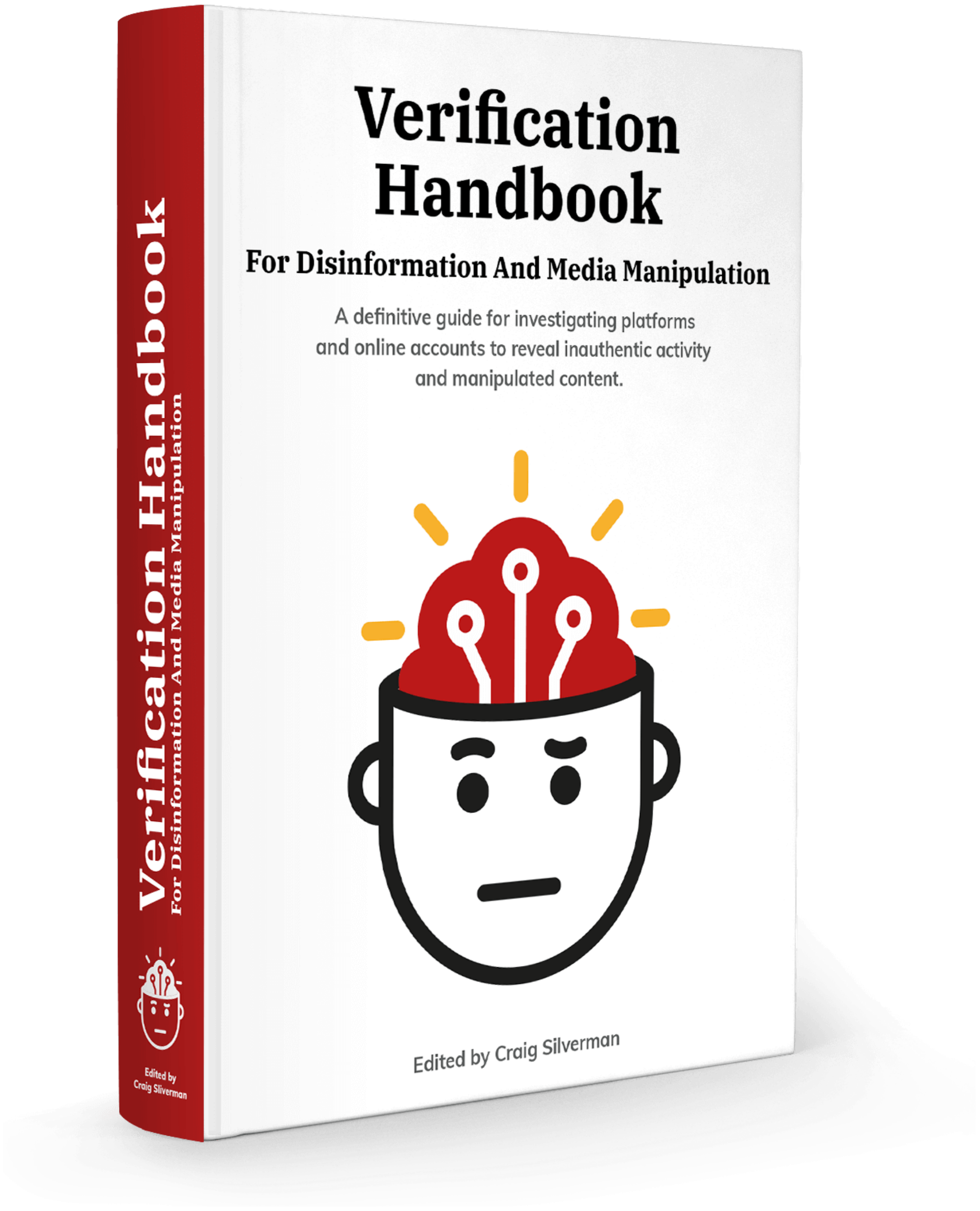
Verification Handbook


Verification Handbook
For Disinformation And Media Manipulation
A definitive guide for investigating platforms and online accounts to reveal inauthentic activity and manipulated content.
Edited by Craig Silverman
Verification Handbook For Disinformation And Media Manipulation
Edited by Craig Silverman

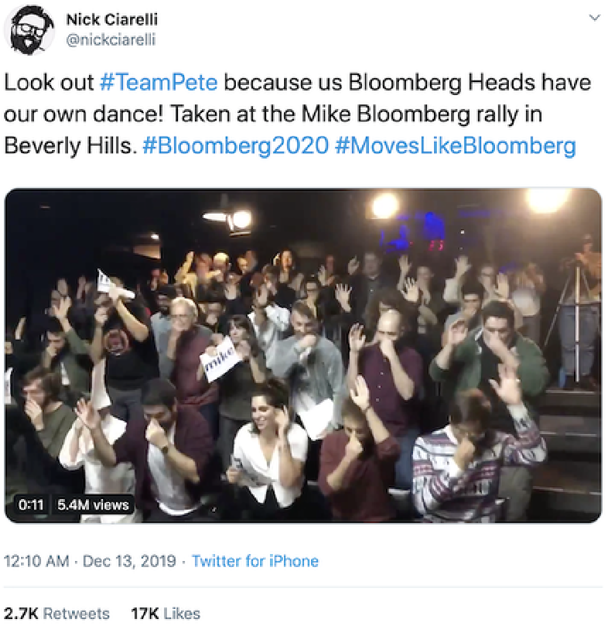
Chapters

Investigating Disinformation and Media Manipulation

Written by: Craig Silverman

***Craig Silverman** is the media editor of BuzzFeed News, where he leads a global beat covering platforms, online misinformation and media manipulation. He previously edited the "Verification Handbook" and the "Verification Handbook for Investigative Reporting," and is the author of, "Lies, Damn Lies, and Viral Content: How News Websites Spread (and Debunk) Online Rumors, Unverified Claims and Misinformation."*

In December 2019, Twitter user @NickCiarelli shared a video he said showed a dance routine being adopted by supporters of Michael Bloomberg's presidential campaign. The video's lackluster enthusiasm and choreography immediately helped it rack up retweets and likes, mostly from people who delighted in mocking it. The video eventually attracted more than 5 million views on Twitter.



Nick Ciarelli
@nickciarelli

Look out #TeamPete because us Bloomberg Heads have our own dance! Taken at the Mike Bloomberg rally in Beverly Hills. #Bloomberg2020 #MovesLikeBloomberg

12:10 AM · Dec 13, 2019 · Twitter for iPhone

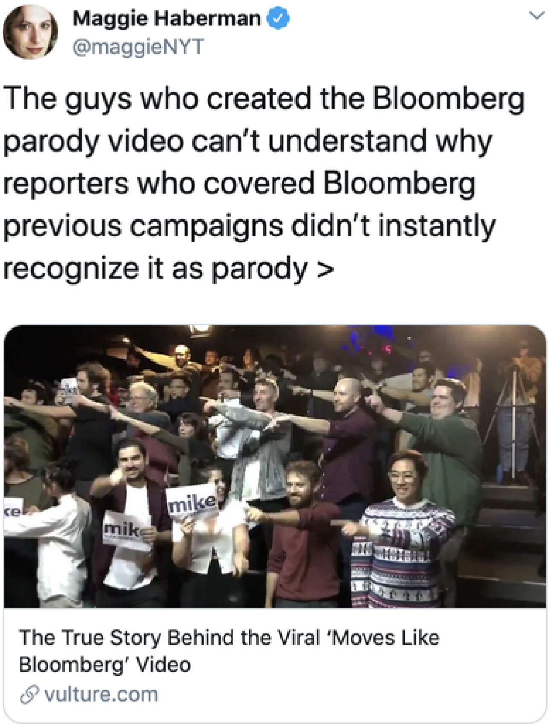
2.7K Retweets 17K Likes

Ciarelli's Twitter bio said he was an intern for the Bloomberg campaign, and his subsequent tweets included proof points such as a screenshot of an email from an alleged Bloomberg campaign staffer approving budget for the video.

But a quick Google search of Ciarelli's name showed he's a comedian who has created humor videos in the past. And that email from a Bloomberg staffer? It was sent by Ciarelli's frequent comedic partner, Brad Evans. That information was also just a Google search away.

But in the first minutes and hours, some believed the cringeworthy video was an official Bloomberg production.

Maggie Haberman, a prominent New York Times political reporter, tweeted that journalists who covered Bloomberg's previous mayoral campaigns had reason to not dismiss it right away:



Maggie Haberman
@maggieNYT

The guys who created the Bloomberg parody video can't understand why reporters who covered Bloomberg previous campaigns didn't instantly recognize it as parody >

The True Story Behind the Viral 'Moves Like Bloomberg' Video
vulture.com

Knowledge can take many forms, and in this new digital environment, journalists have to be wary of relying too much on any given source of information — even if it's their own firsthand experience.

Apparently, some reporters who knew Bloomberg and his style of campaigning felt the video could be real. At the same time, journalists who knew nothing about Bloomberg and chose to judge the video by its source could have found the correct answer immediately — in this case, simply Googling the name of the man who shared it.

The point isn't that experience covering Bloomberg is bad. It's that at any given moment we can be led astray by what we think we know. And in some cases our base of knowledge and experience can even be a negative. We can also be fooled by digital signals such as retweets and views, or by efforts to manipulate them.

As the Bloomberg video showed, it takes little effort to create misleading signals like a Twitter bio or a screenshot of an email that appears to back up the content and claim. These in turn help content go viral. And the more retweets and likes it racks up, the more those signals will convince some that the video could be real.

Of course, there are far more devious examples than this one. Unlike Ciarelli, the people behind information operations and disinformation campaigns rarely reveal the ruse. But this case study shows how confusing and frustrating it is for everyone, journalists included, to navigate an information environment filled with easily manipulated signals of quality and trust.

Trust is the foundation of society. It informs and lubricates all transactions and is key to human connection and relationships. But it's dangerous to operate with default trust in our digital environment.

If your default is to trust that the Twitter accounts retweeting a video are all amplifying it organically, you will get gamed. If you trust that the reviews on a product are all from real customers, you'll waste your money. If you trust that every news article in your news feed represent an unbiased collection of what you most need to see, you will end up misinformed.

This reality is important for every person to recognize, but it's essential for journalists. We are being targeted by coordinated and well-funded campaigns to capture our attention, trick us into amplifying messages, and bend us to the will of states and other powerful forces.

The good news is this creates an opportunity — and imperative — for investigation.

This handbook draws on the knowledge and experience of top journalists and researchers to provide guidance on how to execute investigations of digital media manipulation, disinformation and information operations.

We are operating in a complex and rapidly evolving information ecosystem. It requires an equally evolving approach built on testing our assumptions, tracking and anticipating adversaries, and applying the best of open-source investigation and traditional reporting techniques. The vulnerabilities in our digital and data-driven world require journalists to question and scrutinize every aspect of it and apply our skills to help guide the public to accurate, trustworthy information. It also requires journalists to think about how we can unwittingly give oxygen to bad actors and campaigns designed to exploit us, and rush to point fingers at state actors when the evidence does not support it.

The goal of this handbook is to equip journalists with the skills and techniques needed to do this work effectively and responsibly. It also offers basic grounding in the theory, context and mindset that enable journalists to deliver work of high quality that informs the public, exposes bad actors, and helps improve our information environment. But the first thing to understand is that hands-on knowledge and tools are useless unless you approach this work with the right mindset.

This means understanding that everything in the digital environment can be gamed and manipulated, and to recognize the wide variety of people and entities with incentive to do so. The beauty of this environment is that there is often, though not always, a trail of data, interactions, connections and other digital breadcrumbs to follow. And much of it can be publicly available if you know where and how to look.

Investigating the digital means taking nothing at face value. It means understanding that things which appear to be quantifiable and data-driven — likes, shares, retweets, traffic, product reviews, advertising clicks — are easily and often manipulated. It means recognizing that journalists are a key focus of media manipulation and information operations, both in terms of being targeted and attacked, as well as being seen as a key channel to spread mis- and disinformation. And it means equipping yourself and your colleagues with the mindset, techniques and tools necessary to ensure that you're offering trusted, accurate information — and not amplifying falsehoods, manipulated content or troll campaigns.

At the core of the mindset is the digital investigation paradox: By trusting nothing at first, we can engage in work that reveals what we should and should not trust. And it enables us to produce work that the communities we serve are willing and able to trust.

Along with that, there are some fundamentals that you will see emphasized repeatedly in chapters and case studies:

- **Think like an adversary.** Each new feature of a platform or digital service can be exploited in some way. It's critical to put yourself in the shoes of someone looking to manipulate the environment for ideological, political, financial or other reasons. When you look at digital content and messages, you should consider the motivations driving its creation and propagation. It's also essential to stay abreast of the latest techniques being used by bad actors, digital marketers and others whose livelihood relies on finding new ways to gain attention and earn revenue the digital environment.

- **Focus on actors, content, behavior and networks.** The goal is to analyze the actors, content and behavior and how they are to document how they might be working in unison as a network. By comparing and contrasting these four things with each other, you can begin to understand what you're seeing. As you'll see in multiple chapters and case studies, a fundamental approach is to start with one piece of content or an entity such as a website and pivot on it to identify a larger network through behavior and other connections. This can involve examining the flow of content and actors across platforms, and occasionally into different languages.

- **Monitor and collect.** The best way to identify media manipulation and disinformation is to look for it all the time. Ongoing monitoring and tracking of known actors, topics and communities of interest is essential. Keep and organize what you find, whether in spreadsheets, screenshot folders or by using paid tools like Hunchly.

- **Be careful with attribution.** It’s sometimes impossible to say exactly who’s behind a particular account, piece of content, or a larger information operation. One reason is that actors with different motives can behave in similar ways, and produce or amplify the same kind of content. Even the platforms themselves — which have far better access to data and more resources — make attribution mistakes. The most successful and compelling evidence usually combines digital proof with information from inside sources — an ideal mix of online and traditional investigative work. That’s becoming even more difficult as state actors and others evolve and find new ways to hide their fingerprints. Attribution is difficult; getting it wrong will undermine all of the careful work that led up to it.

Finally, a note on the two handbooks that preceded this edition. This work builds on the foundations of the first edition of the Verification Handbook and the Verification Handbook for Investigative Reporting. Each offers fundamental skills for monitoring social media, verifying images, video and social media accounts, and using search engines to identify people, companies and other entities.

Many of the chapters and case studies in this handbook are written with the assumption that readers possess the basic knowledge laid out in these previous publications, particularly the first handbook. If you are struggling to follow along, I encourage you to start with the first handbook.

Now, let’s get to work.

The Age of Information Disorder

Written by: Claire Wardle

***Claire Wardle** leads the strategic direction and research for First Draft, a global nonprofit that supports journalists, academics and technologists working to address challenges relating to trust and truth in the digital age. She has been a Fellow at the Shorenstein Center for Media, Politics and Public Policy at Harvard's Kennedy School, the Research Director at the Tow Center for Digital Journalism at Columbia University's Graduate School of Journalism and head of social media for UNHR, the United Nations Refugee Agency.*

As we all know, lies, rumors and propaganda are not new concepts. Humans have always had the ability to be deceptive, and there are some glorious historical examples of times when fabricated content was used to mislead the public, destabilize governments or send stock markets soaring. What's new now is the ease with which anyone can create compelling false and misleading content, and the speed with which that content can ricochet around the world.

We've always understood that there was complexity in deception. One size does not fit all. For example, a white lie told to keep the peace during a family argument is not the same as a misleading statement by a politician trying to win over more voters. A state-sponsored propaganda campaign is not the same as a conspiracy about the moon landing.

Unfortunately, over the past few years, anything that might fall into the categories described here has been labeled "fake news," a simple term that has taken off globally, often with no need for translation.

I say unfortunate, because it is woefully inadequate to describe the complexity we're seeing. Most content that is deceptive in some way does not even masquerade as news. It is memes, videos, images or coordinated activity on Twitter, YouTube, Facebook or Instagram. And most of it isn't fake; it's misleading or, more frequently, genuine, but used out of context.

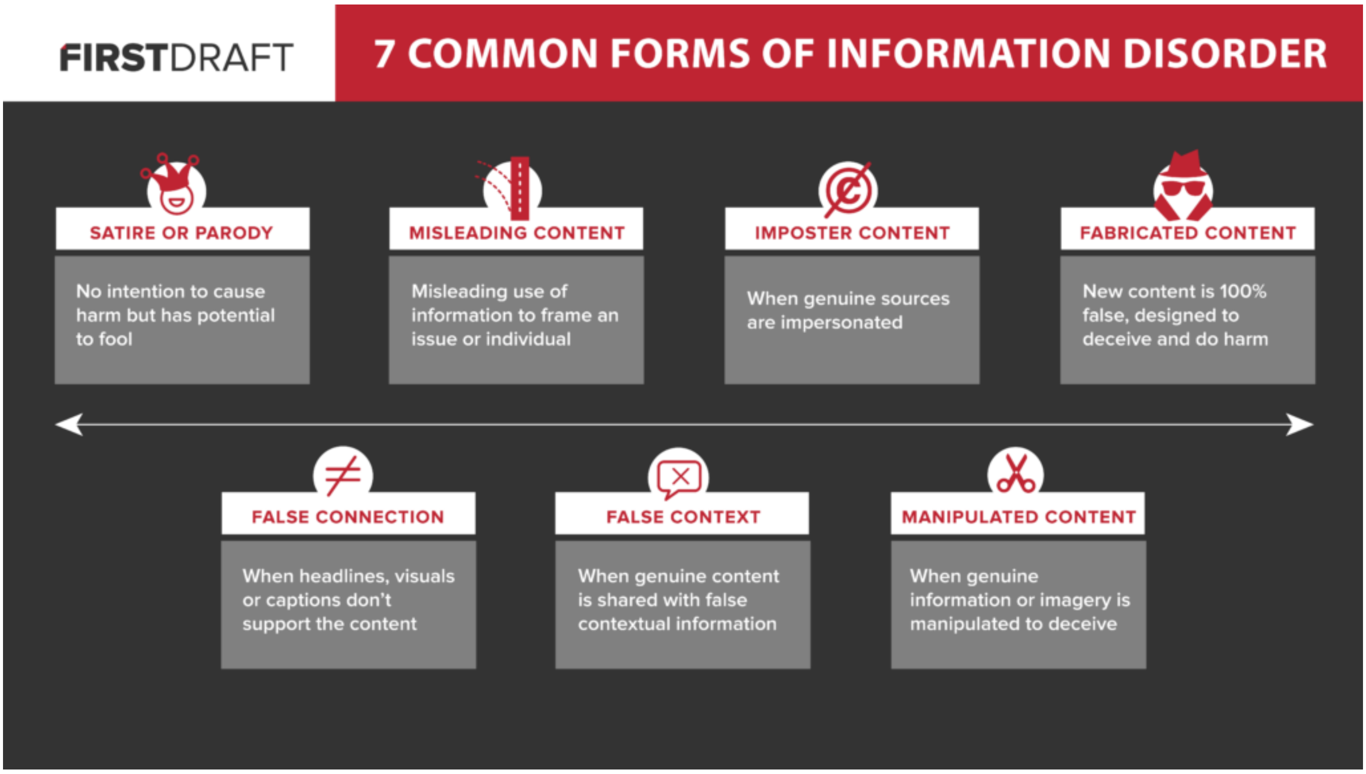
The most impactful disinformation is that which has a kernel of truth to it: taking something that is true and mislabeling it, or sharing something as new when actually it's three years old.

Perhaps most problematic is that the term fake news has been weaponized, mostly by politicians and their supporters to attack the professional news media around the world.

My frustration at the phrase led me to coin the term "information disorder" with my co-author Hossein Derakhshan. We wrote a report in 2017 entitled "Information Disorder," and explored the challenges of the terminology that exists on this topic. In this chapter, I will explain some of the key definitional aspects to understanding this subject, and critically talking about it.

7 Types of Information Disorder

Back in 2017, I created the following typology to underscore the different types of information disorder that exist.



FIRSTDRAFT

7 COMMON FORMS OF INFORMATION DISORDER

SATIRE OR PARODY

No intention to cause harm but has potential to fool

MISLEADING CONTENT

Misleading use of information to frame an issue or individual

IMPOSTER CONTENT

When genuine sources are impersonated

FABRICATED CONTENT

New content is 100% false, designed to deceive and do harm

FALSE CONNECTION

When headlines, visuals or captions don't support the content

FALSE CONTEXT

When genuine content is shared with false contextual information

MANIPULATED CONTENT

When genuine information or imagery is manipulated to deceive

Satire/Parody

Understandably, many people have pushed back against my including satire in this typology, and I certainly struggled with including this category. But unfortunately, agents of disinformation deliberately label content as satire to ensure that it will not be "fact-checked," and as a way of excusing any harm that comes from the content. In an informational ecosystem, where context and cues, or mental shortcuts (heuristics) have been stripped away, satirical content is more likely to confuse the reader. An American might know that The Onion is a satirical site, but did you know that, according to Wikipedia, there are 57 satirical news websites globally? If you don't know the website is satirical, and it's speeding past you on a Facebook feed, it's easy to be fooled.

Recently, Facebook took the decision not to fact-check satire, but those who work in this space know how the satire label is used as a deliberate ploy. In fact, in August 2019, the U.S. debunking organization Snopes wrote a piece about why they fact-check satire. Content purporting to be satire will evade the fact-checkers, and frequently over time, the original context gets lost: people share and re-share not realizing the content is satire and believing that it is true.

False Connection

This is old-fashioned clickbait: the technique of making claims about content via a sensational headline, only to find the headline is horribly disconnected from the actual article or piece of content. While it's easy for the news media to think about the problem of disinformation as being caused by bad actors, I argue that it's important to recognize that poor practices within journalism add to the challenges of information disorder.

Misleading Content

This is something that has always been a problem in journalism and politics. Whether it's the selection of a partial segment from a quote, creating statistics that support a particular claim but don't take into account how the data set was created, or cropping a photo to frame an event in a particular way, these types of misleading practices are certainly not new.

False Context

This is the category where we see the most content: It almost always occurs when genuine imagery is re-shared as new. It often happens during a breaking news event when old imagery is re-shared, but it also happens when old news articles are re-shared as new, when the headline still potentially fits with contemporary events.

Imposter Content

This is when the logo of a well-known brand or name is used alongside false content. This tactic is strategic because it plays on the importance of heuristics. One of the most powerful ways we judge content is if it has been created by an organization or person that we already trust. So by taking a trusted news organization's logo and adding it to a photo or a video, you're automatically increasing the chance that people will trust the content without checking.

Manipulated Content

This is when genuine content is tampered with or doctored in some way. The video of Nancy Pelosi from May 2019 is an example of this. The Speaker of the U.S. House of Representatives was filmed giving a speech. Just a few hours later, a video emerged of her speaking that made her sound drunk. The video had been slowed down, and by doing so, it made it appear like she was slurring her words. This is a powerful tactic, because it's based on genuine footage. If people know she gave that speech with that backdrop, it makes them more trusting of the output.

Fabricated Content

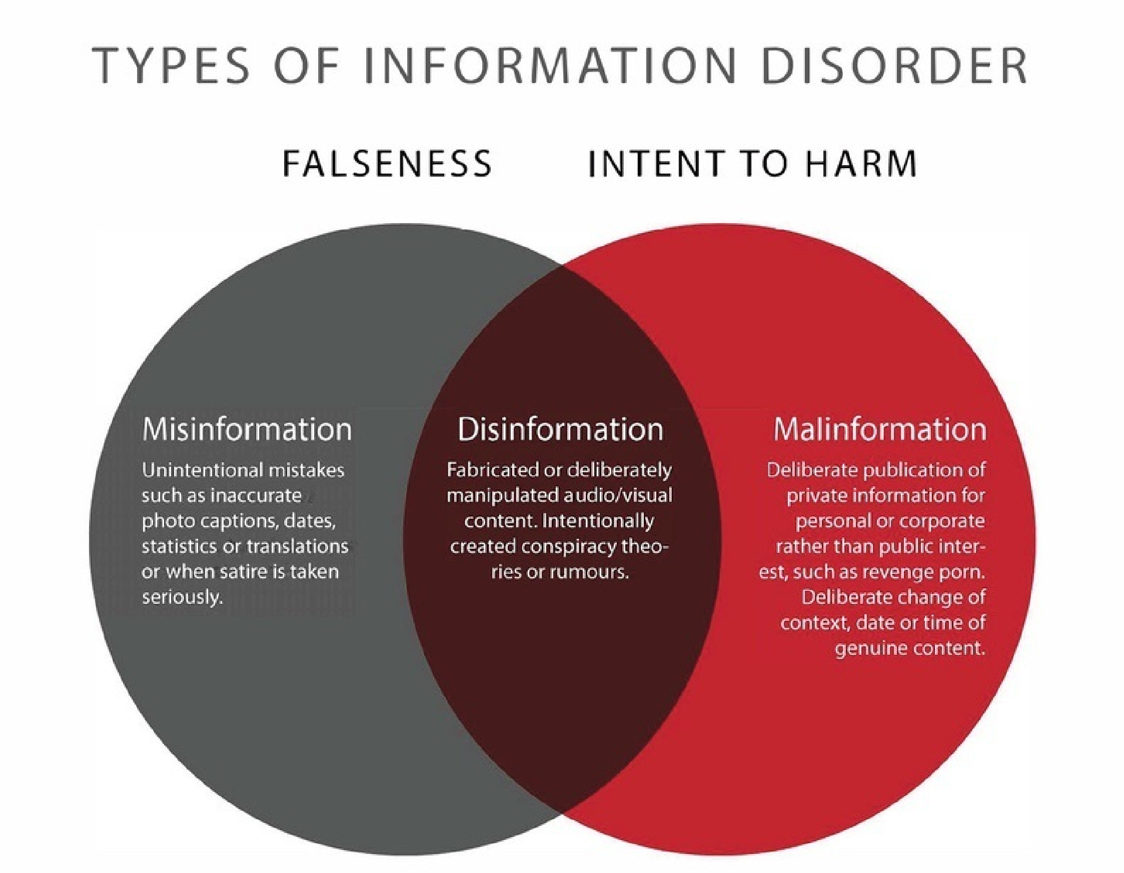
This category is for when content is 100% fabricated. This might be making a completely new fake social media account and spreading new content from it. This category includes deepfakes, where artificial intelligence is used to manufacture a video or audio file in which someone is made to say or do something that they never did.

Understanding Intent and Motivation

These types are useful for explaining the complexity of the polluted information environment, but it doesn't tackle the question of intent. This is a crucial part of understanding this phenomenon.

To do that, Derakhshan and I created this Venn diagram as a way of explaining the difference between misinformation, disinformation and a third term we created, malinformation. Misinformation and disinformation are both examples of false content. But disinformation is created and shared by people who hope to do harm, whether that's financial, reputational, political or physical harm. Misinformation is also false, but people who share the content don't realize it's false. This is often the case during breaking news events when people share rumors or old photos not realizing that they're not connected to the events.

Malinformation is genuine information, but the people who share it are trying to cause harm. The leaking of Hillary Clinton's emails during the 2016 U.S. presidential election is an example of that. So is sharing revenge porn.

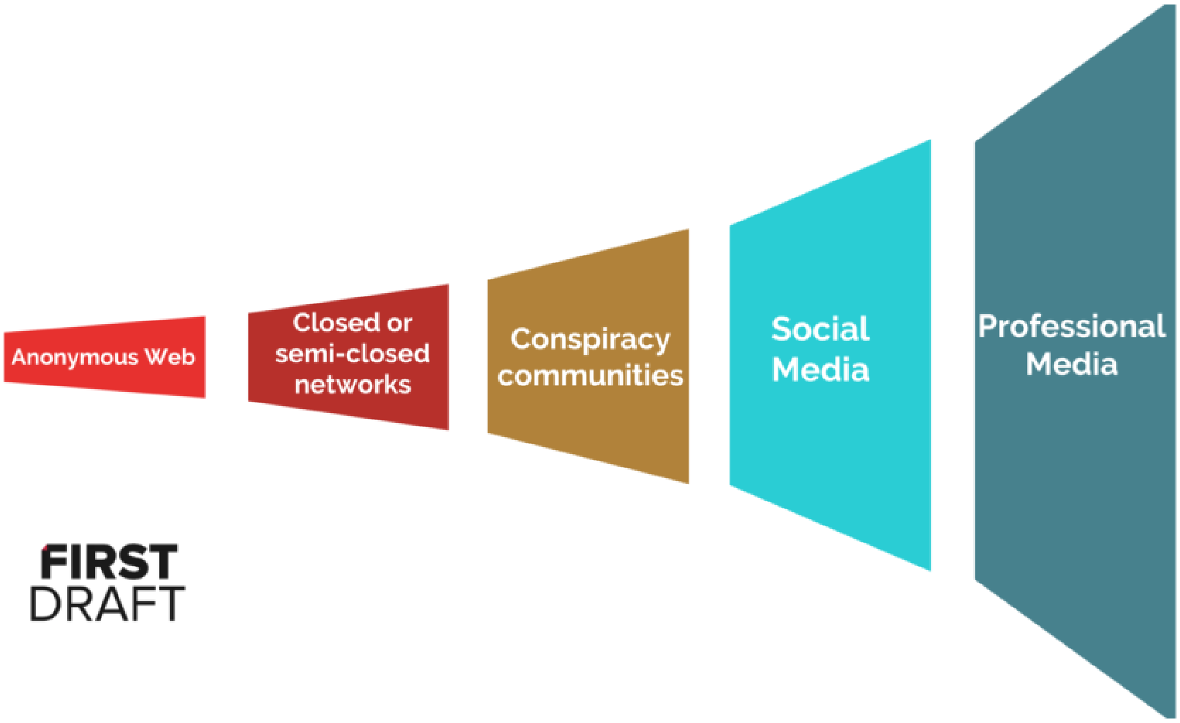

TYPES OF INFORMATION DISORDER
FALSENESS
INTENT TO HARM
Misinformation
Unintentional mistakes such as inaccurate photo captions, dates, statistics or translations or when satire is taken seriously.
Disinformation
Fabricated or deliberately manipulated audio/visual content. Intentionally created conspiracy theories or rumours.
Malinformation
Deliberate publication of private information for personal or corporate rather than public interest, such as revenge porn. Deliberate change of context, date or time of genuine content.

