit den gleichen oder ähnlichen Techniken „erfunden" wurden. So erzeugen Forensikprogramme wie FaceForensics++ Fälschungen, indem sie bereits existierende Werkzeuge für Deepfakes benutzen, damit große Mengen gefälschter Bilder erstellen und diese als Trainingsdaten für ihre Algorithmen verwenden, um damit Fälschungserkennungen durchzuführen. Das wiederum bedeutet, dass sie bei *neueren* Fälschungsmethoden und -techniken möglicherweise nicht sehr effektiv sind, da sie an ihnen nicht trainieren konnten.

Ein zentraler Punkt ist außerdem, dass jeder Hinweis auf synthetische Inhalte gegengeprüft und mit anderen Verifikationsmethoden bestätigt werden sollte. Deepfakes und synthetische Medien entwickeln sich schnell, die Technologien dafür werden immer breiter verfügbar, sie werden kommerzialisiert und einfacher zu benutzen. Sie benötigen schon heute weniger Quellen als Rohmaterial zur Erstellung einer Fälschung, als man vielleicht denkt.

Während neue Technologien zur Erkennung aufkommen und in Plattformen und in Werkzeuge für OSINT-Journalisten und -Rechercheure integriert werden, bleibt der beste Weg zur Verifizierung ein sicherer Umgang mit bereits vorhandenen Verfahren für Bild/Video und die Ergänzung durch forensische Werkzeuge, die Bildmanipulationen erkennen können. Dem menschlichen Auge allein zu vertrauen, ist keine verlässliche Strategie!

7. IN GESCHLOSSENEN GRUPPEN UND MESSENGERN RECHERCHIEREN UND DARÜBER BERICHTEN

von: Claire Wardle
deutsche Bearbeitung: Marcus Engert

Claire Wardle *ist für die strategische Ausrichtung und Forschung von First Draft verantwortlich. First Draft ist eine internationale gemeinnützige Nichtregierungsorganisation, die Journalisten, Wissenschaftler und Technikexperten unterstützt, die sich mit Vertrauen und Wahrheit im digitalen Zeitalter beschäftigen. Sie war Fellow am Shorenstein Center for Media, Politics and Public Policy an der Kennedy School in Harvard, Forschungsdirektorin am Tow Center for Digital Journalism an der Columbia University und Leiterin des Bereichs Social Media für UNHCR, das Flüchtlingshilfswerk der Vereinten Nationen.*

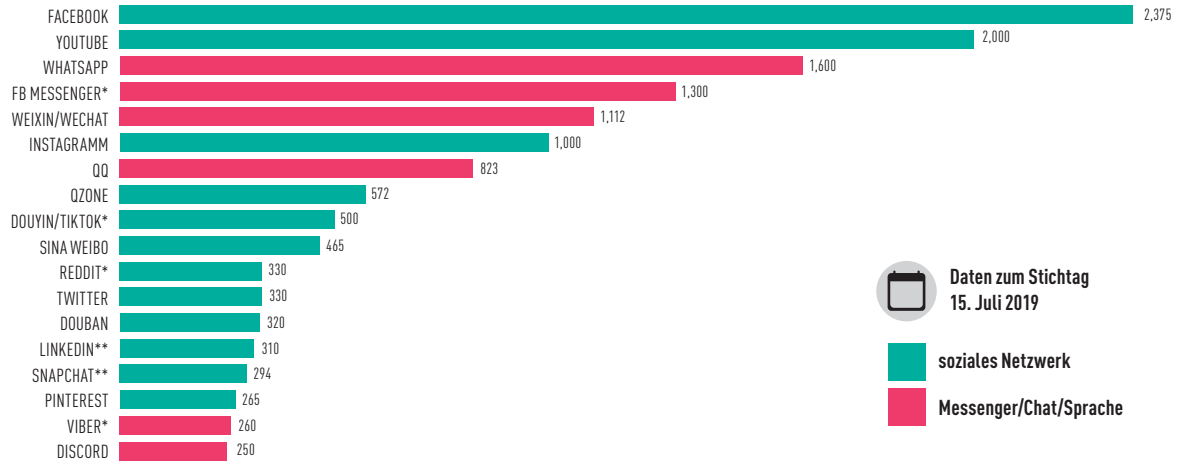
Im März 2019 sprach Mark Zuckerberg über den „Schwenk hin zur Privatsphäre" von Facebook. Er meinte damit, dass das soziale Netzwerk die Facebook-Gruppen in den Mittelpunkt rücken wollte, als Reaktion darauf, dass die Menschen zunehmend nicht mehr vor einer breiten Öffentlichkeit, sondern mit einer kleineren Anzahl von Menschen in privaten Räumen kommunizierten. In den letzten Jahren wurde denjenigen von uns, die in diesem Bereich arbeiten, die Bedeutung kleinerer Gruppen für die soziale Kommunikation immer deutlicher. In diesem Kapitel werde ich daher die verschiedenen Plattformen und Anwendungen erklären, etwas zu den Herausforderungen bei der Beobachtung dieser Räume sagen und schließen mit einer Diskussion über die ethischen Fragen dieser Art von Arbeit.

VERSCHIEDENE PLATTFORMEN, VERSCHIEDENE ANWENDUNGEN

Jüngste Erhebungen von We Are Social zeigen die anhaltende Dominanz von Facebook und YouTube, aber die sich daran anschließenden drei Services sind WhatsApp, der Facebook Messenger und WeChat – und damit Chat- bzw. Messengerdienste.

Aktive Nutzer der wichtigsten Plattformen

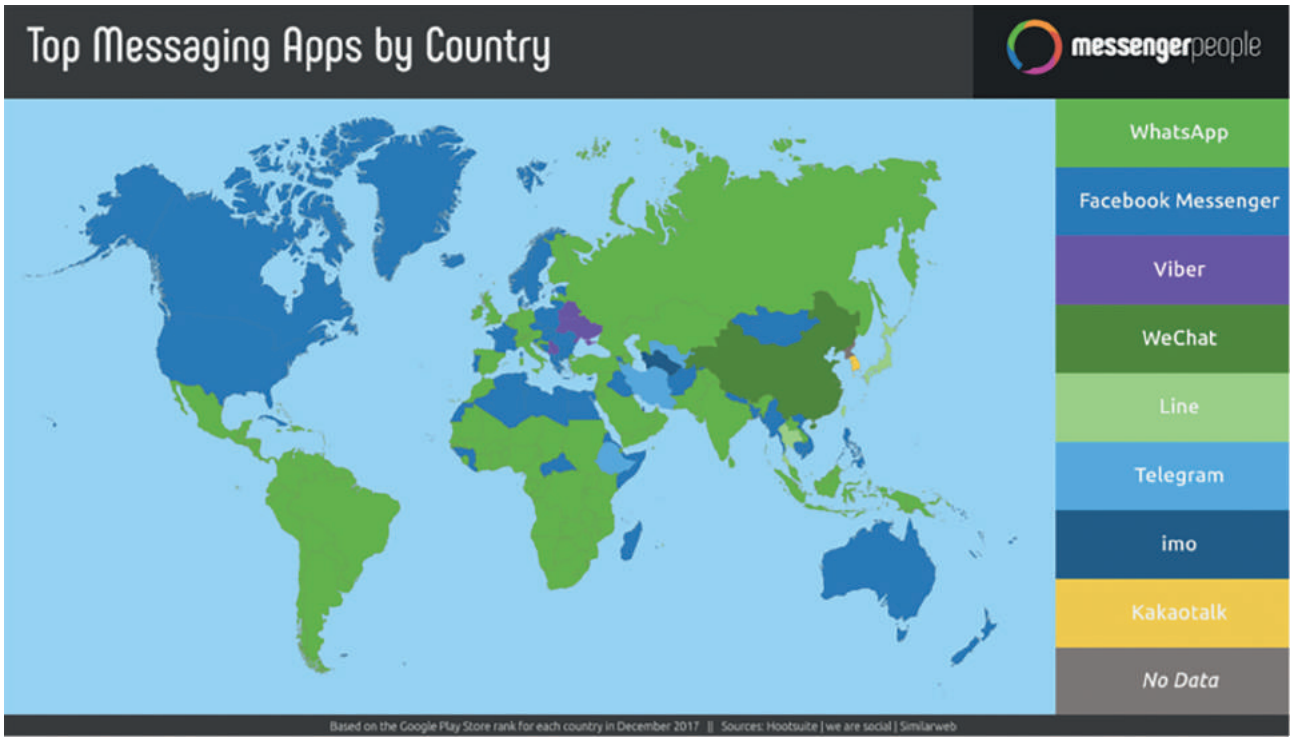
basierend auf monatlichen aktiven Nutzern, aktiven Profilen oder monatlichen eindeutigen Besuchern pro Plattform, in Millionen



Quellen: Kepios Analysis, Geschäftsberichte und Mitteilungen der Unternehmen, Werbebuchungs-Tools der Plattform selbst
Hinweise: Plattformen mit (*) haben in den vergangenen 12 Monaten keine neuen Nutzerzahlen veröffentlicht, daher können diese Zahlen ungenauer sein. (**) Diese Plattformen veröffentlichen keine Daten zu monatlich aktiven Nutzern. Zahlen für Linkedin basieren auf monatlichen eindeutigen Webseitenbesuchern in Dezember 2018, via Similarweb. Zahlen für Snapchat extrapoliert aus Daten von Techchrunch (Juni 2017).

Hootsuite we are social

In vielen Regionen der ganzen Welt sind mittlerweile Chat-Apps zur wichtigsten Nachrichtenquelle für viele Verbraucher geworden, insbesondere WhatsApp. Besonders auffällig ist das in Brasilien, Indien und Spanien. Ohne Zweifel sind WhatsApp und der Facebook Messenger weltweit beliebt, aber in einigen Ländern dominieren Alternativen. Im Iran zum Beispiel ist es Telegram. In Japan ist es Line, in Südkorea KakaoTalk und in China WeChat.



Die Grafik zeigt, in welchem Land der Erde welche Chat-App Marktführer ist.

All diese Seiten und Dienste unterscheiden sich nur gering im Hinblick auf Verschlüsselung, Gruppen, Funktionalitäten zum Senden von Mitteilungen an eine große Zahl von Menschen und zusätzliche Optionen wie zum Beispiel Kaufüptionen innerhalb der App.

Geschlossene Facebook-Gruppen

Es gibt drei Arten von Facebook-Gruppen: offene, geschlossene und private.

- Offene Gruppen können von jedermann über die Suche gefunden werden, und jeder kann ihnen beitreten.
- Geschlossene Gruppen lassen sich über die Suche finden, aber die Aufnahme in die Gruppe muss von einem Gruppeninhaber oder -administrator bestätigt werden.
- Geschlossene Gruppen lassen sich nicht über die Suche finden. Man muss für einen Beitritt eingeladen werden.

Immer mehr Menschen finden sich in Facebook-Gruppen zusammen, zum Teil weil ihnen der Facebook-Algorithmus solche vorschlägt, aber auch, weil sich Menschen eher dazu entscheiden, Zeit mit anderen Menschen zu verbringen, die sie bereits kennen oder mit denen sie Haltungen oder Interessen teilen.

Discord

Erhebungen von Statista aus dem Juli 2019 zufolge hatte Discord zu diesem Zeitpunkt 250 Millionen monatlich aktive Nutzer (zum Vergleich: Snapchat hatte 294 Millionen, Viber 260 Millionen und Telegram 200 Millionen). Discord ist vor allem in der Computerspielszene beliebt, wurde in den letzten Jahren aber auch zu einem Sammelpunkt für koordinierte Desinformationskampagnen. Eine Gemeinsamkeit von Discord und einigen geschlossenen Facebook-Gruppen ist, dass man vor der Aufnahme Fragen beantworten muss. Je nachdem, was der Gruppeninhaber dort eingestellt hat, können diese sich um den Beruf, die Religion, politische Ansichten oder die persönliche Sicht auf bestimmte soziale Fragen drehen.

VERSCHLÜSSELUNG, GRUPPEN UND KANÄLE

Ein Grund, warum diese Plattformen und Apps so beliebt wurden, ist, dass sie unterschiedliche Stufen von Verschlüsselung anbieten. Standardmäßige Ende-zu-Ende-Verschlüsselung bieten WhatsApp und Viber an, was sie zu sicheren Kommunikationskanälen macht. Andere wie Telegram, der Facebook Messenger oder Line verschlüsseln Unterhaltungen, wenn der Nutzer das entsprechend aktiviert. Bestimmte Apps haben Gruppen oder Kanäle, mit denen Informationen an eine große Zahl von Menschen gesendet werden können. Bei WhatsApp liegt die maximale Gruppengröße bei 256, im Facebook Messenger bei 250. Bei Telegram kann eine Gruppe privat oder öffentlich suchbar sein und 200 Mitglieder haben. Sobald diese Zahl erreicht ist, lässt sie sich in eine sogenannte Supergruppe umwandeln, der dann bis zu 75.000 Menschen beitreten können. Außerdem bietet Telegram auch Kanäle an, mit denen öffentliche Nachrichten an eine große Zahl von Nutzer geschickt werden können, wobei diese in einem Kanal darauf dann nicht antworten können: Man kann einem Kanal beitreten und sehen, was dort geteilt wird, aber etwas eigenes absenden kann man nicht.

LAUFENDE BEOBACHTUNG

Es gibt keinen Zweifel daran, dass Falschinformationen sich über geschlossene Gruppen und Apps verbreiten. Ob dort mehr oder weniger Falschinformationen kursieren als auf den klassischen Plattformen ist schwer zu sagen, da es keine Möglichkeit gibt, zu sehen bzw. zu vergleichen, was dort geteilt wird. Dass es aber ein Problem ist, das wissen wir, und besorgniserregende Ereignisse aus Indien, Frankreich und Indonesien haben das mehrfach belegt. Während der Amokläufe in El Paso und in Dayton im August 2019 haben sich in den USA Gerüchte und Falschbehauptungen über Telegram und über den Facebook Messenger verbreitet. Die Frage ist also, ob Journalisten, Wissenschaftler, Faktenchecker, Menschen aus dem Gesundheitswesen, humanitäre Helfer in geschlossenen Gruppen sein und diese beobachten sollten. Und wenn sie in diesen Gruppen sein sollten, wie sollte ihre Arbeit aussehen, damit sie ethisch einwandfrei und sicher vonstatten geht?

Obschon diese Arbeit mit erheblichen Herausforderungen verbunden ist, möglich wäre es. Gleichzeitig nutzen viele Menschen diese Apps aus genau dem Grund, dass sie darin nicht überwacht werden. Sie benutzen sie, weil sie verschlüsselt sind. Sie erwarten ein gewisses Maß an Privatsphäre. Das muss für jeden, der in diesen Räumen arbeitet, von zentraler Bedeutung sein. Auch wenn wir diesen Räumen beitreten und sie überwachen können: Es ist von größter Wichtigkeit, sich der Verantwortung bewusst zu sein, die wir gegenüber den Teilnehmern in diesen Gruppen haben, denen oft nicht klar ist, was technisch möglich ist.

TECHNIKEN FÜR DIE SUCHE

Die Suche nach diesen Gruppen kann schwierig sein, da es für jede Gruppe unterschiedliche Protokolle gibt. Für Facebook-Gruppen können Sie innerhalb der Facebook-Suche nach Themen suchen und die Ergebnisse nach Gruppen filtern. Wer fortgeschrittenere Suchoperatoren verwenden möchten, kann bei Google nach einem Schlagwort suchen und dort site:facebook.com/groups anfügen.

Bei Telegram lässt sich in der App nach Gruppen suchen, allerdings nur auf Android-Telefonen, nicht auf iPhones. Es gibt zudem Anwendungen, die man auf dem Computer nutzen kann, wie https://www.telegram-group.com/. Für Discord gibt es Seiten wie https://disboard.org/search

DISKUSSIONEN ÜBER DAS BEITRETEN UND MITMACHEN

Wie bereits erwähnt, werden bei einigen dieser Gruppen Fragen gestellt, bevor man ihnen beitreten darf. Bevor man auf sie antwortet, sollte man mit einem zuständigen Redakteur darüber sprechen, wie man diese Fragen beantworten wird. Soll man ehrlich Auskunft darüber geben, wer man ist und warum man der Gruppe beitreten will? Gibt es eine Möglichkeit, der Gruppe beizutreten und gleichzeitig absichtlich vage zu bleiben? Wenn nicht, wie gewichtig sind die Rechtfertigungsgründe, die eigene Identität zu verbergen? (Ein Beispiel könnte sein, dass eine Offenlegung der eigenen Identität als Journalist die Sicherheit für sich selbst oder andere gefährden könnte.) Wenn der Zugang gestattet wird: Will man sich in irgendeiner Weise einbringen oder nur „lauern", um Informationen zu finden, die man danach anderswo bestätigen kann?

ENTSCHEIDUNGEN ÜBER DAS AUTOMATISCHE SAMMELN VON INHALTEN AUS GRUPPEN

Man kann „offene" Gruppen finden, indem man nach Links sucht, die auf anderen Seiten oder in anderen Gruppen veröffentlicht wurden. Diese erscheinen dann möglicherweise auch in Suchmaschinen. In dem Fall ist es auch möglich, Computer den Inhalt dieser Gruppen automatisch sammeln zu lassen. Forscher, die Wahlen in Brasilien und Indien beobachten, haben dies bereits getan, und mir wurde von anderen Organisationen berichtet, die ähnliche Arbeit leisten.

Ein solches Vorgehen ermöglicht es, mehrere Gruppen gleichzeitig zu beobachten, was sonst oft nicht möglich ist. Wichtig ist, dass nur ein kleiner Anteil der Gruppen auf diese Weise auffindbar ist. Zudem handelt es sich dabei in der Regel um Gruppen, die es darauf anlegen, gefunden zu werden, weil sie eine breite Mitgliedschaft anstreben, und die daher nicht für alle Gruppen repräsentativ sind. Das wirft für mich persönlich auch ethische Fragen auf. Es gibt jedoch Leitplanken: die Daten sichern, sie nicht mit anderen teilen und Nachrichten anonymisieren. Wir brauchen branchenübergreifende Standards für diese Art von Arbeit.

TIPLINES/HOTLINES FÜR TIPPS

Eine andere Technik besteht darin, eine Rufnummer einzurichten, mit der man die Allgemeinheit ermutigt, dort Hinweise und Tipps zu hinterlassen. Der Schlüssel zum Erfolg einer solchen Tipline ist ein einfacher, klarer, konkreter Aufruf zum Handeln, der erklärt, was gesucht wird und wie die Inhalte verwendet werden sollen. Geht es Ihnen nur darum, Themen zu beobachten, oder wollen Sie baldmöglichst eine Richtigstellung zu kursierenden falschen Informationen veröffentlichen? Auf die ethischen Fragen zurückkommend, die die Arbeit mit geschlossenen Gruppen so sehr beeinflussen: Es ist wichtig, nicht einfach nur alles „mitzunehmen", sondern auszuwählen. Auch wenn man die Ethik einmal außen vor lässt, so zeigen alle Untersuchungen, dass ein Publikum deutlich weniger geneigt ist, weiterhin Tipps einzusenden, wenn es nicht weiß, wie diese verwendet werden. Menschen sind eher bereit zu helfen, wenn sie das Gefühl haben, wie Partner behandelt zu werden.

Der andere Aspekt, der hier natürlich auf der Hand liegt, ist die Tatsache, dass es extrem einfach ist, eine Tipline durch das Einreichen von Falschinformationen zu sabotieren, von einem Einzelnen oder auch von einer Gruppe, die mehrfach den gleichen Inhalt einsendet und damit den Eindruck entstehen lässt, das Problem sei größer, als es de facto ist.

ETHIK DER BERICHTERSTATTUNG AUS GESCHLOSSENEN MESSAGING-GRUPPEN

Sobald Inhalte gefunden sind, stellt sich die Frage, wie man darüber berichten kann. Sollte man transparent machen, wie sie gefunden wurden? In ihren Gruppenregeln bitten viele Gruppen darum, das, was in einer Gruppe diskutiert wird, nicht außerhalb der Gruppe zu verbreiten. Wenn in der Gruppe eine Menge Desinformation geteilt wird, welche Auswirkungen hat es dann, darüber zu berichten? Lässt sich extern bestätigen, was man in anderen Gruppen oder Online-Räumen gefunden hat? Kann man die eigene Sicherheit gewährleisten? Gilt das Gleiche für Kollegen und Familien? Denken Sie daran, dass Doxing (auch: Doxxing – öffentliche Aufrufe zum Zusammentragen von Informationen über eine Person mit dem Ziel, diese anschließend zu veröffentlichen und so Einschüchterung zu erzeugen oder auch die öffentliche Identifizierung bislang anonymer Personen herbeizuführen) für manche Akteure in diesem Feld Teil des „Spiels" ist.

SCHLUSSFOLGERUNGEN

Die Berichterstattung über geschlossene Gruppen in sozialen Netzwerken und Messengern und aus ihnen heraus ist voller Herausforderungen. Gleichzeitig werden diese Gruppen als Quellen, als Räume, in denen Informationen ausgetauscht werden, immer wichtiger. Denken Sie in einem ersten Schritt daher über die in diesem Kapitel umrissenen Fragen nach. Sprechen Sie mit Kollegen, Redakteuren, Managern – und wenn es in Ihrer Redaktion keine Richtlinien für diese Art der Berichterstattung gibt, beginnen Sie mit der Arbeit daran. Es gibt dafür noch keine Standardregeln. Es hängt von der Story, der Plattform, dem Reporter, den redaktionellen Richtlinien einer Redaktion ab. Wichtig ist nur, dass alle Einzelheiten ausgewogen berücksichtigt werden, bevor man mit dieser Art der Berichterstattung beginnt.

7 a. Fallbeispiel: Bolsonaro im Krankenhaus

von: Sérgio Lüdtke
deutsche Bearbeitung: Marcus Engert

Sérgio Lüdtke *ist Journalist und Redakteur von Projeto Comprova, einem Zusammenschluss von 24 Medienorganisationen, die kollaborativ zu Gerüchten über politische Belange in Brasilien recherchieren. 2018 untersuchte Comprova verdächtige Inhalte, die in sozialen Medien und Chat-Apps über die brasilianischen Präsidentschaftswahlen verbreitet wurden.*

Am 6. September 2018, einen Monat vor der brasilianischen Präsidentschaftswahl, hielt der rechtsextreme Kandidat Jair Bolsonaro eine Wahlkampfveranstaltung in der Innenstadt von Juiz de Fora ab, einer Stadt mit 560.000 Einwohnern, 200 Kilometer von Rio de Janeiro entfernt. Eine Woche vorher war Bolsonaro in der ersten Vorrunde der brasilianischen Präsidentschaftswahlen in Führung gegangen. Er übernahm Platz 1, nachdem die Kandidatur des ehemaligen Präsidenten Luiz Inácio Lula da Silva vom Obersten Wahlgericht für ungültig erklärt worden war.

In den Hochrechnungen für die Stichwahlen hingegen unterlag Bolsonaro gegen drei der vier nächstplatzierten Kandidaten in den Umfragen. Die Situation für ihn war kritisch, da er täglich nur zwei Neun-Sekunden-Blöcke in den freien Wahlsendungen im Fernsehen bekam. Nach den brasilianischen Wahlgesetzen müssen Radio- und Fernsehsender den politischen Parteien freie Zeit geben, um ihre Positionen öffentlich präsentieren zu dürfen. Die Zeit dafür wird nach der Anzahl der Sitze berechnet, die eine Partei bei der letzten Wahl zum Repräsentantenhaus errungen hat. Bolsonaros Mangel an Sitzen bedeutete also sehr wenig Sendezeit. Infolgedessen war er auf seine Anhänger in sozialen Netzwerken angewiesen und musste direkten Kontakt zu den Wählern auf der Straße aufnehmen.

In Juiz de Fora wie auch in anderen Städten, die er zuvor besuchte, nahm Bolsonaro an einem Marsch teil, bei dem er von seinen Anhängern auf den Schultern getragen wurde. Er wurde von einer Schar von Bewunderern begleitet, als der Marsch plötzlich unterbrochen wurde. In der Mitte der Menge erhob ein Mann die Hand – und stach auf den Kandidaten ein. Das Messer verursachte eine tiefe Wunde in Bolsonaros Bauch. Und es öffnete eine Büchse der Pandora in sozialen Netzwerken.

Gerüchte und Verschwörungstheorien begannen sich zu verbreiten. Einige beschuldigen den Attentäter, Verbindungen zur Partei der ehemaligen Präsidentin Dilma Rousseff zu haben, die 2016 aus dem Amt entfernt wurde. Gefälschte Fotos zeigten den Attentäter neben Amtsvorgänger Lula stehen. Dass er Verbindungen zur linken Partei Partido Socialismo e Liberdade (PSOL) hatte und dass sich seine Anwälte weigerten, zu sagen, wer ihre Honorare zahlte, brachte die Gerüchteküche endgültig zum Überkochen.

Gleichzeitig nahmen Videos und Nachrichten in den sozialen Medien zu, die versuchten, Bolsonaros Position zu untergraben. Einige der teils böswilligen Beiträge behaupteten, die Messerstecherei sei inszeniert gewesen, tatsächlich sei Bolsonaro zu einer Krebsbehandlung im Krankenhaus gewesen, und die veröffentlichten Fotos, die die Operation zeigten, seien gefälscht.

Die Messerstecherei war für Bolsonaro ein Grund, sich aus den Wahlkampfaktivitäten zurückzuziehen, brachte ihm aber eine bessere Position in den Umfragen ein. (Letztendlich gewann Bolsonaro die Wahl.) Am 19. September, fast zwei Wochen nach dem Angriff, entdeckte Eleições sem Fake – ein Programm der Universität von Minas Gerais, das WhatsApp-Gruppen beobachtet – eine Tonaufnahme, die die Runde machte. Die Audio-Aufzeichnung wurde von 16 der fast 300 WhatsApp-Gruppen, die das Projekt beobachtete, geteilt; einige davon waren Bolsonaro-Anhänger.

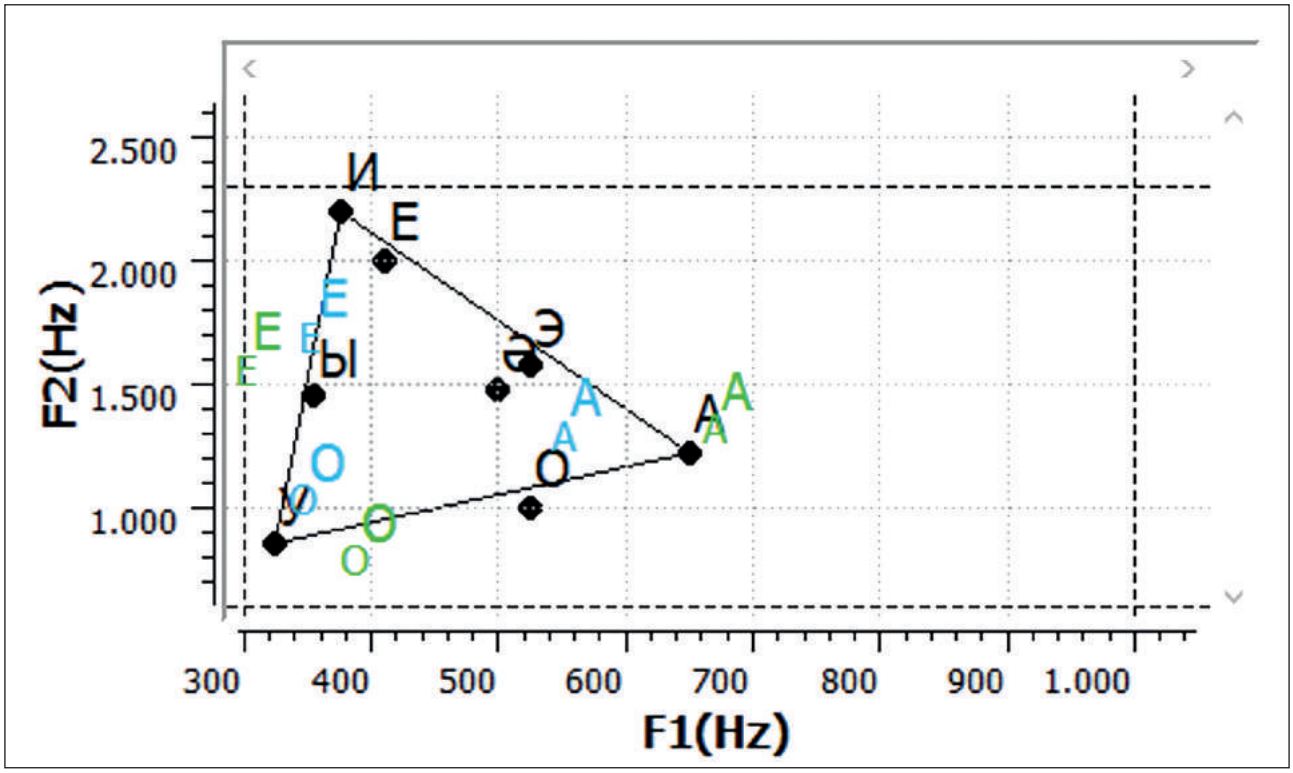
Am gleichen Tag begann unsere Organisation, Comprova, ebenfalls über WhatsApp, Anfragen von Lesern zu erhalten, die um Überprüfung der Echtheit der Aufzeichnung baten. Die Tonaufnahme war etwa eine Minute lang. Zu hören war die Stimme eines wütenden Mannes, die der von Bolsonaro ähnelte, der mit jemandem stritt, der dem Anschein nach Bolsonaros Sohn Eduardo war. Er beschwerte sich darüber, im Krankenhaus festgehalten zu werden. Es ist zu hören, wie er sagt, er könne „dieses Theater" nicht länger ertragen, was darauf hindeutete, dass alles nur inszeniert sei. An diesem Tag war Bolsonaro noch Patient des Albert-Einstein-Krankenhauses in São Paulo. Im medizinischen Bericht hieß es, er habe er kein Fieber, werde intravenös ernährt und seine Darmfunktion sei wiederhergestellt.

Comprova konnte die Originalquelle der Aufnahme nicht ermitteln. Die Datei verbreitete sich hauptsächlich über WhatsApp, zu einer Zeit, als Dateien noch an bis zu 20 Chats weitergesendet werden konnten. Dadurch konnte sie sich schnell verbreiten und bald auch in andere soziale Netzwerke gelangen. Das machte es schier unmöglich, sie zur ursprünglichen Quelle zurückzuverfolgen. (WhatsApp hat mittlerweile die Anzahl der Gruppen eingeschränkt, an die man eine Nachricht weiterleiten kann.)

Da es Comprova nicht möglich war, die Quelle der Aufnahme zu identifizieren, konzentrierten wir uns auf konventionellere Recherchemethoden und baten das Instituto Brasileiro de Perícia (das brasilianische Institut für Gerichtsmedizin) um ein Gutachten. Experten verglichen die Aufnahme, die sich über WhatsApp verbreitete, mit einer Aufnahme von Bolsonaro aus einem Interview im April 2018. Sie kamen zu dem Schluss, dass die Stimme auf der Aufnahme nicht die Stimme Bolsonaros war.

Die Experten führten eine qualitative Analyse der Stimme, der Sprache und der Sprachmerkmale des Mannes durch, der in der Aufnahme zu hören war. Dann verglichen sie diese Parameter mit den beiden Stimmproben. In dieser Analyse untersuchten sie Vokal- und Konsonantenmuster, Sprachrhythmus und -geschwindigkeit, Intonationsmuster, Stimmqualität und Eigenheiten des Sprechers sowie die Verwendung bestimmter Wörter und grammatikalischer Regeln.

Das untere Bild zeigt zum Beispiel eine Frequenzanalyse von „Formanten“. So heißen jene Tonhöhen, die durch Vibrationen des Vokaltrakts erzeugt werden, dem Hohlraum, in dem der am Kehlkopf erzeugte Schall gefiltert wird. Die Luft im Inneren des Vokaltraktes schwingt in verschiedenen Tonhöhen, je nach Größe und Form der Öffnung. Das Bild zeigt eine Frequenzanalyse der Formanten mit den Vokalen „a“, „e“ und „o“. Die grünen Vokale entsprechen dem Audio-Muster, das sich über WhatsApp verteilte, und die blauen Vokale entsprechen einem Muster, das aus einem Interview stammt, welches Bolsonaro einige Tage vor dem Angriff gegeben hatte.



Grüne Vokale = Audio-Material aus WhatsApp, blaue Vokale = Audio-Material aus einem Interview von Bolsonaro einige Tage vor dem Angriff.

Eine zusätzliche Analyse ergab, dass der Sprecher im WhatsApp-Audio einen typischen Akzent aus dem ländlichen Raum des Bundesstaates São Paulo aufwies. Dieser kam jedoch in Bolsonaros Sprachmustern nicht vor. In den verglichenen Samples wurden Unterschiede in Resonanz, Artikulation, Sprechgeschwindigkeit und phonetischer Abweichung festgestellt.

Comprova konsultierte einen zweiten Experten. Auch der kam zu dem Schluss, dass sich die Stimme in der Aufnahme in mehreren Punkten von der Stimme Bolsonaros unterschied. Der Tonfall der Stimme schien ihm ein wenig schärfer zu sein als der von Bolsonaro. Er stellte zudem fest, dass das Sprechtempo schneller war als in einem anderen Video, das Bolsonaro im Krankenhaus aufgenommen hatte.

Was den Verdacht einer Fälschung zusätzlich erhärtete, war die schlechte Qualität der Aufnahme. Nach Ansicht erfahrener Experten ist dies ein typischer Trick bei Fälschungen: Eine geringere Qualität von Audios, Videos und Fotos erschwert deren Analyse.

Was Bolsonaros Reaktion betrifft, so haben seine Söhne Flavio und Carlos die Aufnahme in sozialen Netzwerken als „Fake News" bezeichnet. Würde man sie heute verbreiten, so würden vermutlich weniger Menschen glauben, dass die Stimme Bolsonaro gehört. Vor der Wahl hingegen, als man ihn täglich nur für 18 Sekunden im Fernsehen zu sehen bekam und er die Wahlkampfdebatten infolge der Krankenhausaufenthalte und Behandlungen verpasste, war die Stimme des derzeitigen Präsidenten nicht so bekannt. Das schuf eine Gelegenheit, mit einer falschen Tonaufnahme viele zu täuschen.

Mehr als ein Jahr später ist es immer noch schwer zu verstehen, warum auch Gruppen, die für Bolsonaro waren oder für seinen Sieg kämpften, diese Tonaufnahme teilten, die, wenn sie sich als authentisch erwiesen hätte, seine Chancen auf den Wahlsieg hätte zerstören können. Vermutlich werden wir das nie erfahren. Gerade deswegen aber ist die Datei ein eindrücklicher Beleg dafür, dass sich Inhalte in sozialen Medien rasend schnell verbreiten, wenn sie nur explosiv genug sind.

8. WIE MAN WEBSITES UNTERSUCHT

von: Craig Silverman
deutsche Bearbeitung: Marcus Engert

Craig Silverman *ist als Medienredakteur von BuzzFeed News für einen weltweiten Themenbereich zuständig, von Plattformen über Online-Falschinformationen bis hin zu Medienmanipulation. Er ist Herausgeber des „Verification Handbook" und des „Verification Handbook for Investigative Reporting" und Autor des Buches „Lies, Damn Lies, and Viral Content: How News Websites Spread (and Debunk) Online Rumors, Unverified Claims and Misinformation".*

Websites werden auch von denjenigen benutzt, die Medien und Publikum manipulieren wollen, um Einkünfte zu erzielen, E-Mails und andere persönliche Informationen zu sammeln, oder schlicht, um im digitalen Raum einen Brückenkopf aufzubauen, von dem aus man dann weiter agiert. Journalistinnen und Journalisten müssen also lernen, wie man eine Web-Präsenz untersucht und wie man eine solche Untersuchung, wenn möglich, in eine größere Recherche einbindet, die Social-Media-Accounts, Apps, Unternehmen und anderes einschließen kann. Denken Sie daran, dass Texte, Bilder oder die gesamte Website selbst im Laufe der Zeit verschwinden können – vor allem, wenn Sie anfangen, die Leute zu kontaktieren und Fragen zu stellen. Die Wayback Machine, mit der sich Websites archivieren und später wieder aufrufen lassen, auch wenn sie zwischenzeitlich verändert wurden, sollte zur Routine in Ihrer Arbeitsweise werden. Kann eine Seite dort nicht ordentlich gespeichert werden, benutzen Sie Dienste wie archive.today. Das ermöglicht es, in Ihrem Text für Belege auf die archivierte Seite zu verlinken, statt einer Seite noch Publikum zu verschaffen, die womöglich Falsch- oder Desinformationen verbreitet. (Auch Hunchly ist ein großartiges gebührenpflichtiges Werkzeug zur automatischen Erstellung eines persönlichen Archivs: Es archiviert die Seiten, die man besucht, in Echtzeit und während man daran arbeitet.) Diese Archivierungswerkzeuge sind zudem wichtig, um zu untersuchen, wie eine Website in der Vergangenheit ausgesehen hat. Die Installation der Browser-Erweiterung von Wayback Machine zu archivieren ist daher nicht die schlechteste Idee, da sie es einfacher macht, Seiten zu archivieren und sich frühere Versionen davon anzusehen.

Eine weitere nützliche Browser-Erweiterung ist Ghostery, die einem die auf einer Website installierten Tracker anzeigt. (Tracker analysieren das Verhalten der Besucher auf einer Website.) Auf diese Weise lässt sich schnell erkennen, ob eine Website Google Analytics und/oder Google AdSense-IDs verwendet, was bei einer der in diesem Kapitel beschriebenen Recherchemethoden hilfreich ist.

Wir schauen uns in diesem Kapitel vier Kategorien an, die man bei der Untersuchung einer Website analysieren kann: den Inhalt, den Code, die Analytics sowie die Registrierung.

INHALT

In der Regel erzählen Websites von selbst ein wenig darüber, was sie sind. Ob auf einer speziellen Info-Seite, einer Beschreibung in der Fußzeile oder an anderer Stelle – solche Beschreibungen sind ein guter Ausgangspunkt. Gleichzeitig kann bereits hier ein Mangel an klaren Informationen ein Hinweis darauf sein, dass die Seite in Eile erstellt wurde oder versucht, Einzelheiten über Eigentümer und Zweck zu verbergen. Hier einige Punkte, auf die man achten kann:

- Werden Eigentümer, Verantwortliche oder eine Firma genannt? Interessant kann auch sein, wenn es einen Bereich „Über diese Webseite" oder Ähnliches eben nicht gibt.
- Gibt es einen Bereich zu urheberrechtlichen Fragen und wenn ja, wird dort eine Person, eine Firma oder eine andere Website genannt?
- Werden in einer Datenschutzrichtlinie oder in Allgemeinen Geschäftsbedingungen Namen, Adressen oder Unternehmen aufgeführt? Unterscheiden sich diese Namen oder Firmen von dem, was in der Fußzeile, auf der Info-Seite oder an anderen Stellen auf der Website aufgeführt ist?
- Wenn die Website Artikel veröffentlicht, prüfen Sie, ob es eine Verfasserzeile gibt und ob es sich dabei um anklickbare Links handelt. Wenn ja, sehen Sie nach, ob diese zu einer Autorenseite mit weiteren Informationen führen, zum Beispiel zu einer Biographie oder Links zu Social-Media-Profilen des Autors.

- Verlinkt die Website auf Profile in den sozialen Netzwerken? Dies könnte auch in Form von kleinen Symbolen am oberen, unteren oder seitlichen Rand der Homepage oder in Form eines kleinen eingebetteten Kastens erfolgen, mit dem man dazu einlädt, die Facebook-Seite mit „Gefällt mir" zu markieren. Wenn die Seite Symbole für Plattformen wie Facebook und Twitter zeigt, fahren Sie mit der Maus darüber und schauen Sie unten links in Ihrem Browser-Fenster, ob dort eine Website angezeigt wird, zu der die Symbole führen. Oft wird sich für eine hastig erstellte Website nicht die Mühe gemacht, die genauen Profilnamen in Vorlagen zur Erstellung einer Website einzutragen. In diesem Fall wird der Link einfach als facebook.com/ ohne Benutzernamen angezeigt. Andersherum kann auch der Name eines Social-Media-Profils zu einem Hinweis werden.
- Listet die Seite irgendwelche Produkte, Kunden, Empfehlungen oder andere Menschen oder Unternehmen auf, zu denen eine Verbindung besteht und wo es sich lohnen könnte, sich das genauer anzuschauen?
- Graben Sie tiefer als nur auf der Startseite. Klicken Sie sich durch sämtliche Menü-Punkte. Scrollen Sie bis ganz zum Ende. Durchsuchen Sie Fußzeilen nach weiteren Unterseiten, bei denen man etwas finden könnte.

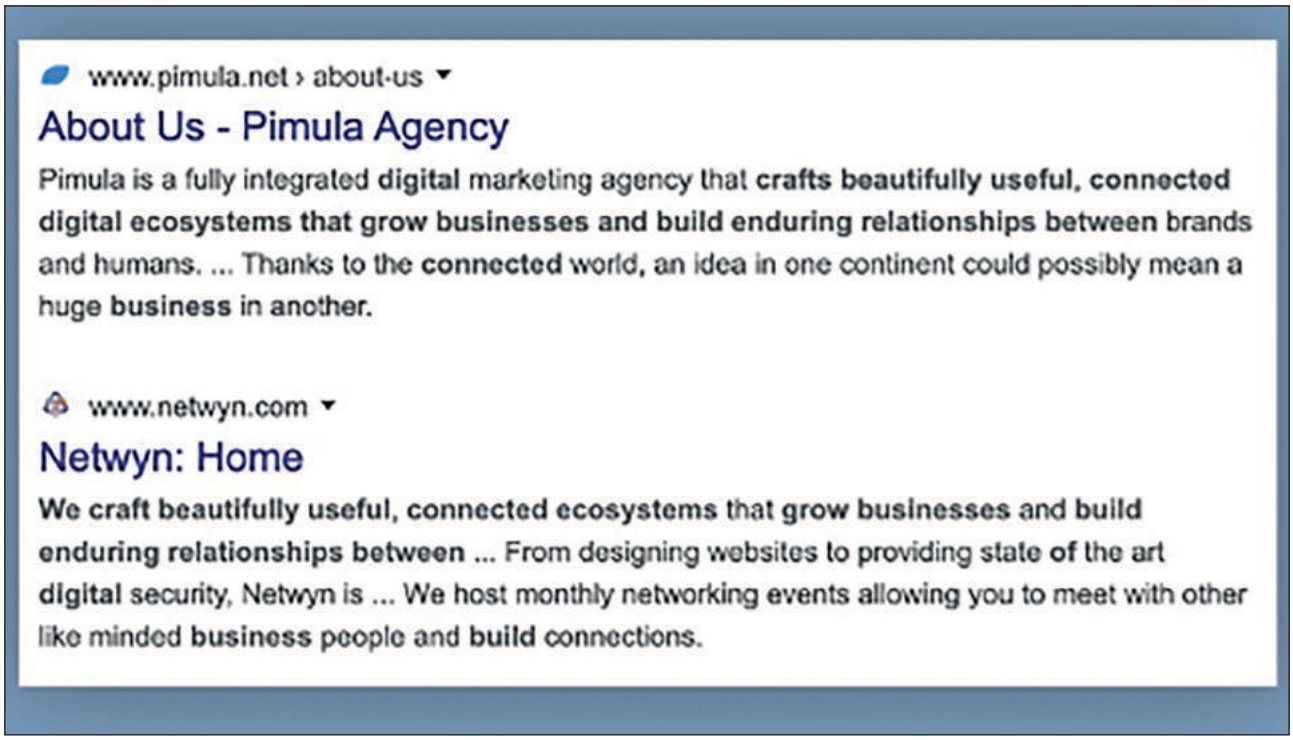
Ein wichtiger Teil der Recherche ist es, die Inhalte daraufhin zu prüfen, ob sie originär sind. Findet sich der Text im „Über diese Seite"-Bereich so auch auf anderen Seiten? Wurden andere Texte der Seite anderswo kopiert? Verbreitet die Website falsche oder irreführende Informationen oder trägt sie zur Durchsetzung einer bestimmten Agenda bei?

2018 recherchierte ich zu einem groß angelegten Betrug mit einem digitalen Werbeprogramm, an dem mobile Apps und Websites mit Inhalten ebenso beteiligt waren wie Briefkastenfirmen, falsche Mitarbeiter und erfundene Unternehmen. Schließlich fand ich mehr als 35 Websites, die mit dem Betrugssystem in Verbindung standen. Ein Weg, viele dieser Seiten zu finden, war, den Text auf der Info-Seite einer Website zu kopieren und in das Google-Suchfeld einzufügen. Ich fand sofort etwa 20 Websites mit genau dem gleichen Text:

Auf den Info-Seiten vieler am Betrug beteiligten Seiten stand wortgleich der gleiche Text. Er lautete: „Wir bringen unsere Services in mehr als vier Millionen Haushalte mit Set-Top-Boxen, und wir bieten mobile Video-Services an, die mehr als zehn Millionen Abonnenten erreichen."

Die Leute, die das Programm betrieben, hatten professionell Websites für ihre Scheinfirmen erstellt, damit diese legitim wirkten. Dies sollte potentielle Kunden des Werbe-Netzwerks schnell überzeugen, falls diese sich vor einer Buchung die Portale anschauen. Dazu gehörte zum Beispiel eine Firma namens Atoses. Auf ihrer Homepage waren mehrere Mitarbeiter mit Profilbildern aufgeführt. Eine Bilder-Rückwärtssuche mit diesen Bildern über Yandex (die beste Bildsuche für Gesichter) ergab schnell, dass es sich bei mehreren von ihnen um Archivbilder handelte:

Am Fußende seiner Website hatte Atoses diesen Text eingebaut: „We craft beautifully useful, connected ecosystems that grow businesses and build enduring relationships between online media and users." – was auf Deutsch so viel heißt wie: „Wir gestalten wunderbar nützliche, vernetzte Ökosysteme, die Unternehmen wachsen lassen und dauerhafte Beziehungen zwischen Online-Medien und Nutzern aufbauen." Und genau der gleiche Text erschien auf den Websites von gleich zwei anderen angeblichen Marketingagenturen:

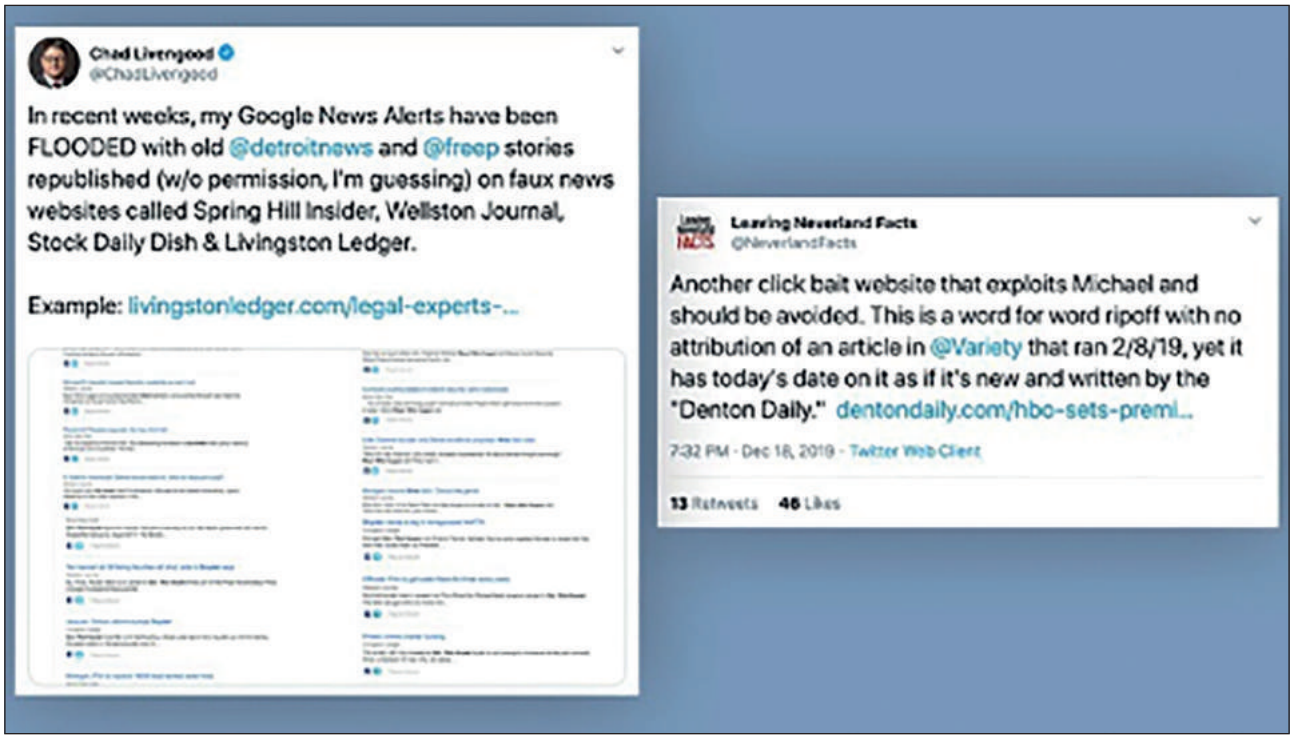


Wenn eine Firma Archivbilder für angebliche Mitarbeiter und geklaute Texte auf ihrer Website verwendet, dann können Sie sicher sein: Sie ist nicht, was sie vorgibt zu sein.

Manchmal ist es auch eine gute Idee, Textteile aus Nachrichtenartikeln auf journalistischen Websites zu kopieren und diese Textausschnitte mit einer Suchmaschine zu suchen. Manchmal steckt hinter einer Website, die vorgibt, eine zuverlässige Nachrichtenquelle zu sein, auch nur eine Seite, die andere plagiiert. 2019 stieß ich auf eine Website namens forbesbusinessinsider.com. Die Seite gab vor, eine Nachrichtenseite über die Technologiebranche zu sein. In Wirklichkeit handelte es sich um ein systematisches Massenplagiat von Artikeln aus einer Vielzahl von Quellen, darunter, witzigerweise, ein Artikel über gefälschte lokale Websites, den ich selbst geschrieben hatte.

Ein weiterer grundlegender Schritt besteht darin, die Adresse einer Website zu kopieren und sie über Google zu suchen. Nehmen wir zum Beispiel forbesbusinessinsider.com. Mit der Suche erhält man einen Eindruck davon, wie viele Unterseiten der Seite durch Suchmaschinen bereits indexiert und erfasst wurden, und man findet andere Personen oder Websites, die diese Seite erwähnen oder verlinken. Man kann auch prüfen, ob die Seite bereits bei Google News gelistet wurde, indem man die Startseite von Google News öffnet und dort forbesbusinessinsider.com in das Suchfeld eingibt.

Was ebenfalls aufschlussreich sein kann: den Link zur Seite in die Suchmasken von Twitter oder Facebook einzugeben. So lässt sich schauen, ob Menschen die Seite verbreiten und wenn ja, wer und wie. In einer Recherche bin ich einmal über die Seite dentondaily.com gestolpert. Auf der Startseite fanden sich nur ein paar Artikel aus 2020, aber in Twitter konnte ich sehen, dass die Seite schon vorher gestohlene Artikel beworben und verbreitet hat, was Menschen bemerkten und sich darüber beschwerten. Diese älteren Texte waren zwischenzeitlich von der Seite entfernt worden, aber die Tweets mit den Beschwerden waren Belege für dieses frühere Verhalten.



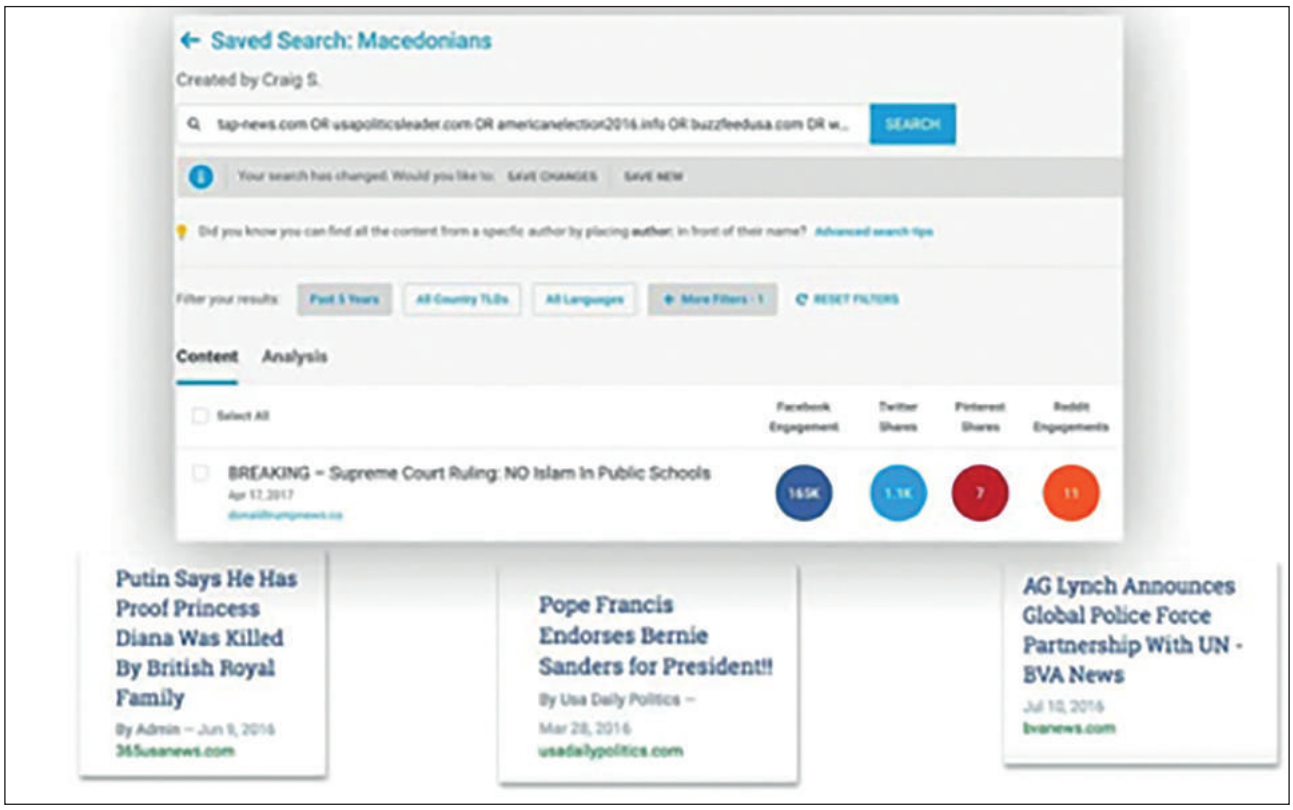
Twitter-Nutzer beschweren sich über Plagiate durch die Seite dentondaily.com. Der links abgebildete Nutzer schreibt, er sei durch einen Suchauftrag auf Google (Google News Alert) darauf aufmerksam geworden. Der rechte Nutzer schreibt von einer „wortwörtlichen Übernahme ohne Quellenangabe".

Sobald Sie sich in den Inhalt einer Website eingearbeitet haben, ist es an der Zeit zu verstehen, wie ihre Inhalte sich verbreiten. Dazu schauen wir uns zwei Hilfsmittel an: BuzzSumo und CrowdTangle.

Im Jahr 2016 arbeitete ich mit dem Rechercheur Lawrence Alexander zusammen, um amerikanische Politik-Nachrichtenseiten zu überprüfen, die von Übersee aus betrieben werden. Schon bald lag unsere ganze Aufmerksamkeit auf Seiten, die aus Veles, einer Stadt in Nordmazedonien, zu kommen schienen. Mit Hilfe der Informationen aus der Domainregistrierung (mehr dazu weiter unten) konnten wir mehr als 100 US-amerikanische Politik-Seiten identifizieren, die von dort kamen.

Ich wollte ein Gefühl dafür bekommen, wie erfolgreich deren Inhalte waren und welche Art von Geschichten sie veröffentlichten. Ich nahm also die Adressen mehrerer Websites, die mir am aktivsten wirkten, und legte mit ihnen eine Suche in BuzzSumo an. BuzzSumo kann eine Liste der Inhalte auf einer Website anzeigen und diese nach den Interaktionen, die diese Inhalte auf Facebook, Twitter, Pinterest und Reddit erhalten haben, sortieren. (Man kann BuzzSumo als kostenlose Version nutzen, allerdings liefert das kostenpflichtige Produkt weitaus mehr Ergebnisse.)