These terms matter, as intent is part of how we should understand a particular piece of information. There are three main motivations for creating false and misleading content. The first is political, whether foreign or domestic politics. It might be a case of a foreign government's attempting to interfere with the election of another country. It might be domestic, where one campaign engages in "dirty" tactics to smear their opponent. The second is financial. It is possible to make money from advertising on your site. If you have a sensational, false article with a hyperbolic headline, as long as you can get people to click on your URL, you can make money. People on both sides of the political spectrum have talked about how they created fabricated "news" sites to drive clicks and therefore revenue. Finally, there are social and psychological factors. Some people are motivated simply by the desire to cause trouble and to see what they can get away with; to see if they can fool journalists, to create an event on Facebook that drives people out on the streets to protest, to bully and harass women. Others end up sharing misinformation, for no other reason than their desire to present a particular identity. For example, someone who says, "I don't care if this isn't true, I just want to underline to my friends on Facebook, how much I hate [insert candidate name]."

The Trumpet of Amplification

To truly understand this wider ecosystem, we need to see how intertwined it all is. Too often, someone sees a piece of misleading or false content somewhere, and believes it was created there. Unfortunately, those who are most effective when it comes to disinformation understand how to take advantage of its fragmented nature.

Remember also, that if rumors, conspiracies or false content weren't shared, they would do no harm. It's the sharing that is so damaging. I therefore created this image, which I call the Trumpet of Amplification, as a way of describing how agents of disinformation use coordination to move information through the ecosystem.



Too often, content is posted in spaces like 4Chan or Discord (an app used by gamers to communicate). These spaces are anonymous and allow people to post without recourse. Often these spaces are used to share specific details about coordination, such as "we're going to try to get this particular hashtag to trend," or "use this meme to respond to today's events on Facebook."

The coordination often then moves into large Twitter DM groups or WhatsApp groups, where nodes within a network spread content to a wider group of people. It might then move into communities on sites like Gab, Reddit or YouTube. From there, the content will often be shared into more mainstream sites like Facebook, Instagram or Twitter.

From there, it will often get picked up by the professional media, either because they don't realize the provenance of the content and decide to use it in their reporting, without sufficient checks, or they decide to debunk the content. Either way, the agents of disinformation see it as a success. Poor headlines where the rumor or misleading claim is reported, or debunks where the false content is embedded in the story, play into the original plan: to drive amplification, to fan the rumor with oxygen.

At First Draft, we talk about the concept of the tipping point. For journalists, reporting on falsehoods too early provides additional and potentially damaging oxygen to a rumor. Reporting too late means it has taken hold and there is little that can be done. Working out that tipping point is challenging. It differs by location, topic and platform.

Conclusion

Language matters. This phenomenon is complex and the words we use makes a difference. We already have academic research that shows that increasingly audiences equate the description "fake news" with poor reporting practices from the professional media.

Describing everything as disinformation, when it might not actually be false content, or is being shared unknowingly by people who don't think is false, are other crucial elements of understanding what is happening.

We live in an age of information disorder. It is creating new challenges for journalists, researchers and information professionals. To report or not to report? How to word headlines? How to debunk videos and images effectively? How to know when to debunk? How does one measure the tipping point? They are all new challenges that exist today for those working in the information environment. It’s complicated.

The Lifecycle of Media Manipulation

Written by: Joan Donovan

Dr. Joan Donovan *is the Research Director at Harvard Kennedy's Shorenstein Center on Media, Politics and Public Policy*

In an age where a handful of powerful global tech platforms have disrupted the traditional means by which society is informed, media manipulation and disinformation campaigns now challenge all political and social institutions. Hoaxes and fabrications are propagated by a mixed group of political operatives, brands, social movements and unaffiliated "trolls" who have developed and refined new techniques to influence public conversation, wreaking havoc on a local, national and global scale. There's widespread agreement that media manipulation and disinformation are important problems facing society. But defining, detecting, documenting and debunking disinformation and media manipulation remains difficult, especially as attacks cross professional sectors such as journalism, law and technology. Therefore, understanding media manipulation as a patterned activity is an essential first step in working to investigate, expose and mitigate them.

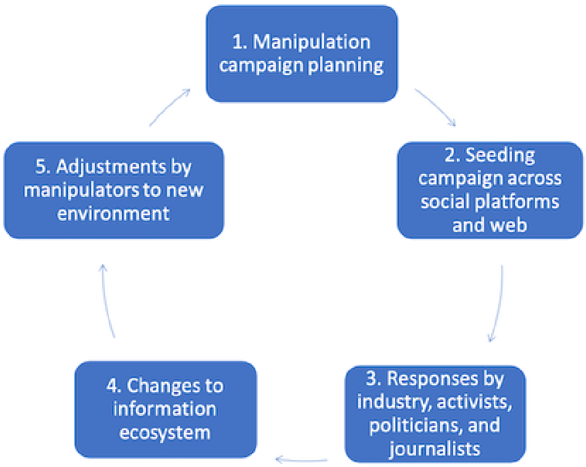
Defining media manipulation and disinformation

To define media manipulation, we first split the term in two parts. In its most general form, *media is an artifact of communication*. Examples include text, images, audio and video in material and digital mediums. When studying media, any relic can be used as recorded evidence of an event. Crucially, media is created by individuals for the purpose of communicating. In this way, media conveys some meaning across individuals, but interpreting that meaning is always relational and situated within a context of distribution.

To claim media is manipulated is to go beyond simply saying that media is fashioned by individuals to transmit some intended meaning. The Merriam-Webster dictionary defines manipulation as "to change by artful or unfair means so as to serve one's purpose." While it can sometimes be difficult to know the exact purpose a single artifact was created to serve, investigators can determine the who, what, where and how of its communication to help determine if manipulative tactics were used as part of the distribution process. Manipulation tactics can include cloaking one's identity or the source of the artifact, editing to conceal or change the meaning or context of an artifact, and tricking algorithms by using artificial coordination, such as bots or spamming tools.

In this context, disinformation is a subgenre of media manipulation, and refers to the creation and distribution of intentionally false information for political ends. Technologists, experts, academics, journalists and policymakers must agree on the distinctive category of disinformation because efforts to fight against disinformation require the cooperation of these groups.

For our part, the Technology and Social Change research team (TaSC) at Harvard Kennedy School's Shorenstein Center is using a case study approach to map the life cycle of media manipulation campaigns. This methodological approach seeks to analyze the order, scale and scope of manipulation campaigns by following media artifacts through space and time, drawing together multiple relationships to sort through the tangled mess. As part of this work, we've developed an overview of the life cycle of a media manipulation campaign, which is useful for journalists as they attempt to identify, track and expose media manipulation and disinformation.

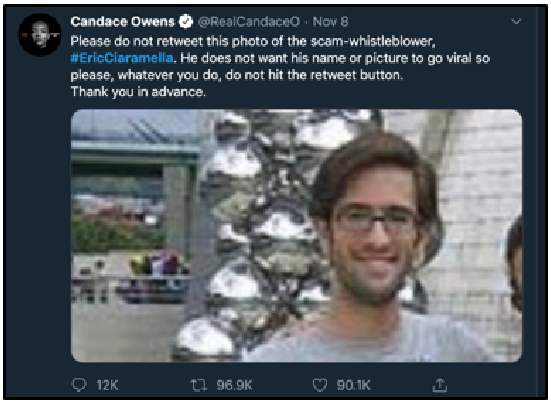

1. Manipulation campaign planning
2. Seeding campaign across social platforms and web
3. Responses by industry, activists, politicians, and journalists
4. Changes to information ecosystem
5. Adjustments by manipulators to new environment

Life Cycle of a Media Manipulation Campaign

The life cycle has five points of action, where the tactics of media manipulators can be documented using qualitative and quantitative methods. Note that most manipulation campaigns are not "discovered" in this order. Instead when researching, look for any one of these points of action and then trace the campaign backward and forward through the life cycle.

Case study: 'Blow the Whistle'

Let's examine the social media activity around the whistleblower complaint made about the activity of President Donald Trump related to Ukraine to see how a media manipulation campaign unfolds, and how ethical action by journalists and platforms early in the life cycle can help thwart manipulation efforts.



Candace Owens @RealCandaceO · Nov 8

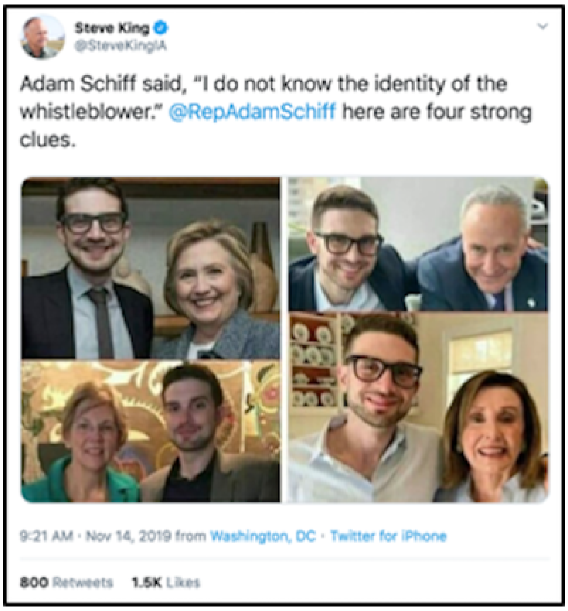
Please do not retweet this photo of the scam-whistleblower, #EricCiaramella. He does not want his name or picture to go viral so please, whatever you do, do not hit the retweet button.
Thank you in advance.

12K 96.9K 90.1K

Planning and Seeding (Stages 1 & 2) — In the conspiracy theory media ecosystem, the whistleblower's identity is already known and his name is circulating on blogs, Twitter, Facebook, YouTube videos and discussion forums. Importantly, unique names can substitute for keywords and hashtags, which function as discrete searchable data points. There was a concerted push to spread the alleged name and the person's photo. Yet, the name seems to be locked in this online media echo chamber of right-wing and conspiracy accounts and entities. Even with this coordinated effort by conspiracy-themed influencers to push the alleged whistleblower's name into the mainstream, they were not able to break out of their own filter bubbles. Why is that?

Responses by journalists, activists etc. (Stage 3) — In contrast, leftist and centrist media did not print the name of the alleged whistleblower or amplify claims that he was outed. Mainstream media outlets refrained from calling attention to the circulation of this person's name in the social media ecosystem, even though it's a newsworthy story for reporters on the tech and politics beat. Those that did cover it often emphasized how the act of circulating this name was an attempt to manipulate the discussion around the whistleblower's complaint, and avoided spreading the name. This is due in large part to the ethics of journalism, where reporters have a special duty to protect the anonymity of sources, which extends to whistleblowers.

Changes to information ecosystem (Stage 4) — While mainstream journalists were omitting his name, the alleged name of the whistleblower, "Eric Ciaramella," is a unique keyword. This meant that people who searched for it could pull up a wide variety of content rooted in the conspiracy-influenced point of view. In addition to ethical journalists effectively turning down a story that could attract significant traffic, each platform company began actively moderating content that used the alleged whistleblower's name as a keyword. YouTube and Facebook removed content that used his name, while Twitter prevented his name from trending. Google's search did allow for his name to be queried and returned thousands of links to conspiracy blogs.



Steve King
@SteveKingIA

Adam Schiff said, "I do not know the identity of the whistleblower." @RepAdamSchiff here are four strong clues.

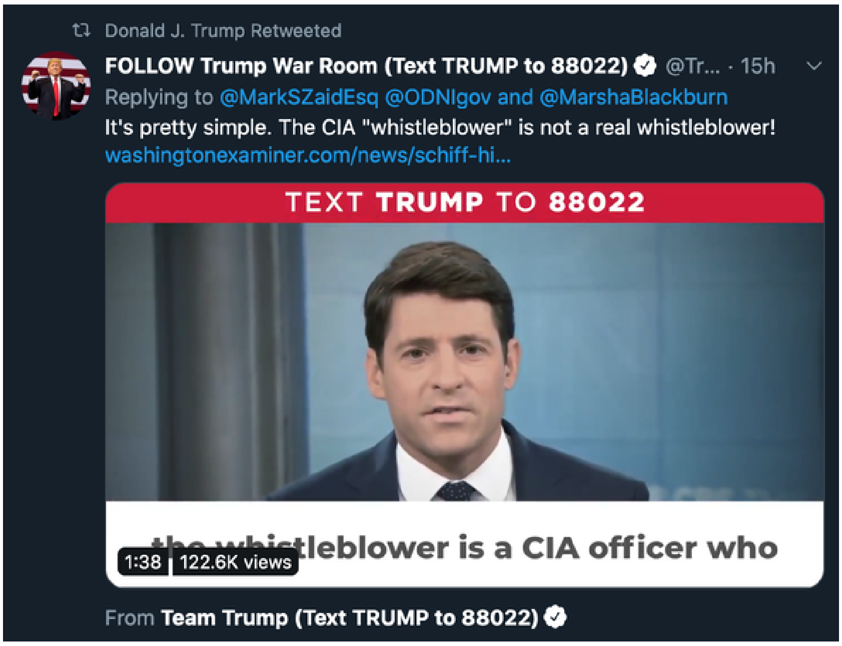
9:21 AM · Nov 14, 2019 from Washington, DC · Twitter for iPhone

800 Retweets 1.5K Likes

Adjustments by manipulators (Stage 5) — Manipulators were aggravated by these attempts to prevent the spread of misinformation and changed their tactics. Instead of pushing content with the alleged whistleblower's name, manipulators began circulating images of a different white man (with glasses and a beard) that resembled the image they previously circulated with his name. These new images were coupled with a "deep state" conspiracy narrative that the whistleblower was a friend of establishment Democrats, and therefore had partisan motives. However, this was an image of Alexander Soros, the son of billionaire investor and philanthropist George Soros, a frequent target of conspiracies.

When that failed to generate media attention, President Trump's Twitter account, @RealDonaldTrump, retweeted an article giving the alleged whistleblower's name, emphasizing that "The CIA whistleblower is not a real whistleblower!" to his 68 million followers. The original tweet came from @TrumpWarRoom, which is his campaign's official and verified account. A cascade of media coverage followed, including many major mainstream outlets, all of which took pains to remove or cover the alleged whistleblower's name. Many people called on social media for the whistleblower to testify in the Senate impeachment hearings, where his name was invoked alongside other important potential witnesses, broadening the possibility that others will stumble on it when searching for other names. And thus begins a new cycle of media manipulation.

Queries for the name of the whistleblower are on the rise and conspiracies abound on blogs about his personal and professional motivations for informing on Trump's activities. Journalists reporting on these tweets oscillate between discussions of witness intimidation, citing that an act like this can deter future whistleblowers, while also tipping into lurid curiosity by reporting on the gossip surrounding Trump's motive for outing the alleged whistleblower. As such, it is laudable that some media organizations are trying to hold elites to account, but the task is impossible without the platform companies' addressing how their products have become useful political tools for media manipulation and spreading disinformation.



Donald J. Trump Retweeted

FOLLOW Trump War Room (Text TRUMP to 88022) @Tr... · 15h

Replying to @MarkSZaidEsq @ODNIgov and @MarshaBlackburn

It's pretty simple. The CIA "whistleblower" is not a real whistleblower!
washingtonexaminer.com/news/schiff-hi...

From Team Trump (Text TRUMP to 88022)

Documenting the life cycle

Media manipulators attempted to "trade up the chain" by seeding a name and photos on social media in order to eventually cause large, legitimate media to amplify it, where platforms would allow it to trend and become easily discoverable. But decisions and actions by platforms and journalists meant the attempt to push the alleged identity of the whistleblower into mainstream consciousness largely failed until a newsworthy figure pushed the issue. While many media organizations strive to abide by ethical guidelines, social media has become a weapon of the already powerful to set media agendas and drive dangerous conspiracies.

Generally speaking though, this case study is a significant improvement over prior efforts to stop the spread of disinformation, where journalists amplified disinformation campaigns as they tried to debunk them, and platform companies felt no duty to provide accurate information to audiences. This overall shift is promising, but accountability for elites is still lacking. For journalists and researchers alike, the stakes of detecting, documenting and debunking media manipulation campaigns are high. In this hyperpartisan moment, any claim to name a disinformation campaign may also bring hordes of trolls and unwanted attention. Grappling with the content and the context of disinformation requires us all to forensically document with rigor how campaigns start, change and end. And to recognize that every perceived ending of a campaign may very well be a new beginning.

1. Investigating Social Media Accounts

Written by: Brandy Zadrozny

Brandy Zadrozny *is an investigative reporter for NBC News, where she mostly covers misinformation, disinformation and extremism on the internet.*

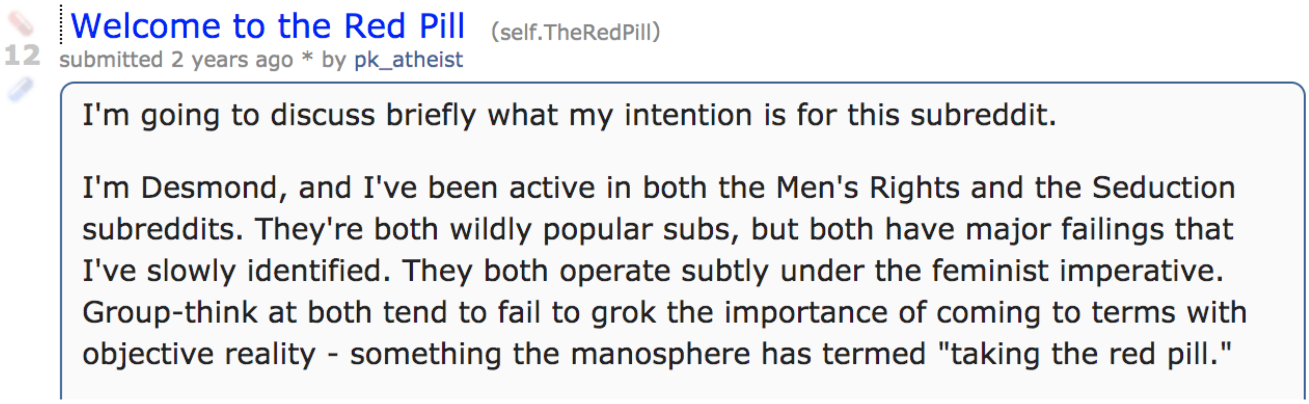
Nearly every story I report involves social media sleuthing. From profile backgrounding to breaking news to longer investigations, social media platforms offer some of the best ways to learn about a subject's real life — their family, friends, jobs, personal politics and associations — as well as a window into secret thoughts and hidden online identities.

It's an incredible time to be a journalist; people increasingly live their lives online and tools to find and search a subject's social profiles are ubiquitous. At the same time, both normal folks and bad actors are getting smarter about hiding their tracks. Meanwhile, social media platforms like Facebook have reacted to negative press about privacy breaches and harmful ideologies spread on their platform by closing down the tools that journalists and researchers have become reliant on to uncover stories and identify people.

In the following chapter I'll show some core approaches for investigating social accounts. The tools are the ones currently in my rotation, but before long they'll be killed by Facebook or replaced by something better. The reporters who are best at this work have their own processes and gadgets to get there, but really, as in any brand of reporting, obsession and (virtual) shoe leather yield the best results. Be prepared to read thousands of tweets, click until the end of the Google results, and dive down a social media rabbit hole if you want to collect the tiny biographical clues that will help you answer the question, "Who is this?"

Usernames

A username is sometimes all we have, which is fine, because it's almost always where we start. Such was the case of a then-New Hampshire Republican state representative who built one of Reddit's most popular and odious men's communities. The investigation behind the unmasking of the architect of Reddit's The Red Pill, now a quarantined community, started with the username "pk_atheist."

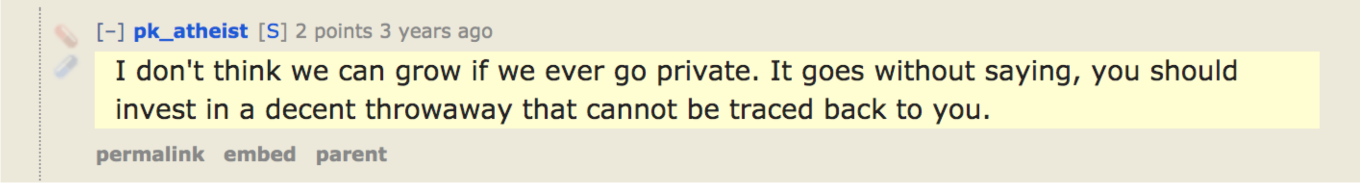


Welcome to the Red Pill (self.TheRedPill)
submitted 2 years ago * by pk_atheist

I'm going to discuss briefly what my intention is for this subreddit.

I'm Desmond, and I've been active in both the Men's Rights and the Seduction subreddits. They're both wildly popular subs, but both have major failings that I've slowly identified. They both operate subtly under the feminist imperative. Group-think at both tend to fail to grok the importance of coming to terms with objective reality - something the manosphere has termed "taking the red pill."

Some people hold on to usernames, using them with minimal variations, across various platforms and email providers. The more security-focused, like the New Hampshire state representative, create and ditch usernames with each new endeavor.



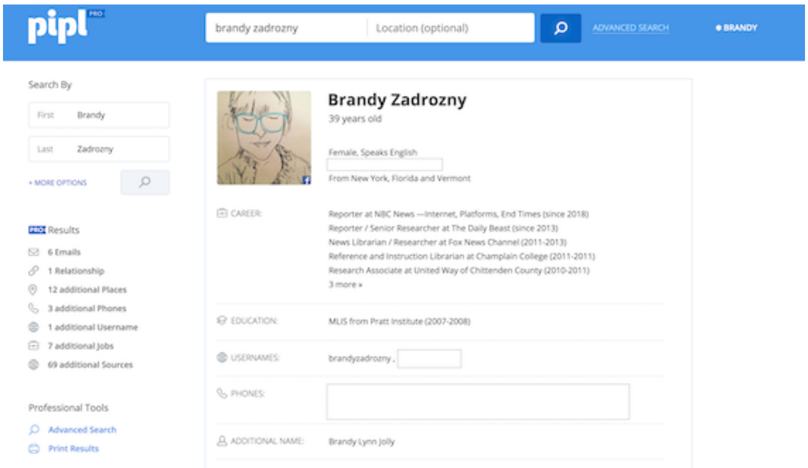
[–] **pk_atheist** [S] 2 points 3 years ago

I don't think we can grow if we ever go private. It goes without saying, you should invest in a decent throwaway that cannot be traced back to you.

Whatever the case, there are a few sites that you should feed the username you're searching into.

First, I plug the username into Google. People — especially younger ones who eschew the larger social platforms — tend to leave a trail even in more unexpected places, including comment sections, reviews and forums, that can lead you to information and other accounts.

Along with a Google search, use proprietary services. They cost money and depending on your newsroom's budget, you may or may not have access. Most shops have Nexis, which is great for public records and court documents but sadly lacking in the email/username department. It's also useful for researching people only in the United States. Pipl and Skopenow are among the best tools I've found for cross referencing "real world" information like phone numbers and property records with online records like emails and usernames, and both work globally. These paid search engines often provide phone and property records, but they can also identify Facebook and LinkedIn profiles that remain even after an account has been closed. They also connect accounts that people have largely forgotten about, such as old blogs and even Amazon wish lists — a gold mine for learning about what a person reads, buys and wants. You also get a lot of false positives with these, so I tend to start my investigation with their results and continue with other means of verification.



brandy zadrozny | Location (optional)

Search By
First Brandy
Last Zadrozny

PRO Results
6 Emails
1 Relationship
12 additional Places
3 additional Phones
1 additional Username
7 additional Jobs
69 additional Sources

Professional Tools
Advanced Search
Print Results

Brandy Zadrozny
39 years old

Female, Speaks English
From New York, Florida and Vermont

CAREER: Reporter at NBC News —Internet, Platforms, End Times (since 2018)
Reporter / Senior Researcher at The Daily Beast (since 2013)
News Librarian / Researcher at Fox News Channel (2011-2013)
Reference and Instruction Librarian at Champlain College (2011-2011)
Research Associate at United Way of Chittenden County (2010-2011)
3 more »

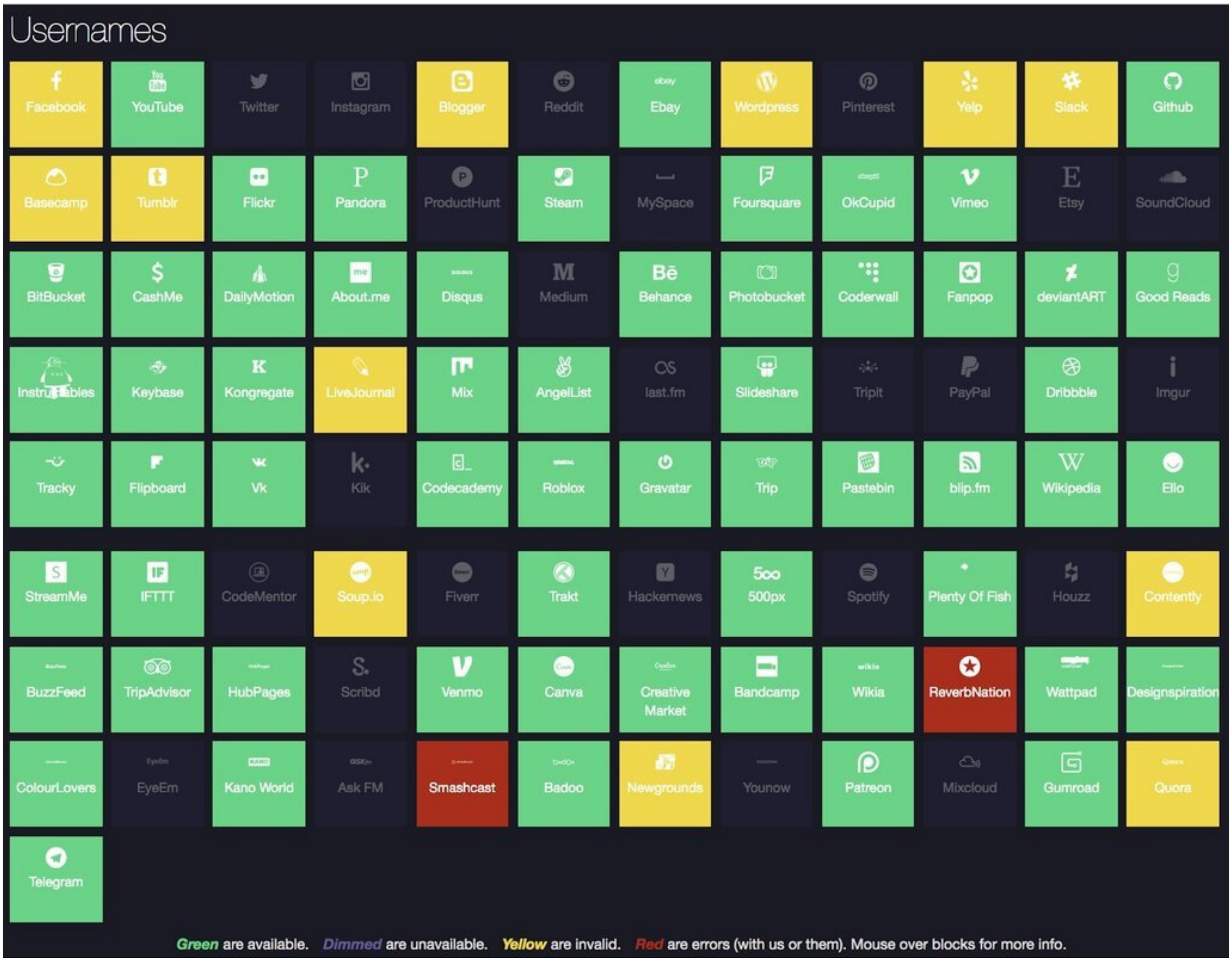
EDUCATION: MLIS from Pratt Institute (2007-2008)

USERNAMES: brandyzadrozny ,

PHONES:

ADDITIONAL NAME: Brandy Lynn Jolly

When I find a username or email I think might belong to my subject, I plug it into an online tool like namechk or namecheckr that looks for username availability across multiple platforms. These tools are designed to be an easy way for marketers to see if a given username they're planning to register is available across platforms. But they're also useful for checking whether a username you're investigating also exists elsewhere. Obviously, just because a username has been registered on multiple platforms doesn't mean these accounts all belong to the same person. But it's a great starting point to looking across platforms.



Usernames

Facebook	YouTube	Twitter	Instagram	Blogger	Reddit	Ebay	Wordpress	Pinterest	Yelp	Slack	Github
Basecamp	Tumblr	Flickr	Pandora	ProductHunt	Steam	MySpace	Foursquare	OkCupid	Vimeo	Etsy	SoundCloud
BitBucket	CashMe	DailyMotion	About.me	Disqus	Medium	Behance	Photobucket	Coderwall	Fanpop	deviantART	Good Reads
Instructables	Keybase	Kongregate	LiveJournal	Mix	AngelList	last.fm	Slideshare	Tripit	PayPal	Dribbble	Imgur
Tracky	Flipboard	Vk	Kik	Codecademy	Roblox	Gravatar	Trip	Pastebin	blip.fm	Wikipedia	Ello
StreamMe	IFTTT	CodeMentor	Soup.io	Fiverr	Trakt	Hackernews	500px	Spotify	Plenty Of Fish	Houzz	Contently
BuzzFeed	TripAdvisor	HubPages	Scribd	Venmo	Canva	Creative Market	Bandcamp	Wikia	ReverbNation	Wattpad	Designspiration
ColourLovers	EyeEm	Kano World	Ask FM	Smashcast	Badoo	Newgrounds	Younow	Patreon	Mixcloud	Gumroad	Quora
Telegram											

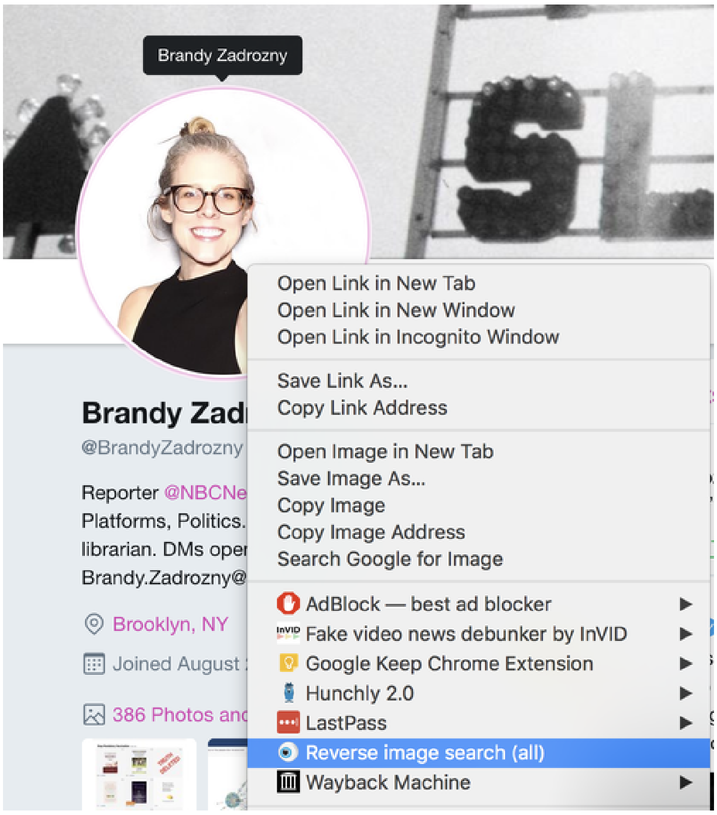
Green are available. *Dimmed* are unavailable. *Yellow* are invalid. *Red* are errors (with us or them). Mouse over blocks for more info.

For further username checking, there's haveibeenpwned.com and Dehashed.com, which search data breaches for user information and can be a quick way to validate an email address and provide new leads.

Photos

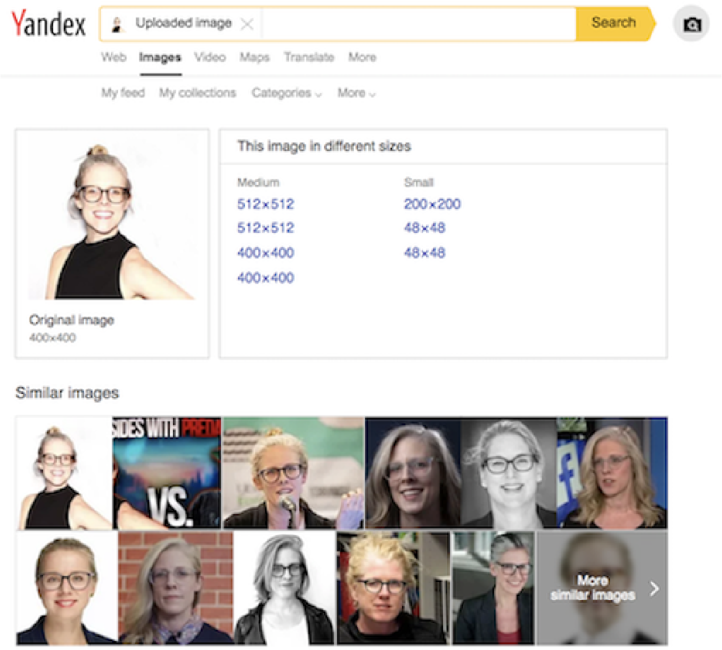
A username isn't always enough to go on, and nothing persuades like a picture. Profile photos are another way to verify the identity of a person across different accounts.

Google's reverse image search is fine, but often other search engines — especially Russia's Yandex — may deliver better results. I use the Reveye Chrome extension, which allows me to right click on an image and search for its match across multiple platforms including Google, Bing, Yandex and Tineye. The Search by Image extension also has a neat capture function that allows you to search from an image within an image.



Brandy Zadı
@BrandyZadrozny

Reporter @NBCNe
Platforms, Politics.
librarian. DMs oper
Brandy.Zadrozny@



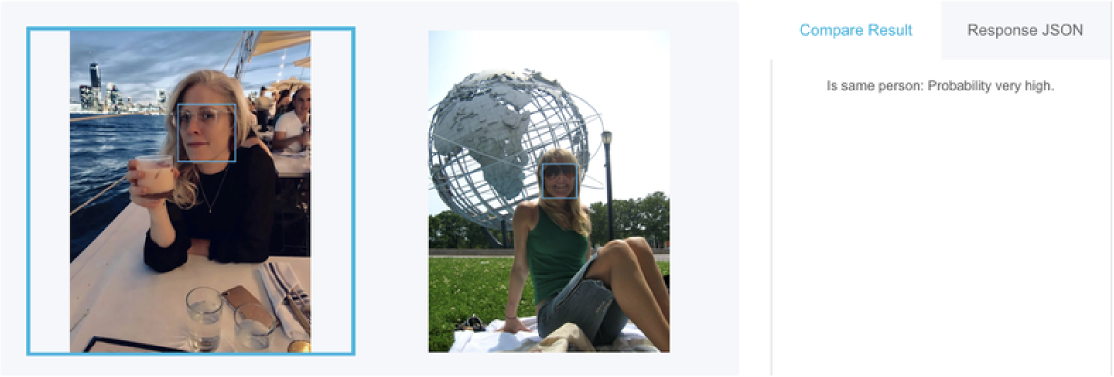
This image in different sizes

Medium	Small
512×512	200×200
512×512	48×48
400×400	48×48
400×400	

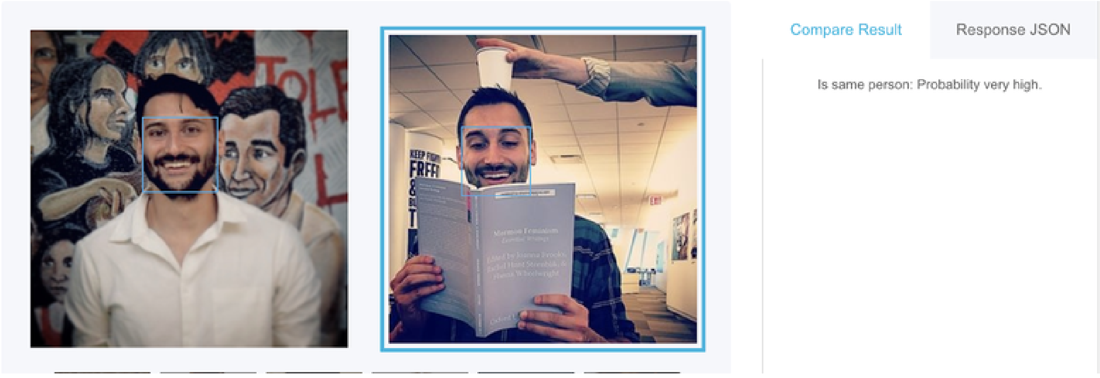
There are problems with reverse image searching, of course. The search engines referenced above do a poor job finding images across Twitter and are all but useless for turning up results from sites like Instagram and Facebook.

What I'm most often looking at are different images of people. I can't count how many times I've squinted at my monitor and asked my colleagues, "Is this the same person?"

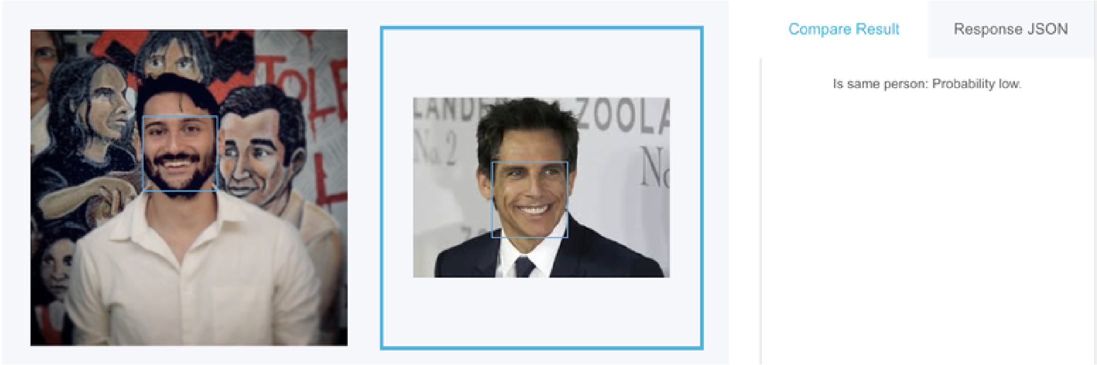
I just don't trust my eyes. Identifying characteristics across photos like moles or facial hair or features is helpful; lately, I also like to check it with a facial recognition tool like Face++, which allows you to upload two photos and then gives a probability that those belong to the same person. In these examples, the tool was able to positively identify me in photos 10 years apart. It also identified my colleague Ben across social media profile pics on Twitter and Facebook while correctly noting that he is not, in fact, Ben Stiller.



Is same person: Probability very high.



Is same person: Probability very high.

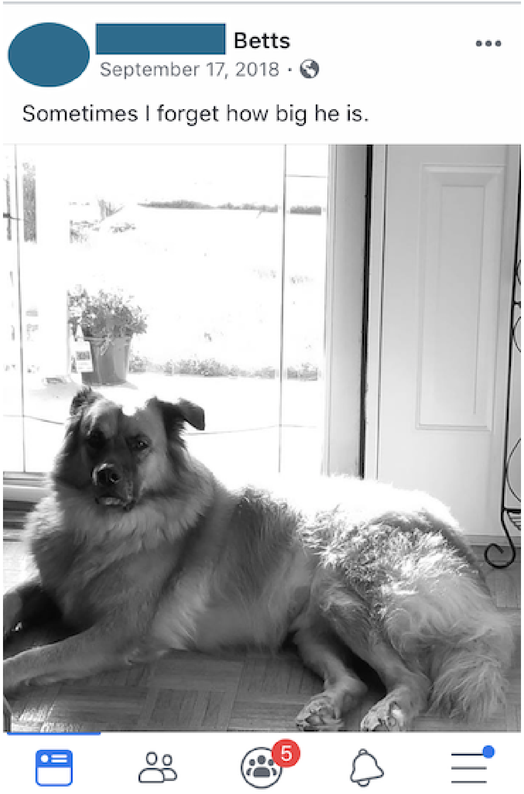


Is same person: Probability low.

If you're chasing trolls or scammers, you might find they've put more effort into obscuring their profile photo, or they may use fake photos. That's when editing the photo and flipping it might help reverse engineer their process.

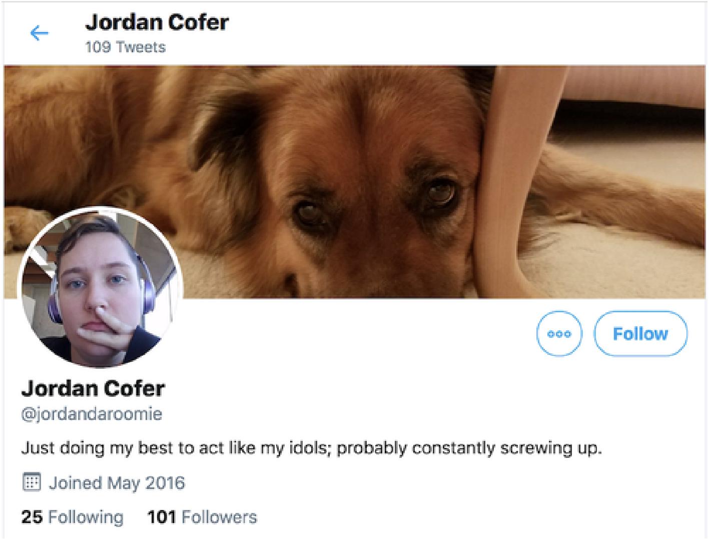
It's not just profile photos that can be signposts, however. As people become more aware of and concerned with their own privacy and that of their family, they're still inclined to share photos of things they're proud of. I've identified people by connecting photos of things like cars, homes or pets. In this sense, photos become a means to connect accounts and the people behind them to one another, enabling you to build out the network around your target. This is a core practice when investigating social media accounts.

For example, we were looking to confirm the social accounts of a man who shot and killed nine people outside of a bar in Dayton, Ohio. His Twitter account offered clues to his political ideology but his handle, @iamthespookster, was unique and didn't resemble his real name, which had been released by authorities. The fact that one of his victims was his sibling, a transgender man whose name was not in public records and hadn't come out to the wider world yet, further complicated identifying the key figures. But throughout his and his family's profiles were pictures of a dog, a pet that appeared as the banner image of his transgender brother's unreported account.



[redacted] Betts
September 17, 2018 ·

Sometimes I forget how big he is.



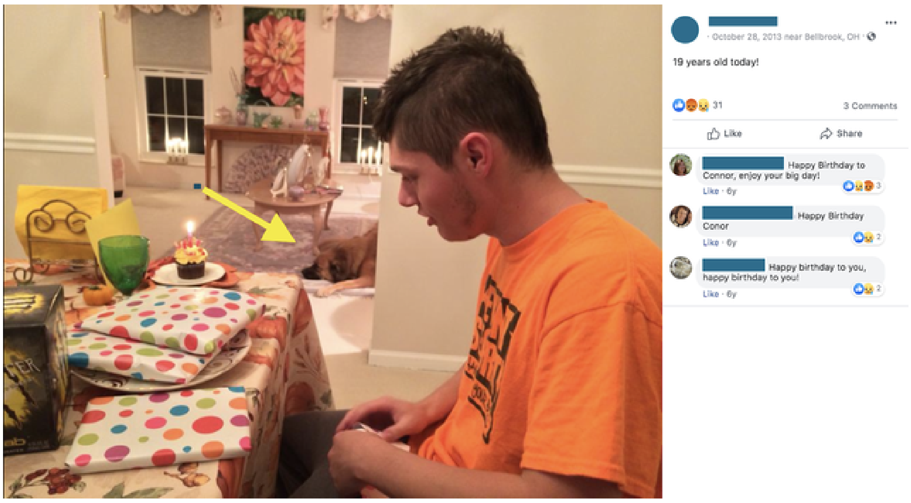
Jordan Cofer
109 Tweets

Jordan Cofer
@jordandaroomie

Just doing my best to act like my idols; probably constantly screwing up.

Joined May 2016

25 Following 101 Followers



[redacted]
· October 28, 2013 near Bellbrook, OH ·

19 years old today!

31 3 Comments

[redacted] Happy Birthday to Connor, enjoy your big day!
Like · 6y 3

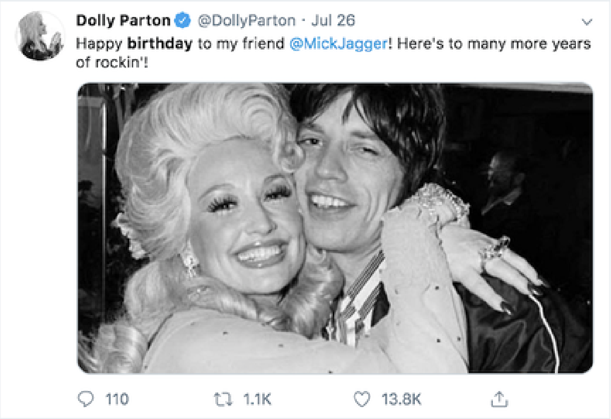
[redacted] Happy Birthday Conor
Like · 6y 2

[redacted] Happy birthday to you, happy birthday to you!
Like · 6y 2

The dog wasn't the only helpful detail in the previous image. That image came from the Ohio shooter's father, and helped us verify his personal accounts and those belonging to his family.

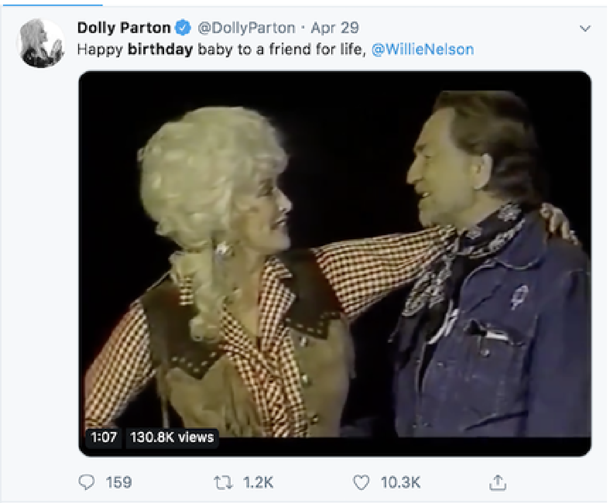
If you have an account on Facebook or Twitter, I can probably tell you the day you were born, even if you don't share it on your profile or post about it yourself. Since a date of birth is often one of the first identifying pieces of police-provided information in breaking news situations, a reliable way to verify a social media account is by scrolling to the month and day in question on a suspected account and looking out for birthday wishes. Even if their own pages are empty, often moms and dads (like Connor Betts' above) will post about their children's birthdays.

The same is true for Twitter, because who doesn't love a birthday?



Dolly Parton @DollyParton · Jul 26

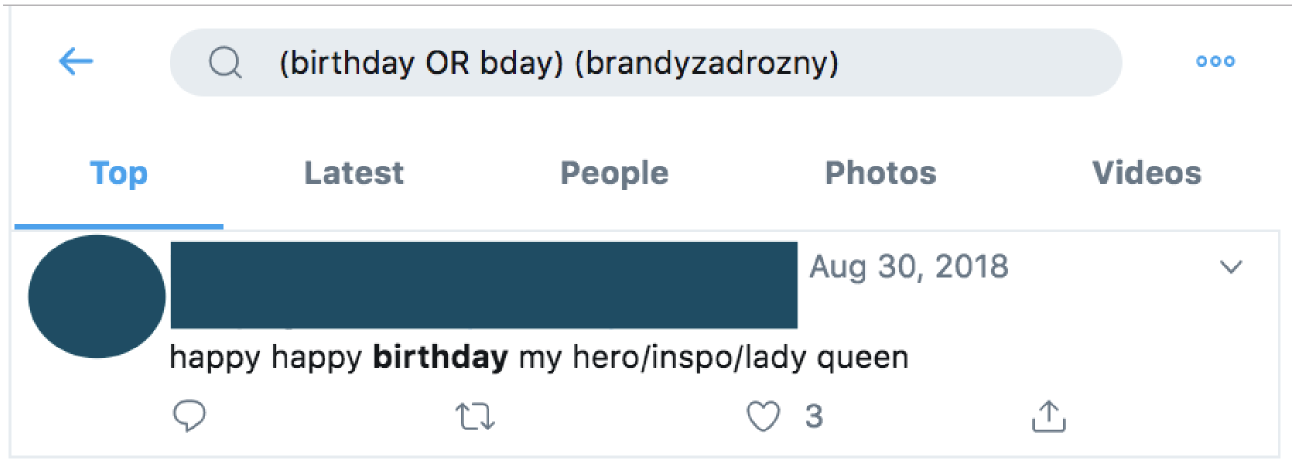
Happy **birthday** to my friend @MickJagger! Here's to many more years of rockin'!



Dolly Parton @DollyParton · Apr 29

Happy **birthday** baby to a friend for life, @WillieNelson

But it's even easier to find an identifying post on Twitter, because its advanced search tool is among the best offered by social platforms. Although I rarely announce my birthday, if ever, I was able to find a birthday tweet from a loving colleague who outed me.



(birthday OR bday) (brandyzadrozny)

[redacted] [redacted] Aug 30, 2018

happy happy **birthday** my hero/inspo/lady queen

Birthdays are just one example. Weddings, funerals, holidays, anniversaries, graduations — nearly every major life marker is celebrated on social media. These provide an opening for searching and investigating an account.

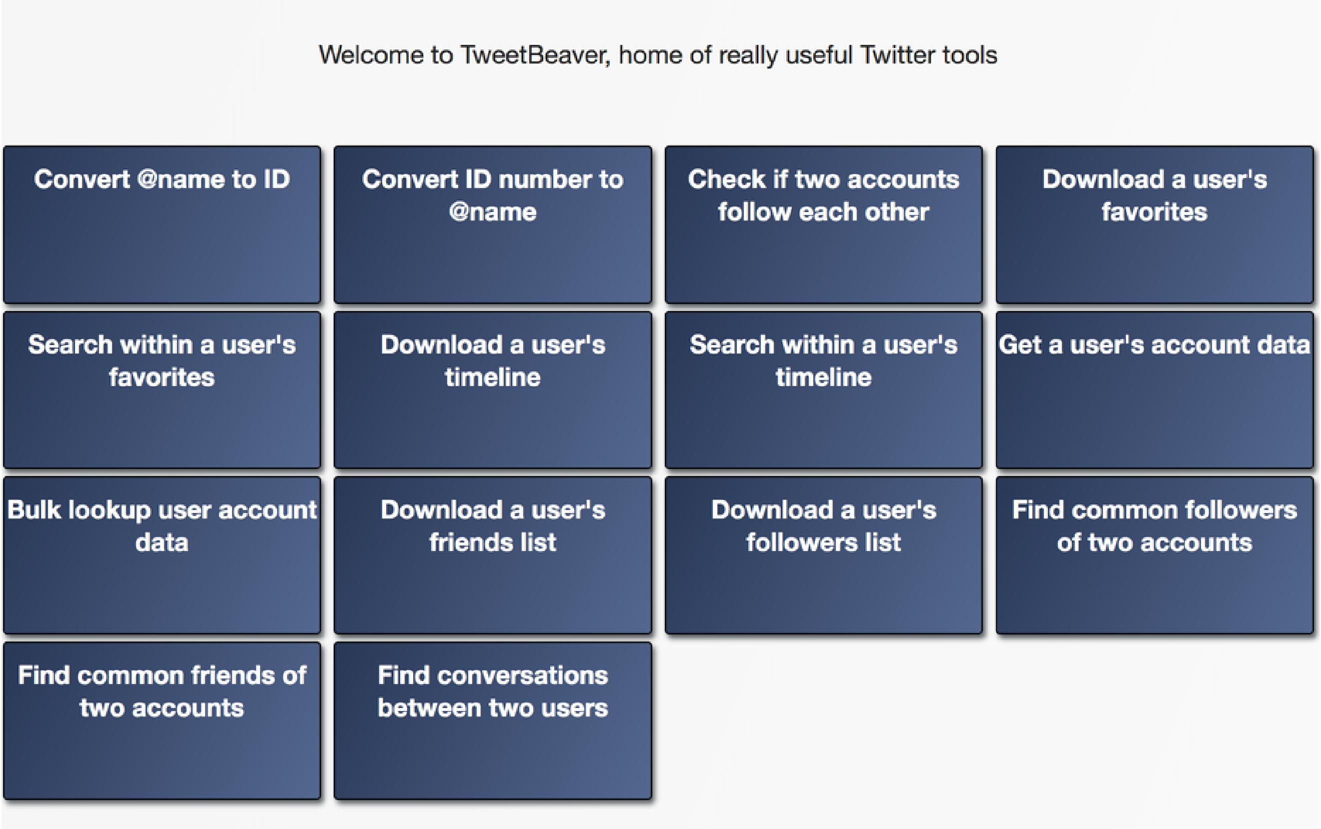
You can search for these keywords and by other filters with Facebook search tools. They don't get as much mileage as they did before the platform's pivot to privacy, but they exist. One of my favorites is whopostedwhat.com.

Relationships

You can judge a person by the company they keep on social media. We can tell a lot about a person's life and leanings by examining the people with whom they interact online.

When I first joined Twitter, I made my husband and best friend sign up too, just so they could follow me. I think about that when I'm looking into accounts for work. The platforms don't want you to be alone, either, so when you first open an account, an algorithm powers up. Influenced by the contacts list in your phone, your appearance in the contact lists of existing accounts, your location and other factors, a platform will suggest accounts to follow.

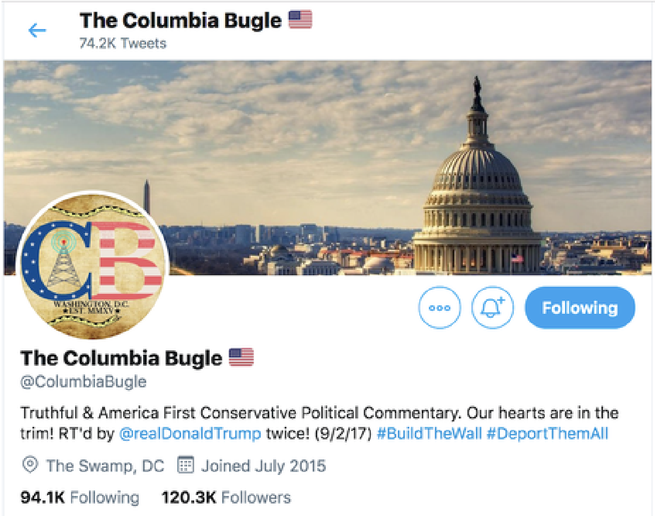
Because of that truth, it's always illuminating to look at an account's earliest followers and friends. TweetBeaver is a good tool for investigating the connections between large accounts and for downloading things like timelines and favorites of smaller accounts. For larger datasets, I rely on a developer with API access.



Welcome to TweetBeaver, home of really useful Twitter tools

Convert @name to ID
Convert ID number to @name
Check if two accounts follow each other
Download a user's favorites
Search within a user's favorites
Download a user's timeline
Search within a user's timeline
Get a user's account data
Bulk lookup user account data
Download a user's friends list
Download a user's followers list
Find common followers of two accounts
Find common friends of two accounts
Find conversations between two users

Let's take The Columbia Bugle, a popular far-right anonymous Twitter account that boasts that it was retweeted twice by Donald Trump's account.



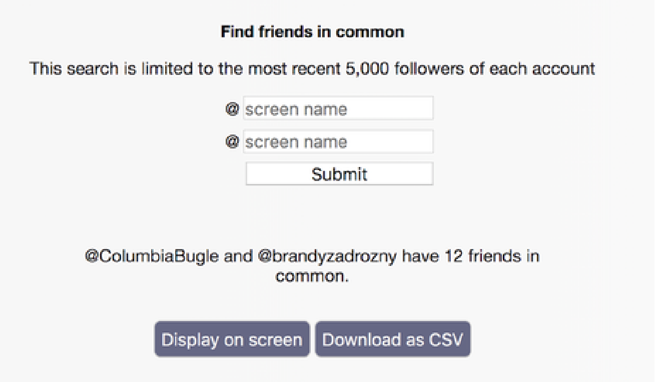
The Columbia Bugle 🇺🇸
74.2K Tweets

The Columbia Bugle 🇺🇸
@ColumbiaBugle

Truthful & America First Conservative Political Commentary. Our hearts are in the trim! RT'd by @realDonaldTrump twice! (9/2/17) #BuildTheWall #DeportThemAll

The Swamp, DC Joined July 2015

94.1K Following 120.3K Followers



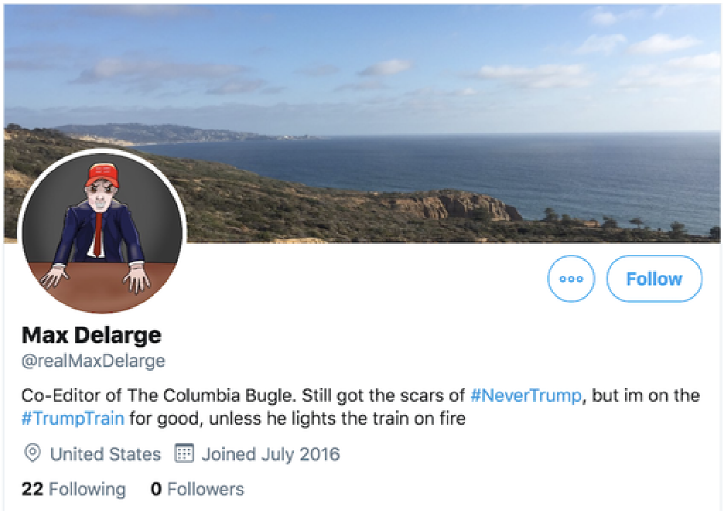
Find friends in common

This search is limited to the most recent 5,000 followers of each account

@ screen name
@ screen name
Submit

@ColumbiaBugle and @brandyzadrozny have 12 friends in common.

The earliest follows of Max Delarge, an account claiming to be the editor of The Columbia Bugle, are San Diego-specific news sources and San Diego-specific sports accounts. Since many of Columbia Bugle's tweets include videos from San Diego Trump rallies and events at the University of California, San Diego, we can be fairly confident that the person behind the account lives near San Diego.

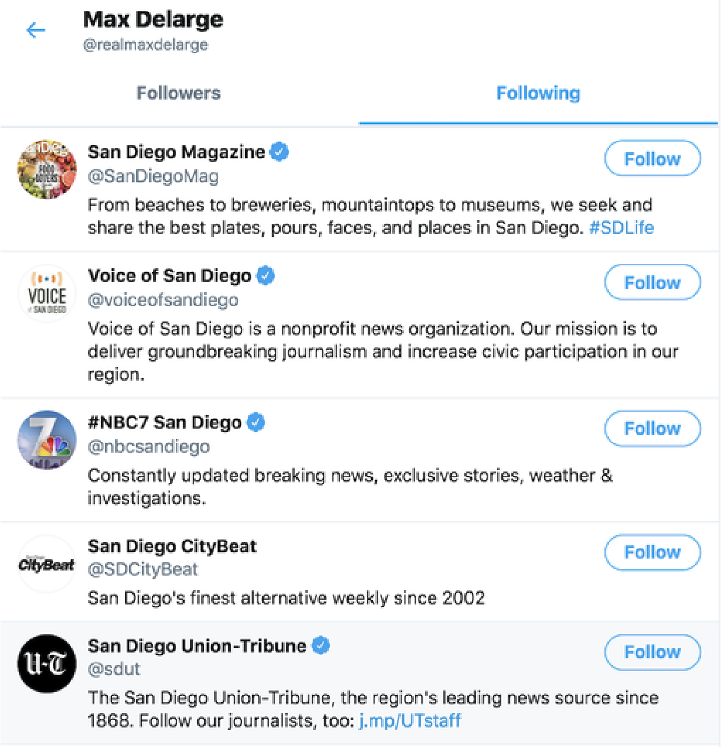


Max Delarge
@realMaxDelarge

Co-Editor of The Columbia Bugle. Still got the scars of #NeverTrump, but im on the #TrumpTrain for good, unless he lights the train on fire

United States Joined July 2016

22 Following 0 Followers



Max Delarge
@realmaxdelarge

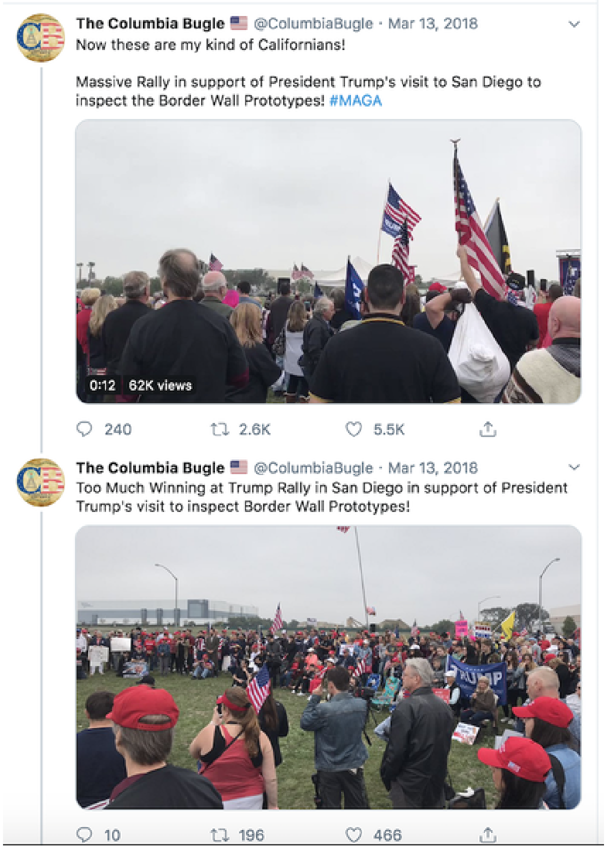
San Diego Magazine @SanDiegoMag
From beaches to breweries, mountaintops to museums, we seek and share the best plates, pours, faces, and places in San Diego. #SDLife

Voice of San Diego @voiceofsandiego
Voice of San Diego is a nonprofit news organization. Our mission is to deliver groundbreaking journalism and increase civic participation in our region.