Es war sofort klar, dass die Artikel mit den meisten Interaktionen auf Facebook völlig falsche Dinge verbreiteten. Damit hatten wir wichtige Informationen und einen ganz anderen Blickwinkel, der sich von dem früherer Recherchen unterschied. Die Abbildung unten zeigt die Basis-Ergebnisseite von BuzzSumo, auf der Facebook-, Twitter-, Pinterest- und Reddit-Interaktionen für eine bestimmte Seite sowie einige Beispiele für falsche Berichte aus dem Jahr 2016 aufgelistet sind:

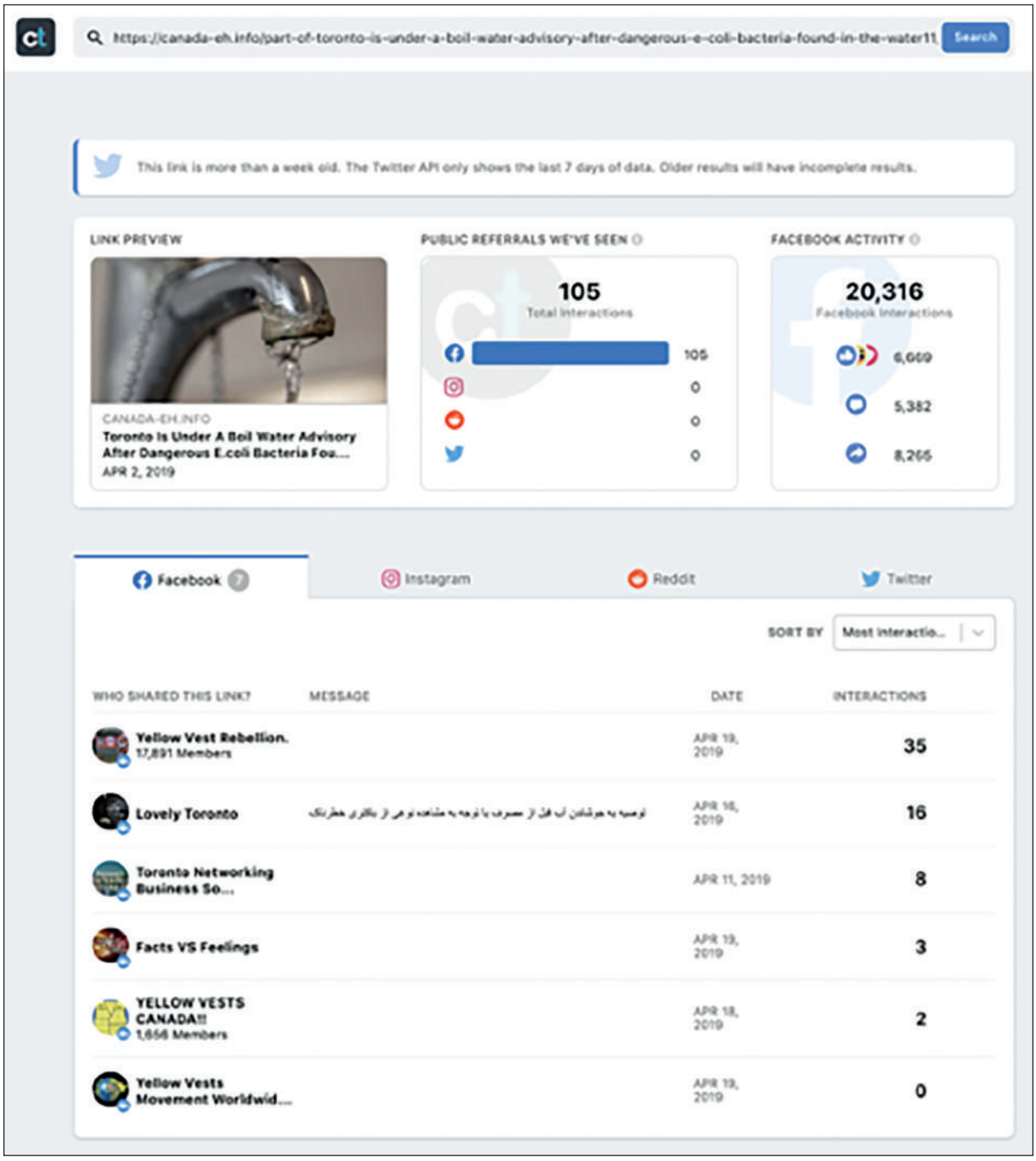


Drei Überschriften zu Artikeln aus 2016, die komplett falsch sind: „Putin sagt, er habe Beweise, dass Prinzessin Diana von der britischen Königsfamilie umgebracht wurde", „Papst Franziskus unterstützt Bernie Sanders als Präsident" und „Generalstaatsanwalt Lynch kündigt weltweite Polizei-Partnerschaft mit den Vereinten Nationen an".

Eine weitere Möglichkeit, um festzustellen, wie sich der Inhalt einer Website auf Facebook, Twitter, Instagram und Reddit verbreitet, ist CrowdTangle. Man kann die kostenlose Browser-Erweiterung CrowdTangle installieren oder die webbasierte Suche verwenden. Beide haben die gleichen Funktionen, wir arbeiten hier mit der Online-Suche. (Beide sind kostenlos, aber Sie benötigen ein Facebook-Konto für den Zugriff.)

Der Hauptunterschied zwischen BuzzSumo und CrowdTangle besteht darin, dass Sie die URL einer Website in BuzzSumo eingeben können und automatisch den (in sozialen Netzwerken) erfolgreichsten Inhalt dieser Website angezeigt bekommen. CrowdTangle kommt quasi aus der anderen Richtung: Man kann damit einen speziellen Link, nicht eine ganze Seite, untersuchen, und sieht, wo dieser Link sich wie und durch wen am erfolgreichsten verbreitet hat. Wenn Sie also buzzfeednews.com in CrowdTangle eingeben, werden Sie die Statistiken für genau diese Seite sehen, während BuzzSumo die gesamte Domain nach allen Inhalten durchsuchen und diese einzelnen Links dann auflisten wird. Ein weiterer Unterschied besteht darin, dass das Link-Suchtool und die Erweiterung von CrowdTangle für Twitter nur über die letzten sieben Tage hinweg funktioniert. In BuzzSumo besteht diese zeitliche Einschränkung nicht.

Als Beispiel habe ich den Link zu einer alten, falschen Geschichte, wonach in Toronto Wasser abgekocht werden müsse, in die CrowdTangle-Linksuche eingegeben. (Auf der Website wurde die Geschichte später gelöscht, zum Zeitpunkt der Erstellung dieses Kapitels war der Link selbst aber noch aktiv.) CrowdTangle gibt an, dass dieser Link seit seiner Veröffentlichung mehr als 20.000 Reaktionen, Kommentare und Weiterverteilungen auf Facebook erhalten hat. Wir sehen auch einige jener Seiten und öffentlichen Gruppen, die den Link verbreitet haben, ähnlich für Instagram, Reddit und Twitter. (Erinnerung: für Twitter nur die Ergebnisse über die letzten sieben Tage.)



Auffällig scheint zu sein, dass die hohe Anzahl der insgesamt gefundenen Facebook-Interaktionen nicht zu den kleinen Zahlen hinter den Seiten und Gruppen passen. Das hängt zumindest teilweise damit zusammen, dass einige der großen Seiten, die den Link nach seiner ersten Veröffentlichung verbreiteten, später von Facebook entfernt wurden. Das ist eine nützliche Erinnerung daran, dass CrowdTangle nur Daten von einerseits aktiven und andererseits öffentlichen Konten auswertet.

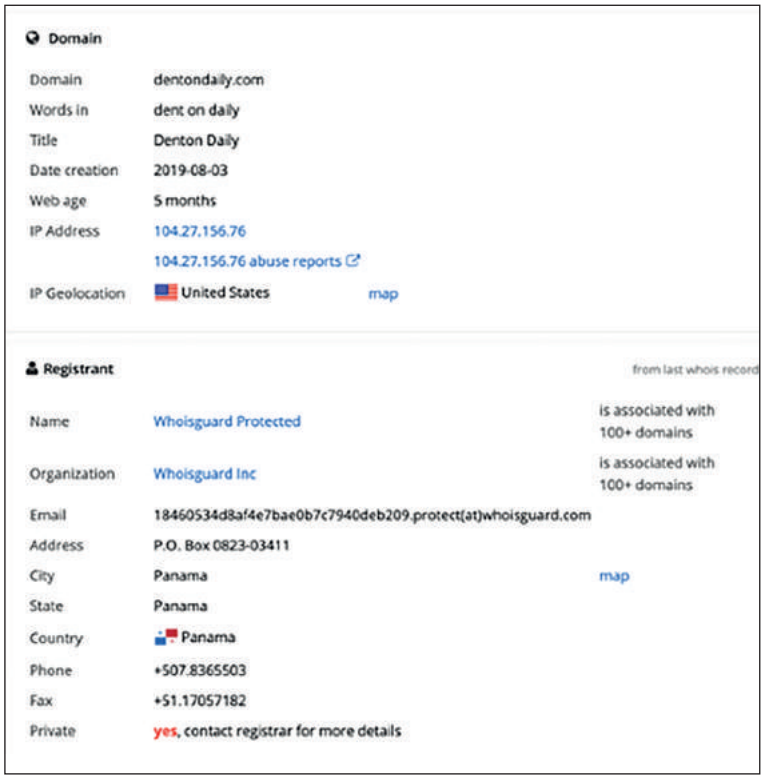
Das Ergebnis ist also eine Auswahl, aber dennoch eine äußerst nützliche, da sich oft eine klare Verbindung zwischen bestimmten Social-Media-Konten und einer Website zeigen lässt. Wenn immer wieder dieselbe Facebook-Seite konsistent – oder sogar ausschließlich – Inhalte von einer bestimmten Website teilt, dann kann das ein Hinweis darauf sein, dass sie von denselben Personen betrieben wird. Der nächste Schritt ist dann, sich in diese Seite einzuarbeiten, um Informationen aus der Facebook-Seite mit jenen auf der Website zu vergleichen und möglicherweise die beteiligten Personen und ihre Motivationen zu identifizieren. Einige der in CrowdTangle aufgelisteten Ergebnisse können auch von Personen stammen, die den Artikel in einer Facebook-Gruppe teilen. Notieren Sie also das Konto, das den Link verteilt hat, und prüfen Sie, ob es noch andere Inhalte der Website verbreitet hat. Auch hier könnte eine Verbindung bestehen.

REGISTRIERUNG

Jeder Domainname im Web ist Teil einer zentralen Datenbank, in der Basisinformationen über seine Anmeldung und Geschichte hinterlegt sind. In manchen Fällen haben wir Glück und können darin Informationen über die Person oder Organisation finden, die für die Registrierung einer Domain bezahlt hat. Man kann diese Informationen mit einer sogenannten Whois-Suche abrufen, die von vielen kostenlosen Tools angeboten wird. Es gibt auch eine Handvoll sehr guter entweder kostenloser oder preisgünstiger Anbieter, die zusätzliche Informationen abrufen können, zum Beispiel wer eine Domain früher besessen hat, auf welchen Servern sie gehostet wurde und andere nützliche Details.

Ein Wermutstropfen ist, dass es relativ günstig ist, für den Schutz der persönlichen Daten zu bezahlen, wenn man eine Domain registriert. Darum lautet bei Whois-Abfragen das Ergebnis oft „Registration Private", „WhoisGuard Protected" oder „Perfect Privacy LLC". Auch wenn wir für Registrant in dem Fall keine Infos bekommen, sehen wir immer noch das Datum der letzten Registrierung der Domain, den Zeitpunkt des Ablaufs dieser Registrierung und die IP-Adresse im Internet, unter der die Website verwaltet wird. Einer der besten kostenlosen Anbieter, um eine Domain und ihre Geschichte zu untersuchen, ist DomainBigData. Man kann dort neben dem Link selbst auch nach E-Mail-Adressen oder den Namen von Firmen oder Personen suchen. Andere erschwingliche Anbieter, die man sich einmal als Lesezeichen hinterlegen sollte, sind DNSlytics, Security Trails und Whoisology. Einer der besten Anbieter ist DomainTools, dessen *Iris Investigation Platform* sehr gut, aber auch verhältnismäßig teuer ist.

Wenn wir nun dentondaily.com bei DomainBigData eingeben, dann sehen wir: Die Registrierungsdaten sind anonym, unter Registrant steht der Eintrag „Whoisguard Protected". Was wir aber erkennen können ist, dass der letzte Registrierungseintrag im August 2019 erfolgte.



Domain

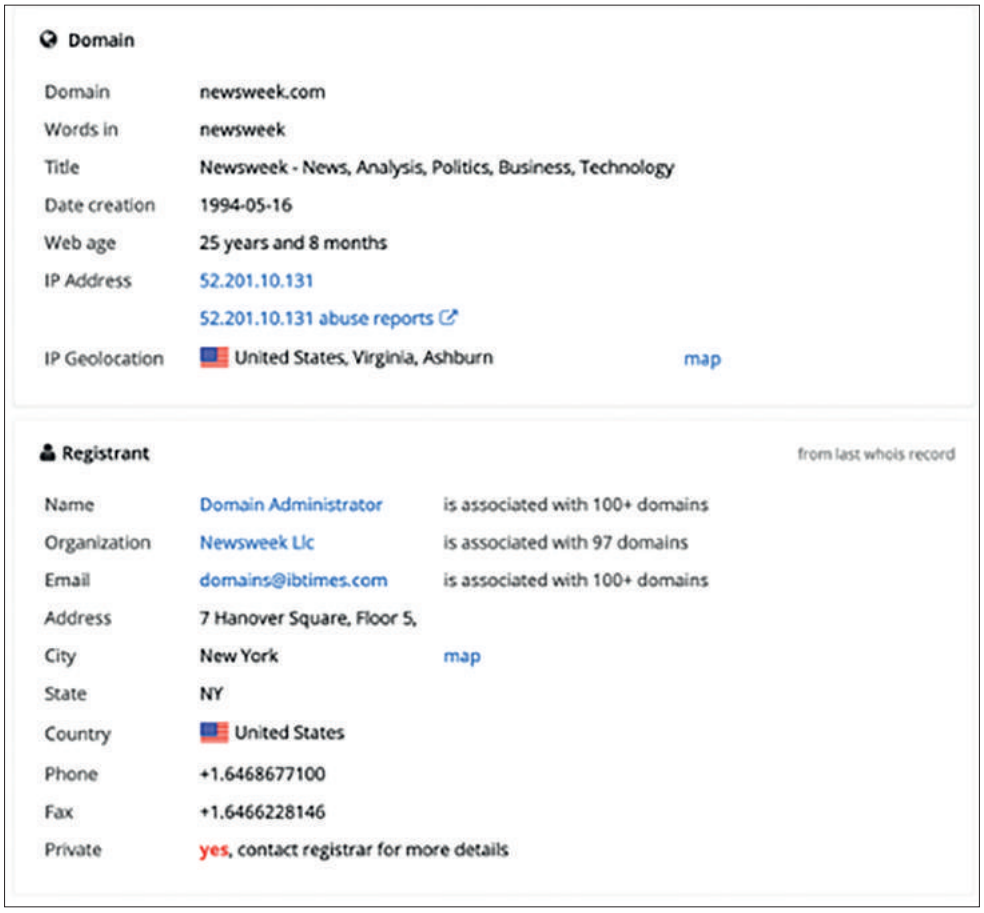
Domain	dentondaily.com
Words in	dent on daily
Title	Denton Daily
Date creation	2019-08-03
Web age	5 months
IP Address	104.27.156.76
	104.27.156.76 abuse reports
IP Geolocation	United States map

Registrant from last whois record

Name	Whoisguard Protected	is associated with 100+ domains
Organization	Whoisguard Inc	is associated with 100+ domains
Email	18460534d8af4e7bae0b7c7940deb209.protect(at)whoisguard.com	
Address	P.O. Box 0823-03411	
City	Panama	map
State	Panama	
Country	Panama	
Phone	+507.8365503	
Fax	+51.17057182	
Private	yes, contact registrar for more details	

Für dentondaily.com können wir einsehen, dass die Seite am 3. August 2019 auf einen Inhaber in Panama registriert wurde. Als Name ist unter Registrant nur „Whoisguard Protected" hinterlegt, was bedeutet: Der Inhaber hat dafür gezahlt, anonym zu bleiben.

Um noch ein Beispiel zu untersuchen, geben wir einfach einmal newsweek.com bei DomainBigData ein. Sofort ist zu erkennen: Der Inhaber hat nicht dafür bezahlt, anonym zu bleiben. Es gibt einen Namen einer Firma, eine Anschrift, eine Telefon- und eine Fax-Nummer.



Domain

Domain	newsweek.com
Words in	newsweek
Title	Newsweek - News, Analysis, Politics, Business, Technology
Date creation	1994-05-16
Web age	25 years and 8 months
IP Address	52.201.10.131
	52.201.10.131 abuse reports
IP Geolocation	United States, Virginia, Ashburn map

Registrant from last whois record

Name	Domain Administrator	is associated with 100+ domains
Organization	Newsweek Llc	is associated with 97 domains
Email	domains@ibtimes.com	is associated with 100+ domains
Address	7 Hanover Square, Floor 5,	
City	New York	map
State	NY	
Country	United States	
Phone	+1.6468677100	
Fax	+1.6466228146	
Private	yes, contact registrar for more details	

Im zweiten Beispiel ist auch zu sehen, dass die Domain seit Mai 2014 bei diesem Eigentümer liegt und dass die Seite im Moment auf der IP-Adresse 52.201.10.13 abgelegt ist.

An dieser Stelle sollte einem auffallen, dass der Name der Firma, die E-Mail-Adresse und die IP-Adresse farbig hervorgehoben sind – denn sie sind Links. Ein Klick darauf zeigt uns an, welche anderen Domains diese Registrierungsdaten mit unserer Domain hier teilen. Im Laufe einer Recherche sind solche Verbindungen extrem wichtig, daher sollte immer im Blick behalten werden, zu versuchen, andere Domains zu finden, die auf die gleiche Person oder Firma registriert sind. Zu den IP-Adressen ist wichtig zu wissen: Es können Websites auf dem gleichen Server liegen, die schlicht nichts miteinander zu tun haben. Das geschieht in der Regel, weil zwei Menschen sich unabhängig voneinander für den gleichen Anbieter entschieden haben, bei dem sie ihre Website erworben haben. Als Grundregel kann man deshalb vielleicht festhalten: Je weniger Websites auf einem Server liegen, desto wahrscheinlicher ist, dass es eine Verbindung gibt. Aber: Sicher ist das nicht. Wenn Sie auf einem Server hunderte und aberhunderte Websites sehen, dann haben die vermutlich keine Verbindung, was die Eigentümer betrifft. Aber wenn, sagen wir, nur neun Websites auf einem Server sind, und die Seite, die Sie interessiert, hat die Registrierungsdaten anonymisiert, dann kann es sich lohnen, eine Whois-Suche mit den anderen acht Seiten durchzuführen, um zu überprüfen, ob einige von ihnen womöglich den gleichen Eigentümer eingetragen haben, und um zu schauen, ob sich dazu Informationen finden lassen. Menschen bezahlen manchmal für Anonymisierung bei der Registrierung von einigen Websites und tun das bei anderen Seiten nicht.

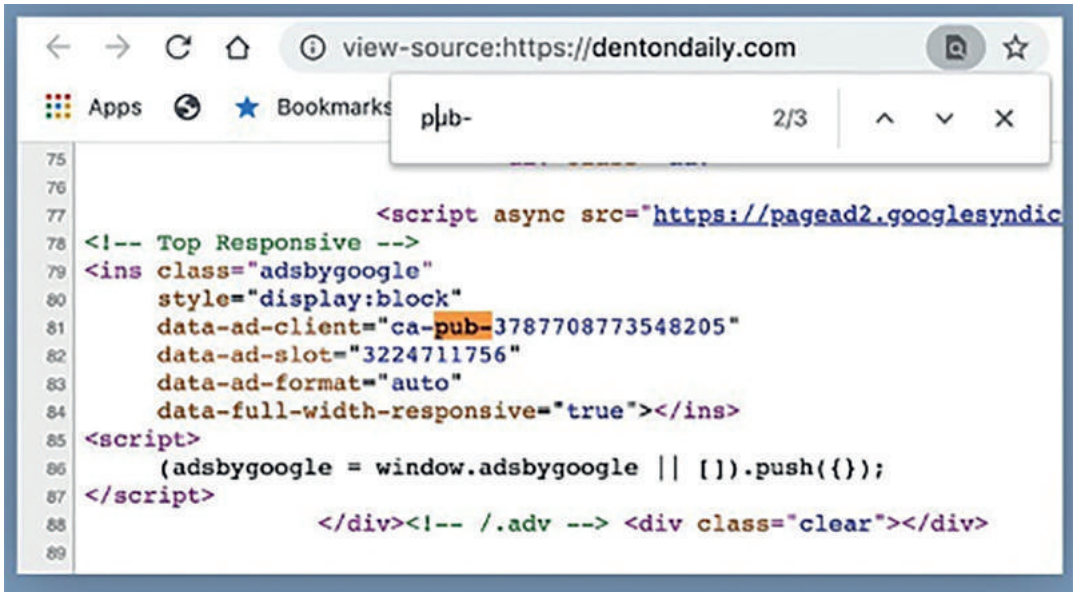
Seiten über die IP-Adresse, den Inhalt und/oder die Registrierungsdaten in Verbindung zu bringen, ist ein vielversprechender Weg, um Netzwerke offenzulegen und die Akteure hinter ihnen zu finden. Schauen wir uns nun einen anderen Weg dafür an: über den Quellcode einer Website.

CODE UND ANALYTICS

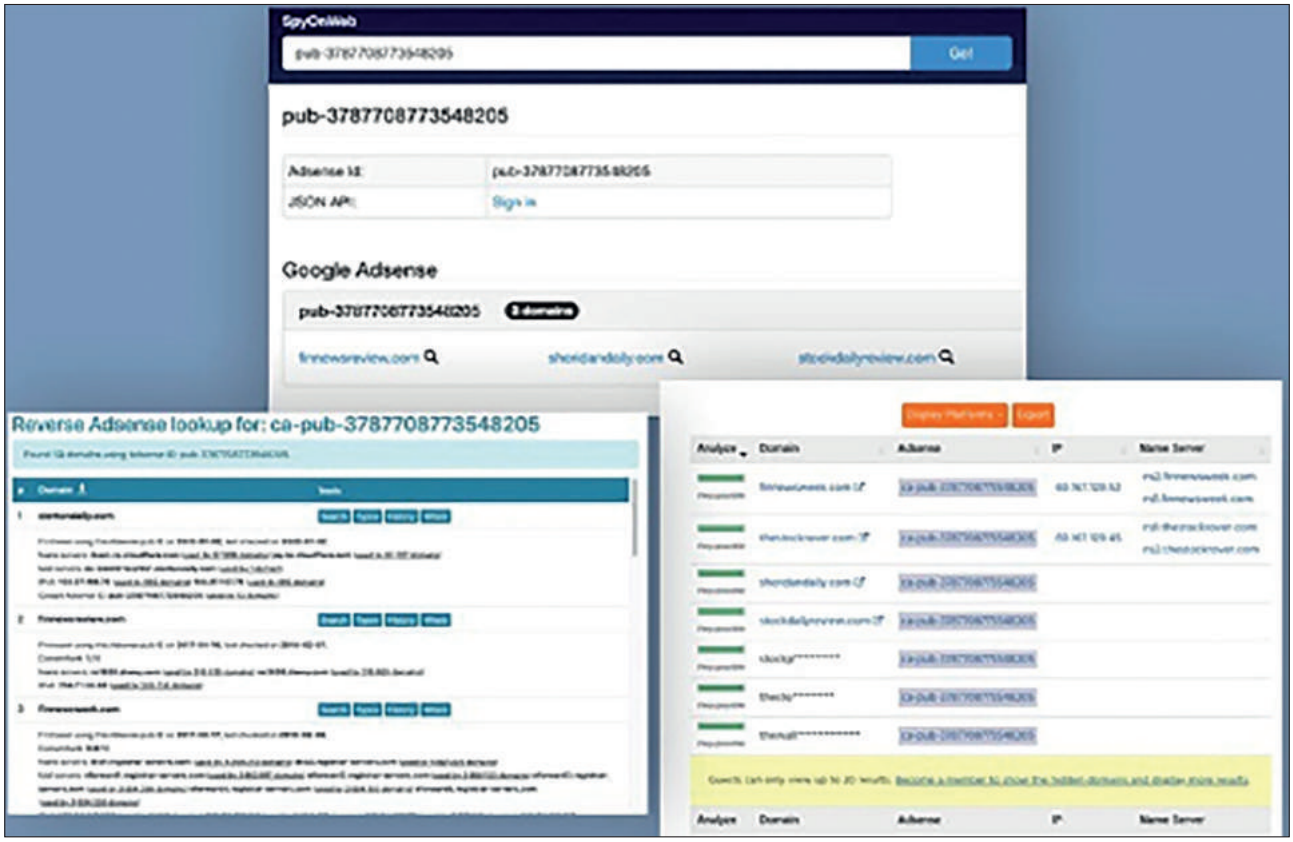
Dieser Ansatz, der zuerst von Lawrence Alexander von Bellingcat beschrieben wurde, beginnt damit, dass man sich den Quellcode einer Website anschaut, um zu prüfen, ob man darin einen Code für Google Analytics und/oder Google AdSense finden kann. Es handelt sich dabei um weitverbreitete Dienste von Google, die es im ersten Fall, einem Websitebetreiber erlauben, Statistiken über die eigene Seite zu erheben, und im zweiten Fall, Geld mit Werbung auf der Seite zu verdienen. Einmal eingebaut bekommt jede Website eine eindeutige ID, die mit dem Analytics- oder AdSense-Konto des Eigentümers verknüpft ist. Betreibt jemand mehrere Websites, wird häufig dasselbe Analytics- oder AdSense-Konto verwendet, um die Seiten (oder die Einnahmen) möglichst effektiv zu verwalten. Dies bietet Rechercheuren die Möglichkeit, eine Verbindung zwischen scheinbar unverbundenen Websites aufzuzeigen, einfach indem man in zwei Quellcodes von zwei Websites die gleiche Google-ID findet. Was sich kompliziert anhört, ist glücklicherweise sehr einfach.

Zuerst rufen Sie die Seite auf, die Sie untersuchen. Wir bleiben hier einmal bei dentondaily.com. (Wie der Quellcode bei anderen Browsern eingesehen werden kann, steht hier: https://praxistipps.chip.de/quelltext-einer-website-anzeigen_1280) Im Browser Chrome klicken Sie, wenn Sie mit einem Mac arbeiten, danach auf „Anzeigen“, dann auf „Entwickler“ und im sich dort öffnenden Menü auf „Quelltext anzeigen“. Auf einem PC drücken Sie stattdessen ctrl-U.

Alle Google Analytics-IDs beginnen mit „ua-“, gefolgt von einer Zahlenfolge. AdSense-IDs haben „pub-“ vor der Zahlenfolge stehen. Und danach braucht man im Quellcode einfach nur zu suchen. Auf einem Mac drücken Sie cmd-F; auf einem PC ctrl-F. Daraufhin wird ein kleines Suchfeld angezeigt. Geben Sie „ua-“ oder „pub-“ dort ein, und Sie sehen farbig markiert alle Suchtreffer im Code.



Sobald Sie auf diese Weise eine ID gefunden haben, kopieren Sie sie und suchen Sie anschließend mit Diensten wie SpyOnWeb, DNSlytics, NerdyData or AnalyzeID danach, welche Seiten diese ID noch benutzen. Mitunter zeigen verschiedene Dienste unterschiedliche Ergebnisse (nicht) an, daher sollte man eine ID ausführlich testen und die Ergebnisse vergleichen. In diesem Beispiel hier hat SpyOnWeb drei Websites gefunden, die die gleiche AdSense-ID verwenden, DNSlytics und AnalyzeID hingegen fanden einige Seiten mehr.