#NBC7 San Diego @nbcsandiego
Constantly updated breaking news, exclusive stories, weather & investigations.

San Diego CityBeat @SDCityBeat
San Diego's finest alternative weekly since 2002

San Diego Union-Tribune @sdut
The San Diego Union-Tribune, the region's leading news source since 1868. Follow our journalists, too: j.mp/UTstaff

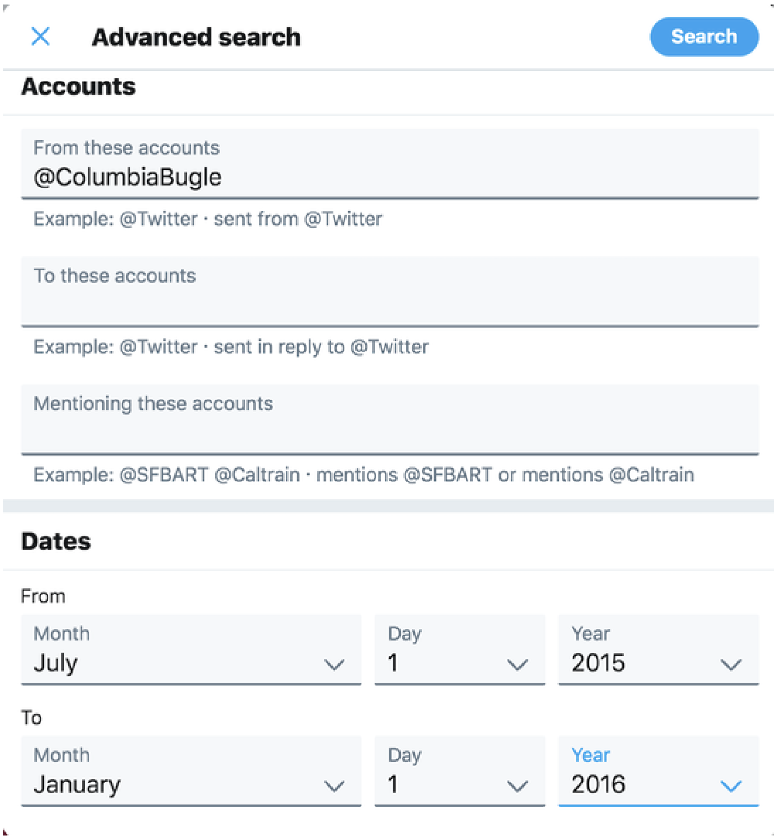


The Columbia Bugle 🇺🇸 @ColumbiaBugle · Mar 13, 2018
Now these are my kind of Californians!

Massive Rally in support of President Trump's visit to San Diego to inspect the Border Wall Prototypes! #MAGA

The Columbia Bugle 🇺🇸 @ColumbiaBugle · Mar 13, 2018
Too Much Winning at Trump Rally in San Diego in support of President Trump's visit to inspect Border Wall Prototypes!

With a new investigation, I like to start at the beginning of someone's Twitter history and work forward in time. You can get there by hand, with an assist from an autoscroller chrome extension, or you can use Twitter's advanced search to limit the time frame to the first few months of an account's existence.



Advanced search

Search

Accounts

From these accounts
@ColumbiaBugle

Example: @Twitter · sent from @Twitter

To these accounts

Example: @Twitter · sent in reply to @Twitter

Mentioning these accounts

Example: @SFBART @Caltrain · mentions @SFBART or mentions @Caltrain

Dates

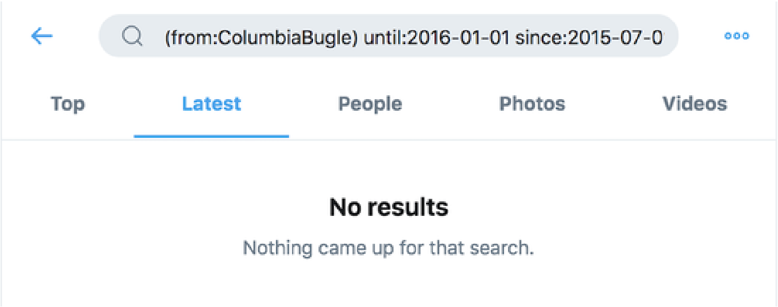
From

Month: July
Day: 1
Year: 2015

To

Month: January
Day: 1
Year: 2016

Curiously, the first six months of this account shows zero tweets.

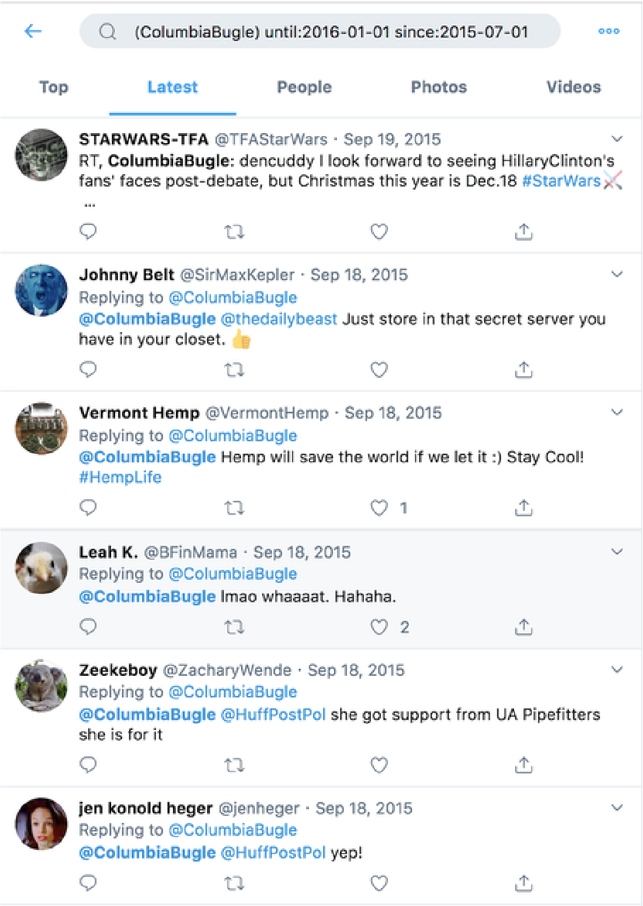


(from:ColumbiaBugle) until:2016-01-01 since:2015-07-0

No results

Nothing came up for that search.

This suggests that the person behind The Columbia Bugle might have deleted his earlier tweets. To find out why that might be, I can tweak my search. Instead of tweets *from* the account I'll look for any tweets *mentioning* The Columbia Bugle.



(ColumbiaBugle) until:2016-01-01 since:2015-07-01

STARWARS-TFA @TFAStarWars · Sep 19, 2015
RT, **ColumbiaBugle:** dencuddy I look forward to seeing HillaryClinton's fans' faces post-debate, but Christmas this year is Dec.18 #StarWars⚔ ...

Johnny Belt @SirMaxKepler · Sep 18, 2015
Replying to @ColumbiaBugle
@ColumbiaBugle @thedailybeast Just store in that secret server you have in your closet. 👍

Vermont Hemp @VermontHemp · Sep 18, 2015
Replying to @ColumbiaBugle
@ColumbiaBugle Hemp will save the world if we let it :) Stay Cool! #HempLife
♡ 1

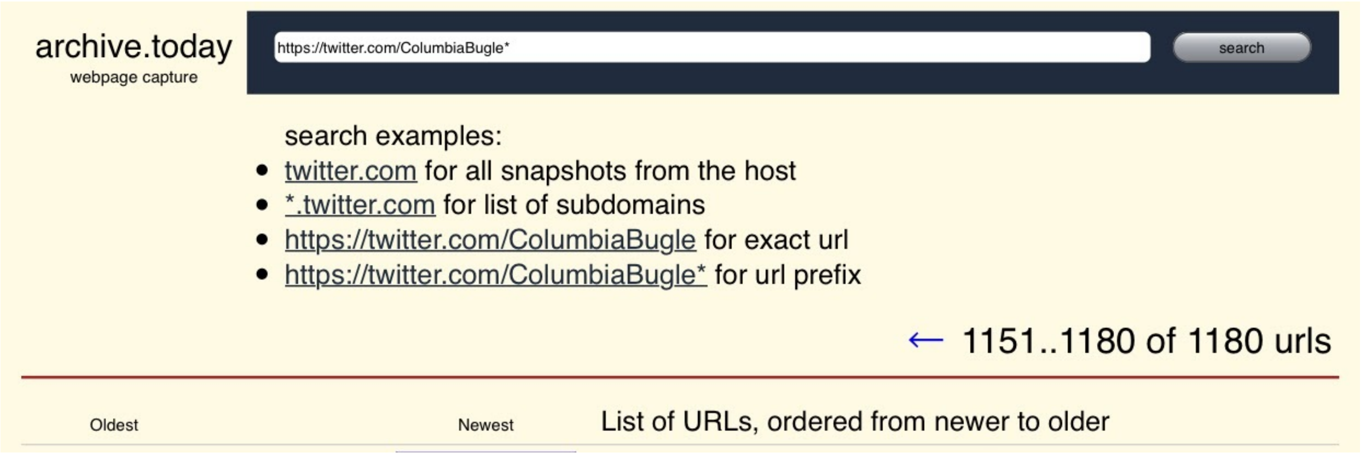
Leah K. @BFinMama · Sep 18, 2015
Replying to @ColumbiaBugle
@ColumbiaBugle lmao whaaaat. Hahaha.
♡ 2

Zeekeboy @ZacharyWende · Sep 18, 2015
Replying to @ColumbiaBugle
@ColumbiaBugle @HuffPostPol she got support from UA Pipefitters she is for it

jen konold heger @jenheger · Sep 18, 2015
Replying to @ColumbiaBugle
@ColumbiaBugle @HuffPostPol yep!

These conversations confirm that ColumbiaBugle erased its first year of tweets, but doesn't tell us why and the first accounts that the account interacted with don't offer many clues.

To find recently deleted tweets, you can search Google's cache; older deleted tweets can also sometimes be accessed in the Internet Archive's Wayback Machine or another archive. The manual archive site archive.is turns up several deleted tweets from where ColumbiaBugle participated in an event where college students wrote pro-Trump messages on their campuses. To see all the tweets someone may have archived from that account, as I did to find this tweet, you can search by URL prefix, using an asterisk after the account name like this:



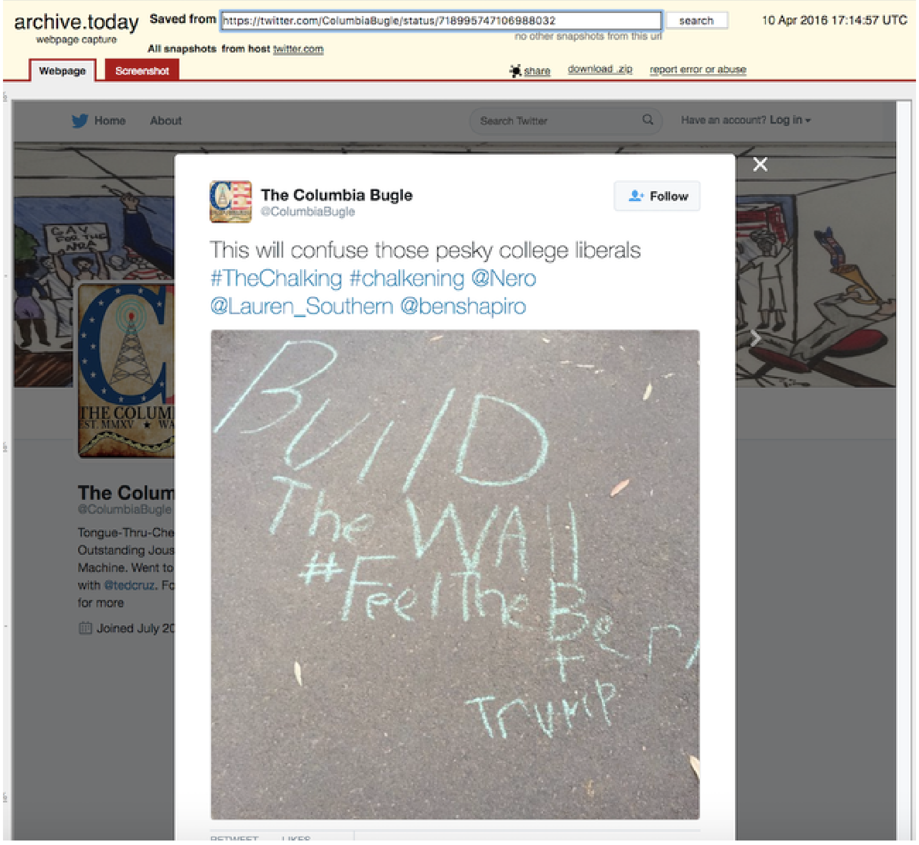
https://twitter.com/ColumbiaBugle*

search examples:

- twitter.com for all snapshots from the host
- *.twitter.com for list of subdomains
- https://twitter.com/ColumbiaBugle for exact url
- https://twitter.com/ColumbiaBugle* for url prefix

← 1151..1180 of 1180 urls

List of URLs, ordered from newer to older



archive.today webpage capture Saved from https://twitter.com/ColumbiaBugle/status/718995747106988032 search 10 Apr 2016 17:14:57 UTC

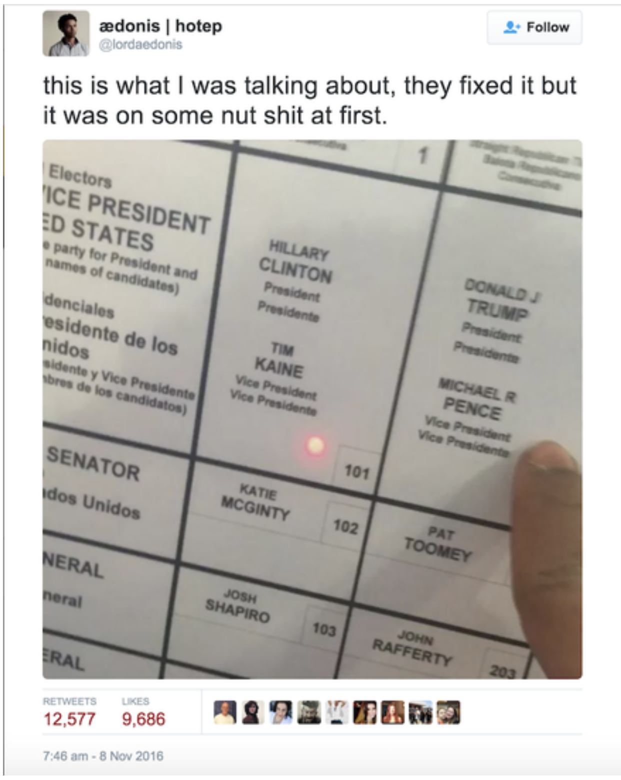
no other snapshots from this url

All snapshots from host twitter.com

The Columbia Bugle @ColumbiaBugle

This will confuse those pesky college liberals #TheChalking #chalkening @Nero @Lauren_Southern @benshapiro

It's rare for someone to successfully keep their real life separate from their online activities. For example, my NBC News colleague and I told the story of 2016's most viral — and misleading — Election Day voter fraud claim, with an assist from a neighborhood acquaintance of the far-right troll who tweeted it.



ædonis | hotep @lordaedonis

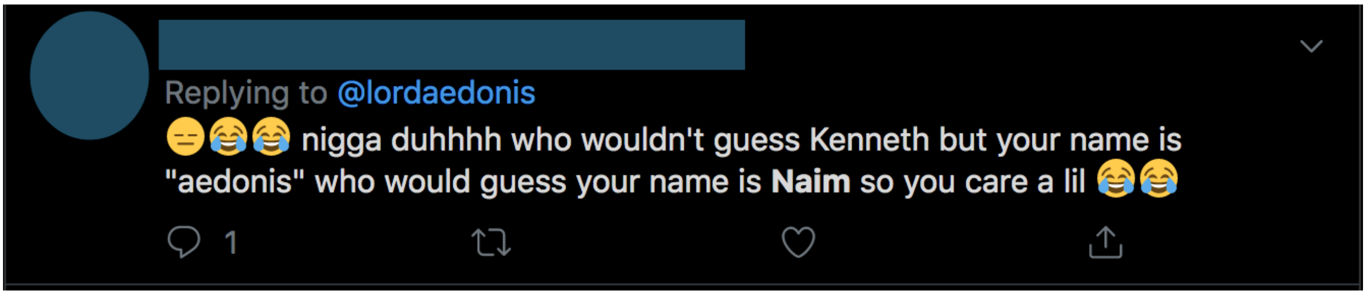
this is what I was talking about, they fixed it but it was on some nut shit at first.

RETWEETS 12,577 LIKES 9,686

7:46 am - 8 Nov 2016

Follow
HILLARY CLINTON President Presidente
TIM KAINE Vice President Vice Presidente
DONALD J. TRUMP President Presidente
MICHAEL R. PENCE Vice President Vice Presidente
SENATOR
KATIE MCGINTY 102
PAT TOOMEY
JOSH SHAPIRO 103
JOHN RAFFERTY 203

Though the tweet originated with a man known to his followers as @lordaedonis, people from his actual neighborhood had responded to past tweets with his real name, which we included in a profile of an attention-hungry entrepreneur whose tweet was spread by a Kremlin-backed Twitter account, and eventually seen by millions and promoted by the soon-to-be president.



[redacted]

Replying to @lordaedonis

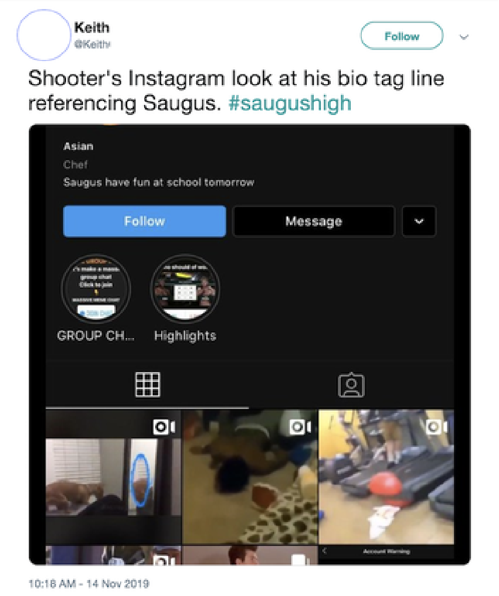
😑😂😂 nigga duhhhh who wouldn't guess Kenneth but your name is "aedonis" who would guess your name is **Naim** so you care a lil 😂😂

My favorite kind of stories are those that reveal the real people behind influential, anonymous social media accounts. These secret accounts are less reliant on the algorithm, and more carefully crafted to be an escape from public life. They allow someone to keep tabs on and communicate with family and friends apart from their public account, or to communicate the ideas and opinions that for personal or political reasons, they dare not say out loud.

Journalist Ashley Feinberg is the fairy godmother of these kind of juicy stories, ones that unmask the alt accounts of prominent figures like James Comey or Mitt Romney. Her secret was simply a matter of finding smaller accounts of family members that Comey and Romney would naturally want to follow, and then scrolling through them until she found an account that seemed inauthentic but whose content and friends/followers network matched that of these real people.

Be wary of fake accounts

Each platform has its own personality, search capabilities and usefulness in different news situations. But a word of caution with social media accounts: The same rule of trust but verify applies. Groups of people revel in tricking journalists. Especially in breaking news situations, fake accounts will always be born, many with ominous or threatening posts meant to attract reporters. This fake Instagram account used the name of a mass shooter and was created after a shooting at Saugus High School in California. It gained attention via screenshots on Twitter, but BuzzFeed News later revealed it did not belong to the shooter.



Keith
@Keith

Shooter's Instagram look at his bio tag line referencing Saugus. #saugushigh

Asian
Chef
Saugus have fun at school tomorrow

10:18 AM - 14 Nov 2019

Confirming a social account with the subject, family and friends, law enforcement and/or social media PR are ways to protect yourself from being duped.

Finally, and perhaps the most important note: There’s no one right order in which to complete these steps. Often, I’m led down rabbit holes and have more tabs open than I’m proud of. Creating a system that you can replicate — whether it’s tracking your steps in a Google doc or letting a paid tool like Hunchly monitor as you search — is the key to clarifying connections between people and the lives they lead online, and turning those conclusions into stories.

1a. Case Study: How investigating a set of Facebook accounts revealed a coordinated effort to spread propaganda in the Philippines

Written by: Vernise Tantuco

and Gemma Bagayaua-Mendoza

A professional journalist for roughly 20 years, ***Gemma Bagayaua-Mendoza*** *is the head of research and strategy at Rappler. She leads the fact-check unit as well as Rappler's research into online disinformation and misinformation.*

Vernise Tantuco *is a member of Rappler's research team, where she works on fact checks and studies disinformation networks in the Philippines.*

In the fall of 2016, John Victorino, an investment analyst, sent Rappler a list of what he said was 26 suspicious Facebook accounts from the Philippines. We began investigating and monitoring the accounts, and quickly found the details listed in their profiles were false. Over the course of weeks of investigation, these 26 accounts led us to uncover a much more extensive network of pages, groups and accounts.

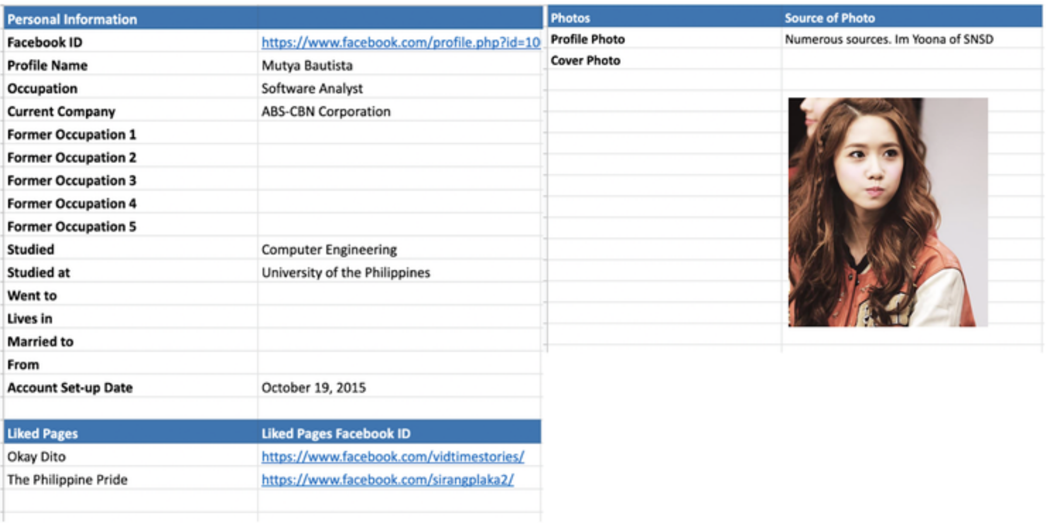
These accounts, along with a set of pages and groups they were connected to, were eventually removed by Facebook. They also inspired Rappler to create Sharktank, a tool for monitoring how information flows on Facebook. That work formed the basis of a series of investigative stories about how propaganda and information operations on Facebook affect democracy in the Philippines. The series included an investigation into the activities of the 26 fake accounts, and kicked off our continued coverage of how Facebook has been weaponized in the Philippines to spread political disinformation, harass people and undermine democracy in the country.

This case study examines how we investigated the original 26 accounts and used them to uncover much larger networks.

Verifying identities, exposing sockpuppets

Our first step in investigating the set of accounts was to try to verify if they were connected to real people. This part required good old fashioned fact-checking and began with our creating spreadsheets to track details related to the accounts, including the personal details they listed, the pages they liked and other information.

For example, Facebook user Mutya Bautista described herself as a “software analyst” at ABS-CBN, the Philippines’ largest television network. Rappler checked with ABS-CBN, who confirmed that she did not work for them.



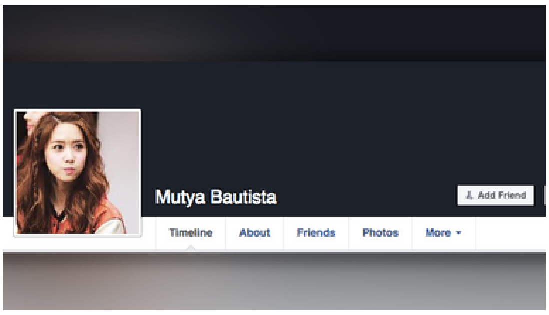
Personal Information	
Facebook ID	https://www.facebook.com/profile.php?id=10
Profile Name	Mutya Bautista
Occupation	Software Analyst
Current Company	ABS-CBN Corporation
Former Occupation 1	
Former Occupation 2	
Former Occupation 3	
Former Occupation 4	
Former Occupation 5	
Studied	Computer Engineering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

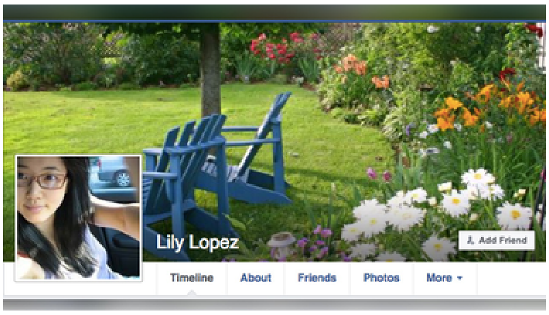
Liked Pages	Liked Pages Facebook ID
Okay Dito	https://www.facebook.com/vidtimestories/
The Philippine Pride	https://www.facebook.com/sirangplaka2/

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

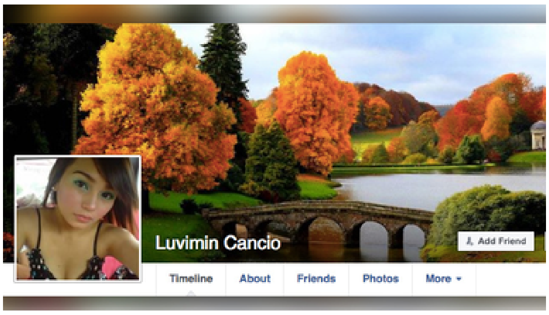
Using reverse image search tools, we found that many of 26 accounts used profile photos of celebrities or personalities.

Bautista, for example, used a picture of Im Yoona of the Korean pop group Girl's Generation. The Lily Lopez account, shown below, used the image of Korean actress Kim Sa-rang.

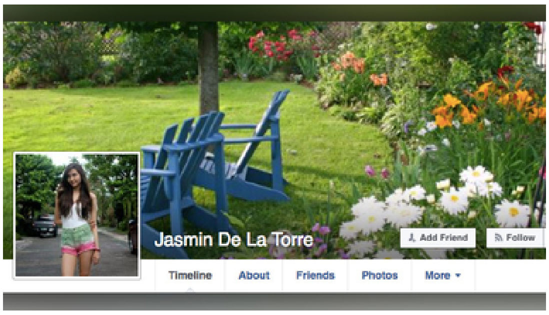


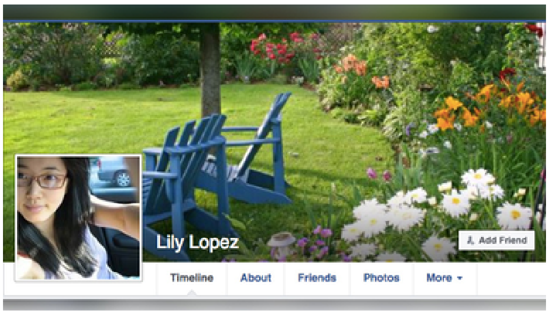


Another account, Luvimin Cancio, used an image from softcorecams.com, a porn site, as its profile photo. We identified this website as the source of the photo through the reverse image search tool TinEye.



The accounts also used similar cover photos on their profiles. Below, the cover photo of the account of Jasmin De La Torre is the same as that of Lily Lopez.



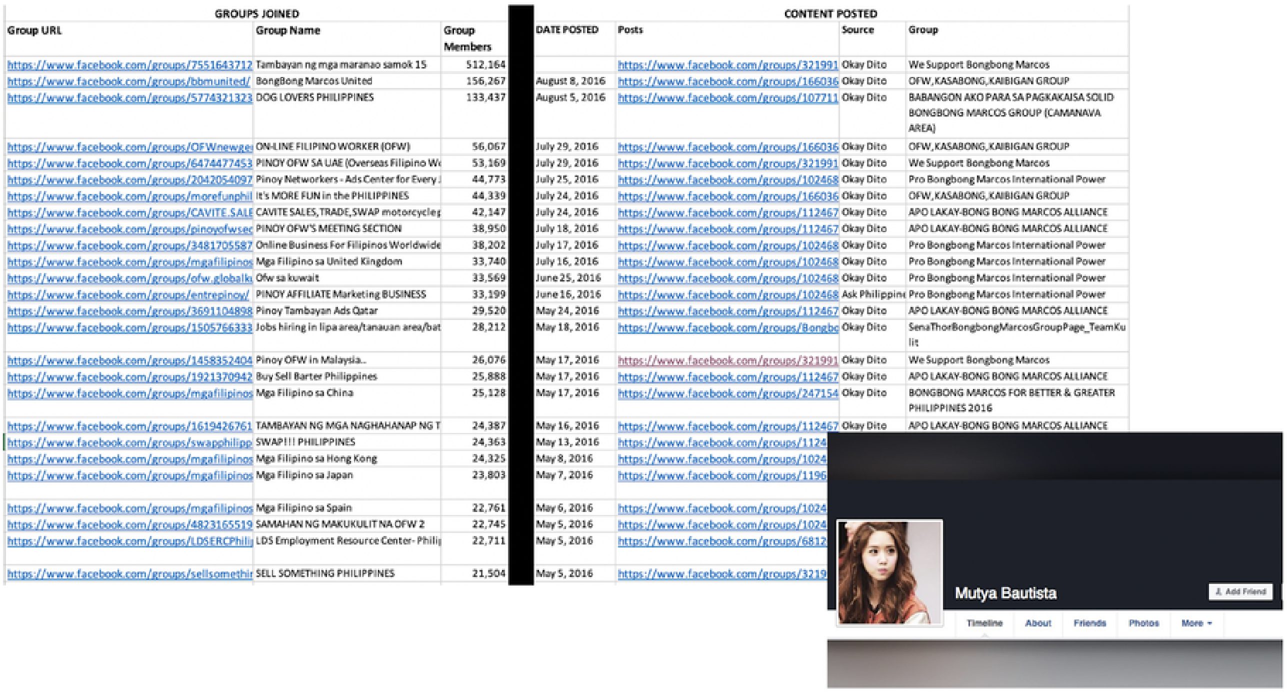


We also noticed one curious thing about the 26 accounts: These users had more groups than friends.

This was unusual because, in the Philippines, most people have friends and family abroad. Facebook basically serves as the communication channel through which people keep in touch with family and friends. So they tend to have many friends as opposed to being members of a huge number of groups.

Bautista's friends list, which was public at the time, showed she had only 17 friends. In fact, each of the 26 accounts that we identified each had fewer than 50 friends when we discovered them in 2016.

Bautista however, was a member of over a hundred groups, including groups campaigning for then-vice presidential candidate Ferdinand Marcos Jr., a number of communities of Filipinos overseas, as well as buy and sell groups, each with members ranging from tens of thousands to hundreds of thousands. Altogether, these groups have over 2.3 million members on Facebook. Below is a list of some of the biggest groups, including their follower counts. Also included is a list of the posts Bautista made to these groups.



GROUPS JOINED				CONTENT POSTED			
Group URL	Group Name	Group Members	[redacted]	DATE POSTED	Posts	Source	Group
https://www.facebook.com/groups/7551643712	Tambayan ng mga maranao samok 15	512,164	[redacted]		https://www.facebook.com/groups/321991	Okay Dito	We Support Bongbong Marcos
https://www.facebook.com/groups/bbmunited/	BongBong Marcos United	156,267	[redacted]	August 8, 2016	https://www.facebook.com/groups/166036	Okay Dito	OFW,KASABONG,KAIBIGAN GROUP
https://www.facebook.com/groups/5774321323	DOG LOVERS PHILIPPINES	133,437	[redacted]	August 5, 2016	https://www.facebook.com/groups/107711	Okay Dito	BABANGON AKO PARA SA PAGKAKAISA SOLID BONGBONG MARCOS GROUP (CAMANAVA AREA)
https://www.facebook.com/groups/OFWnewgei	ON-LINE FILIPINO WORKER (OFW)	56,067	[redacted]	July 29, 2016	https://www.facebook.com/groups/166036	Okay Dito	OFW,KASABONG,KAIBIGAN GROUP
https://www.facebook.com/groups/6474477453	PINOY OFW SA UAE (Overseas Filipino W	53,169	[redacted]	July 29, 2016	https://www.facebook.com/groups/321991	Okay Dito	We Support Bongbong Marcos
https://www.facebook.com/groups/2042054097	Pinoy Networkers - Ads Center for Every	44,773	[redacted]	July 25, 2016	https://www.facebook.com/groups/102468	Okay Dito	Pro Bongbong Marcos International Power
https://www.facebook.com/groups/morefunphil	It's MORE FUN in the PHILIPPINES	44,339	[redacted]	July 24, 2016	https://www.facebook.com/groups/166036	Okay Dito	OFW,KASABONG,KAIBIGAN GROUP
https://www.facebook.com/groups/CAVITE.SALE	CAVITE SALES,TRADE,SWAP motorcycle	42,147	[redacted]	July 24, 2016	https://www.facebook.com/groups/112467	Okay Dito	APO LAKAY-BONG BONG MARCOS ALLIANCE
https://www.facebook.com/groups/pinoyofwsec	PINOY OFW'S MEETING SECTION	38,950	[redacted]	July 18, 2016	https://www.facebook.com/groups/112467	Okay Dito	APO LAKAY-BONG BONG MARCOS ALLIANCE
https://www.facebook.com/groups/3481705587	Online Business For Filipinos Worldwide	38,202	[redacted]	July 17, 2016	https://www.facebook.com/groups/102468	Okay Dito	Pro Bongbong Marcos International Power
https://www.facebook.com/groups/mgafilipinos	Mga Filipino sa United Kingdom	33,740	[redacted]	July 16, 2016	https://www.facebook.com/groups/102468	Okay Dito	Pro Bongbong Marcos International Power
https://www.facebook.com/groups/ofw.globalk	Ofw sa kuwait	33,569	[redacted]	June 25, 2016	https://www.facebook.com/groups/102468	Okay Dito	Pro Bongbong Marcos International Power
https://www.facebook.com/groups/entrepinoy/	PINOY AFFILIATE Marketing BUSINESS	33,199	[redacted]	June 16, 2016	https://www.facebook.com/groups/102468	Ask Philippine	Pro Bongbong Marcos International Power
https://www.facebook.com/groups/3691104898	Pinoy Tambayan Ads Qatar	29,520	[redacted]	May 24, 2016	https://www.facebook.com/groups/112467	Okay Dito	APO LAKAY-BONG BONG MARCOS ALLIANCE
https://www.facebook.com/groups/1505766333	Jobs hiring in lipa area/tanauan area/bat	28,212	[redacted]	May 18, 2016	https://www.facebook.com/groups/Bongbo	Okay Dito	SenaThorBongbongMarcosGroupPage_TeamKulit
https://www.facebook.com/groups/1458352404	Pinoy OFW in Malaysia..	26,076	[redacted]	May 17, 2016	https://www.facebook.com/groups/321991	Okay Dito	We Support Bongbong Marcos
https://www.facebook.com/groups/1921370942	Buy Sell Barter Philippines	25,888	[redacted]	May 17, 2016	https://www.facebook.com/groups/112467	Okay Dito	APO LAKAY-BONG BONG MARCOS ALLIANCE
https://www.facebook.com/groups/mgafilipinos	Mga Filipino sa China	25,128	[redacted]	May 17, 2016	https://www.facebook.com/groups/247154	Okay Dito	BONGBONG MARCOS FOR BETTER & GREATER PHILIPPINES 2016
https://www.facebook.com/groups/1619426761	TAMBAYAN NG MGA NAGHAHANAP NG T	24,387	[redacted]	May 16, 2016	https://www.facebook.com/groups/112467	Okay Dito	APO LAKAY-BONG BONG MARCOS ALLIANCE
https://www.facebook.com/groups/swapphilipp	SWAP!!! PHILIPPINES	24,363	[redacted]	May 13, 2016	https://www.facebook.com/groups/1124		
https://www.facebook.com/groups/mgafilipinos	Mga Filipino sa Hong Kong	24,325	[redacted]	May 8, 2016	https://www.facebook.com/groups/1024		
https://www.facebook.com/groups/mgafilipinos	Mga Filipino sa Japan	23,803	[redacted]	May 7, 2016	https://www.facebook.com/groups/1196		
https://www.facebook.com/groups/mgafilipinos	Mga Filipino sa Spain	22,761	[redacted]	May 6, 2016	https://www.facebook.com/groups/1024		
https://www.facebook.com/groups/4823165519	SAMAHAN NG MAKUKULIT NA OFW 2	22,745	[redacted]	May 5, 2016	https://www.facebook.com/groups/1024		
https://www.facebook.com/groups/LDSERCPhili	LDS Employment Resource Center- Phili	22,711	[redacted]	May 5, 2016	https://www.facebook.com/groups/6812		
https://www.facebook.com/groups/sellsomethi	SELL SOMETHING PHILIPPINES	21,504	[redacted]	May 5, 2016	https://www.facebook.com/groups/3219		

Mutya Bautista

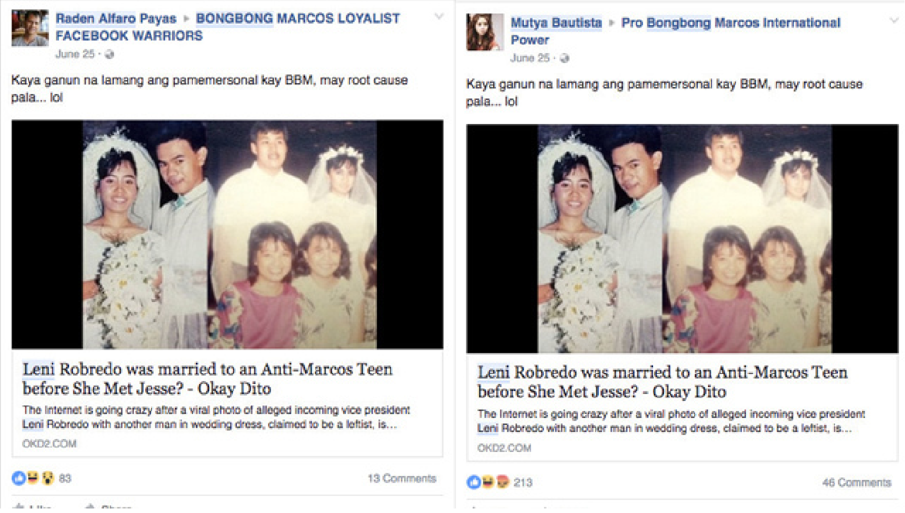
By combining all of these observations and associated data, we concluded that the accounts were sockpuppets: fictional identities created to bolster a particular point of view.

Pro-Marcos network

We could see from the dates associated with the first profile photos and early posts of these 26 accounts that they appeared to have been created in the last quarter of 2015, leading up to the May 2016 elections. We also found that they consistently promoted content that denied the widely documented martial law abuses that took place in the 1970s under the Marcos regime. The accounts also attacked the rivals of the former dictator's son, vice presidential candidate Ferdinand "Bongbong" Marcos Jr.

In the example below, user Mutya Bautista shared a now-debunked claim that Bongbong's rival — then-newly proclaimed vice president Leni Robredo — was previously married to an activist before she married her second husband, the late Interior and local government secretary Jesse Robredo. Bautista posted the story headlined "Leni Robredo was married to an anti-Marcos teen before she met Jesse?" to the group "Pro Bongbong Marcos International Power," with the comment: "*Kaya ganun na lamang ang pamemersonal kay [Bongbong Marcos], may root cause pala.*" ("That's why it's personal against [Bongbong Marcos], there's a root cause.")

Another suspicious account with the name Raden Alfaro Payas shared the same article to the group "Bongbong Marcos loyalist Facebook warriors" with the exact same caption — word for word, down to the last punctuation mark — on the same day.



Raden Alfaro Payas ▸ BONGBONG MARCOS LOYALIST FACEBOOK WARRIORS
June 25 ·

Kaya ganun na lamang ang pamemersonal kay BBM, may root cause pala... lol

Leni Robredo was married to an Anti-Marcos Teen before She Met Jesse? - Okay Dito
The Internet is going crazy after a viral photo of alleged incoming vice president Leni Robredo with another man in wedding dress, claimed to be a leftist, is...
OKD2.COM

83 13 Comments

Mutya Bautista ▸ Pro Bongbong Marcos International Power
June 25 ·

Kaya ganun na lamang ang pamemersonal kay BBM, may root cause pala... lol

Leni Robredo was married to an Anti-Marcos Teen before She Met Jesse? - Okay Dito
The Internet is going crazy after a viral photo of alleged incoming vice president Leni Robredo with another man in wedding dress, claimed to be a leftist, is...
OKD2.COM

213 46 Comments

Fake accounts are often used to spam groups with links, and you can sometimes catch them reusing the same text when they do it. At the time, it was possible to use Facebook Graph search to look at the public posts of users in groups. However, Facebook closed off many Graph search features in 2019, including this function. As a result, it's now necessary to go into groups and search to see what specific users have been sharing.

Connected websites

By analyzing what content the accounts shared, we were able to see that the 26 sockpuppets were promoting the same websites: Okay Dito (OKD2.com), Ask Philippines (askphilippines.com) and why0why.com, among others.

OKD2.com has published a number of hoaxes and other propaganda material favoring the Marcos family and President Rodrigo Duterte. It now masquerades as a classified ads site. But in September 2016 we found that content from the site was shared 11,900 times on Facebook, thanks in part to the sockpuppets.

Through these websites, Rappler eventually traced the potential puppet master of the 26 accounts: someone named Raden Alfaro Payas.

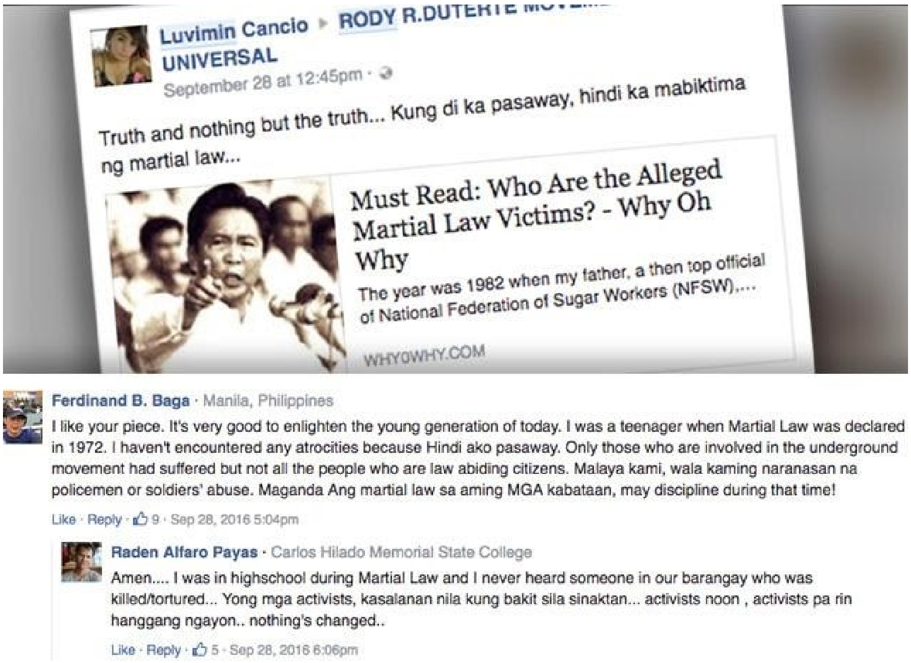
Tracking the puppeteers

Like many sites that Rappler monitors, OKD2.com's current domain registration records are private. The site also does not disclose its authors or owners, and has no contact information other than a web form.

Fortunately, we were able to use historical domain records to identify a person associated with the site. Using domaintools.com, we could see that as of July 2015, OKD2.com was registered in the name of one Raden Payas, a resident of Tanauan City, Batangas. We also found that OKD2.com shared the same Google AdSense ID as other websites, such as askphilippines.com and why0why.com, that the 26 accounts were sharing. We identified the AdSense IDs on these sites by viewing the source code of pages on them and looking for a series of numbers that began with the letters "ca-pub-." Each Google AdSense account is given a unique ID that begins with "ca-pub-," and each page of a site that is linked to an account will have this code on it.

Along with the domain record, we also saw that one of the 26 accounts was called Raden Alfaro Payas (Unofficial). We also found another account in his name with the username "realradenpayas," which interacted with some of the sockpuppets.

For example, he commented on a post from Luvimin Cancio that linked to a story denying the martial law atrocities under Marcos. The "real" Payas account said he was in high school during the martial law years and he "never heard" of anybody being killed or tortured.



Luvimin Cancio ▸ RODY R.DUTERTE MOVE... UNIVERSAL

September 28 at 12:45pm ·

Truth and nothing but the truth... Kung di ka pasaway, hindi ka mabiktima ng martial law...

Must Read: Who Are the Alleged Martial Law Victims? - Why Oh Why

The year was 1982 when my father, a then top official of National Federation of Sugar Workers (NFSW),...

WHYOWHY.COM

Ferdinand B. Baga · Manila, Philippines

I like your piece. It's very good to enlighten the young generation of today. I was a teenager when Martial Law was declared in 1972. I haven't encountered any atrocities because Hindi ako pasaway. Only those who are involved in the underground movement had suffered but not all the people who are law abiding citizens. Malaya kami, wala kaming naranasan na policemen or soldiers' abuse. Maganda Ang martial law sa aming MGA kabataan, may discipline during that time!

Like · Reply · 9 · Sep 28, 2016 5:04pm

Raden Alfaro Payas · Carlos Hilado Memorial State College

Amen.... I was in highschool during Martial Law and I never heard someone in our barangay who was killed/tortured... Yong mga activists, kasalanan nila kung bakit sila sinaktan... activists noon , activists pa rin hanggang ngayon.. nothing's changed..

Like · Reply · 5 · Sep 28, 2016 6:06pm

Jump-starting the Sharktank

These 26 fake accounts and their reach inspired Rappler to create its Sharktank database and automate data collection from public Facebook groups and pages. As of August 2019, Rappler has tracked roughly 40,000 pages with millions of followers.

What began as an investigation into a set of suspicious accounts turned into a continuing study of a network of thousands of fake and real accounts, groups and pages that spread disinformation and propaganda, distorting and politics and weakening the democracy of a nation.

1b. Case Study: How we proved that the biggest Black Lives Matter page on Facebook was fake

Written by **Donie O'Sullivan**

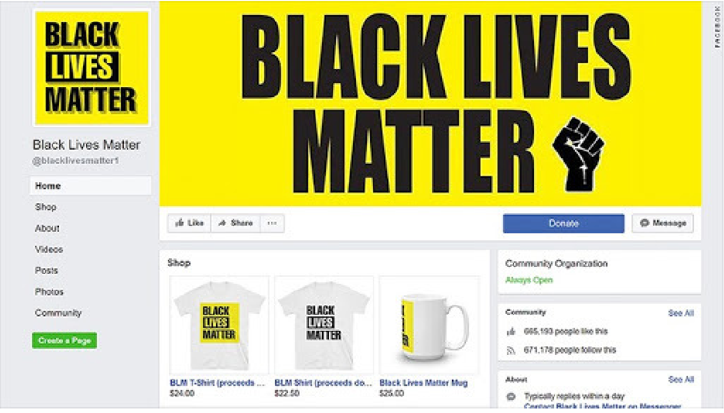
***Donie O'Sullivan** is a CNN reporter covering the intersection of technology and politics. He is part of the CNN Business team and works closely with CNN's investigative unit tracking and identifying online disinformation campaigns targeting the American electorate.*

In the summer and fall of 2017, as the world began learning the details of Russia's expansive effort to influence American voters through social media, it became clear that African Americans and the Black Lives Matter movement were among the main targets of the Kremlin's campaign to sow division.

My colleagues at CNN and I spent months reporting how Russia had been behind some of the biggest Black Lives Matter (BLM) accounts on social media. As I spoke to BLM activists, I would sometimes be asked, "Do you know who runs *the biggest* Black Lives Matter page on Facebook?"

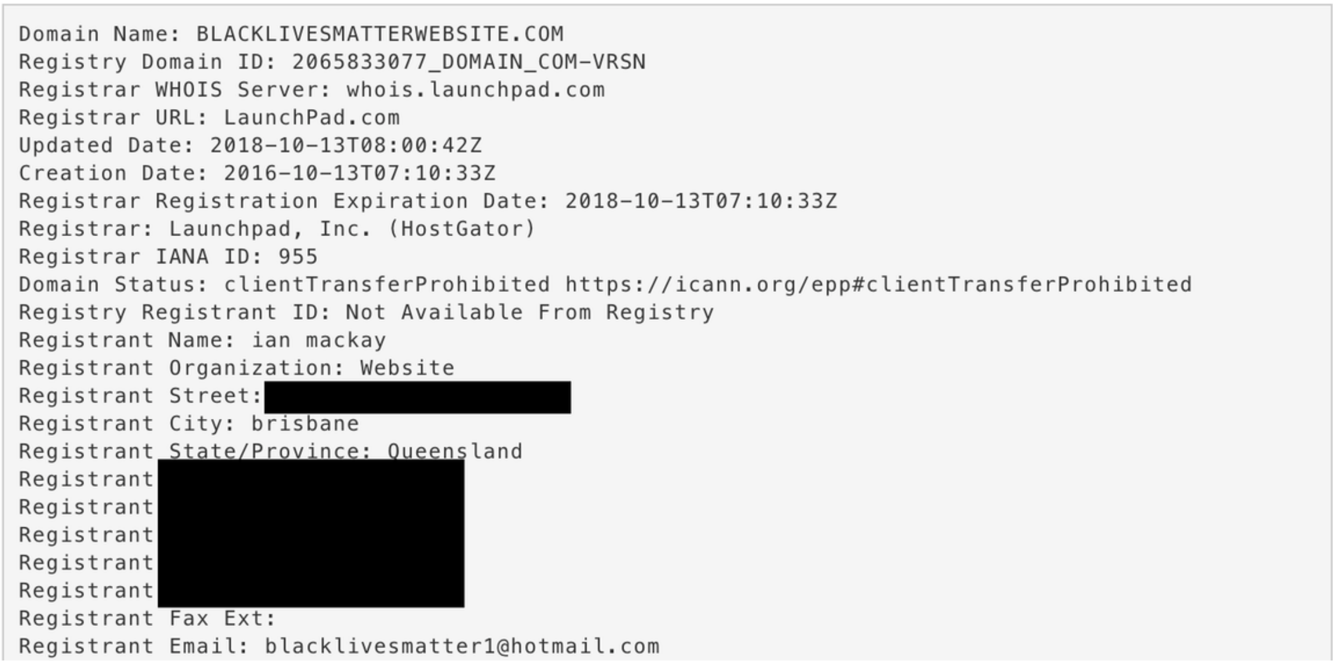
Incredibly, no one — including the most prominent BLM activists in the country and organizers on the ground — knew the answer. Some had understandably suspected the page might be run from Russia. But our investigation found it wasn't Russian, or American — it was run by a white man in Australia.

The page, simply titled "Black Lives Matter," looked legitimate. As of April 2018 it had almost 700,000 followers. It consistently shared links to stories about police brutality and inequality; it ran online fundraisers; it even had an online store that sold BLM merchandise.



It's not unusual for a page that size to be run anonymously. Some activists don't want to put their names on a page and risk attracting attention from trolls or scrutiny from law enforcement looking to shut down protests. Outside the U.S., the ability for activists to run pages anonymously has been critical to digital activism and key to some movements. (It was also precisely what Russia exploited, adding to suspicions that this BLM was connected.)

Around the time I began paying attention to this mysterious page, Jeremy Massler, a freelance investigator and incredible online sleuth, reached out with a tip. Massler had looked at the domain registration records of websites that the huge BLM Facebook page was consistently linking to. Although the domains had been registered privately, he found one of them had, for a period in 2016, belonged to a person in Perth, Australia, named Ian MacKay — a white man.



```
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
```

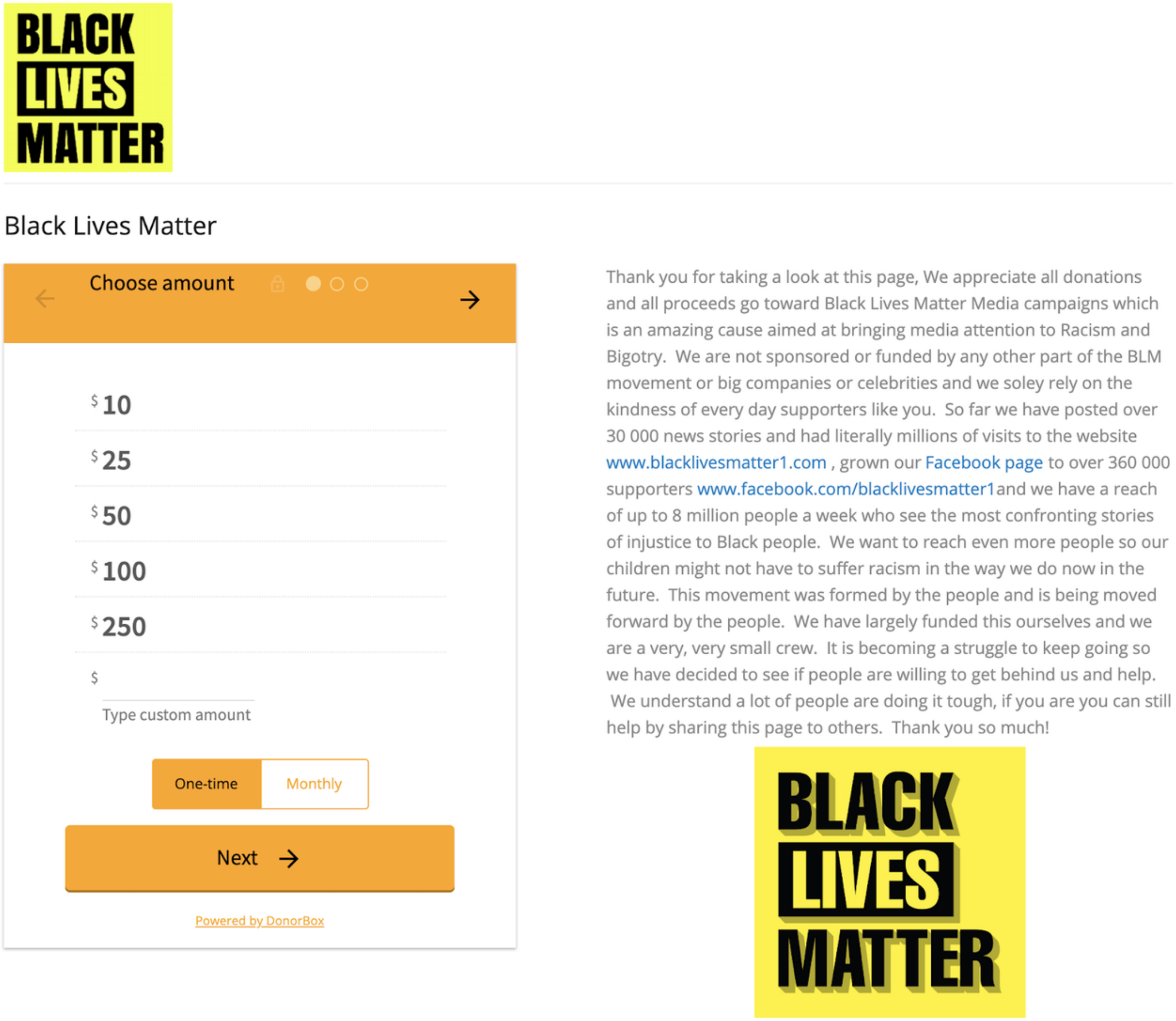
Massler contacted MacKay, who told him he bought and sold domains as a hobby and had nothing to do with the Facebook page. It was the same excuse MacKay, a middle-aged union official, gave me when I reached him by phone a few months later. But by that time we'd found that MacKay had registered dozens of website names, many relating to black activism.

Despite my concerns about the page and the fact that several activists told me they were suspicious of it, I didn't find MacKay's explanation unbelievable on its face. Domain names can be valuable, and people buy and sell them all the time. The fact he had also registered and sold domains that were not related to black activism made his case even more credible. But then something strange happened. A few minutes after I spoke to MacKay, the Facebook page came down. It hadn't been taken down by Facebook, but by whoever was running it — and it hadn't been deleted, only temporarily removed.

That seemed suspicious, so Massler and I began to dig more.

The Facebook page, which came back online in the weeks after my call with MacKay, had during its lifetime promoted fundraising campaigns ostensibly for BLM causes.

In one instance, it claimed to be raising money for activists in Memphis, Tennessee. But when I spoke to activists there, no one knew anything about the fundraiser or where the money might have gone. Other activists even told us that, suspecting it was a scam, they had reported the page to Facebook. But the company hadn't taken any action.



Black Lives Matter

Thank you for taking a look at this page, We appreciate all donations and all proceeds go toward Black Lives Matter Media campaigns which is an amazing cause aimed at bringing media attention to Racism and Bigotry. We are not sponsored or funded by any other part of the BLM movement or big companies or celebrities and we soley rely on the kindness of every day supporters like you. So far we have posted over 30 000 news stories and had literally millions of visits to the website www.blacklivesmatter1.com , grown our Facebook page to over 360 000 supporters www.facebook.com/blacklivesmatter1and we have a reach of up to 8 million people a week who see the most confronting stories of injustice to Black people. We want to reach even more people so our children might not have to suffer racism in the way we do now in the future. This movement was formed by the people and is being moved forward by the people. We have largely funded this ourselves and we are a very, very small crew. It is becoming a struggle to keep going so we have decided to see if people are willing to get behind us and help. We understand a lot of people are doing it tough, if you are you can still help by sharing this page to others. Thank you so much!

As I started to contact the multiple online payment and fundraising platforms the page had used, those companies began removing the fundraisers, saying they had broken their rules. Citing user privacy, none of the payment companies provided me with information on the record about where the money was going. This is a common challenge. Citing their privacy policies, platforms and digital services rarely reveal the names or contact information of account holders to the press.

I later learned from a source familiar with some of the payments processed that at least one account was tied to an Australian bank account and IP address. Another source told me that around $100,000 had been raised. Developing sources at tech companies who are willing to tell you more information than the company will say on the record is becoming increasingly important as many stories cannot be uncovered purely using open source information as scammers and bad actors become more sophisticated.

I brought this information to Facebook to comment for the story and told them I had evidence the page was linked to Australia, that payment companies had removed the campaigns after they investigated, and that we knew some of the money was going to Australia. A Facebook spokesperson said the social media platform's investigation "didn't show anything that violated our Community Standards."

It wasn't until shortly before publishing our story — and only after I raised my concern about Facebook's investigation and its spokesperson's response to a more senior Facebook employee — that Facebook took action and removed the page.