Es kann vorkommen, dass eine Website eine ID in der Vergangenheit genutzt hat, diese aktuell aber nicht mehr benutzt. Daher ist es wichtig, die Suche im Quellcode mit allen Seiten durchzuführen, auf die man gestoßen ist. Hilfreich kann auch sein, dass sowohl die AdSense- als auch die Analytics-ID in älteren Versionen der Seite, die in der Wayback Machine archiviert wurden, noch vorhanden sein können. Wenn also auf einer aktuellen Version einer Seite keine IDs zu finden sind, lohnt es sich möglicherweise, sich durch die archivierten älteren Versionen zu arbeiten. All diese Dienste sind kostenlos. Mitunter werden aber Gebühren fällig, wenn man ausführlichere Ergebnisse erhalten möchte, vor allem dann, wenn die ID, die man untersucht, auf einer sehr großen Anzahl von Websites verwendet wird.

Noch ein letztes Wort zur Untersuchung von Quellcodes: Es lohnt sich, einmal in Ruhe durch den ganzen Code zu gehen, auch wenn man von HTML, JavaScript, PHP oder anderen gängigen Programmiersprachen nichts versteht. So wird manchmal vergessen, den Titel einer Unterseite oder Website zu ändern, wenn für eine neue Seite die gleiche Designvorlage wiederverwendet wird. Auch solch einfache Fehler können wertvolle Anknüpfungspunkte bieten.

Als ich den Anzeigenbetrug mit Scheinfirmen wie Atoses untersuchte, stieß ich unter anderem auf eine Firma namens FLY Apps. Ich schaute mir den Quellcode ihres nur aus einer Seite bestehenden Internetauftritts an, und ziemlich weit oben im Code stach mir das Wort „Loocrum“ ins Auge (Hervorhebung von mir):

Dieses Wort zu googlen brachte mich zu einer Firma namens Loocrum, die genau das gleiche Websitedesign wie FLY Apps verwendete und teilweise den gleichen Inhalt hatte. Eine Whois-Suche ergab, dass die E-Mail-Adresse, die zur Registrierung von loocrum.com verwendet wurde, auch zur Registrierung anderer Tarnfirmen verwendet worden war – Firmen, die ich zuvor innerhalb des Betrugsschemas gefunden hatte. Diese Verbindung zwischen FLY Apps und Loocrum lieferte einen wichtigen und entscheidenden Beleg dafür, dass vier Männer, die FLY Apps betrieben, auch mit diesem von mir untersuchten Gesamtsystem in Verbindung standen. Und diese wichtige Verbindung kam nur zutage, weil ich etwas ziellos durch den Quellcode blätterte und nach Worten im Klartext suchte, die dort fehl am Platz schienen.

FAZIT

Auch mit all den oben genannten Ansätzen und Werkzeugen werden Sie oft genug das Gefühl haben, in eine Sackgasse zu geraten. Oft aber gibt es eine andere Möglichkeit, aus dieser wieder herauszufinden und auf einer Website Verbindungen oder Zugänge zum nächsten Schritt in der Recherche zu finden. Darum: Klicken Sie auf jeden Link. Studieren Sie den gesamten Inhalt. Lesen Sie den Quellcode. Prüfen Sie, wer auf der Seite zitiert oder auf wen verwiesen wird. Prüfen Sie, wer die Seite verbreitet. Und untersuchen Sie wirklich alles, was Ihnen sonst noch einfällt, um herauszufinden, was dort vor sich geht.

9. WERBUNG IN SOZIALEN NETZWERKEN UNTERSUCHEN

von: Johanna Wild
deutsche Bearbeitung: Marcus Engert

Johanna Wild *ist eine auf öffentliche Quellen spezialisierte Rechercheurin bei Bellingcat, wo sie sich auf die Entwicklung von Techniken und Werkzeugen für digitale Recherchen spezialisiert hat. Sie kommt aus dem Journalismus und hat in der Vergangenheit mit Journalistinnen und Journalisten in (ehemaligen) Konfliktgebieten gearbeitet. In Ostafrika hat sie für Voice of America Journalisten bei der Produktion von Sendungen unterstützt.*

Die Anzeigen, die Ihnen in den sozialen Medien angezeigt werden, und die Anzeigen, die jemand, der im Bus oder in der Bahn neben Ihnen sitzt, angezeigt bekommt, sind nicht die gleichen. Je nach Standort, Geschlecht, Alter und den Dingen, die Ihnen gefallen oder die Sie online geteilt haben, bekommen Sie möglicherweise Anzeigen für luxuriöse Feriensuiten in Málaga angezeigt und Ihr Nachbar bekommt Anzeigen für japanische Handyspiele.

Microtargeting – das heißt die Kategorisierung der Nutzer in eng zugeschnittene Einzel-Zielgruppen, um ihnen Anzeigen zu zeigen, die zu ihren Lebensumständen und Interessen passen – ist zu einem wichtigen Thema bei Wahlen geworden. Die Sorge dahinter ist, dass gezielte Kampagnen mit Anzeigen, die Angst oder Hass schüren oder falsche Informationen verbreiten, bestimmte Teile der Bevölkerung ansprechen könnten – und dass verschiedenen Gruppen auf verschiedenen Kanälen verschiedene Botschaften präsentiert werden. Normalerweise werden Anzeigen von Politikern, die in sozialen Netzwerken geschaltet werden, keiner Faktenprüfung unterzogen. Facebook hat beispielsweise im Januar 2020 bekräftigt, dass man weiterhin jede politische Werbung zulassen will, solange die sich an die Regularien des Netzwerks hält. Das bedeutet: Bestimmte Nutzergruppen können mit Anzeigen angesprochen werden, die Falschinformationen zu wichtigen politischen oder sozialen Themen enthalten.

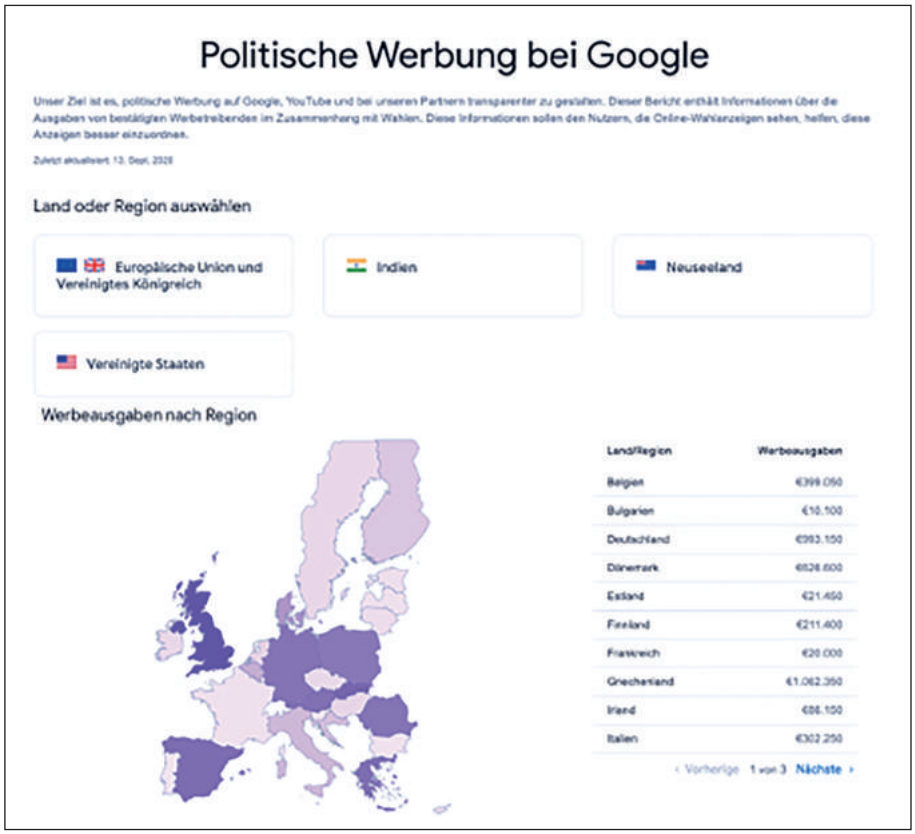
Bis vor kurzem war es für Journalistinnenen und Journalisten und Forscherinnen und Forscher annähernd unmöglich, Einblicke in die Anzeigen zu erhalten, die sich an verschiedene Nutzergruppen richten. Als Reaktion auf die öffentliche Kritik an solch mangelnder Transparenz schufen mehrere soziale Netzwerke Anzeigenbibliotheken, die es jedem ermöglichen, Informationen über die auf der jeweiligen Plattform veröffentlichten Anzeigen einzusehen – egal ob man zur Zielgruppe einer Anzeige gehört hätte oder nicht.

Speziell der Bibliothek von Facebook wurde vorgeworfen, nicht zuverlässig alle gebuchten Anzeigen anzuzeigen. Wenn Sie also diese Bibliotheken nutzen, nehmen Sie sich ruhig etwas Zeit, um zu prüfen, ob alle Anzeigen, die Sie sehen, wenn Sie durch das Netzwerk scrollen, dort auch zu finden sind.

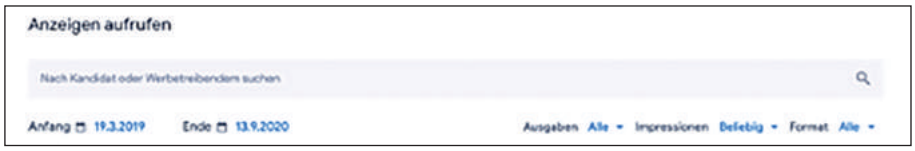
Anzeigenbibliotheken sind nichtsdestotrotz ein wichtiger Schritt zu mehr Transparenz und bieten spannende neue Möglichkeiten, digitale Anzeigen zu untersuchen. Die folgenden Techniken helfen beim Einstieg in die Recherchen zu Anzeigen, die auf wichtigen Plattformen wie Google, Twitter und Facebook geschaltet wurden.

GOOGLE

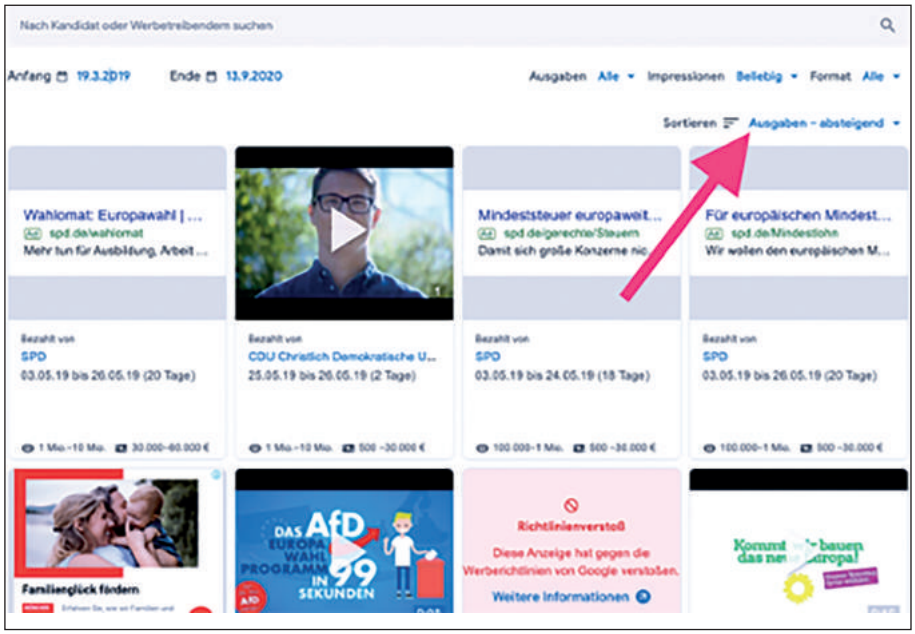
Googles Übersicht über politische Werbung ist gut in seinem Transparenzbericht versteckt. Über diesen Link gelangen Sie zum Bereich über politische Werbung, der Informationen über Google- und YouTube-Anzeigen aus der Europäischen Union, Indien und den Vereinigten Staaten enthält. Klicken Sie auf eine dieser Regionen, erhalten Sie eine Liste der Länder, die dort geschalteten Anzeigen und die dafür ausgegebenen Summen seit dem Start des Berichts.



Per Klick auf eines der Länder kommt man zu der jeweiligen Datenbank.

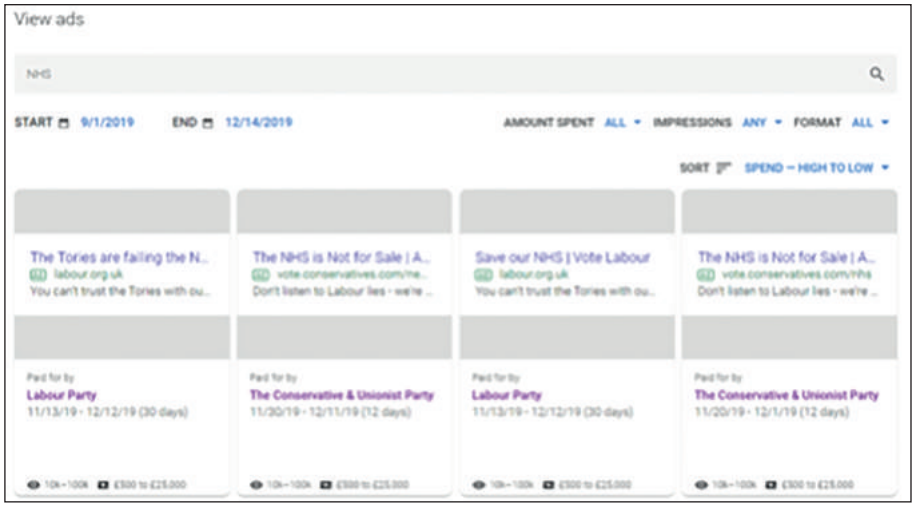


Die Ergebnisse lassen sich anschließend filtern: nach Datum, Höhe der Ausgaben, Häufigkeit, mit der eine Anzeige den Benutzern angezeigt wird (Impressionen), und dem Format der Anzeige, also ob Sie video-, bild- oder textbasierte Anzeigen aufgelistet bekommen möchten. Es ist auch ziemlich einfach, die größten Spender zu finden. Wenn Sie zum Beispiel die größten politischen Werbekampagnen sehen möchten, die in Deutschland seit dem Start des Berichts bis Januar 2020 bei Google gebucht wurden, ändern Sie einfach die Sortierung in „Ausgaben – absteigend“:

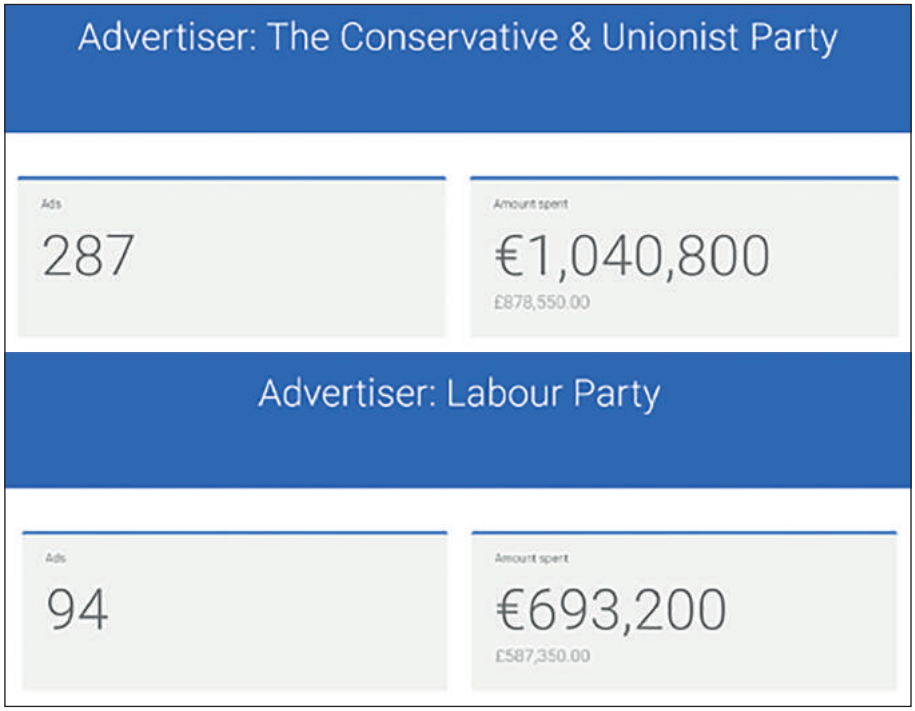


Wenig überraschend finden die größten Anzeigenkäufe kurz vor Wahlen oder an Wahltagen selbst statt, in unserem Fall: zu den Europawahlen (23. bis 26. Mai 2019). Wir können auch sehen, dass die CDU, die AfD und die Grünen Anzeigen für YouTube-Videos buchten, die SPD hingegen einzelne Themen und Stichworte bewarb.

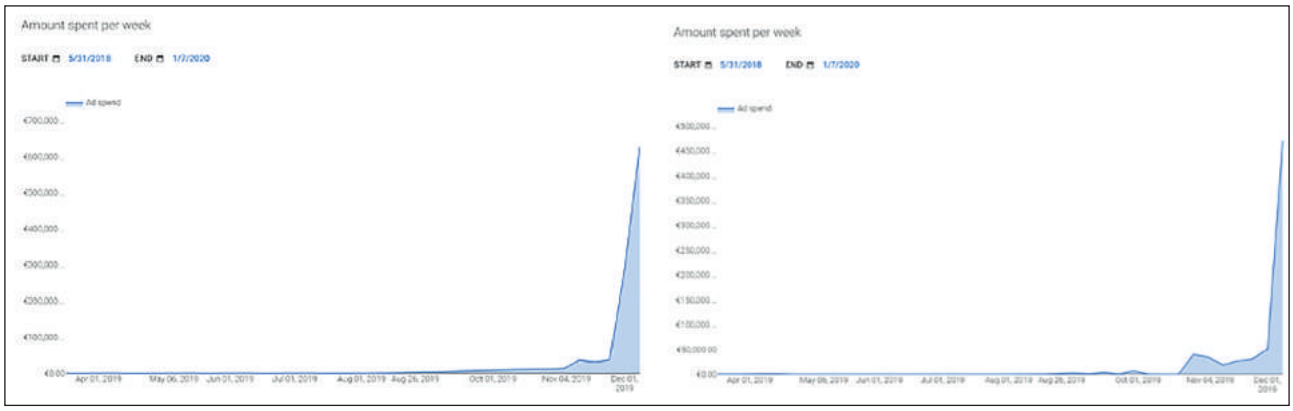
Man kann zudem nach Schlüsselwörtern suchen. Gibt man in der Übersicht der britischen Wahlwerbungsausgaben zum Beispiel NHS (für National Health Service, die staatliche britische Gesundheitsversorgung) ein, so sieht man, dass die Labour Party und die Konservativen im November und Dezember 2019 Google-Suchanzeigen kauften, um die Pläne der jeweils anderen Partei für den NHS zu kritisieren.



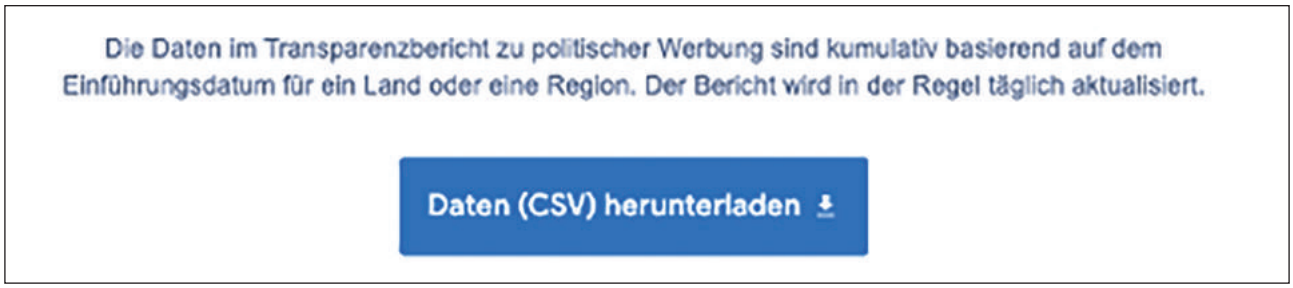
Indem man auf den Namen des Werbetreibenden klickt, kann man auch sehen, welche Gesamtsumme dieser Akteur bei Google seit Beginn des Berichts für politische Werbung ausgegeben hat. Für die beiden führenden britischen Parteien sah das im Januar 2020 zum Beispiel so aus:



Außerdem kann man sich auf einem Zeitstrahl anzeigen lassen, wann in welcher Größenordnung welche Summen gebucht wurden – hier der Verlauf für die britischen Konservativen (erste Grafik) und die Labour Party (zweite Grafik):



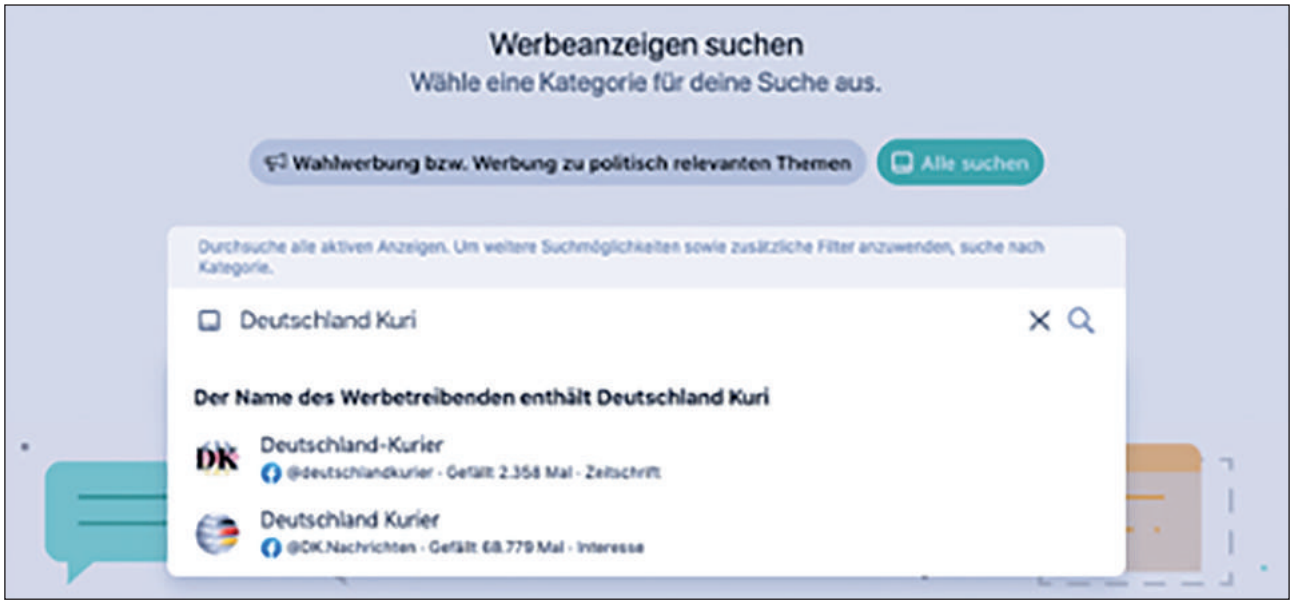
Wollen Sie genauer analysieren und tiefer in die Datenbank für politische Werbebuchungen einsteigen, scrollen Sie ans Ende der Seite. Dort gibt es die Option, die Rohdaten in einem Tabellenformat herunterzuladen:



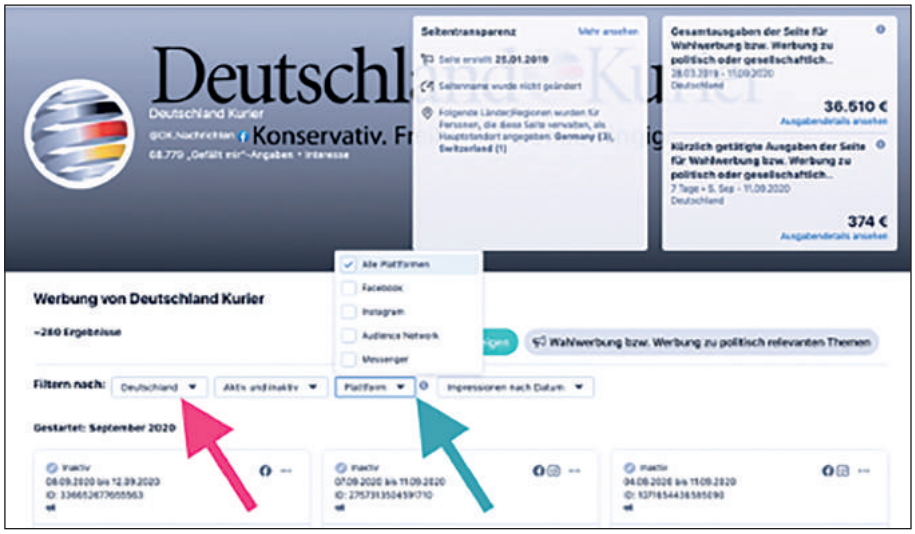
Diese Daten lassen sich in ein Tabellenprogramm wie Excel, Numbers, OpenOffice Calc oder Google Tabellen importieren, so dass man anschließend dort weiter filtern oder analysieren kann.

FACEBOOK

Die Übersicht über politische Werbung von Facebook ist in zwei Teile gegliedert: „Alle suchen" und „Wahlwerbung bzw. Werbung zu politisch relevanten Themen". Wenn Sie auf „Alle suchen" klicken, können Sie nach bestimmten Werbetreibenden nur anhand des Namens suchen, nicht über Stichwörter. Wollen Sie beispielsweise Anzeigen vom Deutschland Kurier sehen, einer Publikation, die häufig Inhalte zur Unterstützung der AfD veröffentlicht, können Sie hier diesen Namen eingeben und Facebook wird Ihnen Seiten vorschlagen, in denen dieser Name vorkommt:



Die Ergebnisse zeigen uns, dass seit dem Start des Berichts auf Facebook über den Deutschland Kurier Werbeanzeigen im Wert von 36.510 Euro bei Facebook gebucht wurden:

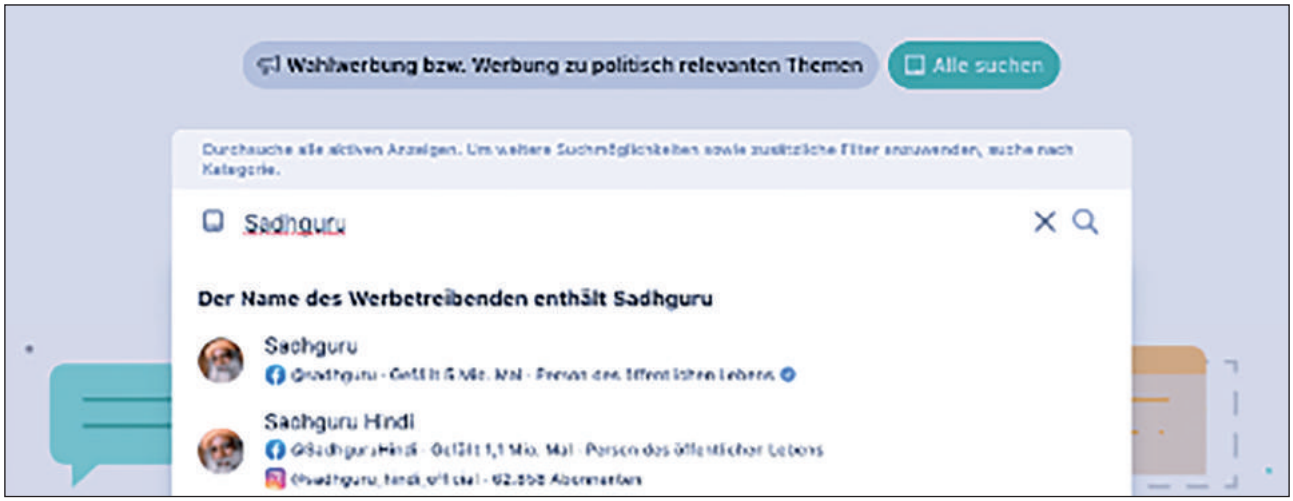


Achten Sie darauf, dass Sie bei den Filtern für die Ergebnisse das richtige Land (oder: „alle") für Ihre Recherche eingestellt haben und dass Sie gegebenfalls auswählen, ob Sie alle zu Facebook gehörenden Services in die Analyse einbezogen wissen wollen oder zum Beispiel nur Instagram oder den Messenger. (Das Audience Network ist ein Werbenetzwerk von Facebook, welches Werbung in mobilen Apps und Websites außerhalb von Facebooks eigenen Diensten platziert.) In den meisten Fällen wird die beste Wahl sein, alle Plattformen auszuwählen, um eine möglichst breite Übersicht zu bekommen. Haben Sie eine konkrete Werbung vor Augen, können Sie auf „Anzeigendetails ansehen" klicken und bekommen dann weiterführende Informationen zu dieser speziellen Buchung:

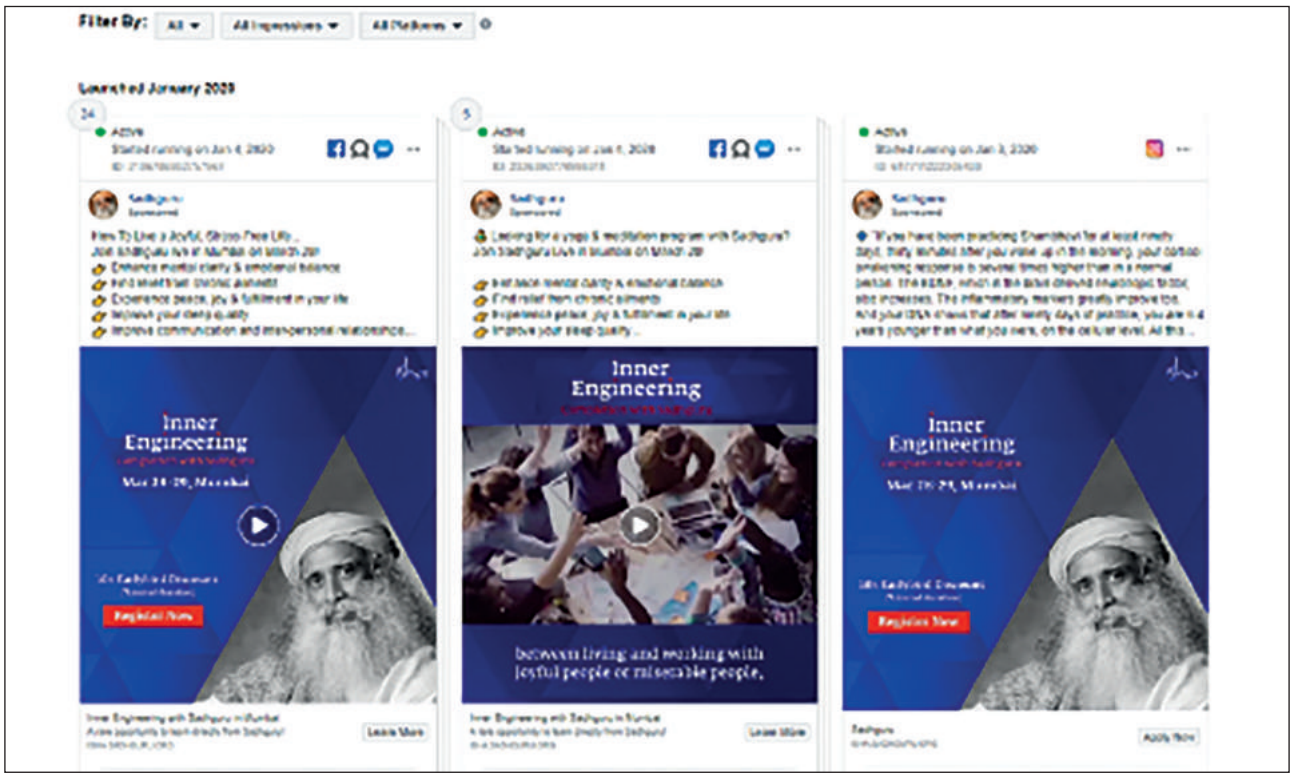
In diesem Fall gab der Deutschland Kurier weniger als 100 Euro für eine Anzeige aus, die mit einem Zitat von Albert Einstein für Misstrauen gegenüber der „Obrigkeit" wirbt, die Anzeige wurde 10.000 bis 15.000 Mal ausgespielt und erreichte überwiegend Männer über 45 Jahre.

Die zweite Option für die Suche in Facebooks Werbebibliothek ist die Auswahl „Wahlwerbung bzw. Werbung zu politisch relevanten Themen", die eine Datenbank mit „Anzeigen von Kandidaten für öffentliche Ämter sowie Anzeigen zu Wahlen, Bürgerinitiativen oder gesellschaftlichen Themen" darstellt. Der große Vorteil dieser Option ist, dass Sie nach jedem beliebigen Schlüsselwort suchen können. Schauen wir uns auch dazu ein Beispiel an.

Sadhguru ist der Name eines bekannten indischen Spirituellen, der von sich selbst sagt, er sei mit keiner politischen Partei verbunden. Er erklärte, er sehe es als seine Pflicht an, jede aktuelle Regierung zu unterstützen, „ihr Bestes zu tun". Wenn wir seinen Namen in den Abschnitt „Alle Anzeigen" eingeben, schlägt Facebook uns die persönliche Facebook-Seite des Sadhguru vor.



Hier sehen wir eine Auswahl unpolitischer Anzeigen, die Sadhguru veröffentlicht hat und in denen er für seine Yoga- und Meditationskurse wirbt.



Führen wir die gleiche Suche durch, diesmal allerdings in der Datenbank „Wahlwerbung bzw. Werbung zu politisch relevanten Themen“ und ohne die von Facebook vorgeschlagene Seite zu akzeptieren:

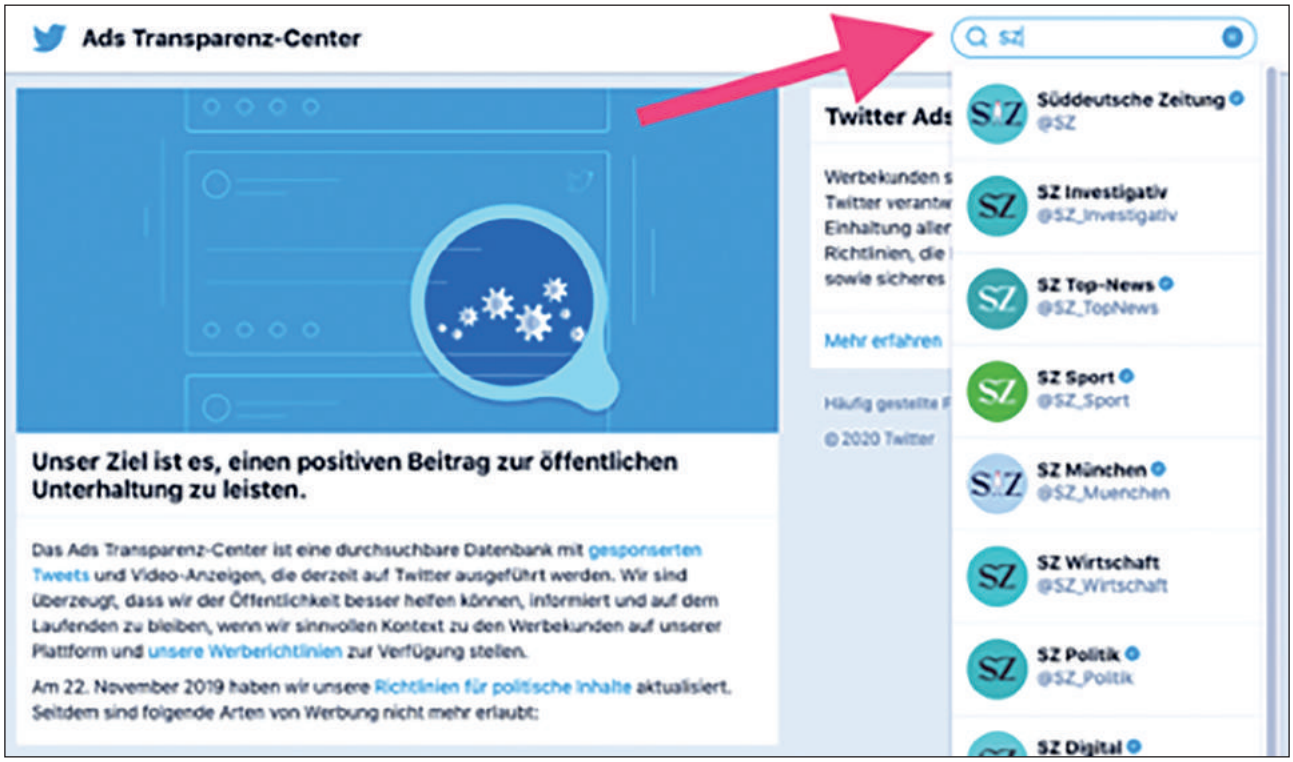
Die Ergebnisse ändern sich drastisch. Wir sehen jetzt eine Zusammenstellung von Anzeigen, die den Namen Sadhguru enthalten und auch von anderen Konten veröffentlicht wurden.

Eine Anzeige stammt von der in Indien regierenden nationalistischen Partei BJP. Sie zeigt ein Video, in dem Sadhguru seine Unterstützung für das umstrittene Staatsbürgerschaftsgesetz der BJP erklärt. Der Gesetzentwurf macht es nichtregistrierten Einwanderern aus einigen Nachbarländern Indiens leichter, die indische Staatsbürgerschaft zu erlangen, gewährt Muslimen allerdings nicht die gleichen Möglichkeiten. Die Anzeige gibt einen Hinweis auf eine mögliche Beziehung zwischen Sadhguru und der BJP, ein Thema, das in Indien weithin diskutiert wird.

Dieses Beispiel zeigt, wie man die Anzeigenbibliothek von Facebook nutzen kann, um den eigenen Recherchen wichtige Informationen hinzuzufügen. Interessant ist übrigens auch der Facebook-Werbebericht, der fortlaufend aktualisiert wird und Gesamtausgaben, Ausgaben bestimmter Werbekunden und Daten zu Ausgaben in bestimmten Regionen öffentlich macht.

TWITTER

Ende 2019 entschied Twitter, keine politische Werbung auf seiner Plattform mehr zu verbreiten. Das Ads Transparenz-Center zu nutzen ist allerdings weiter möglich, um sich über unpolitische Anzeigen der letzten sieben Tage zu informieren. Es gibt hier keine Suchfunktion, was die Suche umständlich macht. Um eine solche zu starten, geben Sie im Fenster oben rechts einen Benutzer- oder Profilnamen ein.



Wenn es von diesem Benutzer in den letzten sieben Tagen Anzeigen gab, werden sie nun angezeigt.

Man kann so sehen, wofür zum Beispiel Redaktionen Aufmerksamkeit generieren wollen. Die von Twitter bereitgestellten Informationen geben leider nichts darüber preis, wann die bezahlte Anzeige tatsächlich ausgespielt wurde.

Um die eigenen Recherchen zu beschleunigen, können Sie sich eines kleinen Tricks bedienen. Nachdem Sie eine Suche durchgeführt haben, werfen Sie einen Blick auf die URL in Ihrem Browser:

Der Link hat immer die gleiche Struktur, mit dem Namen des Benutzerkontos am Ende. Und diesen Namen kann man einfach löschen und durch einen anderen ersetzen:

Wenn Sie nun diese Seite neu laden, dann sehen Sie die Werbe-Informationen von Twitter für Bellingcat. Hat das Konto in den vergangenen sieben Tagen keine bezahlten Anzeigen gebucht, bekommen Sie die Nachricht „Dieser Account hat in den letzten sieben Tagen keine Anzeigen gesponsert" angezeigt. Da man leider nur die vergangenen sieben Tage auswerten kann, ist das Beste, was Sie tun können, eine solche Suche regelmäßig und immer wieder durchzuführen, um neue Werbe-Anzeigen früh zu finden.

SNAPCHAT

Die „Snap political ads library" ist die Bibliothek politischer Werbe-Anzeigen von Snapchat. Diese sind definiert als „Anzeigen zu Themen oder Organisationen, die auf lokaler, nationaler oder globaler Ebene Gegenstand von Debatten sind oder die von öffentlicher Bedeutung sind". Dazu zählen zum Beispiel Themen wie Einwanderung, Bildung oder Schusswaffen. Wer die Bibliothek aufruft, sieht zunächst eine Auswahl verschiedener Jahre:

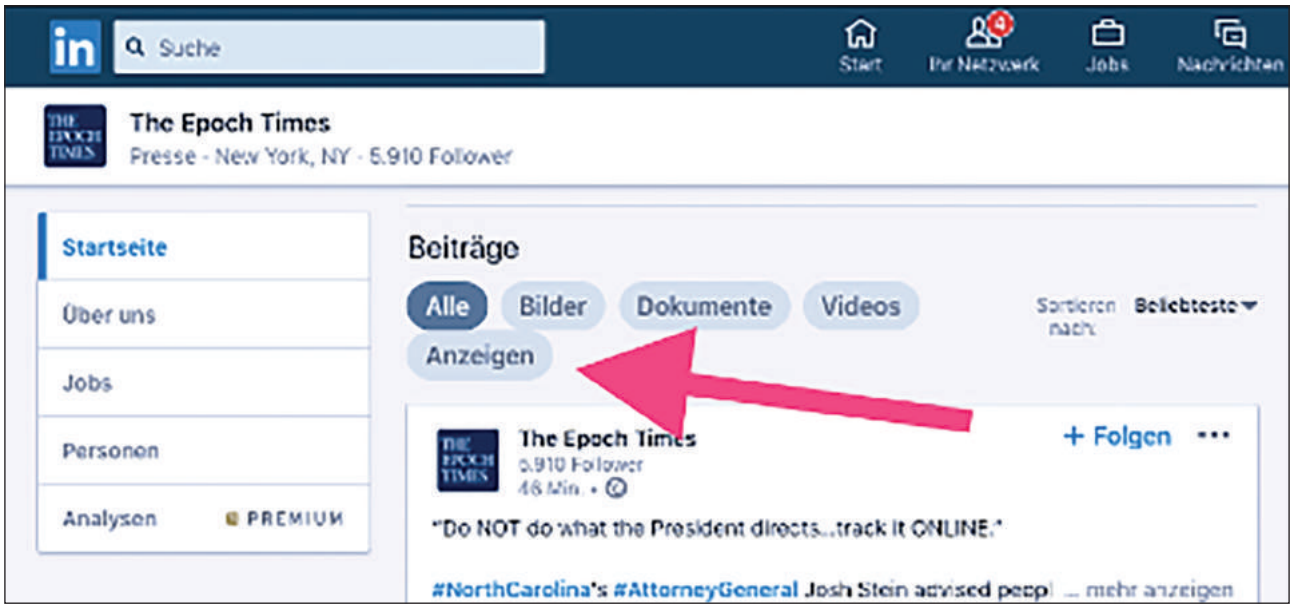
Nach dem Klick auf ein Jahr lässt sich eine Tabelle herunterladen, die alle verfügbaren Informationen über Anzeigen aus diesem Jahr auflistet. Der Inhalt der Tabelle sieht auf den ersten Blick nicht sehr aufregend aus, aber tatsächlich ist er es! Jede Zeile steht für eine Anzeige und zeigt, wer die Anzeige geschaltet hat, wie viel Geld dafür ausgegeben wurde und sogar, welche Merkmale für die gewünschte Zielgruppe ausgewählt wurden.

In diesem Beispiel wollte der Werbetreibende dem Eintrag in der Tabelle zufolge folgende Zielgruppen ansprechen: „Abenteurer, Kunst- und Kulturliebhaber, Strandbesucher und Surfer, Schönheitskönigin-nen, Bücherwürmer und Leseratten, Sammler, Feinschmecker, Hipster und Trendsetter, Zuschauer politischer Nachrichten, Outdoor- und Naturliebhaber, Haustier- und Tierliebhaber, Philanthropen, Weltreisende, Lebensstil von Frauen [sic]". Andere Plattformen machen diese Art von detaillierten Informationen über die Zielrichtung in ihren Anzeigenbibliotheken nicht öffentlich.

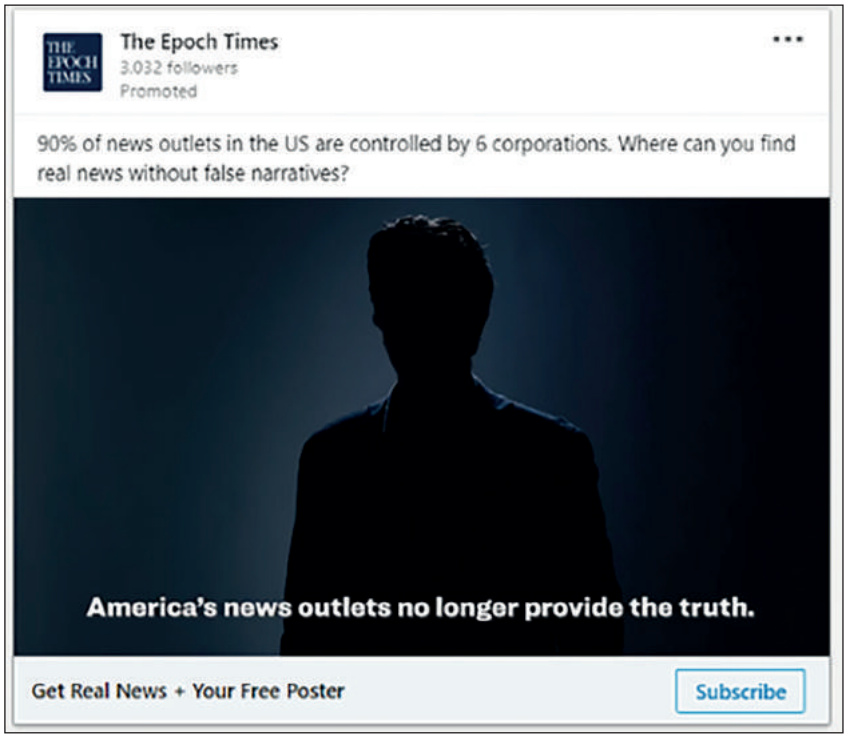
In der Tabelle findet sich auch ein Link, über den Sie die eigentliche Anzeige sehen können. In diesem Beispiel war das eine Botschaft, die Menschen ermutigen soll, kostenlose Regenbogenfahnen zur Unterstützung einer bevorstehenden Abstimmung in der Schweiz zu bestellen, die sich auf den Schutz vor Diskriminierung von LGBT-Personen bezieht.

LINKEDIN

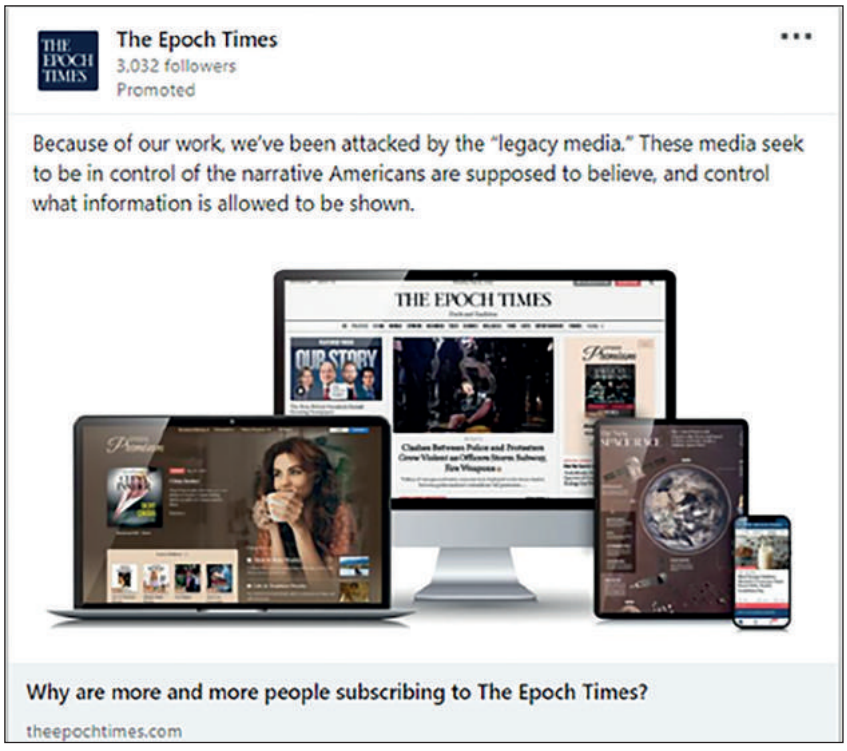
LinkedIn erlaubt auf seiner Seite keine politische Werbung und hat keine Werbebibliothek. Glücklicherweise gibt es eine andere Möglichkeit, sich einen Einblick in Werbung eines bestimmten Unternehmens auf der Plattform zu verschaffen. Öffnen Sie die Seite und scrollen Sie ein wenig nach unten, bis der erste Beitrag angezeigt wird. Dort haben Sie nun die Auswahl, sich nur Anzeigen des Unternehmens anzuschauen:



Klicken Sie auf diese Registerkarte und LinkedIn zeigt Ihnen eine Liste aller Anzeigen, die von diesem Unternehmen in den letzten sechs Monaten veröffentlicht wurden. Mit Hilfe dieser Funktion konnte man sehen, dass die Epoch Times noch Anzeigen auf LinkedIn veröffentlicht, nachdem ihr dies auf Facebook verboten wurde. Diese beiden Anzeigen des Unternehmens behaupten, dass „Amerikas Nachrichtenagenturen nicht mehr die Wahrheit liefern" und sie stellen die Epoch Times als „unabhängige" und „unparteiische" Redaktion dar.

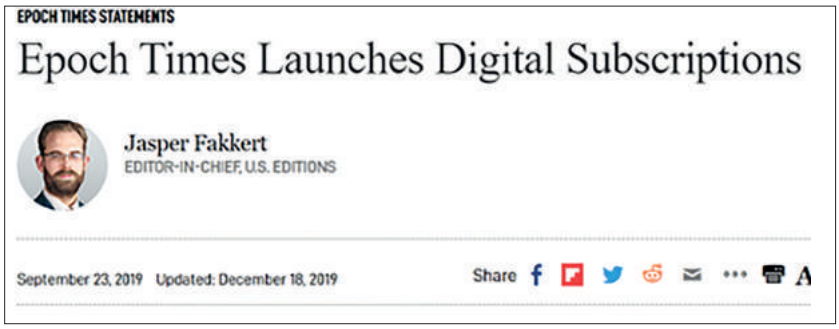


Im Beitrag steht: „90 % der Redaktionen in den USA werden von sechs Firmen kontrolliert. Wo finden Sie echte Nachrichten ohne falsche Erzählungen?" – In dem Bild darunter steht: „Amerikas Redaktionen liefern nicht länger die Wahrheit."



Der Beitrag behauptet: „Wegen unserer Arbeit werden wir von den ‚Altmedien' attackiert. Diese Medien wollen die Kontrolle haben über die Narrative, die die Amerikaner glauben sollen, und sie wollen kontrollieren, welche Informationen gezeigt werden dürfen."

Die genauen Veröffentlichungsdaten solcher Beiträge sind nicht sichtbar, aber Sie können auf die Anzeige klicken (dies funktioniert auch, wenn sie bei LinkedIn nicht aktiv ist) und den Beitrag öffnen – manchmal gibt die Zielseite ein konkreteres Datum an. Die erste der hier gezeigten gekauften Werbe-Einblendungen der Epoch Times führte zu einem Text mit dem Datum „23. September 2019“ und „Aktualisiert: 18. Dezember 2019“. Das half abzuschätzen, wann die Werbung online gewesen sein könnte.


EPOCH TIMES STATEMENTS

Epoch Times Launches Digital Subscriptions

Jasper Fakkert
EDITOR-IN-CHIEF, U.S. EDITIONS

September 23, 2019 Updated: December 18, 2019

Share

Der Chefredakteur der Epoch Times veröffentlicht einen Text, mit dem er digitale Abos ankündigt.

Wenn Sie all diese versteckten Funktionen erst einmal kennen, sind Anzeigenbibliotheken eine einfache und leistungsstarke Ergänzung zu Ihrem digitalen Recherchewerkzeugkasten – und sie sind ein wichtiges Element, das Sie überprüfen müssen, wenn Sie eine Person oder Organisation mit einer Präsenz in sozialen Medien untersuchen.

10. AKTEURE ÜBER VERSCHIEDENE PLATTFORMEN HINWEG VERFOLGEN

von: Ben Collins
deutsche Bearbeitung: Marcus Engert

Ben Collins *ist bei NBC News Reporter für Desinformation, Extremismus und das Netz. In den vergangenen fünf Jahren hat er über den Zuwachs von Verschwörungstheorien, organisierten Hass, ausländische Manipulationskampagnen und das Versagen von Plattformen berichtet. Zuvor arbeitete er bei The Daily Beast, wo sein Team die Profile, Gruppen und Veranstaltungen entdeckte, die während der US-Wahl 2016 von der russischen Trollfarm „Internet Research Agency" erstellt wurden.*

Am 3. August 2019 lief Patrick Crusius in einen Supermarkt in El Paso und tötete 22 Menschen in einer von Nationalismus und weißem Überlegenheitsdenken motivierten Schießerei. Doch bevor er den Laden betrat, veröffentlichte er ein Manifest im Diskussionsforum /pol/ auf 8chan.net, einem anonymen „Message Board", vergleichbar mit einem anonymen digitalen schwarzen Brett, das sich in den letzten Jahren zu einem Treffpunkt für weiße Nationalisten entwickelt hat. Die /pol/-Boards auf 4chan und 8chan sind beinahe vollständig unmoderiert, und bis zum Sommer 2019 war 8chan zu einem Treffpunkt gewalttätiger nationalistischer, rassistischer und rechtsextremer Inhalte und Diskussionen geworden.

Unter anderem auch aus diesem Grund alarmierten 8chan-Nutzer bereits das eine oder andere Mal Behörden und Journalisten, wenn ein neues und gewalttätiges Manifest auftauchte – durch das Hinzufügen von Kommentaren unter dem Manifest selbst oder indem sie online Tipps an Medien oder Strafverfolgungsbehörden gaben. Als der El Paso-Schütze zum ersten Mal sein Manifest veröffentlichte – was er zunächst mit einem falschen Anhang tat –, antwortete ein Benutzer darunter „Hallo FBI". Das eigentliche Manifest wurde dann nur wenig später direkt unter dem Kommentar, der das FBI markierte, veröffentlicht. Diese Art der Selbstoffenbarung kann für Journalisten nach solchen Tragödien eine kritische Information sein. Manchmal wechselten Nutzer in solchen Fällen in offenere Teile des Internets wie Reddit und Twitter, um auf verdächtige Beiträge, die vor den Schießereien gemacht wurden, hinzuweisen. Dies ist wichtig, denn es ist leicht, einen relevanten Beitrag oder Kommentar auf 4chan und 8chan zu übersehen.

Anonyme Plattformen wie 4chan und 8chan spielen eine wichtige Rolle bei der Verbreitung von Falsch- und Desinformation und für Trolle, weil dort oft Menschen zusammenarbeiten, um Kampagnen zu initiieren und zu koordinieren. Reddit – ein weiterer beliebter Ort, an dem die Benutzer weitgehend anonym bleiben können – ist Heimat einer großen Vielzahl von Online-Communitys. Bei einigen handelt es sich um stark moderierte „Subreddits", also Unterforen, in denen die Nutzer Geschichten über ihre Hobbys austauschen oder Nachrichten und Ereignisse diskutieren; andere sind im Grunde genommen frei für alle, und in denen bricht der Hass nicht selten ungebremst hervor. Für Journalistinnen und Journalisten ist es wichtig zu wissen, wie sie all diese Gemeinschaften beobachten und wie sie darüber berichten können. Sie müssen sich also mit den Feinheiten der Funktionsweisen vertraut machen.

Nach diesen Vorbemerkungen kommen hier fünf Regeln, die man einhalten sollte, wenn Ereignisse es erforderlich machen, auf 4chan oder 8chan (oder seiner neueren Variante: 8kun) nach Informationen zu suchen:

1. Glauben Sie nichts und vertrauen Sie niemandem auf 4chan/8chan.
2. Glauben Sie nichts und vertrauen Sie niemandem auf 4chan/8chan.
3. Glauben Sie nichts und vertrauen Sie niemandem auf 4chan/8chan.
4. Es kann sein, dass Sie nützliche Informationen (oder sogar Beweise) im Zusammenhang mit einem Verbrechen, einer Troll-Kampagne oder der Verbreitung von Desinformationen auf 4chan/8chan finden.
5. Glauben Sie nichts und vertrauen Sie niemandem auf 4chan/8chan.