The Australian workers' union where MacKay worked launched an investigation of its own after CNN's report. By the end of the week it had fired MacKay and a second official it said was also involved in the scam.

What was particularly notable about this story was the array of techniques that Massler and I used to get it over the line. We relied heavily on archive sites like the Wayback Machine that allowed us to see the look of websites the page had been linking to and the page itself before it came on our radar. This was particularly useful, as after Massler initially contacted MacKay the people behind the page began trying to cover some of their tracks.

We also used services that track domain registrations, including DomainTools.com, to investigate the sites MacKay had registered and also to find his direct contact details. Massler also extensively used Facebook Graph Search (a tool no longer available) to track the fake Facebook profile accounts that had been set up to promote the page in Facebook Groups. Interrogation of open source information and use of online research tools, like those used to access domain records, are vital instruments — but they are not the only ones.

The simple act of picking up the phone to talk to MacKay and developing sources to provide information that would otherwise not be made public — traditional journalism techniques — were critical in exposing this scam.

2. Finding patient zero

Written by **Henk van Ess**

Henk van Ess *is an assessor for Poynter's International Fact-Checking Network. He is obsessed with finding stories in data. Van Ess trains worldwide media professionals in internet research, social media and multimedia. His clients include NBC News, BuzzFeed News, ITV, Global Witness, SRF, Axel Springer, SRF and numerous NGOs and universities. His websites whopostedwhat.com and graph.tips are heavily used to filter social media. He is @henkvaness on Twitter.*

For decades, Canadian flight attendant Gaëtan Dugas was known as "Patient Zero," the first man to bring AIDS to the United States. This distinction, which was reinforced by books, films and countless news reports made him the "arch-villain of an epidemic that would eventually kill more than 700,000 people in North America."

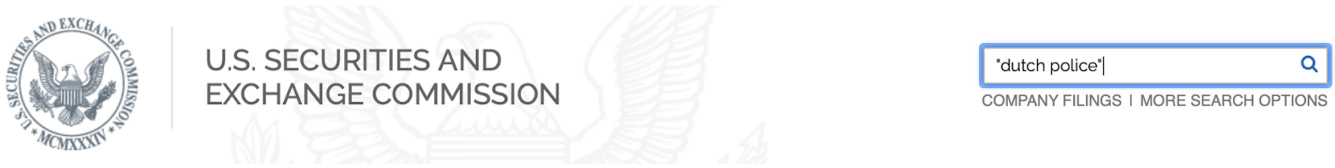
But that was not the case. Bill Darrow, an investigator with the Centers for Disease Control and Prevention, interviewed Dugas and filed him as "Patient O, as in "Out-of-California." It was soon misread as the number 0, setting off a chain reaction of misinformation that persisted until recently.

It's also possible for a journalist to focus on the wrong patient 0 if you don't know how to search properly. This chapter helps you to find primary sources online by getting rid of superficial results and digging deeper.

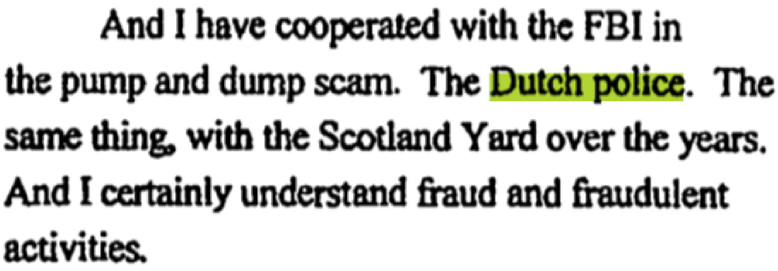
1. Risks of consulting primary sources and how to fix them

Journalists love online primary sources. Firsthand evidence can be found in a newspaper article, a scientific study, a press release, social media or any other possible "patient zero."

Performing a basic keyword search on an official government site can make you think "what you see is what they got." That is often not true. Here is an example. Let's go to the U.S. Securities and Exchange Commission, a source used to find financial information about U.S. citizens as well as businesspeople from all over the world. Let's say we want to find the first occurrence of the phrase "Dutch police" in sec.gov. The built-in search engine of the SEC can help:

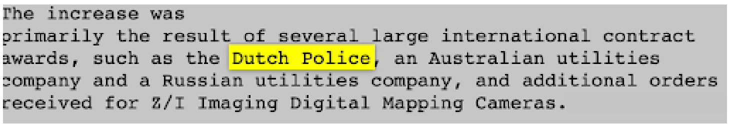


You get just one hit — a document from 2016. So the SEC mentions the Dutch police only once, in 2016, right?



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.

Wrong. The first mention on sec.gov was in 2004, 12 years earlier, in a declassified, encrypted 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.

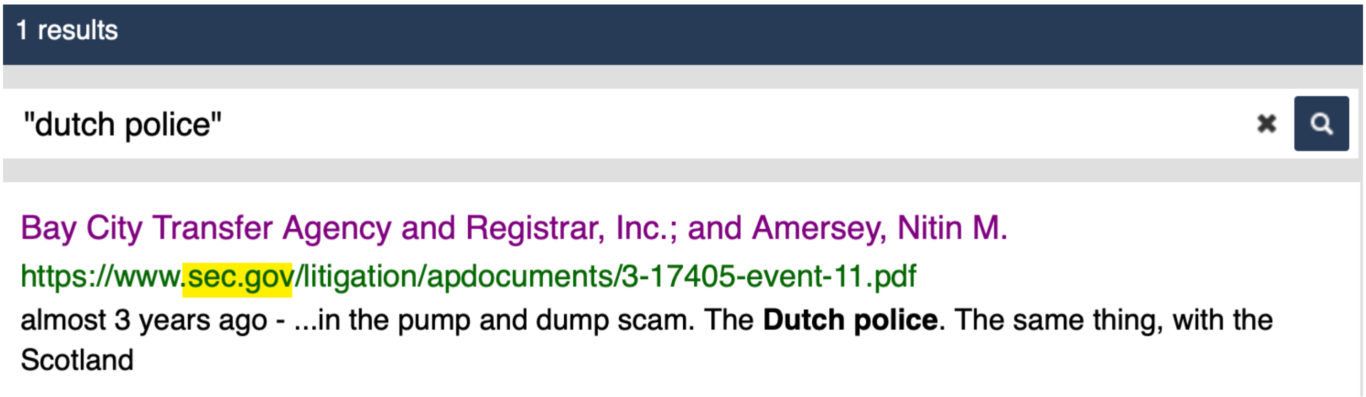
You won't see this in the search results from the search bar on sec.gov, even though this information does come straight from its website. Why the difference?

By default, you should distrust search engines from primary sources. They can give you a false impression of the actual content of the website and its associated databases. The proper way to search is to perform a "primary source check."

Primary source check

Step 1: Look at the failing link

The search result from the SEC provided us with just one source:



1 results

"dutch police"

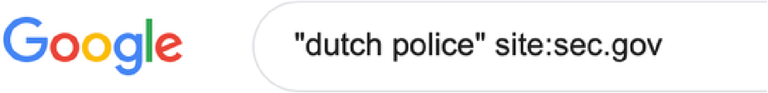
Bay City Transfer Agency and Registrar, Inc.; and Amersey, Nitin M.
https://www.sec.gov/litigation/apdocuments/3-17405-event-11.pdf
almost 3 years ago - ...in the pump and dump scam. The **Dutch police**. The same thing, with the Scotland

Let's work with that disappointment. First, get rid of "https://www," the first part of the link. Watch out for the first backslash after that (/) — in this case it's before the word "litigation/"

That's the part we need: sec.gov

2. Second step: Use "site:"

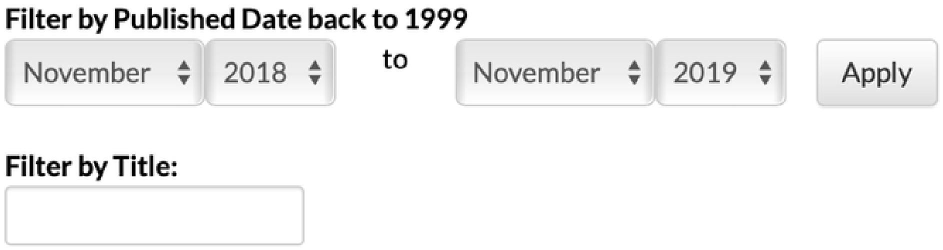
Go to a generic search engine. Start with the query ("Dutch police") and end with "site:" followed directly with the URL (no spaces). This is the formula for finding out if an original source shows you everything:



"dutch police" site:sec.gov

Including specific folders

You can now adapt the "primary source formula" to your needs. Let's go to the press release section of the New Jersey Courts website. Say you want to find out when the Mercer County Bar Association sponsored a Law Day program, but you can't find the primary source in the title of any press release. The "Mercer County Bar Association" is not visible in any title.

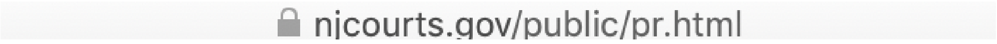


Filter by Published Date back to 1999

November ▴▾ 2018 ▴▾ to November ▴▾ 2019 ▴▾ Apply

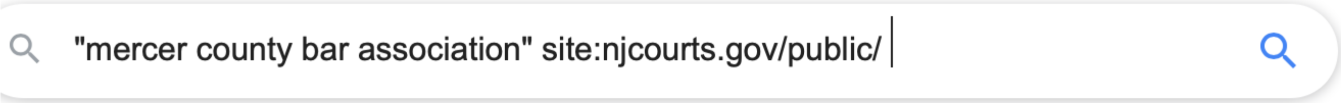
Filter by Title:

Now look at the URL of that page full of poorly indexed press releases:



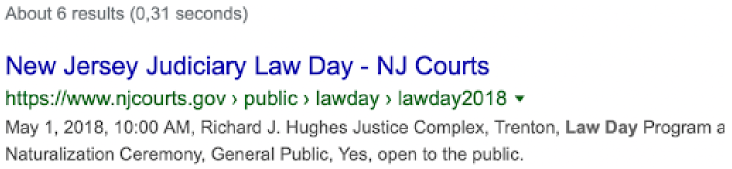
njcourts.gov/public/pr.html

The public relations material is filed away in the folder /public. That should be included in your Google search:



"mercer county bar association" site:njcourts.gov/public/

And there you are:

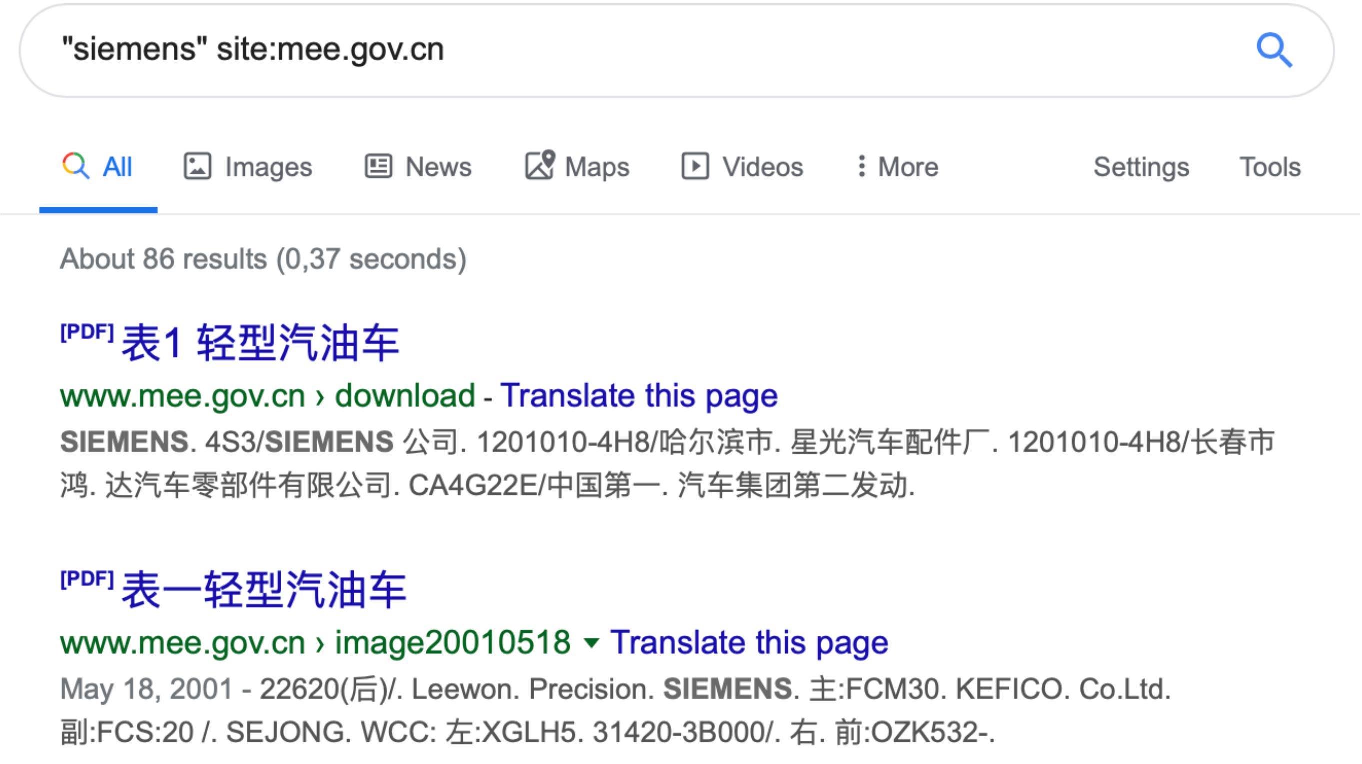


About 6 results (0,31 seconds)

New Jersey Judiciary Law Day - NJ Courts
https://www.njcourts.gov › public › lawday › lawday2018 ▾
May 1, 2018, 10:00 AM, Richard J. Hughes Justice Complex, Trenton, **Law Day** Program a Naturalization Ceremony, General Public, Yes, open to the public.

Predicting folders

China has a Ministry of Ecology and Environment. Do they have English documents about the German company Siemens? With the following formula, you get Chinese and English documents in the search results:



About 86 results (0,37 seconds)

[PDF] 表1 轻型汽油车

www.mee.gov.cn › download - Translate this page

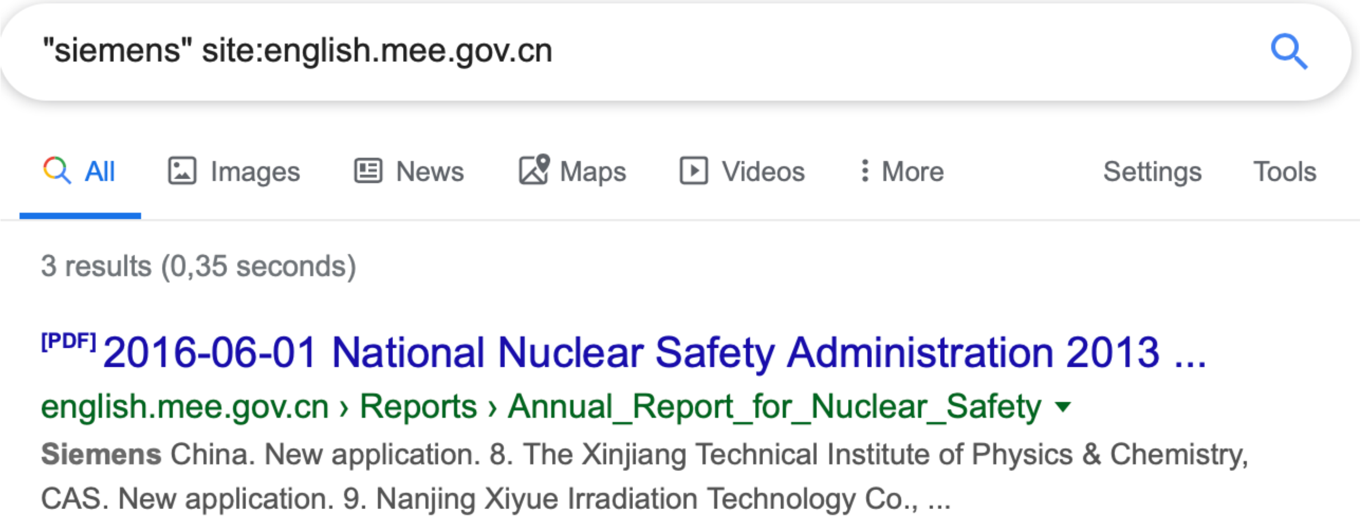
SIEMENS. 4S3/**SIEMENS** 公司. 1201010-4H8/哈尔滨市. 星光汽车配件厂. 1201010-4H8/长春市鸿. 达汽车零部件有限公司. CA4G22E/中国第一. 汽车集团第二发动.

[PDF] 表一轻型汽油车

www.mee.gov.cn › image20010518 ▾ Translate this page

May 18, 2001 - 22620(后)/. Leewon. Precision. **SIEMENS**. 主:FCM30. KEFICO. Co.Ltd. 副:FCS:20 /. SEJONG. WCC: 左:XGLH5. 31420-3B000/. 右. 前:OZK532-.

If you want to filter to see only the English ones, maybe they used the word English in the link? Try it out. It works:



3 results (0,35 seconds)

[PDF] 2016-06-01 National Nuclear Safety Administration 2013 ...

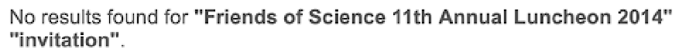
english.mee.gov.cn › Reports › Annual_Report_for_Nuclear_Safety ▾

Siemens China. New application. 8. The Xinjiang Technical Institute of Physics & Chemistry, CAS. New application. 9. Nanjing Xiyue Irradiation Technology Co., ...

2. Following the trail of documents

Sometimes the information we need isn't contained on a webpage, but is actually in a document hosted on a website. Here's how to follow the document trail using Google formulas.

Ross McKitrick is an associate professor in the Economics Department at the University of Guelph, Ontario. Back in 2014, he did a presentation for a climate skeptic group. Let's try to find the invitation for that meeting. We know it was held on May 13, 2014, and was the 11th Annual Luncheon organized by the "Friends of Science (FOS)." If we search Google for these terms we come up empty:



No results found for **"Friends of Science 11th Annual Luncheon 2014" "invitation"**.

Why? Because the word *invitation* is not in many invitations. It's the same with the word *interview*. Many interviews don't contain the word *interview*. Even most maps don't have the word *map* explicitly written on it. My advice? Stop guessing and Go Zen.

Step 1: Establish the document type

Try to find the common denominator of any online invitation. It's often a PDF document. Search for just that with "filetype:pdf" and you might find it.

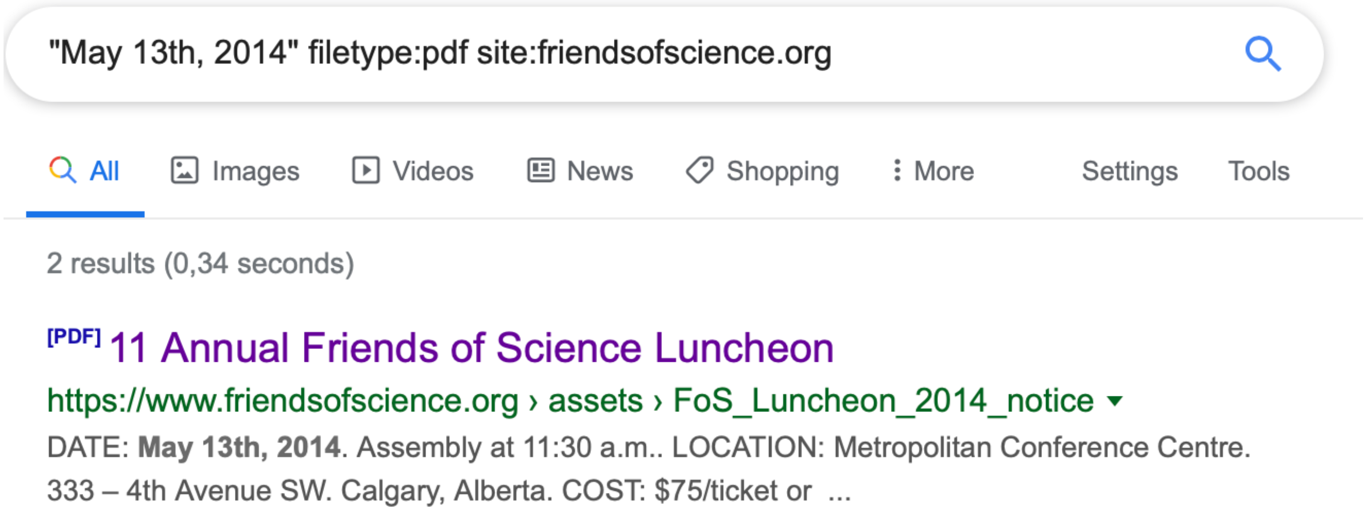
Step 2: Be (climate) neutral

You don't know the exact wording of the invitation. But what you do know, is that the YouTube video was from a May 13, 2014, event. It's feasible that the date is mentioned in the invitation. (Be sure to search for both the cardinal and ordinal forms, May 13 and May 13th.)

Step 3: Who is involved?

We know the organizer is "Friends of Science" and its website is friendsofscience.org.

When you combine all three steps, the query in Google will be:



"May 13th, 2014" filetype:pdf site:friendsofscience.org

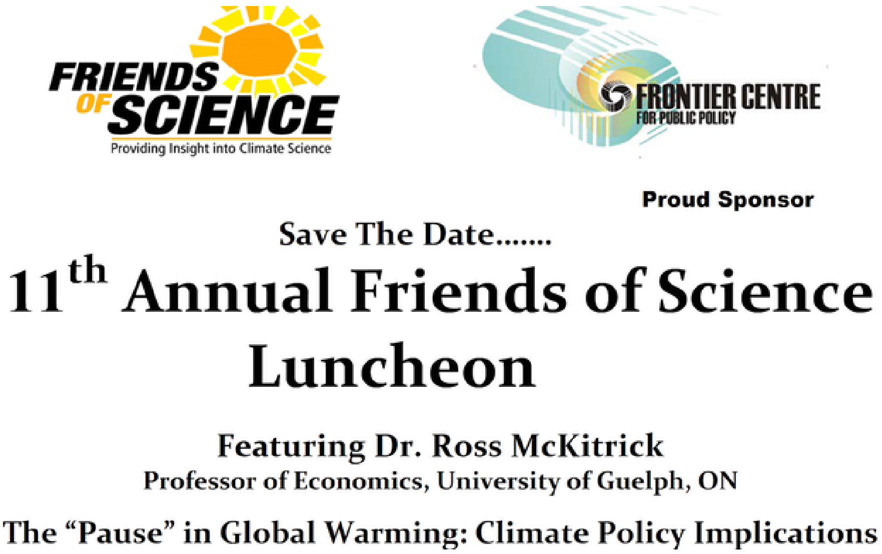
2 results (0,34 seconds)

[PDF] 11 Annual Friends of Science Luncheon

https://www.friendsofscience.org › assets › FoS_Luncheon_2014_notice

DATE: **May 13th, 2014**. Assembly at 11:30 a.m.. LOCATION: Metropolitan Conference Centre. 333 – 4th Avenue SW. Calgary, Alberta. COST: $75/ticket or ...

There it is in the first hit: the invitation for the event.


FRIENDS OF SCIENCE
Providing Insight into Climate Science
FRONTIER CENTRE FOR PUBLIC POLICY

Proud Sponsor

Save The Date.......

11th Annual Friends of Science Luncheon

Featuring Dr. Ross McKitrick

Professor of Economics, University of Guelph, ON

The "Pause" in Global Warming: Climate Policy Implications

The FOS, based in Calgary, is frequently labeled a climate denial group and is funded in part by the oil and gas sector So how would we craft a query to find out more information about it and its network of supporters and funders?

Step 1: Include target

"Friends of Science" results in too many hits, so include also "Calgary."

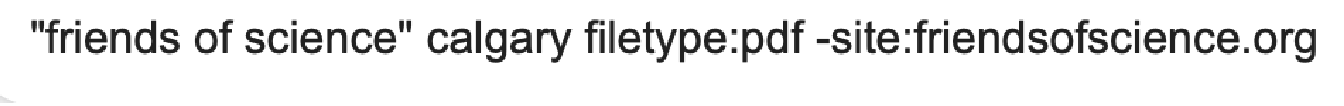
Step 2: Include "filetype"

Go for the next best thing for any official document, "filetype:pdf."

Step 3: Exclude your target's website

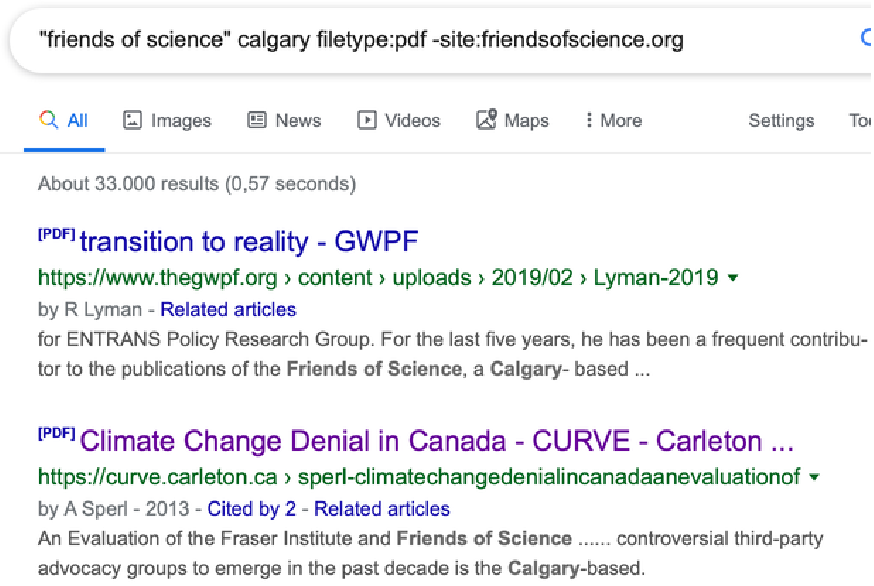
Exclude the target's website Friendsofscience.org by adding "-site:friendsofscience.org." This helps you find information from outside parties.

The full query is:



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

Because you searched for the target in official documents, but not from its own website, you find some brothers in arms and those who are critical of the organization:



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

About 33.000 results (0,57 seconds)

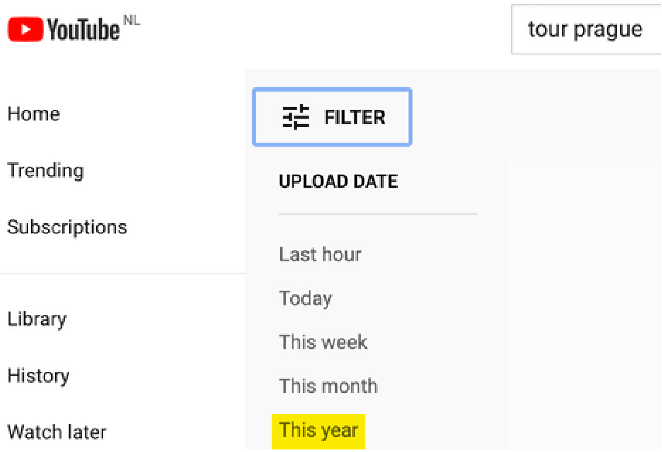
[PDF] transition to reality - GWPF
https://www.thegwpf.org › content › uploads › 2019/02 › Lyman-2019
by R Lyman - Related articles
for ENTRANS Policy Research Group. For the last five years, he has been a frequent contributor to the publications of the **Friends of Science**, a **Calgary**- based ...

[PDF] Climate Change Denial in Canada - CURVE - Carleton ...
https://curve.carleton.ca › sperl-climatechangedenialincanadaanevaluationof
by A Sperl - 2013 - Cited by 2 - Related articles
An Evaluation of the Fraser Institute and **Friends of Science** ...... controversial third-party advocacy groups to emerge in the past decade is the **Calgary**-based.