Ich kann nicht genug betonen, wie wichtig es für Reporter ist, sich an die Regeln 1, 2, 3 und 5 zu halten, selbst wenn dies bedeutet, nicht an das heranzukommen, was hinter Nummer 4 schlummern könnte. Diese Websites sind buchstäblich dazu da, zu trollen, Anspielungen und Unwahrheiten über vermeintliche Feinde zu verbreiten, Lügen über Menschen zu streuen und gelegentlich ein paar quasi-lustige Lügen zu posten, getarnt als wahre Geschichten darüber, wie es ist, ein Teenager zu sein.

Die Skepsis wird durch die Tatsache untermauert, dass diese Adressen von weißen Nationalisten, Incels und anderen verstörten jungen männlichen Schützen als Abladeplatz für deren Manifeste benutzt wurden. Sagen wir es nochmals laut: Wenn es auf 4chan oder 8chan steht (die wir hier ab jetzt der Einfachheit halber stets als 8chan bezeichnen, trotz der nur nominellen Namensänderung in 8kun), ist es sehr wahrscheinlich, dass es sich um eine Lüge handelt. Eine Lüge, die Chaos säen und Reportern Ärger machen soll. Fragen Sie unter einem Beitrag nicht nach weiteren Details. Veröffentlichen Sie am besten gar nichts. Sie werden nur von gestörten Leuten ins Visier genommen werden, die einfach zu viel Zeit haben.

EIN MANIFEST BESTÄTIGEN

Genau deswegen ist es hilfreich, wenn Mitglieder dieser Gruppen sich bemühen, Manifeste oder andere berichtenswerte Inhalte an eine Öffentlichkeit außerhalb dieser Gruppen zu tragen. Durch den Kommentar „Hallo FBI" auf 8chan habe ich von der Existenz des Manifests von El Paso erfahren. Kurz nach den Berichten über die Schießerei durchsuchte ich Twitter mit den Schlüsselwörtern „El Paso 4chan" und „El Paso 8chan". Die Suche nach „[Stadtname] + [8chan oder 4chan oder incels.co]" oder anderen extremistischen Seiten ist durchaus eine nützliche Vorlage für vergleichbare Vorfälle.

Meine Twitter-Suche ergab, dass einige wenige Benutzer Screenshots von Beiträgen des mutmaßlichen Schützen auf 8chan verbreitet hatten, obwohl die meisten diesen Beitrag fälschlicherweise jemandem auf 4chan zuordneten. Also musste ich selbst nach dem Post suchen.

Der schnellste Weg, um einen Beitrag auf 8chan zu finden? Google. Im Chaos nach dem Amoklauf suchte ich nach „site:8ch.net" und fügte als Suchphrase einen Teil eines Satzes aus dem auf Twitter kursierenden angeblichen 8chan-Post des Schützen hinzu. (Hintergrund: 4chan löscht Beiträge nach einer bestimmten Zeitspanne automatisch von seinen Servern, aber es gibt automatische 4chan-Achivierungsseiten. Die wohl vollständigste heißt 4plebs.org. Dort archivierte 4chan-Beiträge können durch einfaches Ersetzen von „4chan" durch „4plebs" in der Google-Suche gefunden werden. Statt nach „boards.4chan.org/pol/13561062.html" sollte dann nach „4plebs.org/pol/13561062.html" gesucht werden). Je nach Vorfall und Szenario kann es hilfreich sein, nach „site:4chan.net + ‚manifesto' OR ‚fbi'" zu suchen und Googles Suchoptionen zu nutzen, um den Zeitrahmen für Suchergebnisse auf die vergangenen 24 Stunden zu beschränken. Andere Nutzer könnten bereits versucht haben, mit ihren Antworten Aufmerksamkeit auf den Schützen zu lenken. Mit meiner ursprünglichen Suchstrategie hatte ich den relevanten 8chan-Beitrag nicht gefunden, weswegen ich zunächst dachte, dass es sich um einen schnell gemachten Schwindel handelte. Aber etwas stimmte nicht. Der in den Screenshots auf Twitter gezeigte Beitrag hatte eine Benutzer-ID und eine Beitragsnummer. Das legte wiederum den Schluss nahe, dass es sich nicht um eine plumpe Fälschung handelte, denn die würde rasend schnell auffliegen. Auf 8chan erhält jeder Beitrag eine eindeutige Benutzer-ID, die algorithmisch generiert und neben dem Datum des Beitrags angezeigt wird. Damit hat jeder Benutzer eine feststehende Nummer, so dass Nutzerinnen und Nutzer in Diskussionen klarmachen können, wem oder worauf sie antworten.

Durch diese Benutzer-ID wusste ich, dass die Person, die als Manifest zunächst eine falsche PDF-Datei mit dem Namen des Schützen veröffentlicht hatte, auch der Benutzer war, der zwei Minuten später das eigentliche Manifest veröffentlichte. Beide Beiträge teilten sich die zufällig erstellte gleiche Benutzer-ID: 58820b.

Neben einer Benutzer-ID befindet sich eine Beitragsnummer, die sozusagen dauerhaft einen eindeutigen Link für jeden neuen Beitrag herstellt. Der Screenshot des auf Twitter veröffentlichten El Paso-Manifests enthielt die Post-ID „No.13561062". Dadurch hätte der Link 8ch.net/pol/res/13561062.html generiert werden müssen. Diese Link-Systematik gilt sowohl für 4chan als auch für 8chan. In diesem Fall gab es den Link nicht. Ich dachte, er sei womöglich gelöscht worden. (Später erfuhr ich, dass der Besitzer von 8chan, Jim Watkins, ihn entfernt hatte, nachdem er auf den Inhalt hingewiesen worden war.) Nun, da der Beitrag weg war, lag meine letzte Hoffnung darauf, dass jemand vorher dessen Wichtigkeit erkannt und ihn archiviert hatte. Glücklicherweise hatte ein 8chan-Benutzer schnell geschaltet und den Beitrag auf der Archivierungsseite archive.is gespeichert. Als ich den Link dort in das Suchfeld einfügte, war klar: Der Beitrag mit dem Manifest war echt. Und jetzt konnte ich ihn sehen.

Doch damit gab es ein neues Problem: die Frage, wann er zum ersten Mal auf 8chan veröffentlicht worden war. Ich brauchte einen genauen Zeitstempel, um bestätigen zu können, dass das Manifest abgeschickt wurde, bevor der El Paso-Schütze seinen Amoklauf begann.

Sowohl 4chan als auch 8chan speichern ihre Zeitstempel nur lokal, so dass es eine komplizierte Aufgabe ist, die Veröffentlichungszeit aus einem anderswo archivierten Beitrag abzuleiten. Doch es gibt einen einigermaßen sicheren Weg, das zu umgehen. Wenn Sie mit der rechten Maustaste auf den Zeitstempel und auf „Quelltext anzeigen" klicken, wird der Quellcode der Seite angezeigt. Es geht um den Abschnitt, der mit ‚<time unixtime='[Zahl]' beginnt. Kopieren Sie diese Zahl und fügen Sie sie in einen Epoch/Unix-Zeitstempel-Konverter wie unixtimestamp.com ein. So erhalten Sie einen sekundengenauen Zeitstempel für den Beitrag in der Zeitzone UTC. Die Umrechnung von UTC-Zeit in El Paso-Zeit ergab, dass das Manifest um 10:15 Uhr Central Time – nur Minuten vor Beginn des Amoklaufs – veröffentlicht wurde. Das half mir, zu bestätigen, dass das Manifest auf 8chan tatsächlich ein Beweisstück in einem gravierenden Fall von rassistischem Inlandsterrorismus war.

AKTEURE ÜBER MEHRERE PLATTFORMEN VERFOLGEN

Im Jahr 2017 tötete Lane Davis – ein ehemaliger „Gamergate-Forscher" (soll heißen: professioneller Internetstalker) für den zwischenzeitlich in Ungnade gefallenen Popstar der Rechten, Milo Yiannopoulos, seinen Vater in seinem eigenen Haus. Davis war in einen Streit mit seinen Eltern geraten, auf dem Notruf war zu hören, wie er kurz vor dem Angriff rechtsextremen Internetjargon von sich gab. Er bezeichnete seine Eltern als „linke Pädophile", bevor sein Vater die Polizei rief, die ihm dabei helfen sollte, den eigenen Sohn, der noch bei seinen Eltern wohnte, rauszuwerfen. Davis hieß online „Seattle4Truth", in YouTube-Videos sprach er häufig von irgendwelchen geheimen Pädophilenringen, von denen er glaubte, sie seien die treibende Kraft hinter dem Liberalismus. Ein unter seinem Namen veröffentlichtes Video auf YouTube trug den Titel „Die tiefe Verbundenheit der progressiven Ideologie mit der Pädophilie".

Das Traumszenario eines Reporters bei der Untersuchung von Online-Extremismus ist ein Täter, der über viele Plattformen hinweg immer den gleichen Benutzernamen benutzt, und das war bei Davis der Fall. Er identifizierte sich selbst als Seattle4Truth auf YouTube und auf Reddit, wo seine Beiträge ein noch verschwörungsverwirrteres Gehirn enthüllten. Wie das rauskam? Indem wir den Nutzernamen Seattle4Truth einfach von Hand in das Schema der Links eingebaut haben, wie sie auf Reddit funktionieren: reddit.com/u/[username].

Einmal gefunden, lassen sich die Beiträge eines Profils nach den neuesten Beiträgen, den beliebtesten Beiträgen und den „umstrittensten" Beiträgen sortieren. Eine Möglichkeit, schnell einen Benutzernamen zu recherchieren, ist die Verwendung des Dienstes Namechk, mit dem man fast 100 Internetplattformen nach einem Benutzernamen durchsuchen kann. Wie weiter unten noch ausgeführt, bedeutet das nicht, dass die gleiche Person hinter diesen Konten steckt, aber es ist eine effiziente Methode, um zu sehen, wo der Benutzername überhaupt verwendet wird. So weiß man anschließend, wo man recherchieren muss. Natürlich kann man auch Benutzernamen googeln, für die man sich interessiert.

Wichtig ist auch, sich bewusst zu sein, in welchen Super-Nischen-Internet-Communitys das Ziel aktiv sein könnte. Der Schütze eines Amoklaufs an einer Schule in New Mexico 2017 wurde von den Nutzern einer Plattform namens KiwiFarms als @satanicdruggie identifiziert – wobei es hauptsächlich um den Kampf gegen Trans-Mobbing geht. Die Nutzer meinten, er sei auf Encyclopedia Dramatica aktiv gewesen, einer Seite, auf der jeder schreiben kann und die eigentlich für Satire gedacht ist, jedoch auch extremistische Inhalte aufweist. Der Amokläufer war dort nicht nur aktiv, er war Systemadministrator, was bedeutet, er hatte Verwaltungsaufgaben und war ein Power-Nutzer. (Wir haben uns von Nutzern der Seite, die über Skype eine reale Beziehung zu ihm aufgebaut hatten, bestätigen lassen, dass es sich um seine Konten handelte.) Eine Google-Suche seines Benutzernamens in Kombination mit dem Suchbefehl „site:encyclopediadramatica.rs + [Benutzername]" ergab, dass er sich auch „Satanic Druggie" nannte, aber außerdem Namen wie „Future School Shooter" und „Adam Lanza", den Namen des Amokläufers an der Sandy Hook Schule, benutzte. Die Geschichte seiner Beiträge im Netz enthüllte eine Besessenheit von Amokläufen an Schulen, die selbst die Polizei nach der Schießerei nicht entdeckt hatte.

Es ist nochmals wichtig zu betonen, dass das Vorhandensein eines Benutzernamens auf verschiedenen Plattformen kein Beleg dafür ist, dass die Konten von ein und derselben Person erstellt wurden. In einem berühmt gewordenen Beispiel behaupteten mit Ian Miles Cheong, Mike Cernovich, InfoWars und GatewayPundit gleich mehrere einschlägig bekannte rechtsextreme Desinformationsgrößen, ein Mann, der bei einem Videospielturnier in Jacksonville zwei Menschen getötet

und zehn weitere verletzt hatte, sei ein Trump-Gegner gewesen. Ihr Grund dafür, das zu tun? Der Schütze, David Katz, benutzte online den Namen „Ravens2012Champs“, und ein Benutzer, der sich auf Reddit deutlich gegen Trump aussprach, hatte einen ähnlichen Namen: „RavenChamps“. Die Berichterstattung darüber war ungenau: Das rechte Portal InfoWars behauptete, er „hasste Trump-Anhänger“. Später stellte sich heraus, dass RavenChamps eine ganz andere Person war, ein Fabrikarbeiter aus Minnesota namens Pavel. „Wissen Sie, ich lebe noch“, schrieb er Stunden nach den Schüssen auf Reddit. (Der tatsächliche Schütze hatte sich nach dem Massaker selbst getötet.)

Man braucht viel mehr als nur einen Benutzernamen, aber er kann ein wichtiger Ausgangspunkt sein, zum Beispiel für Fragen an die Strafverfolgungsbehörden oder um in öffentlichen Quellen zu wühlen oder Menschen anzurufen.

KAMPAGNEN ANNÄHERND IN ECHTZEIT BEOBACHTEN

Desinformations- und Medienmanipulationskampagnen verbreiten sich oft über Reddit und 4chan, und einige lassen sich quasi in Echtzeit verfolgen. Beispielsweise sind Nutzer von 4chan seit Jahren damit beschäftigt, Online-Umfragen zu manipulieren, um bevorzugte Kandidaten zu fördern. Im Jahr 2016 wurden auf 4chan wiederholt Links sowohl zu nationalen als auch zu hyperlokalen Nachrichtenwebsites veröffentlicht, auf denen Umfragen über den bevorzugten Kandidaten der Nutzer, Donald Trump, durchgeführt wurden. Wenn man die Sucheinstellungen von Google so ändert, dass sie nach Beiträgen in der „letzten Stunde“ filtern, und dann nach „site:4chan.org ‚polls‘“ sucht, erhält man in Echtzeit einen ziemlich guten Einblick, welche Umfragen die Nutzer von 4chan manipulieren wollen. Dies hat sich bis in den Wahlkampf der aktuellen Wahlen fortgesetzt. Nutzer von 4chan haben in Umfragen Tulsi Gabbard, die sie als „Mama“ bezeichneten, nach vorn gebracht. Mit Hilfe einer einfachen Google-Suche konnte jeder in Echtzeit sehen, wie es kam, dass die Umfrageergebnisse verschoben wurden, nachdem auf 4chan jemand die anderen Benutzer mit „GIVE HER YOUR POWER“ (Gib ihr deine Macht) dazu aufgefordert hatte.

Aktive Troll-Operationen auf Seiten wie der r/The_Donald-Community von Reddit zu finden, ist sogar noch einfacher, da Reddit eine nützliche „Aufstiegs“-Funktion bietet. Sie zeigt Beiträge an, die gerade stark nachgefragt oder kommentiert werden. Mit der Suche „reddit.com/r/[subreddit-name]/rising“ werden einem jene Ergebnisse angezeigt, die in einem Subreddit, also einer der Gruppen dort, gerade an Fahrt gewinnen.

Sie können sich auch die Beiträge ansehen, die in allen Bereichen von Reddit gerade große Aufmerksamkeit bekommen – über reddit.com/r/all/rising. Die meisten Reddit-Communitys finden sich darin wieder. Einige Subreddits hingegen wurden von Reddit unter Quarantäne gestellt: besonders toxische Gemeinschaften, die zutiefst beleidigende Inhalte produzieren und verbreiten und sich zu Troll-Kampagnen auf andere Gemeinschaften verabreden. Diese werden dort nicht gefunden, sie werden es auch nicht von Google – so ist die von Reddit eingeführte Quarantäne eine effektive Methode, um die Reichweite solcher Kampagnen außerhalb ihrer eigenen Blase zu begrenzen, aber sie erschwert es auch, nachzuverfolgen, wie sich deren Akteure im Moment organisieren.

Unterm Strich ist es eine gute Idee, den aufstrebenden Teil von Gruppen, die für Troll-Kampagnen bekannt sind, während großer politischer Ereignisse, nach Amokläufen oder Wahlen im Auge zu behalten. Die Realität ist leider, dass die Maßnahmen, die die Plattformen unternehmen, um böswilligen Akteuren das Handwerk zu legen, es auch Journalistinnen und Journalisten mitunter schwerer machen, ihre Arbeit zu tun. Technische Werkzeuge können helfen, aber so vieles davon ist Handarbeit. Vieles erfordert Ansätze zur Verifizierung, was Algorithmen nicht können. Letzten Endes kann ein Computer diese Art von Arbeit nicht ersetzen. Es liegt an uns.

11. NETZWERK-ANALYSEN UND ZUSCHREIBUNGEN

von: Ben Nimmo
deutsche Bearbeitung: Marcus Engert

Ben Nimmo *ist Forschungsdirektor bei Graphika und Senior Fellow am Digital Forensic Research Lab des Atlantic Council (Washington, D. C.). Er ist auf die Untersuchung großer plattformübergreifender Informationsoperationen und Methoden der Einflussnahme spezialisiert. Seine Freizeit verbringt der Brite unter Wasser, wo er telefonisch nicht erreichbar ist.*

Wenn man sich mit einer möglichen Informationsoperation befasst, lautet eine der Kernfragen für Wissenschaftler und Journalisten: Wie groß ist die Operation, und wie weit breitet sie sich aus? Das ist etwas anderes als die Auswirkungen einer Operation zu untersuchen. Es geht also darum, die Konten und Standorte zu finden, von denen die Operation betrieben wird.

Für Forscher und Wissenschaftler besteht das Ziel darin, von einer Operation so viel wie möglich ausfindig zu machen, bevor man sie meldet, denn sobald die Operation gemeldet wird, ist zu erwarten, dass die Betreiber abtauchen – und möglicherweise auch, dass sie Konten oder Hinterlassenschaften löschen.

DAS ERSTE GLIED IN DER KETTE

Den Ausgangspunkt zu finden ist am schwersten. Oft beginnt eine Untersuchung mit einem Hinweis eines betroffenen Benutzers oder (seltener) einer Social-Media-Plattform. Die Arbeit des Digital Forensic Research Lab zur Aufdeckung der mutmaßlichen russischen Geheimdienstoperation „Sekundäre Infektion" begann mit einem Hinweis von Facebook, das auf seiner Plattform 21 verdächtige Konten gefunden hatte. Unsere Arbeit erreichte ihren Höhepunkt sechs Monate später, als Graphika, Reuters und Reddit öffentlich machten, dass über diese Operation auf ihren Plattformen Einfluss auf die britischen Wahlen genommen werden sollte. Eine Recherche über Desinformation, die auf amerikanische Veteranen abzielte, begann damit, dass ein Mitarbeiter von Vietnam Veterans of America eine ähnliche Facebook-Seite entdeckte, die doppelt so viele Anhänger hatte wie die Organisation selbst.

Es gibt nicht die eine Regel, um das erste Glied in der Kette zu identifizieren. Die wirksamste Strategie ist die *Suche nach den Unstimmigkeiten*. Das könnte ein Twitter-Account sein, der anscheinend in Tennessee ansässig, aber auf eine russische Mobiltelefonnummer registriert ist; es könnte eine Facebook-Seite sein, deren Betreiber vorgibt, in Niger zu sitzen, aber die Seite von Senegal und Portugal aus verwaltet. Es könnte ein YouTube-Konto mit einer Million Aufrufen sein, das 2019 riesige Mengen an prochinesischen Inhalten veröffentlicht, aber fast alle Aufrufe kamen über Episoden britischer Sitcoms zustande, die schon 2016 hochgeladen wurden. Es könnte sich um eine anonyme Website handeln, die sich auf amerikanische Außenpolitik konzentriert, aber bei der Finanzabteilung des fernöstlichen Militärbezirks der Russischen Föderation registriert ist. Es könnte sich um ein angebliches Interview mit einem Agenten des britischen Geheimdienstes MI6 handeln, der in gestelztem, fast Shakespeare-artigem Englisch spricht. Es könnte sogar ein Twitter-Account sein, der Einladungen zu einer Pornoseite mit unvollständigen Zitaten aus Jane Austens „Verstand und Gefühl" verbreitet.

Der Trick bei solchen Gegebenheiten ist, sich die Zeit zu nehmen, um in Ruhe über sie nachzudenken. Rechercheure und Journalisten werden so oft unter Zeitdruck gesetzt, dass sie schnell etwas mit „das ist aber seltsam" abtun und weitermachen. Aber: Wenn etwas komisch ist, dann ist es das oft aus einem bestimmten Grund. Sich die Zeit nehmen, um zu sagen „Das ist seltsam. Warum ist das so?" kann der erste Schritt sein, um eine neue Operation aufzudecken.

RESSOURCEN, VERHALTEN, INHALT

Sobald eine erste Einheit – wie etwa ein Konto oder eine Website – identifiziert wurde, besteht die Herausforderung darin, herauszufinden, wohin sie zielt. Drei Fragen sind hier entscheidend, dargelegt in Camille François' Desinformations-ABC:

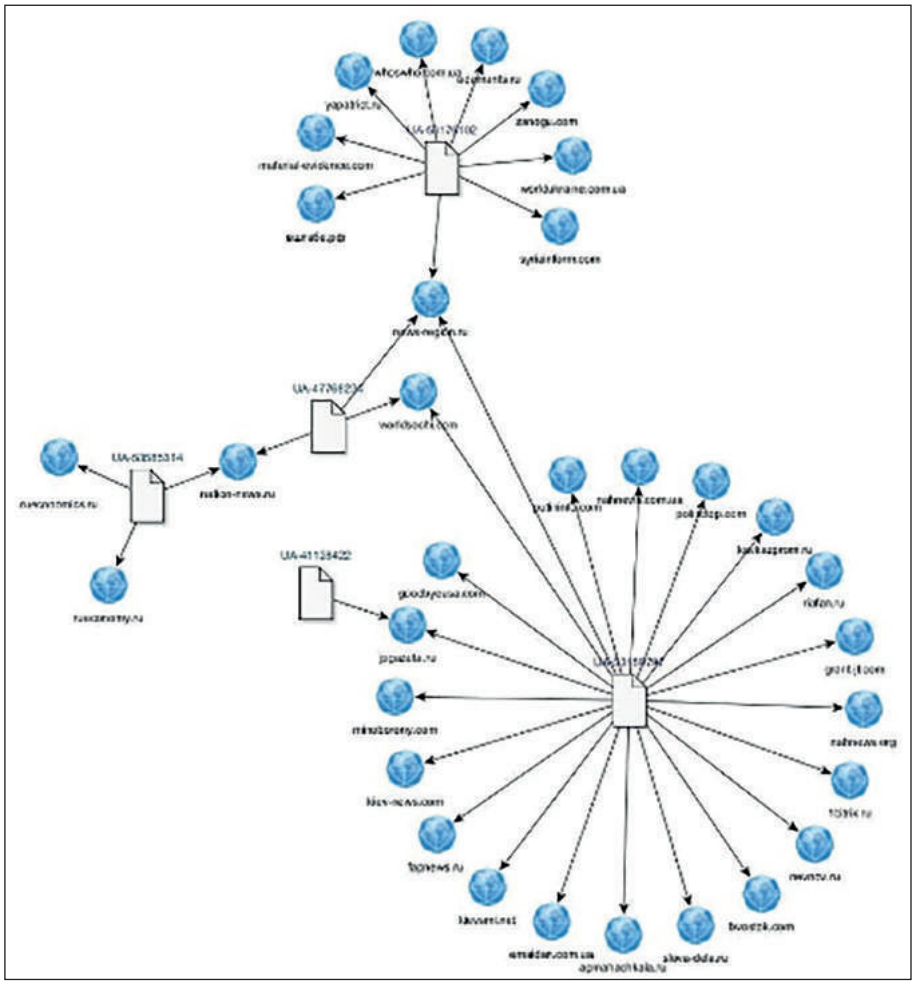
– Welche Informationen über diese erste Einheit sind verfügbar?
– Wie hat sich diese Einheit verhalten?
– Welchen Inhalt hat sie veröffentlicht?

Der erste Schritt besteht darin, so viele Informationen über diese Einheit zu sammeln wie möglich. Falls es sich um eine Website handelt: Wann und von wem wurde sie registriert? Verfügt sie über identifizierbare Merkmale wie einen Google-Analytics-Code oder eine AdSense-Nummer, eine E-Mail-Adresse, über die sie registriert wurde, oder eine Telefonnummer? Diese Fragen können auch anhand zurückliegender WhoIs-Einträge überprüft werden, die man mit Diensten wie lookup.icann.org, domaintools.com, domainbigdata.com oder dem leider fragwürdig benannten spyonweb.com suchen kann.

Screenshot des Anbieters domainbigdata mit einigen Registrierungsdaten für die Website NBeneGroup.com, die vorgab, eine „Jugendanalyse-Gruppe" zu sein und mit Registrierungsdaten, die auf die Finanzabteilung des fernöstlichen Militärbezirks der Russischen Föderation hinweisen.
Quelle: lookup.icann.org

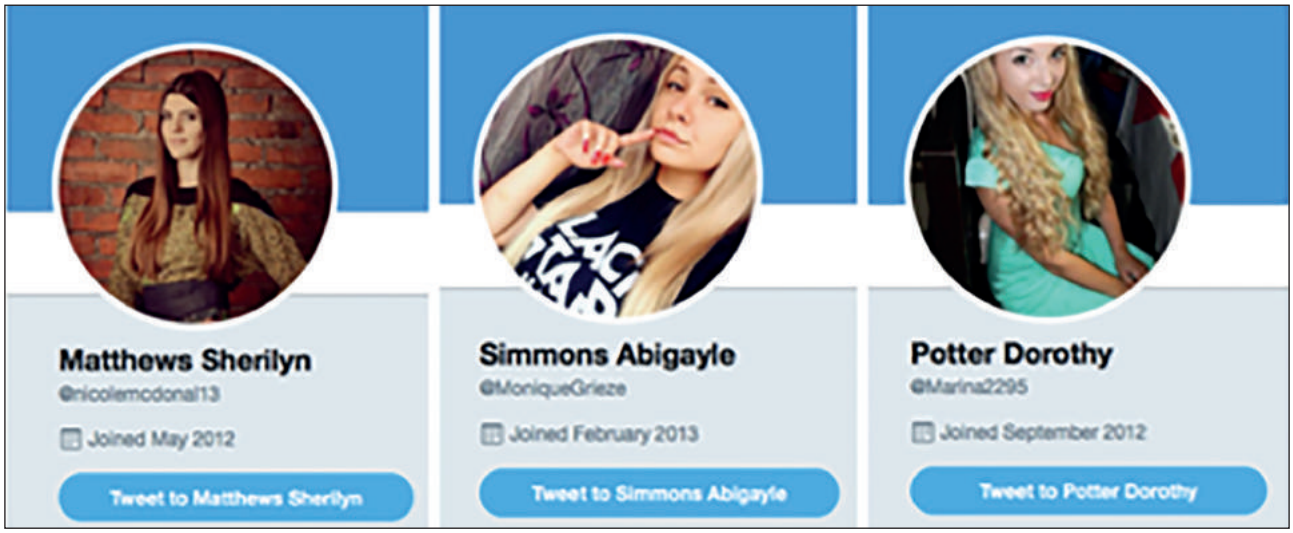
Informationen über eine Website können zur Suche nach weiteren Einheiten verwendet werden. Sowohl domaintools.com als auch spyonweb.com ermöglichen es den Nutzern, mit der IP-Adresse oder einem Google-Analytics-Code weiterzusuchen, was möglicherweise zu anderen damit verknüpften Websites führt (siehe Kapitel 8 dieses Handbuchs) – obwohl bei besseren Informationsoperationen die Registrierungsdaten mittlerweile in der Regel hinter kommerziellen Anonymisierungsdiensten verborgen sind, was die Arbeit erschwert.

Eine der ersten Analysen dieser Art kam 2015 von einem britischen Forscher namens Lawrence Alexander und identifizierte 19 Websites, die von der russischen „Internet Research Agency" betrieben werden, und zwar anhand ihrer Google-Analytics-ID-Nummern. Im August 2018 entlarvte die Sicherheitsfirma FireEye eine großangelegte iranische Operation, indem sie Registrierungsinformationen, inklusive E-Mails, nutzte, um Verbindungen zwischen angeblich nicht miteinander verbundenen Websites aufzuzeigen.



Die von Lawrence Alexander gefundenen Websites und wie sie über gemeinsame Google-Analytics-Codes miteinander verbunden sind.

Handelt es sich bei der ersten untersuchten Einheit um ein Social-Media-Konto, helfen die in den beiden vorangegangenen Kapiteln über Bots und nicht authentische Aktivitäten sowie die in der Untersuchung von Social-Media-Konten gegebenen Hinweise. Wir sollten also fragen: Wann wurde das Konto erstellt? Stimmt sein Nutzername mit dem Profilnamen überein? (Wenn das Profil „@moniquegrieze“ heißt, der Nutzer sich aber „Simmons Abigayle“ nennt, ist es möglich, dass das Konto gekapert wurde oder Teil einer massenhaften Erstellung von Konten war.)



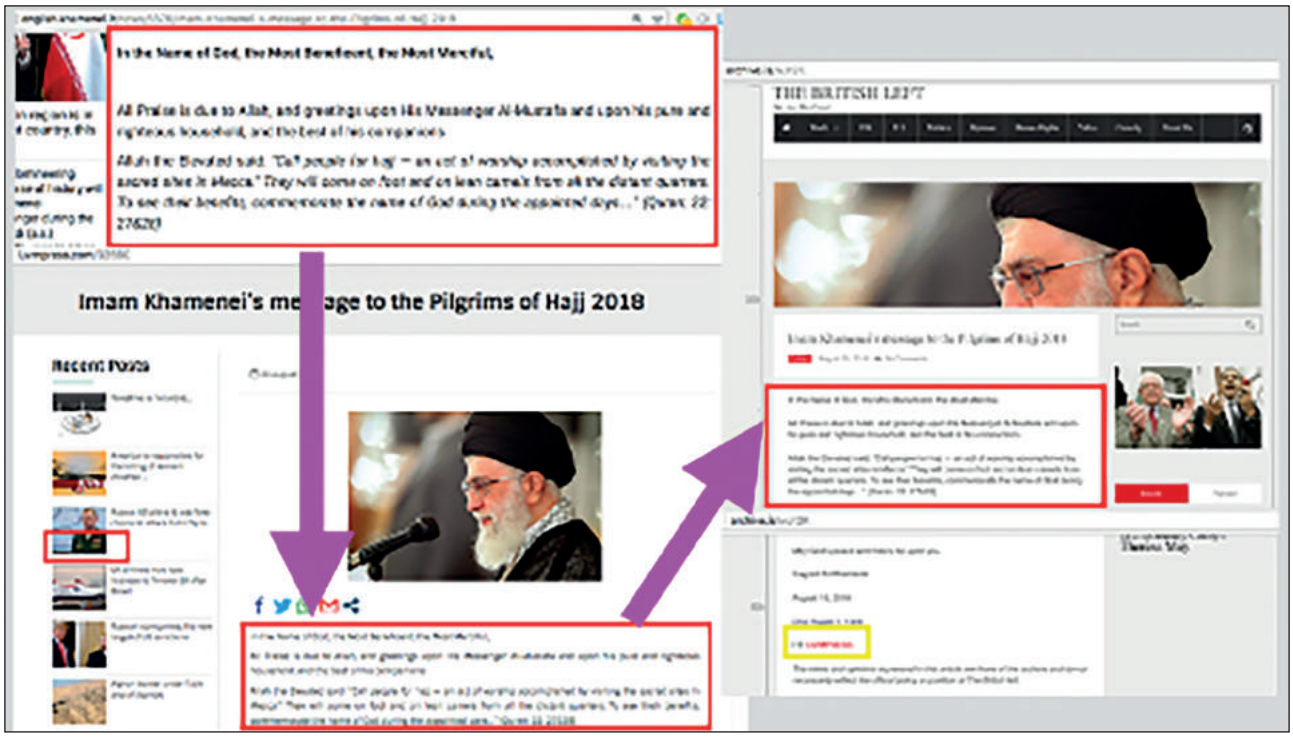
Drei Twitter-Profile, die in eine große Bot-Operation im August 2017 eingebunden waren. Vergleichen Sie die Konten- und Benutzernamen, die sich alle voneinander unterscheiden, was darauf hindeutet, dass die Konten übernommen und umbenannt worden sein könnten.

Finden sich nachprüfbare biographische Angaben oder Links zu anderen Ressourcen auf derselben oder anderen Plattformen? Falls es sich um eine Facebook-Seite oder -Gruppe handelt: Wer verwaltet sie und wo befinden sich die Administratoren? Wem folgen sie, und wer folgt ihnen? Die Facebook-Einstellungen „Seitentransparenz" und „Gruppenmitglieder" können oft wertvolle Anhaltspunkte liefern, ebenso wie Merkmale eines Twitter-Profils wie das Beitrittsdatum und die Gesamtzahl der Tweets und Gefällt-mir-Markierungen. (Auf Facebook und Instagram ist es nicht möglich, das Erstellungsdatum eines Kontos zu sehen, aber das Datum des zuerst hochgeladenen Profilbilds bietet einen guten Anhaltspunkt.)

Website und Facebook-Seite für die angebliche Faktenprüfungs-Website (Es ist falsch – gefälschte Nachrichten aus Mali), aus der hervorgeht, dass sie vorgab, von einer Studentengruppe in Mali betrieben zu werden, in Wirklichkeit aber von Portugal und Senegal aus verwaltet wurde. Bild: DFRLab.

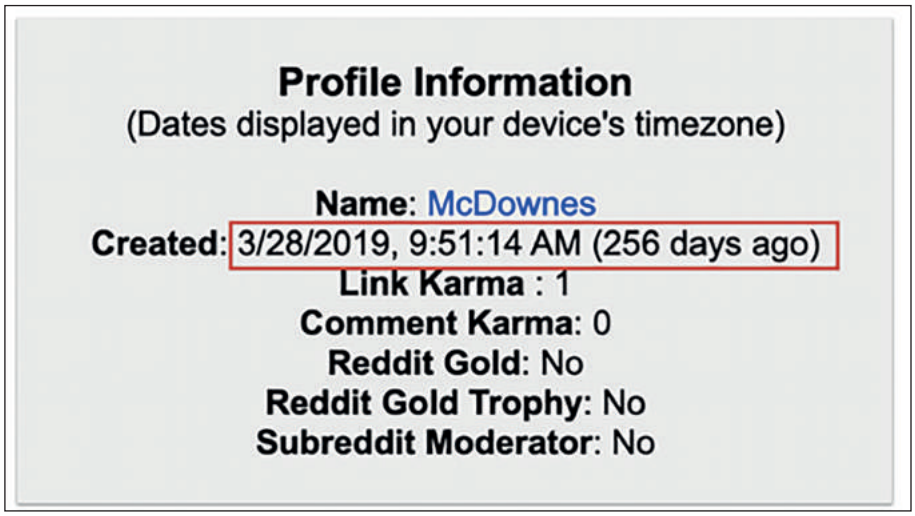
Nachdem die Einzelheiten erfasst worden sind, besteht der nächste Schritt darin, das Verhalten zu charakterisieren. Die Testfrage lautet hier: „Welche Verhaltensmerkmale sind am typischsten für diese Ressource und könnten nützlich sein, um andere Einheiten in derselben Operation zu identifizieren?"

Dies sind weitreichende Fragen, auf die es viele Antworten geben kann, von denen sich einige vielleicht erst in späteren Phasen einer Recherche zeigen. Dazu könnten zum Beispiel YouTube-Kanäle gehören, die zwar westliche Namen und Profilbilder haben, aber politische Videos in chinesischer Sprache verbreiten, unterbrochen von großen Mengen kurzer TikTok-Videos. Dazu könnten Netzwerke von Facebook- oder Twitter-Konten gehören, die immer wieder Links zu derselben Website oder derselben Sammlung von Websites teilen. Es könnten auch Konten enthalten sein, die in ihren Selbstbeschreibungen den gleichen Wortlaut oder ähnliche Variationen desselben Wortlauts verwenden. Dazu könnten vermeintliche „Journalisten" zählen, die keine nachprüfbaren biographischen Angaben haben oder Angaben machen, die als falsch identifiziert werden können. Oder Websites, die den größten Teil ihres Inhalts von anderen Websites plagiieren und nur gelegentlich parteiische, polemische oder irreführende Artikel einstreuen. Es können auch mehrere solcher Faktoren gleichzeitig auftauchen: Die Herausforderung für Forscher besteht darin, eine Kombination von Merkmalen zu finden, die es zulässt, zu sagen: „Diese Ressource ist Teil dieser Operation."



Verhaltensmuster: Ein Artikel, der ursprünglich auf der Website des iranischen geistlichen Oberhaupts Ayatollah Khamenei veröffentlicht wurde, steht wenig später dann ohne Nennung der Quelle bei IUVMpress.com und britishleft.com, zwei Websites eines iranischen Propagandanetzwerks.
Bild von DFRLab

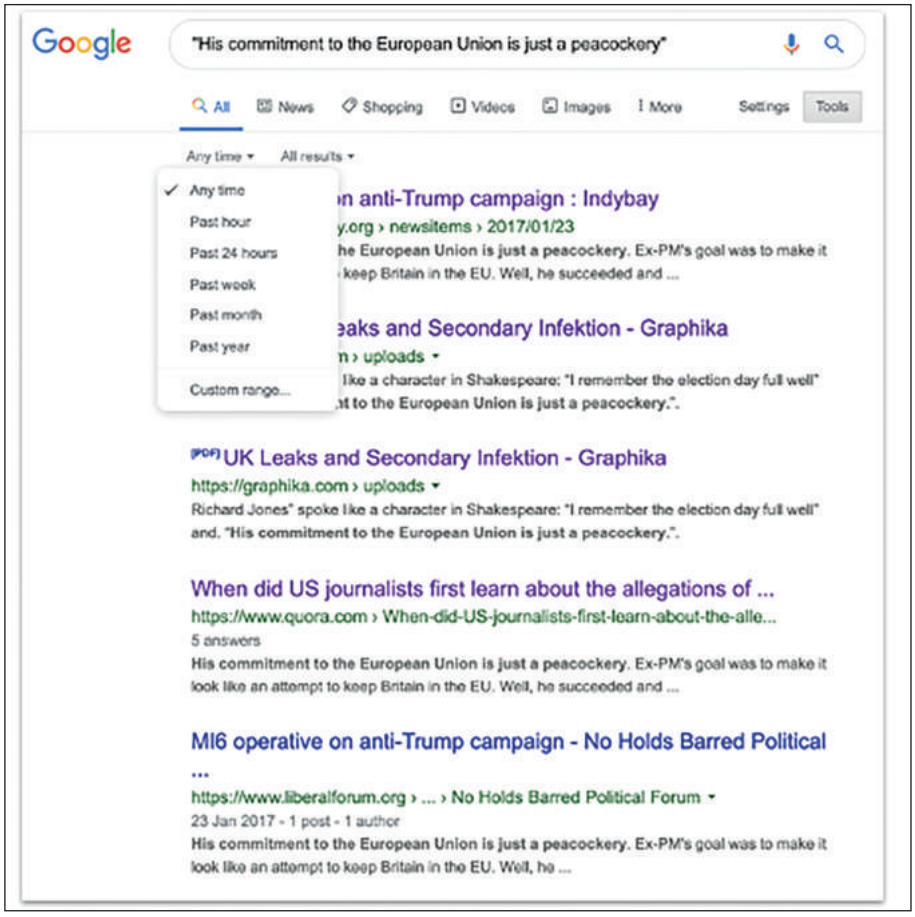
Manchmal kann das Fehlen von Erkennungsmerkmalen selbst ein Erkennungsmerkmal sein. Dies war der Fall bei der von Russland aus geführten Kampagne „Sekundäre Infektion". Dabei wurden hunderte von Konten auf verschiedenen Blogging-Plattformen verwendet, die alle minimale biographische Details enthielten, am Tag ihrer Erstellung einen Artikel veröffentlichten und dann nie wieder verwendet wurden. Dieses Verhaltensmuster war bei so vielen Konten so konsistent, dass während der Untersuchung klar wurde, dass das die Signatur der Operation war. Als kurz vor den britischen Parlamentswahlen im Dezember 2019 Betreiber anonymer Konten anfingen, durchgesickerte US-amerikanisch-britische Handelsdokumente in Umlauf zu bringen, zeigten Graphika und Reuters, dass sie genau dieser Signatur entsprachen. Reddit bestätigte die Analyse.



Reddit-Profil für ein Konto namens „McDownes", das Reddit der russischen Operation „Sekundäre Infektion" zuordnete. Das Konto wurde am 28. März 2019 eingerichtet, veröffentlichte einen Artikel etwas mehr als eine Minute nach seiner Einrichtung und verstummte dann.
Bildquelle: Graphika, Daten von redective.com

Inhaltliche Hinweise können auch dazu beitragen, Ressourcen zu identifizieren, die Teil desselben Netzwerks sind. Wenn eine solche Ressource ein Foto oder Meme mit anderen gemeinsam nutzt, lohnt sich eine Bilder-Rückwärtssuche, um zu prüfen, wo es noch verwendet wurde. Die Browser-Erweiterung von RevEye ist dafür ein besonders nützliches Werkzeug, da sie Rechercheuren die Rückwärtssuche über Google, Yandex, TinEye, Baidu und Bing ermöglicht. Es lohnt sich immer, mehrere Suchmaschinen zu verwenden, da diese oft unterschiedliche Ergebnisse liefern.

Wenn eine Ressource einen Text verteilt, lohnt es sich zu prüfen, wo dieser Text sonst noch erschienen ist. Vor allem bei längeren Texten ist es ratsam, auch einen oder zwei Sätze aus dem dritten oder vierten Absatz oder von weiter unten auszuwählen, da bei Täuschungsoperationen oftmals die Überschriften und Absätze von Artikeln bearbeitet werden, die kopiert worden sind, aber oft wenig Zeit für die Bearbeitung des Textkörpers aufgewendet wird. Fügen Sie den gewählten Abschnitt in Anführungszeichen in eine Suchmaschine ein, so dass Sie nur exakte Übereinstimmungen finden. Über das Menü „Suchfilter" lassen sich die Ergebnisse auch nach Datum sortieren.



Ergebnisse einer Google-Suche nach einer Phrase, die bei einer mutmaßlichen russischen Operation verwendet wurde – unter Nutzung des Google-Tools zur Datumseinschränkung.

Wenn Texte veröffentlicht werden, die Fehler enthalten, hat das für Recherchen einen besonderen Wert, da Fehler naturgemäß auffälliger sind als korrekt geschriebene Wörter. Zum Beispiel bezeichnete ein Artikel eines mutmaßlichen russischen Geheimdienstes die britische Stadt Salisbury, in der der ehemalige russische Agent Sergej Skripal vergiftet wurde, als „Solsbury". Dies ermöglichte eine viel genauere Google-Suche mit weitaus weniger Ergebnissen als eine Suche nach „Skripal" und „Salisbury". Sie brachte weniger Suchergebnisse, diese aber waren für die Recherche weitaus signifikanter.

Wenn Inhalte analysiert werden, ist es besonders wichtig, andere Indikatoren, wie zum Beispiel Verhaltensmuster, zu betrachten, um zu bestätigen, ob eine Ressource Teil einer größeren Operation ist. Es gibt viele legitime Gründe, warum unbedarfte Benutzer Inhalte von Informationsoperationen teilen. Das heißt: Werden von mehreren Nutzern gleiche Inhalte geteilt, ist das zunächst nur ein schwaches Signal und nicht zwingend eine Operation. Zum Beispiel haben viele Benutzer Memes der russischen „Internet Research Agency" geteilt, weil diese tatsächlich virale Qualitäten hatten. Das einfache Teilen von Inhalten allein reicht also nicht aus, um ein aktives Mitglied einer Operation zu identifizieren.

SAMMELN VON BEWEISEN

Informations- und Einflussnahmeoperationen sind komplex und schnelllebig. Eine der frustrierenderen Erfahrungen für einen Forscher, der sich auf die Auswertung öffentlicher Informationen spezialisiert hat (OSINT), ist es, wenn mitten in einer Untersuchung eine Reihe von Ressourcen und Einheiten vom Netz genommen wird. Eine der wichtigsten Regeln bei der Analyse lautet daher, alles zu archivieren, was man findet, denn man bekommt möglicherweise keine zweite Chance. Verschiedene Experten haben unterschiedliche Vorlieben, was das Archivieren betrifft, denn auch die Anforderungen ändern sich von Recherche zu Recherche. Tabellen sind nützlich, um grundlegende Informationen über eine große Anzahl von Einheiten aufzuzeichnen; gemeinsam genutzte, cloudbasierte Speicherplätze sind nützlich, um eine große Anzahl von Screenshots zu speichern. (Wenn Sie Screenshots machen, geben Sie einer Datei sofort einen identifizierbaren Namen: Es ist sehr ärgerlich, unter hunderten Dateien mit dem Namen „Screenshot" die eine richtige finden zu müssen.) Textdokumente mögen geeignet sein, um eine Mischung von Informationen festzuhalten, aber sie werden zu schnell unübersichtlich und unhandlich, wenn es sich um einen größeren Vorgang handelt.

Wofür auch immer Sie sich entscheiden, einige Informationen sollten grundsätzlich immer festgehalten werden. Dazu gehören die Art und Weise, wie etwas gefunden wurde, Name und Link, das Erstellungsdatum (falls bekannt) und Informationen über Follower, Vorlieben und/oder Ansichten. Außerdem eine grundlegende Beschreibung der Einheit (zum Beispiel „arabischsprachiges Pro-Saudi-Konto mit Emma-Watson-Profilbild"), damit Sie sich erinnern können, wenn Sie nach der Betrachtung von 500 anderen Einheiten nochmal darauf zurückkommen. Und wenn Sie im Team arbeiten, halten Sie fest, welches Teammitglied welche Einheit ausgewertet hat.

Links können durch die Nutzung von Archivdiensten wie Wayback Machine oder archive.is dauerhaft gesichert werden. Achten Sie aber darauf, dass Ihre Archivierungen keine echten Benutzer bloßstellen, die möglicherweise unwissentlich mit einer verdächtigen Einheit interagiert haben, und stellen Sie sicher, dass Ihr Archivlink das für Sie relevante Bildmaterial enthält – alternativ: Machen Sie einen Screenshot als Backup. Stellen Sie sicher, dass alle Einheiten an geschützten Orten gespeichert werden, zum Beispiel in kennwortgeschützten Ordnern oder verschlüsselten Dateispeichern. Verfolgen Sie, wer Zugriff hat, und überprüfen Sie den Zugriff regelmäßig. Zuletzt lohnt es sich, einer Einheit einen Vertrauenswert zu geben.

Operationen mit dem Ziel der Einflussnahme finden oft nichtsahnende Nutzer, die ihren Inhalt verstärken: Und genau darum geht es oft. Wie sicher können Sie sich sein, dass das neueste gefundene Profil Teil der Operation ist, die Sie beobachten – und warum? Der Grad des Vertrauens (hoch, mäßig oder niedrig) sollte als separater Eintrag gekennzeichnet werden, und die Gründe (siehe unten) sollten als Anmerkungen hinzugefügt werden.

ZUSCHREIBUNG UND GEWISSHEIT

Die größte Herausforderung bei der Identifizierung einer Informationsoperation besteht darin, sie einem bestimmten Akteur zuzuordnen. In vielen Fällen wird eine genaue Zuschreibung außerhalb der Möglichkeiten dessen liegen, was OSINT-Spezialisten auf der Grundlage öffentlich auffindbarer Informationen leisten können. Das Beste, was erreicht werden kann, ist ein gewisses Maß an Vertrauen, dass eine Operation *wahrscheinlich* von einem bestimmten Akteur geführt wird oder dass verschiedene Einheiten zu einer bestimmten Operation gehören – aber die Feststellung, wer hinter dieser Operation steckt, ist nur mit öffentlichen Informationen selten möglich.

Informationen wie Web-Registrierungen, IP-Adressen und Telefonnummern könnten eine feste Zuordnung liefern, aber sie werden oft anonymisiert und sind nur den Plattformen selbst bekannt. Deshalb ist auch die Kontaktaufnahme mit den Plattformen ein wesentlicher Bestandteil der Recherchen. Da die Plattformen ihre internen Sicherheitsteams aufgestockt haben, sind sie zunehmend bereit, bei Informationsoperationen öffentlich oder vertraulich Hintergründe und Zuschreibungen mitzuteilen. In jüngeren Fällen kam die entscheidende Zuschreibung direkt von den Plattformen, wie zum Beispiel bei der Enthüllung von staatlich unterstützten Informationsoperationen aus China auf Twitter, die auf Hongkong abzielten, oder bei der Enthüllung von Operationen im Zusammenhang mit der saudischen Regierung durch Facebook.

Inhaltliche Hinweise können eine Rolle spielen. Beispielsweise wurden bei einer Operation, die im Oktober 2019 auf Instagram aufgedeckt wurde, Memes veröffentlicht, die fast identisch mit Memes der russischen Internet Research Agency (IRA) waren, bei denen jedoch die Wasserzeichen der IRA entfernt worden waren. Die einzige Möglichkeit, diese Memes zu er-

stellen, bestand darin, die Originalbilder, die die Grundlage für die IRA-Posts bildeten, zu beschaffen und damit die Memes neu zu erstellen. Ironischerweise zeigte gerade dieser Versuch, die Ursprünge der IRA-Beiträge zu verschleiern, dass der Urheber tatsächlich die IRA selbst war.

Letztlich ist Zuschreibung eine Frage von Selbstbeschränkung. Als Wissenschaftler und Rechercheur muss man sich die Frage stellen: „Wie kann ich beweisen, dass diese Operation von der Person, die ich beschuldige, geleitet wurde?" Wer diese Frage nicht mit Sicherheit beantworten kann, sollte sich mit Zuschreibungen zurückhalten. Eine Informationsoperation zu identifizieren und aufzudecken ist eine schwierige und wichtige Arbeit, und eine voreilige oder ungenaue Zuschreibung kann alles untergraben, was zuvor erreicht wurde.

11 a. Fallbeispiel: Die Zuordnung von Endless Mayfly

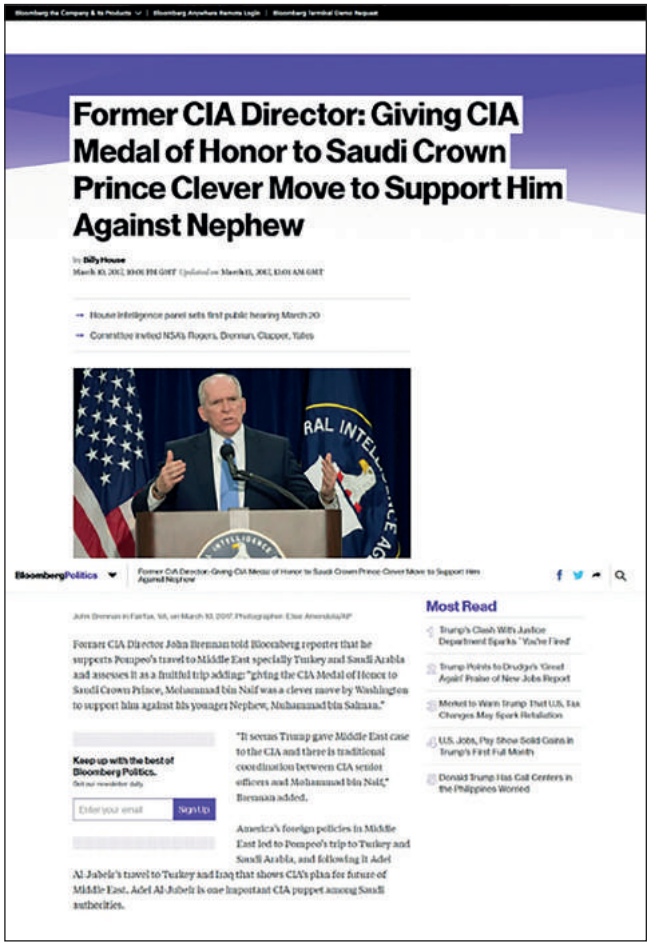
von: Gabrielle Lim
deutsche Bearbeitung: Marcus Engert

Gabrielle Lim *forscht am Technology and Social Change Research Project des Shorenstein Center an der Harvard Kennedy School in Cambridge (USA) und ist Wissenschaftlerin am Citizen Lab. Sie untersucht die Auswirkungen von Zensur und Medienmanipulation auf Sicherheit und Menschenrechte.*

Im April 2017 wurde auf Reddit ein nichtauthentischer Artikel veröffentlicht, der angeblich von der britischen Zeitung *The Independent* stammte und eine Fälschung war. Darin wurde der ehemalige stellvertretende britische Premierminister Nick Clegg falsch mit den Worten zitiert, die damalige Premierministerin Theresa May biedere sich arabischen Regimen an. Aufmerksame Reddit-Nutzer bezeichneten den Beitrag schnell als zweifelhaft und falsch. Er war auf independent.co veröffentlicht worden und nicht auf www.independent.co.uk. Der Absender war eine schwer einzuschätzende Person, die bereits mehrere nichtauthentische Artikel auf Reddit veröffentlicht hatte.

Ausgehend von diesem falschen Artikel, der falschen Web-Adresse und der Person verbrachten die Wissenschaftler am Citizen Lab die nächsten 22 Monate unter anderem damit, das Netzwerk dahinter zu untersuchen – ein Netzwerk einer vielschichtigen Online-Informationsoperation. Das Ziel der Operation, die Endless Mayfly genannt wurde, bestand darin, Journalisten und Aktivisten mit nichtauthentischen Websites zu täuschen, indem die Web-Auftritte von etablierten Medien nachgemacht und dort falsche und hetzerische Informationen verbreitet wurden.

Im Großen und Ganzen war der Plan, einen seriösen Nachrichtensender mit einem nichtauthentischen Artikel zu täuschen, die Wirkung über ein Netzwerk von Websites und gefälschten Twitter-Figuren zu verstärken und den falschen Artikel irgendwann zu löschen oder umzuleiten, nachdem eine gewisse Online-Aufmerksamkeit um ihn herum entstanden war. Hier sehen Sie ein Beispiel für solch einen gefälschten Artikel, der sich als Bloomberg.com angab, allerdings auf bloomberq.com veröffentlicht wurde:



Die Überschrift lautet: „Früherer CIA-Direktor: Dem saudischen Kronprinzen die CIA-Ehrenmedaille zu verleihen, ist ein cleverer Schachzug, um ihn gegen seinen Neffen zu unterstützen"

Und hier ist zu sehen, wie zwei gefälschte Twitter-Personen, die mit Endless Mayfly in Verbindung stehen, eine Kopie der türkischen Nachrichtenwebsite Daily Sabah verbreiten. Bemerkenswert dabei: Die Person rechts nutzt ein Foto der Schauspielerin Elisja Cuthbert als Profilbild.

Zum Zeitpunkt der Veröffentlichung unseres Berichts im Mai 2019 enthielt unser Datensatz 135 nichtauthentische Artikel, 72 Domains, elf Personen, eine gefälschte Organisation und ein proiranisches Veröffentlichungsnetzwerk, das die in den falschen Artikeln enthaltenen Unwahrheiten verstärkte. Am Ende kamen wir mit einiger Sicherheit zu dem Schluss, dass Endless Mayfly eine aus dem Iran gesteuerte Informationsoperation war.

Endless Mayfly zeigt, wie man Analysen des Netzwerks und der Inhalte mit externer Berichterstattung kombinieren kann, um zu einer Zuordnung zu gelangen. Es zeigt auch, wie schwierig es ist, Informationsoperationen einem bestimmten Akteur zuzuordnen, warum mehrere Indikatoren erforderlich sind und wie man ein Vertrauensniveau erreicht, um einen Grad an Sicherheit für die Zuordnung anzugeben. Letztlich ist die Zuschreibung eine schwierige Aufgabe, die oft durch unvollkommene Informationen erschwert wird, es sei denn, Sie bekommen jemanden dazu, seine Beteiligung einzuräumen oder einen anderweitigen Beweis zu erbringen – und beides ist meist nicht der Fall. Aus diesem Grund werden bei der Frage nach der Herkunft von Medienmanipulation oft ein Verdacht, eine Vermutung, eine Wahrscheinlichkeit formuliert und selten eine ultimative Aussage.