3. Filtering social media for primary sources

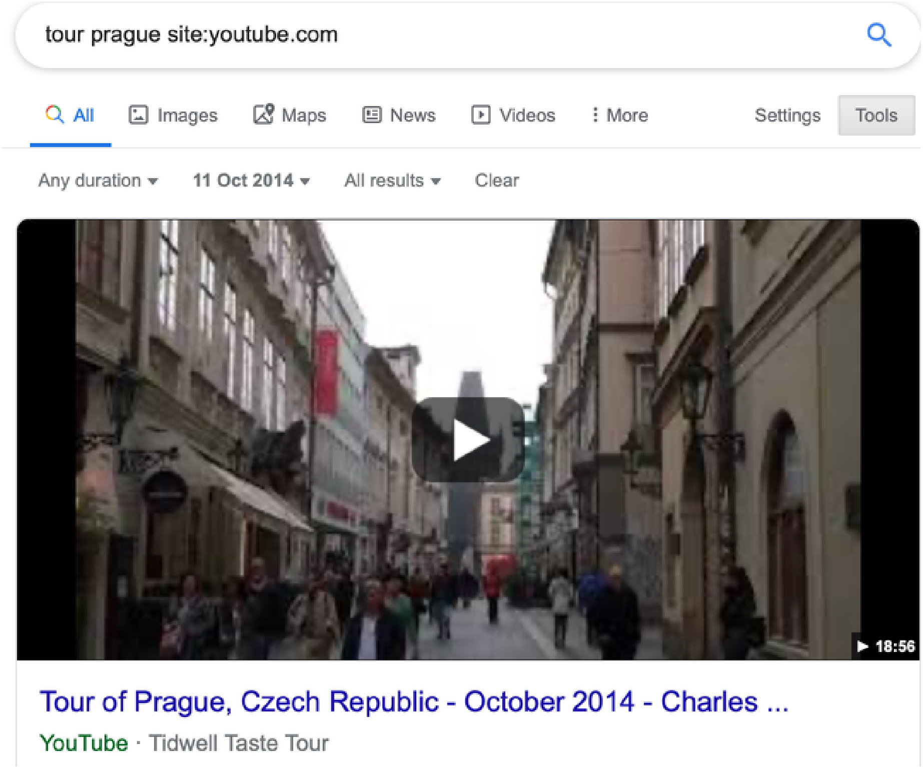
YouTube

YouTube's search tool has a problem: it won't let you filter for videos that are older than one year. If you want to find a video of a tour in Prague from Oct 11, 2014, this is the roadblock you will hit:



tour prague

To solve this, manually enter the preferred date into a Google.com search by using the "Tools" menu on the far right. Then select "Any time" and "Custom Range." Now we get the results we need:



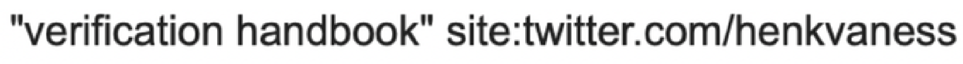
tour prague site:youtube.com

Any duration ▾ 11 Oct 2014 ▾ All results ▾ Clear

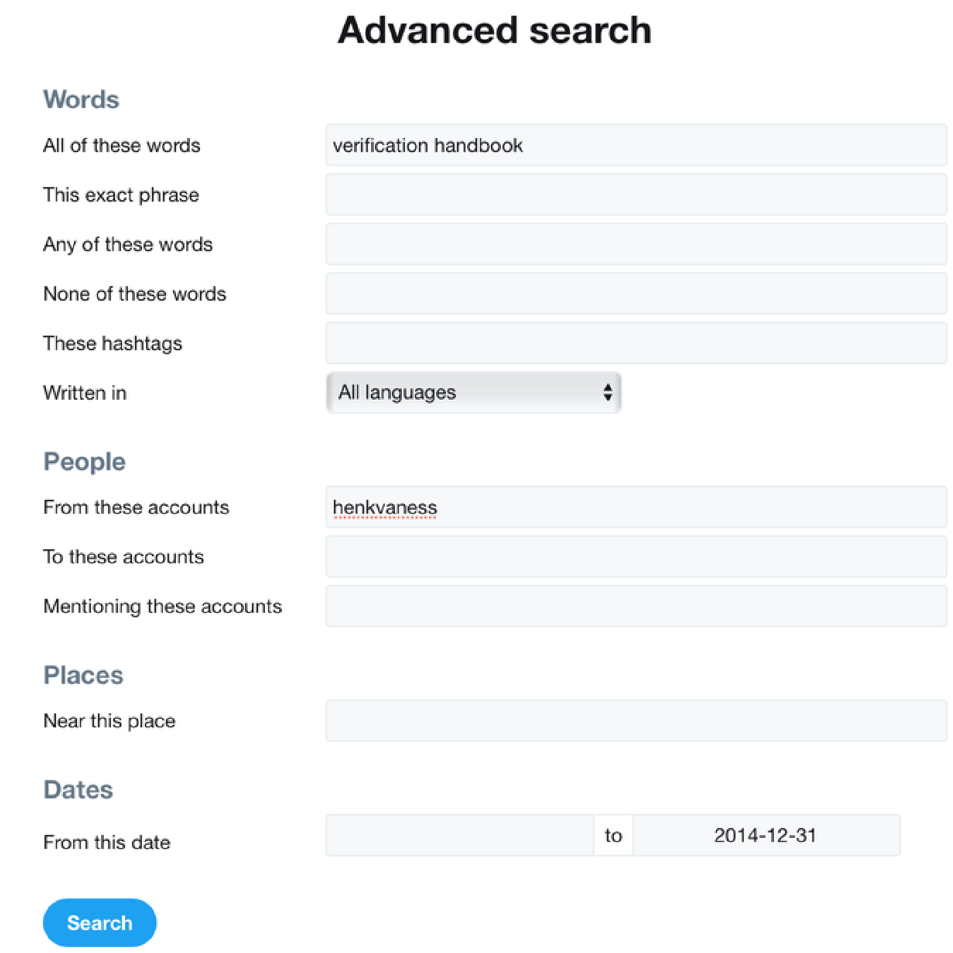
Tour of Prague, Czech Republic - October 2014 - Charles ...
YouTube · Tidwell Taste Tour

Twitter

Despite the power of the “site:” search operator, you’ll be disappointed if you use it in Google to try searching Twitter. For example, we could try this query to find when I tweeted about the Verification Handbook for the first time:


"verification handbook" site:twitter.com/henkvaness

But it returns you only one hit as of this writing. Generic search engines like Google often struggle to deliver quality results from the trillions of posts on Twitter, or on big platforms such as Facebook and Instagram. The answer for Twitter is to use its Advanced Search functionality and add keywords, username and time period, as shown here:



Advanced search

Words

All of these words	verification handbook
This exact phrase	
Any of these words	
None of these words	
These hashtags	
Written in	All languages

People

From these accounts	henkvaness
To these accounts	
Mentioning these accounts	

Places

Near this place	

Dates

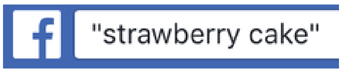
From this date		to	2014-12-31

Don't forget to click on “Latest" on the menu at the top of the search results page so you can view the results in reverse chronological order. By default, Twitter sorts your results by what it considers to be the top tweets.

Facebook

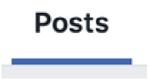
Using “site:” on Facebook is also not ideal, but we can make its native search tool fit our needs. Let’s say for example you want to see posts from March 2019 about strawberry cake from people in Brooklyn. Follow these steps:

Step 1: Type in query

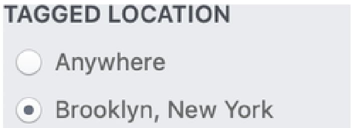


"strawberry cake"

Step 2: Click on posts



Step 3: Define location

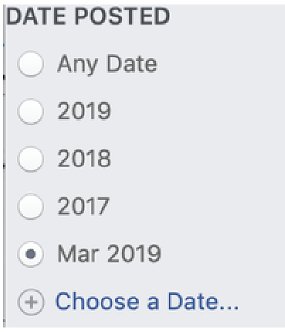


TAGGED LOCATION

○ Anywhere

● Brooklyn, New York

Step 4: Choose a date



DATE POSTED

○ Any Date

○ 2019

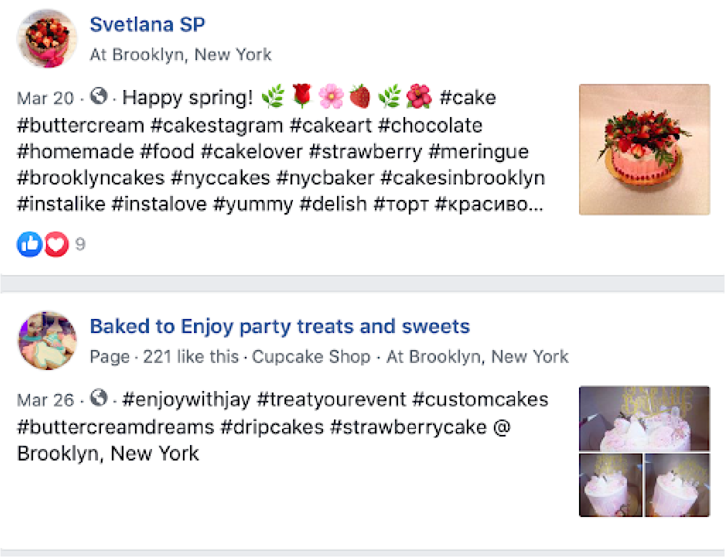
○ 2018

○ 2017

● Mar 2019

⊕ Choose a Date...

And there you are:



Svetlana SP
At Brooklyn, New York

Mar 20 · 🌐 · Happy spring! 🌿🌹🌸🍓🌿🌺 #cake #buttercream #cakestagram #cakeart #chocolate #homemade #food #cakelover #strawberry #meringue #brooklyncakes #nyccakes #nycbaker #cakesinbrooklyn #instalike #instalove #yummy #delish #торт #красиво...

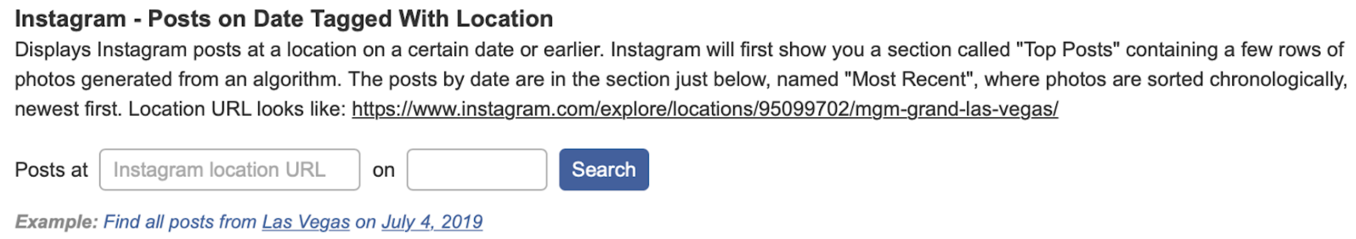
👍❤ 9

Baked to Enjoy party treats and sweets
Page · 221 like this · Cupcake Shop · At Brooklyn, New York

Mar 26 · 🌐 · #enjoywithjay #treatyourevent #customcakes #buttercreamdreams #dripcakes #strawberrycake @ Brooklyn, New York

Instagram

To search Instagram for posts from a specific date in a specific location, you can go to my site, whopostedwhat.com, and fill in your query:



Instagram - Posts on Date Tagged With Location

Displays Instagram posts at a location on a certain date or earlier. Instagram will first show you a section called "Top Posts" containing a few rows of photos generated from an algorithm. The posts by date are in the section just below, named "Most Recent", where photos are sorted chronologically, newest first. Location URL looks like: https://www.instagram.com/explore/locations/95099702/mgm-grand-las-vegas/

Posts at [Instagram location URL] on [ ] Search

Example: Find all posts from Las Vegas on July 4, 2019

3. Spotting bots, cyborgs and inauthentic activity

Written by: Johanna Wild , Charlotte Godart

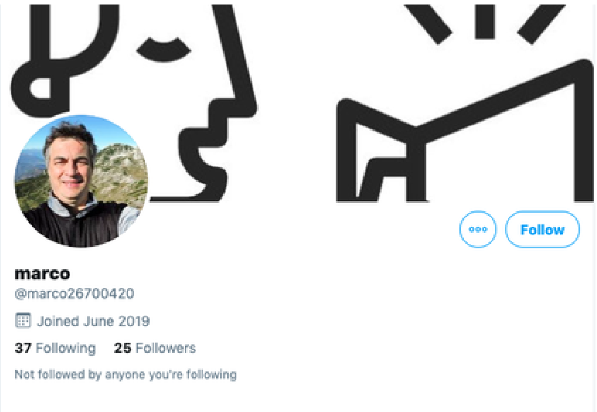
Charlotte Godart *is an investigator and trainer for Bellingcat. Before Bellingcat, she was at the Human Rights Center at UC Berkeley, working within its Investigations Lab, teaching students to conduct open-source research on global conflicts for international humanitarian entities.*

Johanna Wild *is an open-source investigator at Bellingcat, where she also focuses on tech and tool development for digital investigations. She has an online journalism background and previously worked with journalists in (post-)conflict regions. One of her roles was to support journalists in Eastern Africa to produce broadcasts for the Voice of America.*

In late August 2019, Benjamin Strick, a Bellingcat contributor and BBC Africa EYE investigator, was analyzing tweets spreading the hashtags #WestPapua and #FreeWestPapua when he noticed accounts exhibiting abnormal behavior. These accounts were all spreading Indonesian pro-government messages at a moment when the conflict in West Papua was gaining international visibility: A local independence movement had taken to the streets to fight for freedom from Indonesian control, leading to violence between the Indonesian police and protesters.

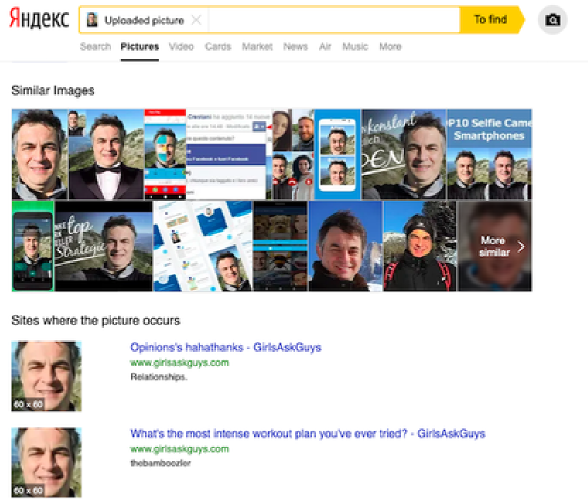
The accounts Strick saw exhibited multiple odd similarities. Soon, he would realize that these were the early indicators of coordinated inauthentic behavior. But at first, he started by noticing the small stuff.

For one, many of the accounts had stolen profile pictures. Take this account for instance, which claimed to be of someone named Marco:



marco
@marco26700420
Joined June 2019
37 Following 25 Followers
Not followed by anyone you're following

Using Yandex's reverse image search tool, Strick found that the account's profile picture had been previously used on other websites under different names. None of the accounts using the photo were for a real person named "Marco." This proved that the accounts were, at the very least, misleading about their true identities.



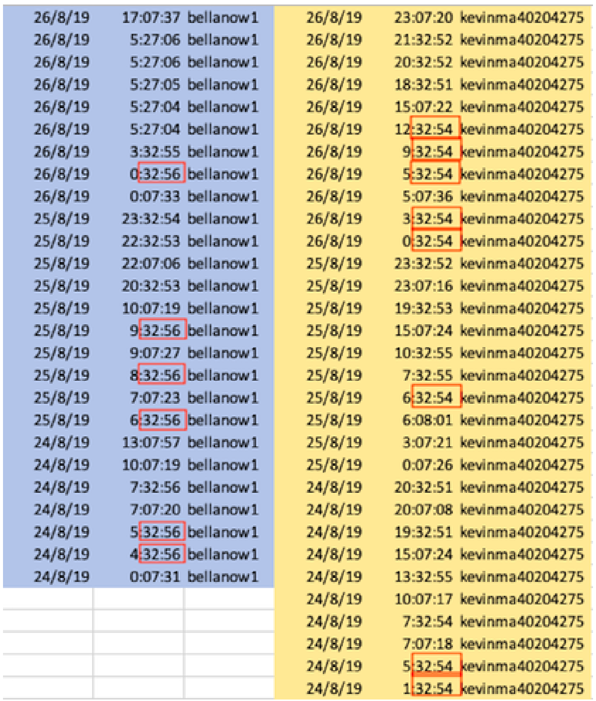
Similar Images

Sites where the picture occurs

Opinions's hahathanks - GirlsAskGuys
www.girlsaskguys.com
Relationships.

What's the most intense workout plan you've ever tried? - GirlsAskGuys
www.girlsaskguys.com
thebamboozler

Beyond faking their identities, Strick also found the accounts published similar or even identical content while often retweeting one another. Even more striking was that some of them showed precise synchronization in the timecode patterns of their tweets. For example, @bellanow1 and @kevinma40204275 mostly published their tweets at minute 7 or minute 32 of any particular hour.



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

It's unlikely that a human would adopt this kind of tweet rhythm. This synchronization across multiple accounts, combined with their misleading photos, suggested the accounts were not linked to real identities, and could be automated. By analyzing suspicious account patterns such as these, Strick eventually concluded that the accounts were part of a pro-Indonesian Twitter bot network that was spreading one-sided and misleading information about the conflict in West Papua. (You can read more about the larger network these accounts were part of in the chapter 11b case study, "Investigating an Information Operation In West Papua.")

What's a bot? The answer is more complicated than you might think

The West Papua case is far from being the only information operation to use social bots. Other operations have been much more widely publicized and criticized, although at their core they contain similarities in how they operate.

A bot is a software application that can automatically perform tasks assigned to it by humans. Whether a bot does good or bad completely depends on the intentions of its “owner.”

The bots most often referred to in public debates are social bots, active on social networks including Facebook, Twitter and LinkedIn. On these platforms, they can be used to spread specific ideological messages, often with the aim to make it look as if there is a groundswell of support for a particular topic, person, piece of content or hashtag.

Social media bots tend to fall into three main categories: the scheduled bot, the watcher bot and the amplifier bot. It’s important to know which kind of bot you’re interested in because each type has a specific purpose. With each purpose comes a different language and communication pattern. In the context of disinformation, we’re most interested in looking into the amplifier bot.

The amplifier bot exists to do exactly what it sounds like: amplify and spread content, with the goal of shaping online public opinion. It can also be used to make individuals and organizations appear to have a larger following than they really do. Its power comes in numbers. A network of amplifier bots can attempt to influence hashtags, spread links or visual content, or gang up to mass spam or harass an individual online in an attempt to discredit them or to make them seem controversial or under siege.

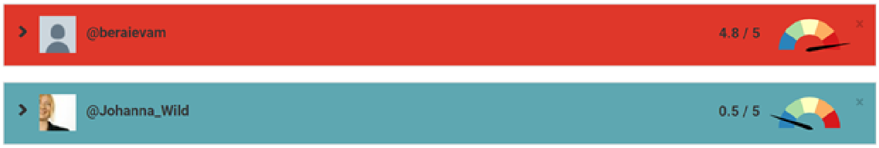
By working together in large numbers, amplifier bots seem more legitimate and therefore help shape the online public opinion landscape. Amplifier bots that spread disinformation do it mainly through hashtag campaigns or by sharing news in the form of links, videos, memes, photos or other content types. Hashtag campaigns involve bots constantly tweeting the same hashtag, or set of hashtags, in coordination. The goal is often to trick Twitter’s trending algorithm into adding a specific hashtag to the trending topics list. An example is “#Hillarysick,” which was propagated widely by bots after Hillary Clinton stumbled in September 2016, shortly before the presidential election. (It’s also important to note that hashtag campaigns don’t require bots, and can be more effective without them. See this investigation of human “hashtag mills” in Pakistan from Dawn.)

Purchasing and creating bots is relatively easy. Countless sites will sell you your own bot army for just a couple of hundred dollars or even less. But a sophisticated, humanlike botnet is much harder to create and maintain.

How to recognize bots

Developers and researchers have created many tools to help assess whether an account might be automated. These tools can be useful in gathering information, but a score from one tool is by no means definitive and should never form the sole basis of any reporting or conclusion.

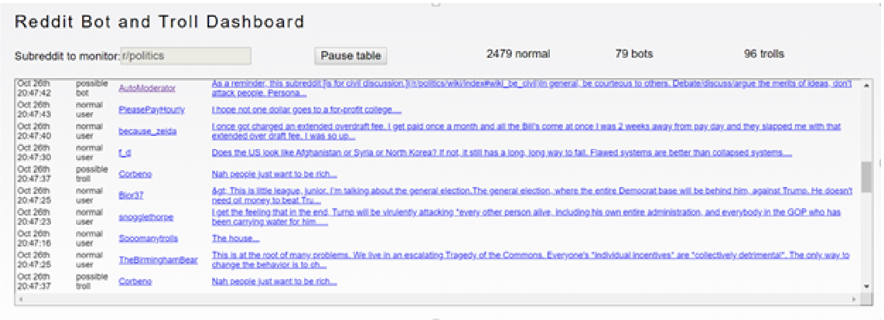
One of the most well-known tools is Botometer, created by researchers at Indiana University. Based on various criteria, it calculates a score for how likely it is that a Twitter account and its followers are bots.



@beraievam 4.8 / 5

@Johanna_Wild 0.5 / 5

For Reddit, Jason Skowronski has created a real-time dashboard. After you set it up for a chosen subreddit, it tries to assess whether the comments were made by bots, trolls or humans.



Reddit Bot and Troll Dashboard

Subreddit to monitor: r/politics

2479 normal

79 bots

96 trolls

Oct 26th 20:47:42	possible bot	AutoModerator	As a reminder, this subreddit [is for civil discussion.](/r/politics/wiki/index#wiki_be_civil) In general, be courteous to others. Debate/discuss/argue the merits of ideas, don't attack people. Persona...
Oct 26th 20:47:43	normal user	PleasePayHourly	I hope not one dollar goes to a for-profit college....
Oct 26th 20:47:40	normal user	because_zelda	I once got charged an extended overdraft fee. I get paid once a month and all the Bill's come at once I was 2 weeks away from pay day and they slapped me with that extended over draft fee. I was so up...
Oct 26th 20:47:30	normal user	f_d	Does the US look like Afghanistan or Syria or North Korea? If not, it still has a long, long way to fall. Flawed systems are better than collapsed systems....
Oct 26th 20:47:37	possible troll	Corbeno	Nah people just want to be rich...
Oct 26th 20:47:25	normal user	Bior37	&gt; This is little league, junior. I'm talking about the general election.The general election, where the entire Democrat base will be behind him, against Trump. He doesn't need oil money to beat Tru...
Oct 26th 20:47:23	normal user	snogglethorpe	I get the feeling that in the end, Tump will be virulently attacking "every other person alive, including his own entire administration, and everybody in the GOP who has been carrying water for him....
Oct 26th 20:47:16	normal user	Sooomanytrolls	The house...
Oct 26th 20:47:25	normal user	TheBirminghamBear	This is at the root of many problems. We live in an escalating Tragedy of the Commons. Everyone's "individual incentives" are "collectively detrimental". The only way to change the behavior is to sh...
Oct 26th 20:47:37	possible troll	Corbeno	Nah people just want to be rich...

While there are exceptions, most publicly available bot detection tools have been created for Twitter. The reason is that many social networks — including Facebook — restrict their APIs (application programming interfaces) in a way that prevents the public from analyzing and using their data to create such public tools.

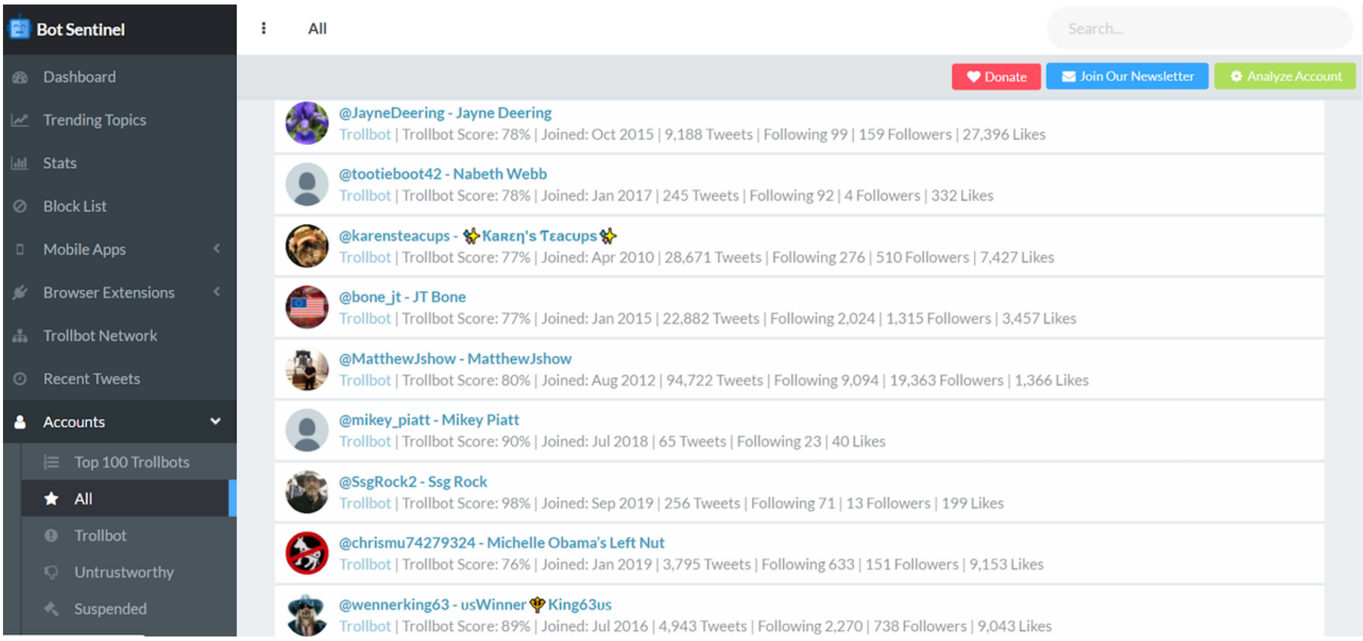
As noted earlier, bot detection tools are a great starting point but they should not be your sole evidence. One reason for their varying degree of accuracy is there is simply no universal list of criteria for recognizing bots with 100% certainty. There's also little agreement about how to classify something as a bot. Researchers at the Oxford Internet Institute's Computational Propaganda Project classify accounts that post more than 50 times a day as having "heavy automation." The Atlantic Council's Digital Forensics Research Lab considers "72 tweets per day (one every ten minutes for twelve hours at a stretch) as suspicious, and over 144 tweets per day as highly suspicious."

It can often be challenging to determine whether a disinformation campaign is conducted by social bots or by humans who are motivated or paid to post large amounts of content about a specific topic. The BBC, for instance, found that accounts who posted similar Facebook messages amplifying favorable content about Boris Johnson in November 2019 were managed by people who pretended to be social bots.

You might also encounter cyborgs, social media accounts that are partly automated and partly managed by humans, which display a combination of natural and inauthentic behavior. Journalists must avoid falsely labeling suspicious accounts as bots without proper evidence and analysis, as a mistaken accusation can undermine your credibility.

One way to deal with these different types of bots, cyborgs and hyperactive human accounts is to focus your investigation on monitoring all inauthentic or bot-like behavior, instead of trying to identify only one type of suspicious account.

For example, Bot Sentinel provides a publicly available database containing (U.S.) Twitter accounts that exhibit suspicious behavior. Their creators decided to collect "accounts that were repeatedly violating Twitter rules" instead of specifically searching for social bots.



@JayneDeering - Jayne Deering
Trollbot | Trollbot Score: 78% | Joined: Oct 2015 | 9,188 Tweets | Following 99 | 159 Followers | 27,396 Likes

@tootieboot42 - Nabeth Webb
Trollbot | Trollbot Score: 78% | Joined: Jan 2017 | 245 Tweets | Following 92 | 4 Followers | 332 Likes

@karensteacups - ✨Καrεη's Τεacups✨
Trollbot | Trollbot Score: 77% | Joined: Apr 2010 | 28,671 Tweets | Following 276 | 510 Followers | 7,427 Likes

@bone_jt - JT Bone
Trollbot | Trollbot Score: 77% | Joined: Jan 2015 | 22,882 Tweets | Following 2,024 | 1,315 Followers | 3,457 Likes

@MatthewJshow - MatthewJshow
Trollbot | Trollbot Score: 80% | Joined: Aug 2012 | 94,722 Tweets | Following 9,094 | 19,363 Followers | 1,366 Likes

@mikey_piatt - Mikey Piatt
Trollbot | Trollbot Score: 90% | Joined: Jul 2018 | 65 Tweets | Following 23 | 40 Likes

@SsgRock2 - Ssg Rock
Trollbot | Trollbot Score: 98% | Joined: Sep 2019 | 256 Tweets | Following 71 | 13 Followers | 199 Likes

@chrismu74279324 - Michelle Obama's Left Nut
Trollbot | Trollbot Score: 76% | Joined: Jan 2019 | 3,795 Tweets | Following 633 | 151 Followers | 9,153 Likes

@wennerking63 - usWinner 🏆 King63us
Trollbot | Trollbot Score: 89% | Joined: Jul 2016 | 4,943 Tweets | Following 2,270 | 738 Followers | 9,043 Likes

Steps to investigate inauthentic behavior

In general, we suggest the following approach for identifying inauthentic and potentially automated behavior on social networks:

1. Manually check the accounts for suspicious behavior.

2. Combine this with the use of tools or more technical network analyses.

3. Investigate their activity, content and network of other accounts they interact with. Combine this with traditional investigation techniques, such as trying to contact them or people they claim to know.

4. Consult with outside experts who specialize in bots and inauthentic activity.

To learn how to manually assess suspicious accounts, it's important to understand the typical warning signs of automated accounts on Twitter, or other social networks.

Every social media bot needs an identity. Bot creators want to make their accounts appear as convincing as possible, but it takes time to set up and maintain credible-looking profiles, in particular if the goal is to run a large bot network. The more accounts someone has, the more time-consuming it is to create and manage them in a way that makes them seem authentic. This is where these accounts slip up. In many cases, their creators do the bare minimum to establish a profile, and a good investigator can detect this.

Here are a few things to look for:

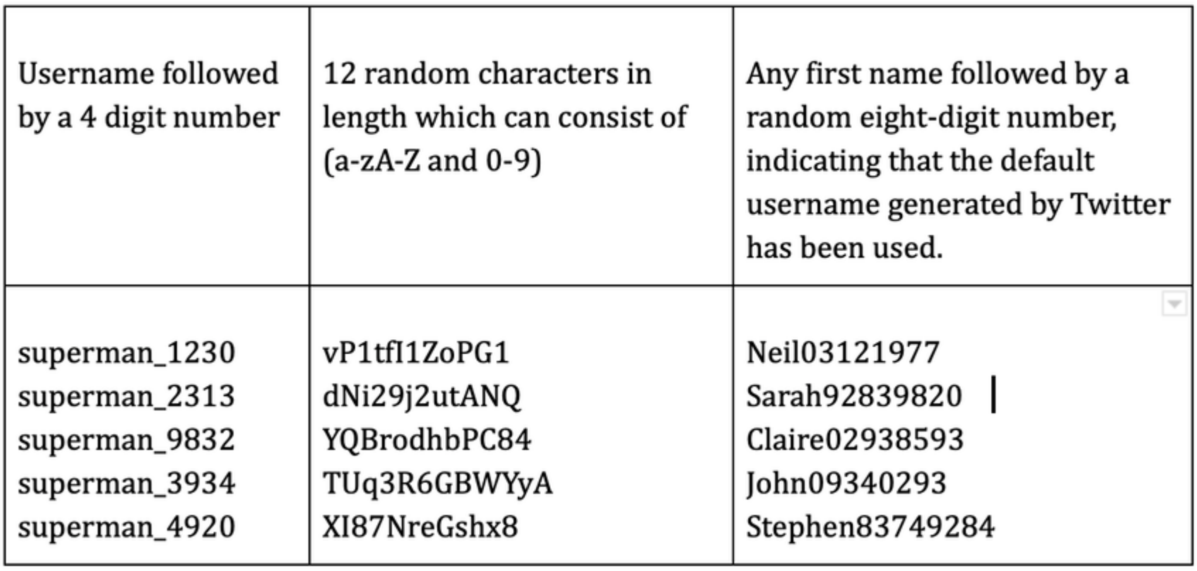
No real profile picture

A stolen profile picture (as seen in Benjamin Strick's West Papua investigation) or no profile picture at all can be an indicator of inauthenticity. Since bot creators want to create many accounts at once, they have to obtain a collection of photos and often copy them from other websites. However, doing so creates inconsistencies. For instance, an account with the profile photo of a male but a username implying that a female is the owner of the account could be a signal that something isn't right. To get around this issue, many bot creators choose cartoons or animals as profile pictures, but again this tactic becomes another pattern to use to detect inauthentic or bot accounts.