MULTIPLE DATENPUNKTE UND ANALYSE „TRIANGULIEREN"

Aufgrund des geheimen Charakters von Informationsoperationen, der Fähigkeit von Akteuren, falsche Fährten zu legen, und des flüchtigen Charakters von Hinweisen sollte die Zuschreibung das Ergebnis einer Kombination von Analyse und Belegen sein. Bei Endless Mayfly kamen wir mit einiger Sicherheit zu dem Schluss, dass es sich um eine Operation aus dem Iran handelte, und zwar aufgrund von Indikatoren, die sich aus drei Arten von Analysen ableiten:

1. Analyse der Narrative
2. Analyse der Netzwerke
3. Analyse der externen Berichterstattung

1. Analyse der Narrative
Anhand von Inhalts- und Diskursanalysen der 135 nichtauthentischen Artikel, die in unserer Untersuchung gesammelt wurden, stellten wir fest, dass sich die dort verbreiteten Narrative mit den Interessen des Irans in Einklang bringen ließen. Jeder Artikel wurde für die Analyse in Kategorien kodiert, die wir nach einer ersten überblickshaften Lektüre aller Artikel festgelegt haben. Es gab zwei Kodierungsrunden: Die erste Runde wurde von zwei Forschern unabhängig voneinander durchgeführt, und eine zweite Runde wurde von denselben Forschern gemeinsam durchgeführt, um etwaige Diskrepanzen zu klären. Diese Tabelle stellt die Ergebnisse unseres Kodierungsprozesses dar.

Kategorie	Anzahl Artikel	Kategorienbeschreibung
Geopolitische Unstimmigkeiten	63 (46,7 %)	Der Artikel beschreibt Ereignisse, Aktionen oder Statements von offiziellen Vertretern gegenüber einem anderen Staat, die als provokativ, feindlich oder als gegen die Interessen dieses Staates gerichtet ausgelegt werden können.
Inländische Unstimmigkeiten	16 (11,9 %)	Der Artikel beschreibt Ereignisse, Aktionen oder Statements von politischen Vertretern, die Zwietracht zwischen verschiedene Parteien oder Akteuren innerhalb eines Staates säen könnten.
Kooperation mit Israel	14 (10,4 %)	Der Artikel beschreibt Ereignisse, Aktionen oder Statements von Politikern oder Regierungsvertretern, die eine Zusammenarbeit zwischen Israel und einem anderen Staat zeigen.
Saudi-Arabien unterstützt Terrorismus	9 (6,7 %)	Der Artikel beschreibt Ereignisse, Aktionen oder Statements, die Saudi-Arabien entweder mit terroristischen Aktivitäten in Verbindung bringen oder unterstellen, dass Saudi-Arabien diese unterstütze.
Andere	5 (3,7 %)	Der Artikel lässt sich keiner der anderen Kategorien zuordnen.
Nicht archiviert	31 (23 %)	Der Artikel kann nicht kodiert werden, da er nicht mehr verfügbar ist und kein Archiv oder Screenshot sowie keine Kopie erstellt wurden, um eine Textanalyse durchführen zu können.
Kopie eines bestehenden Artikels	5 (3,7 %)	Der Artikel ist eine direkte Kopie eines bereits bestehenden Artikels.

Nachdem alle Artikel kodiert waren, konnten wir die am häufigsten von Endless Mayfly verbreiteten Narrative ermitteln. Wir verglichen diese mit unseren vorherigen umfangreichen Recherchen über die Region, über dortige Rivalitäten und Bündnisse, über die geopolitischen Interessen und Bedrohungen und darüber, was dort früher an Informationskontrollen stattgefunden hatte. Dies war notwendig, um die Punkte zu kontextualisieren und die Narrative in einen breiteren politischen Kontext einzuordnen. Mit den Ergebnissen dieser Kodierung kamen wir zu dem Schluss, dass diese Narrative am ehesten den Interessen des Iran dienten.

2. Analyse der Netzwerke

Weiterhin wurde eine Netzwerk-Analyse durchgeführt, um festzustellen, welche Websites oder Plattformen für die Verstärkung der Aufmerksamkeit hinsichtlich der betreffenden Inhalte verantwortlich waren. Bei Endless Mayfly waren zwei Netzwerke an der Verbreitung der unauthentischen Artikel und der darin enthaltenen Falschinformationen beteiligt: ein Netzwerk von proiranischen Websites und ein Cluster von proiranischen Konten auf Twitter. Beide trugen dazu bei, dass wir die Operation Endless Mayfly dem Iran zuschrieben, weil über sie durchweg Geschichten verbreitet wurden, die im Einklang mit der offiziellen iranischen Politik, öffentlichen Erklärungen und den Positionen Irans gegenüber Saudi-Arabien, Israel und den Vereinigten Staaten standen.

Das Netzwerk der Websites — das Netzwerk der Websites bestand aus einer Reihe scheinbar proiranischer Seiten, die sich als unabhängige Redaktionen und Nachrichtenagenturen inszenierten. Insgesamt fanden wir 353 Websites in 132 Domains, die auf die nichtauthentischen Artikel von Endless Mayfly verlinkten oder von diesen verlinkt waren. Dieser Prozess beinhaltete eine Google-Suche aller Web-Adressen der nichtauthentischen Artikel und ihrer Überschriften. Darüber hinaus scannten wir die von den Konten in unserem Netzwerk getwitterten Links und identifizierten Websites, die Verweise zu den Artikeln enthielten.

Im Anschluss an diesen Prozess identifizierten wir die zehn Seiten, die am häufigsten auf die nichtauthentischen Artikel verwiesen. Von diesen zehn Domains hatten acht die gleiche IP-Adresse oder gemeinsame Registrierungsdaten, was darauf hindeutete, dass sie möglicherweise von demselben Akteur verantwortet wurden. Auch der Inhalt dieser Seiten entsprach den offiziellen iranischen Interessen. Beispielsweise stellte die Seite IUVM Press, die 57-mal auf die nichtauthentischen Artikel von Endless Mayfly verwies oder auf sie verlinkte, ein PDF-Dokument mit dem Titel „Statut" online, in dem sich der Autor ausdrücklich gegen „die Aktivitäten und Projekte globaler Arroganzstaaten, den Imperialismus und den Zionismus" positionierte und betonte, dass „der Hauptsitz der Union sich in Teheran befindet – der Hauptstadt der Islamischen Republik Iran".

Das Netzwerk der Konten — ähnlich wie die nichtauthentischen Artikel und das Netzwerk der Websites positionierten sich auch die Konten, die mit Endless Mayfly auf Twitter in Verbindung standen, entschieden kritisch gegenüber Saudi-Arabien, Israel und westlichen Nationen im Allgemeinen. Eine Analyse ihrer Twitter-Aktivitäten ergab, dass sie eine Kombination aus glaubwürdigen und unauthentischen Artikeln verteilten, die politischen Gegenspielern des Iran sehr kritisch gegenüberstanden. So zum Beispiel der Twitter-Account von „Peace, Security, Justice, Community", getarnt als Nichtregierungsorganisation, die durch unsere Untersuchung identifiziert wurde (siehe unten). Er verbreitete nicht nur Inhalte, die gegen Saudi-Arabien, Israel und die USA gerichtet waren, seine Profilbilder nahmen Saudi-Arabien auch ins Visier. Beachten Sie das Fadenkreuz über Saudi-Arabien auf dem Profilfoto und die verwendete Karte. In der Biographie des Berichts werden Saudi-Arabien und die Wahhabiten ausdrücklich als Ursache des Extremismus genannt.

In vergleichbarer Weise erwähnt dieser Tweet von einem anderen zu Endless Mayfly gehörenden Konto namens „Mona A. Rahman" den Journalisten und saudischen Kritiker Ali al-Ahmed, während er gleichzeitig Saudi-Arabiens Kronprinzen Mohammad bin Salman kritisiert.

Im Tweettext steht: „Ich lade Dissidenten ein, sich gegen den mörderischen und barbarischen saudischen Kronprinzen nächsten Monat in London zu versammeln."

3. Analyse der externen Berichterstattung
Wir haben unsere Ergebnisse und Daten auch mit externer Berichterstattung verglichen. Nach einem Hinweis von FireEye im August 2018 deaktivierte Facebook beispielsweise einige Konten und Seiten, die mit dem von Endless Mayfly verwendeten Netzwerk verknüpft waren. In seiner Analyse identifizierte FireEye mehrere Seiten, die Teil des von uns identifizierten Netzwerks aus Websites waren, so zum Beispiel institutomanquehue.org und RPFront.com. Wie wir kam auch FireEye mit einiger Sicherheit zu dem Schluss, dass diese „mutmaßliche Operation zur Einflussnahme" wohl aus dem Iran stammt. Auch Facebook äußerte in einer Erklärung diesen Verdacht.

Ergänzend dazu veröffentlichte Twitter einen Datensatz von mit dem Iran verknüpften Konten, die wegen „koordinierter Manipulation" gesperrt worden waren. Obwohl darin Konten mit weniger als 5.000 Followern anonymisiert wurden, konnten wir ein Endless Mayfly-Konto (@Shammari_Tariq) darin identifizieren. Die Einschätzungen von Twitter, Facebook und FireEye waren zur Untermauerung unserer Hypothese nützlich, da sie Indizien lieferten, die nicht Teil unserer Datenerfassung waren und sich dennoch mit den von uns identifizierten Endless Mayfly-Einheiten überschnitten. Zum Beispiel identifizierte FireEye in seiner Analyse Telefonnummern und Anmelde-Informationen, die mit Twitter-Konten und Websites in Verbindung standen, die an Endless Mayfly beteiligt waren – Belege, die nicht Teil unseres Datensatzes waren. Ebenso lagen Facebook und Twitter vermutlich Konto-Informationen wie zum Beispiel IP-Adressen vor, auf die wir keinen Zugriff hatten. Solche zusätzlichen Datenpunkte, die durch diese externen Berichte identifiziert wurden, trugen so dazu bei, die Aussagekraft der Belege zu erhöhen.

Anmerkungen zum Grad der Sicherheit von Rückschlüssen
Im Fall von Endless Mayfly wiesen die von uns gesammelten Belege – die proiranischen Narrative, die Konten, das Netzwerk aus Websites – auf den Iran als plausible Quelle für die Informationsoperation hin. Dieses Beweismaterial wurde durch glaubwürdige externe Berichte und Untersuchungen von FireEye, Facebook und Twitter ergänzt, die unsere Ergebnisse bestätigten. Jedes einzelne dieser Stücke reichte zwar für sich genommen nicht für eine Zuschreibung aus, gemeinsam aber trugen sie dazu bei, die Hypothese zu erhärten. Dennoch: Trotz der zahlreichen Hinweise, die auf den Iran hindeuteten, hatten wir noch keine endgültigen Beweise. Daher wandten wir einen Deutungsrahmen für Cyber-Zuschreibungen an, der auch von Geheimdienstmitarbeitern genutzt wird. Er nutzt mehrere Indikatoren und Wahrscheinlichkeitsbewertungen (niedrig, mittel, hoch), so dass Analysten ihre Ergebnisse mitteilen und gleichzeitig den Grad ihrer Unsicherheit beziffern können.

Letztendlich kamen wir zu dem Schluss, dass Endless Mayfly eine Operation ist, die mit einiger Berechtigung dem Iran zuzuschreiben ist. Der Direktor der Nationalen Geheimdienste in den USA definiert diese Bewertung so: „Die Informationen sind glaubwürdig und plausibel, aber nicht von ausreichender Qualität oder ausreichend erhärtet, um ein höheres Maß an Sicherheit zu rechtfertigen." Wir haben uns nicht für ein höheres Maß entschieden, weil wir der Meinung waren, wir hätten keine ausreichenden Belege, um eine Operation unter falscher Flagge – also von jemandem, der den Anschein zu erwecken versucht, der Iran stecke hinter dieser Operation – oder von einem Dritten, der mit iranischen Interessen sympathisiert, völlig auszuschließen.

Will man Informationsoperationen wie Endless Mayfly jemandem zuschreiben, wird man sich dabei fast immer auf unvollständige und unvollkommene Informationen stützen müssen. Anmerkungen zum Grad der Sicherheit in solche Zuschreibungen einzubauen, ist daher eine wichtige Komponente – denn sie arbeitet mit mehreren Ebenen von Zurückhaltung. Falsche Zuschreibungen oder ein zu hohes Vertrauensniveau können schlimme Folgen haben, insbesondere wenn sie Einfluss auf die Politik von Staaten oder sogar auf Vergeltungsmaßnahmen nehmen. Um übereilte und falsche Zuschreibungen zu vermeiden, ist es wichtig, mehrere Indikatoren, Beweisarten und Analysewege einzubeziehen und die Sicherheit, mit der die Schlussfolgerungen gezogen werden, auf einem abgestuften Level anzugeben, das alternative Hypothesen und fehlende Daten berücksichtigt.

11 b. Fallbeispiel: Wie wir eine Informationsoperation in West Papua untersuchten

von: Elise Thomas, Benjamin Strick
deutsche Bearbeitung: Marcus Engert

Benjamin Strick *(London) ist ein auf öffentliche Quellen (OSINT) spezialisierter Rechercheur bei der BBC, Autor bei Bellingcat und Dozent für Open-Source-Techniken, orts- und raumbezogene Untersuchungen und Netzwerk-Analysen. Er hat einen juristischen und militärischen Hintergrund und ist auf Geolokalisierungs- und Aufklärungsmethoden im Bereich von Menschenrechten, Konflikten und Privatsphäre spezialisiert.*
Elise Thomas *ist freie Journalistin und Wissenschaftlerin am International Cyber Policy Centre des australischen Strategic Policy Institute. Ihre Arbeiten erschienen in Wired, Foreign Policy, The Daily Beast, The Guardian und anderen Medien. Zuvor arbeitete sie im Koordinationsbüro der Vereinten Nationen für humanitäre Angelegenheiten sowie als Podcast-Autorin und Rechercheurin.*

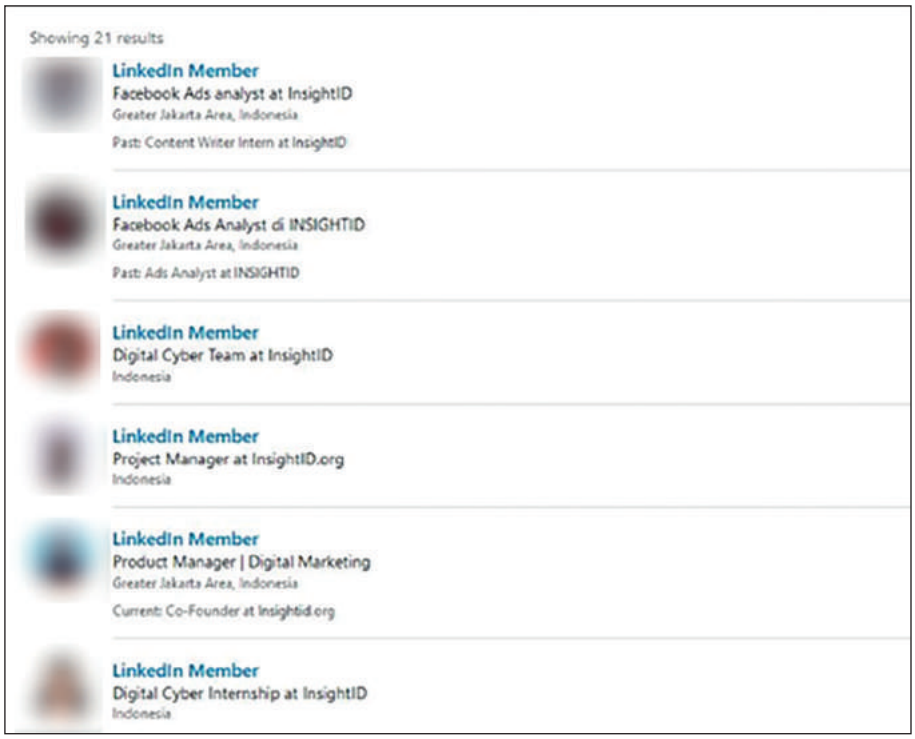
Im August 2019 flammten erneut Separatistenkonflikte in West-Papua auf, einer Provinz, die durch eine umstrittene Entscheidung in den 1960er-Jahren Teil Indonesiens wurde. Seither leidet die Region unter weitreichenden Menschenrechtsverletzungen, die von den indonesischen Behörden begangen wurden, um Dissidenten zu unterdrücken. Der Zugang zu der Region ist stark eingeschränkt, ausländischen Journalisten wurde die Berichterstattung in der Provinz verboten. All dies macht die sozialen Medien zu einer entscheidenden Ressource für die Beobachtung von West-Papua und die Berichterstattung darüber.

Bei dem Versuch, Filmmaterial zu geolokalisieren, das Gewalt in der Stadt FakFak zeigte, bemerkten wir zwei Hashtags, die sich auf Twitter verbreiteten: #WestPapua und #FreeWestPapua.

Die Suche unter diesen Hashtags ergab eine Fülle von gefälschten Konten, die dieselben Videos und denselben Text mit denselben Hashtags automatisch verteilten. Die Betreiber der Konten twitterten und markierten auch Inhalte der anderen mit „Gefällt mir", was den Effekt hatte, dass die Hashtags weiter verstärkt wurden, was ihnen wiederum mehr Aufmerksamkeit verschaffte.

Das Verfahren zur Analyse dieser automatisierten Konten wurde in Kapitel 3 ausführlich beschrieben. Aufbauend auf dieser Arbeit erweiterten wir unsere Recherche: Wir wollten die Personen oder Gruppen, die hinter den Konten stehen, identifizieren. Dabei stießen wir auf eine ähnliche, kleinere und anscheinend unverbundene Kampagne und waren zudem in der Lage, die verantwortliche Person dahinter zu identifizieren. Die Betreiber beider Kampagnen räumten schließlich ihre Beteiligung ein, nachdem sie von der BBC darauf angesprochen worden waren.

Der Umfang der ersten Kampagne und die Tatsache, dass sie über mehrere Plattformen lief, gab uns eine ganze Reihe von Möglichkeiten, erste Anhaltspunkte zu finden, an denen entlang wir recherchieren und Informationen über die Betreiber der Kampagne suchen konnten. Die erste hilfreiche Information waren die Websites, die über das Netzwerk von Twitter- und Facebook-Konten verbreitet wurden. Die Whois-Suche ergab, dass vier der Domains unter einem falschen Namen und einer falschen E-Mail-Adresse, aber mit einer echten Telefonnummer registriert worden waren. Wir gaben die Nummer in WhatsApp ein, um zu sehen, ob sie dort mit einem Konto verbunden war. Das war sie, und mehr noch: Dieser Nutzer hatte auch ein Profilfoto. Mit Hilfe der Rückwärtsbildsuche von Yandex konnten wir das Profilfoto zu Facebook-, LinkedIn- und Freelancer.com-Konten zurückverfolgen und so ein Profilbild der Person finden sowie über LinkedIn auch ihren aktuellen Arbeitsplatz und einige ihrer Arbeitskollegen.



Die Person arbeitete bei einer in Jakarta ansässigen Firma namens InsightID, auf deren Website es hieß, sie biete „integrierte PR- und digitale Marketingprogramme" an. Wir sammelten mit der Zeit auch weitere Hinweise darauf, dass InsightID für das, was wir untersuchten, verantwortlich sein könnte. Auf seiner Website verwies InsightID auf seine Arbeit an der „Papua Program Development Initiative", die „die rasche sozioökonomische Entwicklung Papuas und ihre Herausforderungen untersucht". Ehemalige Mitarbeiter und Praktikanten von InsightID beschrieben die Produktion von Videoinhalten, das Verfassen von Texten und die Übersetzung von Inhalten als Teil ihrer Arbeit am Papua-Entwicklungsprojekt.

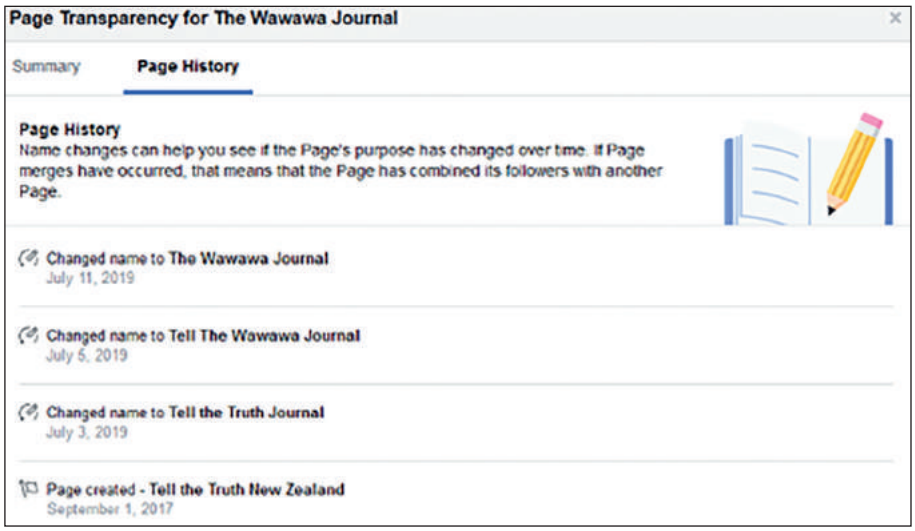
Ein ehemaliger Mitarbeiter gab in seinem LinkedIn-Profil an, dass seine Arbeit auf „West Papuan (Instagram, Facebook, Website)" zu sehen sei. West Papuan war eine von fünf Nachrichten-Websites, die an der Kampagne beteiligt waren. Ein weiterer Mitarbeiter von InsightID richtete ein YouTube-Konto unter seinem eigenen Namen ein, wo er ein Video veröffentlichte, das Teil der Kampagne war. Dieses Video wurde dann auf westpapuan.org eingebettet.

Wir recherchierten weiter unter den Websites und fanden heraus, dass der Mitbegründer von InsightID seine Firmen-E-Mail-Adresse verwendet hatte, um am selben Tag 14 Websites zu registrieren, von denen die meisten eindeutig einen direkten Bezug zu West Papua hatten. Dazu gehörten westpapuafreedom.com, westpapuagenocide.com und westpapuafact.com. Jede neue Information fügte der These, dass InsightID für die Operation verantwortlich war, ein klein wenig mehr Belegkraft hinzu.

Zu diesem Zeitpunkt traten BBC-Journalisten an InsightID heran und baten um einen Kommentar. Obwohl das Unternehmen darauf nicht reagierte, räumte InsightID schließlich seine Verantwortung ein und äußerte in einem Beitrag in den sozialen Medien, dass „unser Inhalt Indonesien gegen die Falscherzählung der Separatistengruppen von Free Papua verteidigt". Wer der Kunde war, der InsightID mit der Durchführung der Informationskampagne beauftragte, konnten wir nicht herausfinden.

Während wir diese größere Operation aufdeckten, untersuchten wir auch ein kleineres Netzwerk von drei Websites, die sich als unabhängige Nachrichtenquellen ausgaben und mit Profilen in sozialen Medien verknüpft waren. Obwohl diese Websites offensichtlich nicht mit der ersten oben genannten Kampagne in Verbindung standen, zielten sie auf die internationale Wahrnehmung der Situation in West-Papua ab und konzentrierten sich auf Zielgruppen in Neuseeland und Australien.

Der Schlüssel zur Identifizierung der verantwortlichen Person war, dass die Facebook-Seite eines dieser Medien, des Wawawawa-Journals, ursprünglich Tell the Truth NZ hieß. Wir konnten das aus der Namensgeschichte der Seite herauslesen. Damit war eine Verbindung zur Website tellthetruthnz.com gegeben, und diese wiederum war registriert auf Muhamad Rosyid Jazuli.



Für Facebook-Seiten lässt sich zurückverfolgen, wann eine Seite ihren Namen geändert hat und wie.

Als BBC-Journalisten ihn mit der Recherche konfrontierten, gab Jazuli zu, der Betreiber der Kampagne zu sein. Er arbeitete mit dem Jenggala Center zusammen, einer vom indonesischen Vizepräsidenten Jusuf Kalla ins Leben gerufenen Organisation. Sie wurde 2014 gegründet, um seine Wiederwahl und die Regierung von Präsident Jokowi zu unterstützen.

Was diese Untersuchung zeigt, ist Folgendes: Das Aufspüren von Informationskampagnen und die Zuordnung zu dafür verantwortlichen Einzelpersonen und Gruppen erfordern nicht unbedingt komplizierte Techniken oder Werkzeuge – man braucht aber Geduld und eine gewisse Portion Glück. Diese Untersuchung stützte sich auf öffentliche Informationen wie Whois-Registrierungsdaten von Websites, Rückwärtssuchen von Bildern, Social-Media-Profile und die Analyse von Website-Quellcodes. Die Tatsache, dass die Kampagne über mehrere Plattformen lief, in Kombination mit den Social-Media- und LinkedIn-Profilen der Mitarbeiter von InsightID, war am Ende ausschlaggebend dafür, dass wir viele kleine Puzzleteile zusammenfügen konnten, um das Gesamtbild zu sehen.

Wenn es eine wichtige Lektion aus diesem Fall gibt, dann ist es die, in Ruhe darüber nachzudenken, wie man Details oder Hinweise von einer Plattform auf eine andere übertragen kann.

ÜBER DIE AUTOREN

Herausgeber: Craig Silverman

Craig Silverman *lebt in Toronto und ist als Medienredakteur von BuzzFeed News für einen weltweiten Themenbereich zuständig, von Plattformen über Falschinformationen im Netz bis hin zu Medienmanipulation. Er ist Herausgeber des „Verification Handbook" und des „Verification Handbook for Investigative Reporting" und Autor des Buches „Lies, Damn Lies, and Viral Content: How News Websites Spread (and Debunk) Online Rumors, Unverified Claims and Misinformation". Seine Arbeit wurde unter anderem mit den Mirror Awards, den U.S. National Press Club Awards, den National Magazine Awards (Canada) und den Digital Publishing Awards ausgezeichnet.*

Deutsche Bearbeitung: Marcus Engert

© BuzzFeed News - by Stefan Beetz

Marcus Engert *ist Investigativ-Journalist und Senior-Reporter im Deutschland-Büro von BuzzFeed News. Er recherchiert unter anderem zu Extremismus, Sicherheitsfragen und Menschenrechtsverletzungen und gehört zum internationalen Rechercheteam, das die FinCen-Files auswertete. Zuvor war er Mitgründer und Chefredakteur von detektor.fm, ein mit dem Deutschen Radiopreis als „Beste Innovation" ausgezeichnetes Podcastlabel und Online-Radio. Marcus Engert wurde unter anderem mit dem Ernst-Schneider-Preis und dem Arthur F. Burns-Fellowship ausgezeichnet, ist IVLP-Alumni und lebt in Leipzig und Berlin.*

IMPRESSUM

Englische Originalfassung

Herausgeber: Craig Silverman

Mitherausgeberin: Claire Wardle

Redaktion: Merrill Perlman

Autoren: Ben Collins, Ben Nimmo, Benjamin Strick, Brandy Zadrozny, Charlotte Godart, Claire Wardle, Craig Silverman, Donie O'Sullivan, Elise Thomas, Farida Vis, Gabrielle Lim, Gemma Bagayaua-Mendoza, Hannah Guy, Henk van Ess, Jane Lytvynenko, Joan Donovan, Johanna Wild, Sam Gregory, Sérgio Lüdtke, Simon Faulkner, Vernise Tantuco

Produktionsmanager: Arne Grauls

Deutsche Bearbeitung: Marcus Engert

Dieses Handbuch wird vom European Journalism Centre herausgegeben und dank der Finanzierung durch die Craig Newmark Philanthropies ermöglicht. Die deutsche Version wurde ermöglicht durch die Landesanstalt für Medien Nordrhein-Westfalen (www.medienanstalt-nrw.de).

Deutsche Ausgabe

Herausgeberin
Landesanstalt für Medien NRW
Sabrina Nennstiel (Leiterin Kommunikation)
Dr. Meike Isenberg (Leiterin Forschung)
Zollhof 2, 40221 Düsseldorf
www.medienanstalt-nrw.de