Automatically created usernames

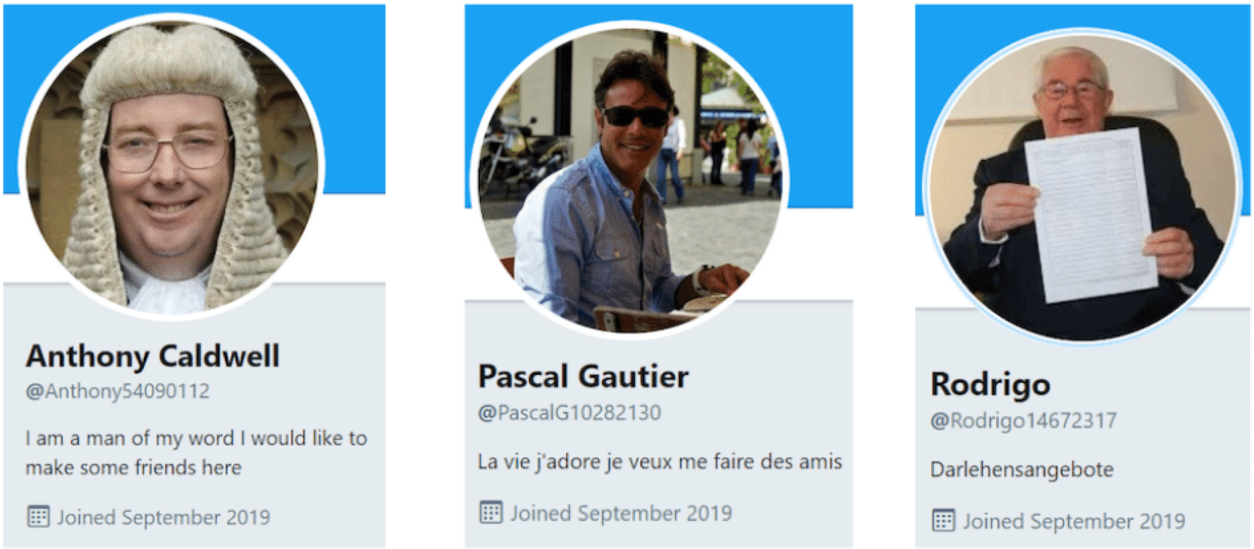
Next, look out for names and usernames. Every Twitter handle is unique, which means the username you want is often already taken. This is an inconvenience to the average person, but becomes a real challenge when you're trying to create 50, 500 or 5,000 accounts in a short period of time.

Bot creators often deploy a strategy to help them easily find unused usernames. Scripts with criteria like the following are used to automatically create usernames:



Username followed by a 4 digit number	12 random characters in length which can consist of (a-zA-Z and 0-9)	Any first name followed by a random eight-digit number, indicating that the default username generated by Twitter has been used.
superman_1230 superman_2313 superman_9832 superman_3934 superman_4920	vP1tfI1ZoPG1 dNi29j2utANQ YQBrodhbPC84 TUq3R6GBWYyA XI87NreGshx8	Neil03121977 Sarah92839820 Claire02938593 John09340293 Stephen83749284

When you notice several Twitter accounts with handles consisting of the same number of characters and digits, you can manually search for more accounts with that pattern in each of the accounts' followers list to potentially identify a network.



Anthony Caldwell
@Anthony54090112
I am a man of my word I would like to make some friends here
Joined September 2019

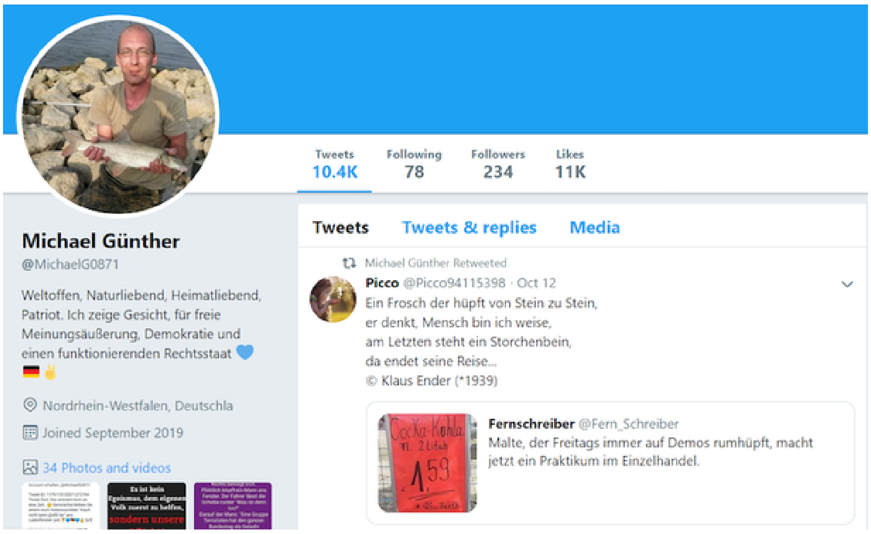
Pascal Gautier
@PascalG10282130
La vie j'adore je veux me faire des amis
Joined September 2019

Rodrigo
@Rodrigo14672317
Darlehensangebote
Joined September 2019

In this example, the accounts have something else in common: They all were created in September 2019. When combined with other signals this can be an indicator that the accounts were all done at the same time by the same person.

Account activity does not fit age

You should become even more suspicious if a new account already has a relatively large number of followers or if it has published a large number of tweets within a short time. The same applies if an older account has very few followers despite being very active.



Tweets	Following	Followers	Likes
10.4K	78	234	11K

Michael Günther
@MichaelG0871

Weltoffen, Naturliebend, Heimatliebend, Patriot. Ich zeige Gesicht, für freie Meinungsäußerung, Demokratie und einen funktionierenden Rechtsstaat 💙
🇩🇪 ✌

Nordrhein-Westfalen, Deutschla

Joined September 2019

34 Photos and videos

Michael Günther Retweeted

Picco @Picco94115398 · Oct 12
Ein Frosch der hüpft von Stein zu Stein,
er denkt, Mensch bin ich weise,
am Letzten steht ein Storchenbein,
da endet seine Reise...
© Klaus Ender (*1939)

Fernschreiber @Fern_Schreiber
Malte, der Freitags immer auf Demos rumhüpft, macht jetzt ein Praktikum im Einzelhandel.

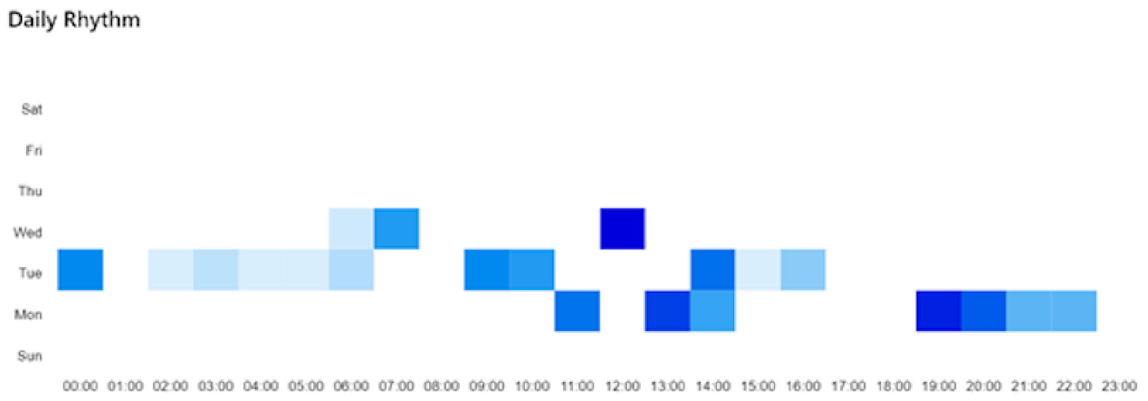
If you come across such an account, analyze the account's tweet activity more deeply. Take the number of tweets located at the top of the page, and divide this by the number of days the account has been active. For example, take an account that has 3,489 tweets as of Nov. 11, 2019, and was created on Aug. 15, 2019. Divide 3,489 by 89 (the days it's been active), and you get 39.2 tweets per day.

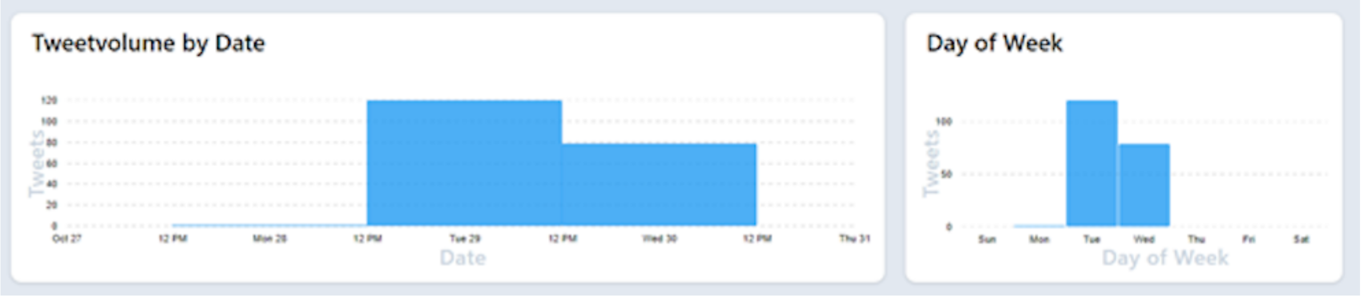
Looking at the tweets made over the lifetime of the account, does the number seem too high, unrealistic or not maintainable?

Suspicious tweet patterns

Another element to examine is tweet rhythm. Humans might show slight preferences for the days and times they usually tweet, but it is unlikely that a person posts consistently only on Monday, Tuesday and Wednesday and is completely silent on all other days of the week over a long period of time.

If you want to see these patterns visualized for one specific account, check out the account analysis tool built by Luca Hammer:

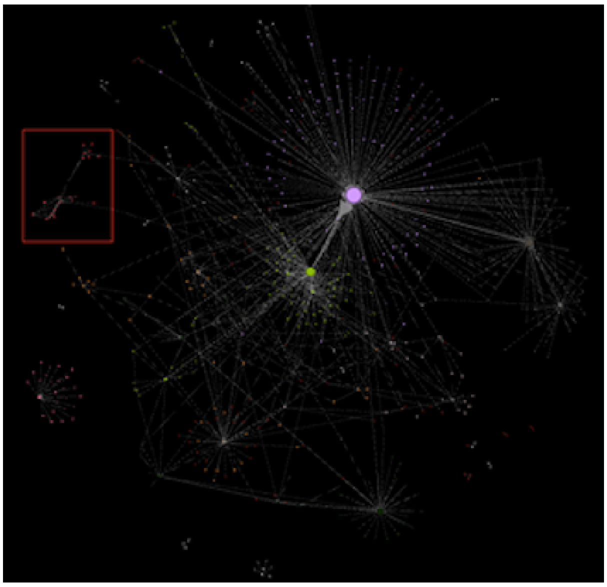




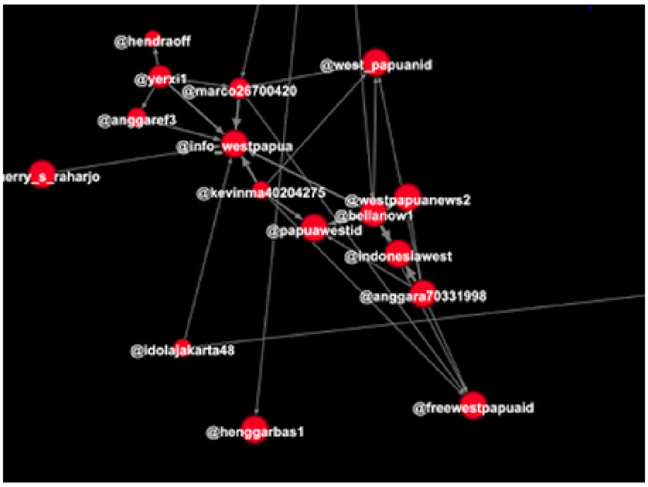
Visualization as part of your investigation

To get a better understanding of the activity of a whole bot network, you can use a visualization platform like Gephi. Bellingcat contributor Benjamin Strick used this tool to analyze the connections between Twitter accounts belonging to a pro-Indonesian bot network.

By looking at the visual representation of the connections between a large number of Twitter accounts, he noticed that the structure on the left side of the picture (in red) stood out.



By zooming in on this area, he could see which Twitter accounts were part of this specific structure.



Each red circle represents a Twitter account and the lines are the relationships between them. Usually, smaller accounts are arranged around a bigger circle in the middle, which means that they all interact with the influential account. The accounts in the structure above, however, did not interact in that way with one another. This led Strick to analyze those abnormal account's behavior.

The future of social bots: Can we out-trick them?

The technology behind social bots has become much more advanced in the last few years, allowing these small software applications to become more adept at simulating human behavior. We are getting to the point where people are predicting that artificial users could engage in sophisticated online communications without their human counterparts realizing that they're actually having a long conversation with a bot.

However, as of now there is no proof that high-level, machine-learning-empowered social bots exist or are being deployed. For now, it seems that many disinformation campaigns are currently still receiving support from less-complex amplifier bots.

"I don't think that there are many sophisticated social bots out there that are able to have real conversations with people and to convince them of certain political positions," said Dr. Ole Pütz, a researcher for the project "Unbiased Bots that Build Bridges" at Bielefeld University in Germany.

According to him, the best way to help the public recognize inauthentic behavior on social networks is to use a detection method that catalogs and weighs all the factors that make an account suspicious. As an example, he says, "This account uses a script to retweet news, it automatically follows others, and that one never uses speech patterns that humans would normally use."

For now, a methodical analysis of account behavior, content, interactions and patterns remains the best approach for identifying inauthentic behavior.

In our case study chapter, we provide a more in-depth and technical explanation of how we analyzed the different factors in a suspicious Twitter network related to the Hong Kong protests.

3a. Case study: Finding evidence of automated Twitter activity during the Hong Kong protests

Written by: Charlotte Godart , Johanna Wild

Charlotte Godart *is an investigator and trainer for Bellingcat. Before Bellingcat, she was at the Human Rights Center at UC Berkeley, working within its Investigations Lab, teaching students to conduct open-source research on global conflicts for international humanitarian entities.*

Johanna Wild *is an open-source investigator at Bellingcat, where she also focuses on tech and tool development for digital investigations. She has an online journalism background and previously worked with journalists in (post-)conflict regions. One of her roles was to support journalists in Eastern Africa to produce broadcasts for the Voice of America.*

In August 2019, Twitter announced the removal of thousands of Twitter accounts it said helped spread disinformation about the Hong Kong protests and were part of "a coordinated state-backed operation." Soon, Facebook and YouTube released statements saying they also removed accounts engaging in coordinated inauthentic behavior about the protests.

Unlike Facebook and YouTube, Twitter released a list of the accounts it removed, offering an opportunity to further investigate the activity. With a participant of a Bellingcat workshop, our team decided to investigate the remaining Twitter content about the protests in Hong Kong to try to identify signs of coordinated inauthentic behavior.

Finding suspicious activity

We started by searching for relevant hashtags about the protests. A simple keyword search for "Hong Kong Riots" brought up many tweets, some containing multiple hashtags.

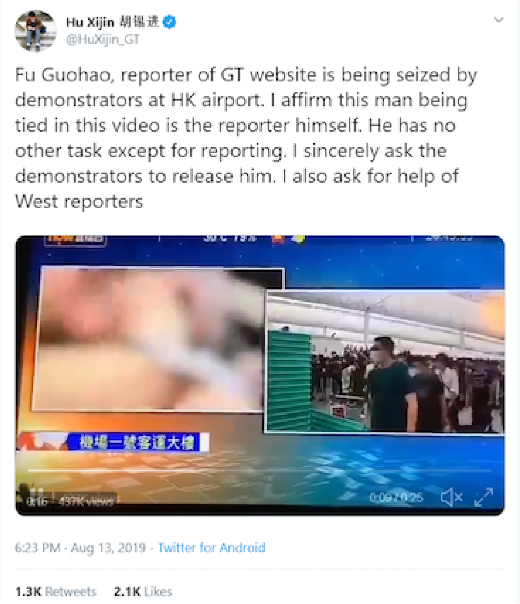
We wanted to focus on pro-China accounts and content, since these were the ones Twitter had already found engaging in inauthentic activity. We tried keyword formulations like:

"Shame on Hong Kong" -police -government

This search yields results that contain the phrase "Shame on Hong Kong" but not the words police or government. The goal was to filter out tweets such as "shame on hong kong police" and keep tweets such as "shame on hong kong protesters." Other search terms were "Hong Kong roaches" and "Hong Kong mobs," which were common descriptors of the protesters by pro-Chinese Twitter accounts.

After using those and other search terms, we examined recent tweets about Hong Kong that received many retweets and likes. You can filter by engagement simply by adding "min_retweets:500" or "min_faves:500" to your search query. This will get only tweets with at least 500 retweets or likes.

We then looked at the Twitter accounts that had interacted with those tweets. For example, there was this tweet from verified user Hu Xijin, editor-in-chief of the Chinese and English editions of the Global Times, a Chinese state-run media outlet:

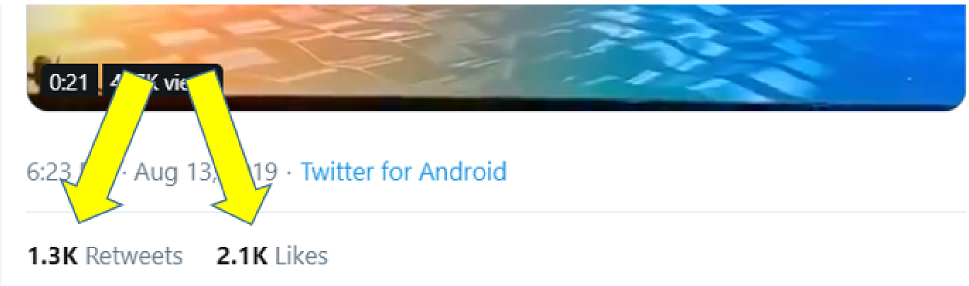

Hu Xijin 胡锡进
@HuXijin_GT

Fu Guohao, reporter of GT website is being seized by demonstrators at HK airport. I affirm this man being tied in this video is the reporter himself. He has no other task except for reporting. I sincerely ask the demonstrators to release him. I also ask for help of West reporters

6:23 PM · Aug 13, 2019 · Twitter for Android

1.3K Retweets 2.1K Likes

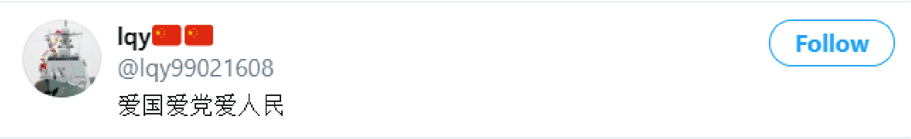
We clicked on the “Retweets” and “Likes” hyperlinks next to each engagement number to display a list of accounts that performed the relevant action.

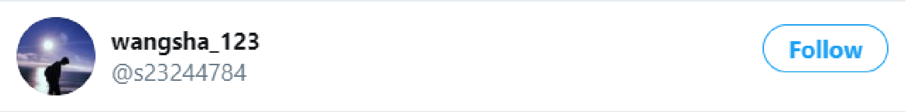


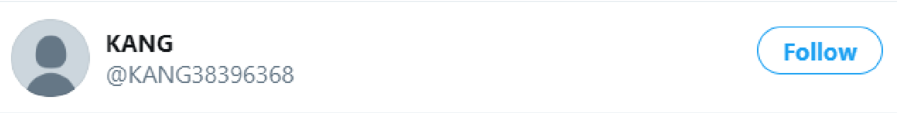
6:23 PM · Aug 13, 2019 · Twitter for Android

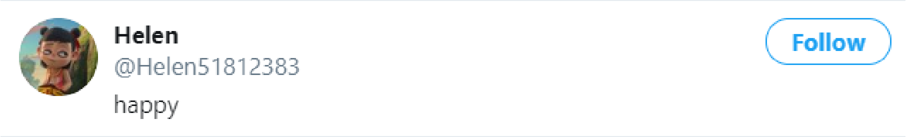
1.3K Retweets 2.1K Likes

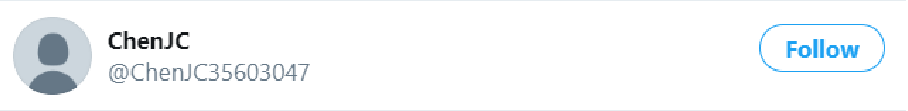
Our hypothesis was that inauthentic pro-China accounts would amplify tweets from prominent Chinese state media personnel. In this case, a lot of usernames stood out because they had an eight-digit number after the name, which indicated that the user had accepted the default username generated by Twitter when they signed up. That warranted further research into their behavior and characteristics.

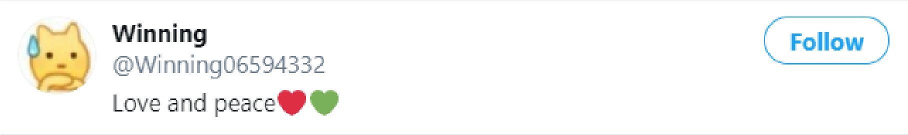

lqy🇨🇳🇨🇳
@lqy99021608
爱国爱党爱人民


wangsha_123
@s23244784


KANG
@KANG38396368


Helen
@Helen51812383
happy

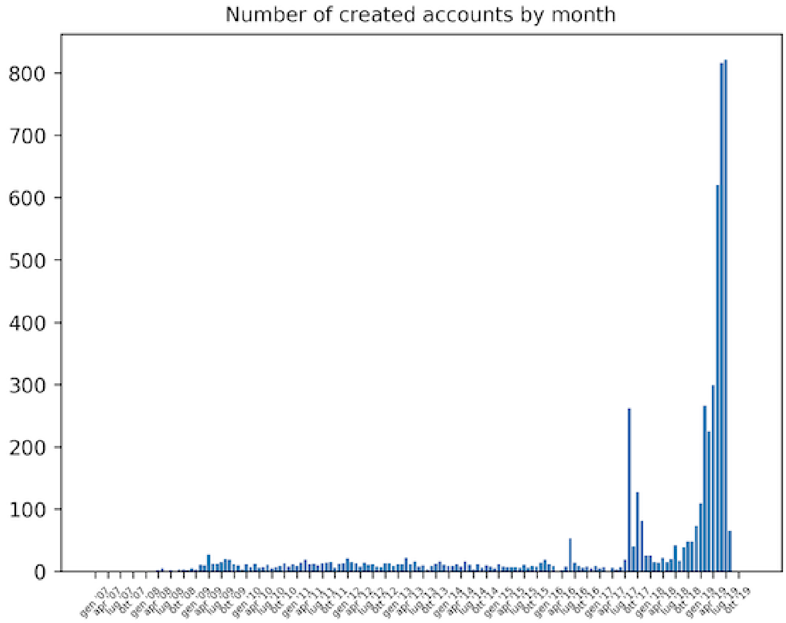

ChenJC
@ChenJC35603047


Winning
@Winning06594332
Love and peace❤️💚

As we examined these accounts, we saw they had a tiny number of followers, were following few accounts, offered no bio, were retweeting other people's tweets and sending almost none of their own, and almost exclusively promoted content in opposition to the protests.

We also noticed that the creation dates for these accounts were very recent, around August 2019. Because Twitter released a list of the pro-China accounts it removed, we could check the creation dates on those accounts and see if they showed a similar trend.

With the help of Luigi Gubello, a coder who is engaged in the online open-source community, we used a simple Python script (you can find the code on his GitHub and more about him here) to identify patterns in the data. The below graph shows that the removed accounts were all created in recent months, which aligned with the characteristics of the set of active accounts we were investigating.

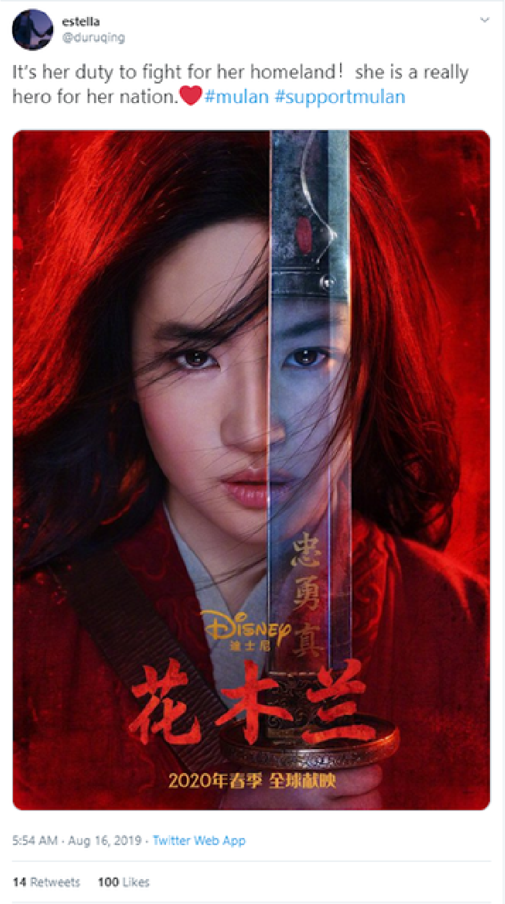


Automating the process

Now that we had identified a sample of tweets that exhibited suspicious characteristics and behavior, we wanted to conduct a much larger analysis. This required some automation. One Bellingcat workshop participant had a background in software development, so he wrote a small piece of JavaScript code — the regular expression (\w+\d{8}) — to perform two functions: extract the usernames of accounts that had retweeted or liked a specific tweet, and then quickly filter the username list so that it focused only on the usernames that matched a pattern. The pattern he filtered for was a name followed by an eight-digit number.

By loading this script in the Chrome developer tools console, which provides web developer tools directly in the browser, it would run in the background whenever he clicked on the "Retweets" or "Likes" hyperlink for a specific tweet. Then it would return results that highlighted usernames fitting the pattern. Go here to see what this looks like.

We could now use his script to examine the accounts interacting with other prominent pro-China tweets. In the midst of the Hong Kong protests, Chinese American actress Liu Yifei shared a Weibo post in support of the police, which led some people on social networks to call for a boycott of her new movie, "Mulan." However, we also noticed that many Twitter accounts supported the actress and her movie using the hashtag #SupportMulan. (CNN also reported on this.) We decided to use the script to examine the users who retweeted or liked the pro-*Mulan* tweets.

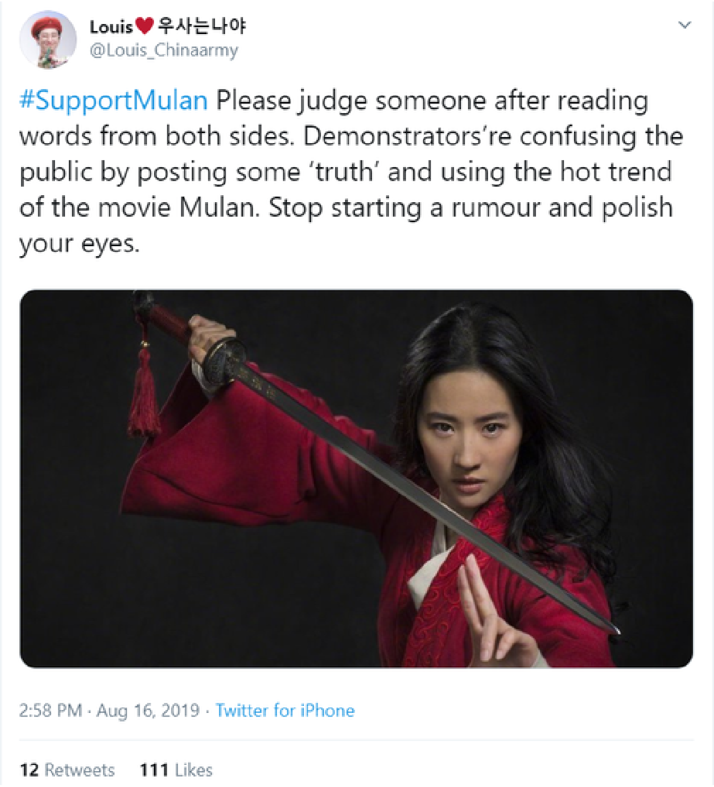


estella
@duruqing

It's her duty to fight for her homeland！ she is a really hero for her nation.❤️#mulan #supportmulan

5:54 AM · Aug 16, 2019 · Twitter Web App

14 Retweets 100 Likes



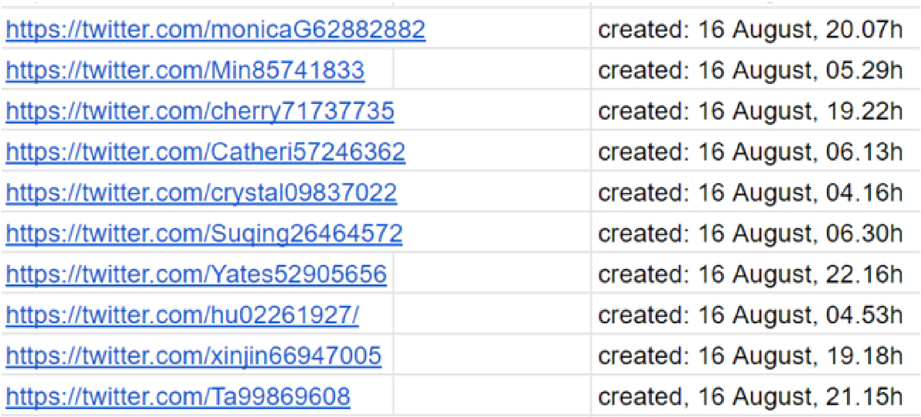
Louis❤우사는나야
@Louis_Chinaarmy

#SupportMulan Please judge someone after reading words from both sides. Demonstrators're confusing the public by posting some 'truth' and using the hot trend of the movie Mulan. Stop starting a rumour and polish your eyes.

2:58 PM · Aug 16, 2019 · Twitter for iPhone

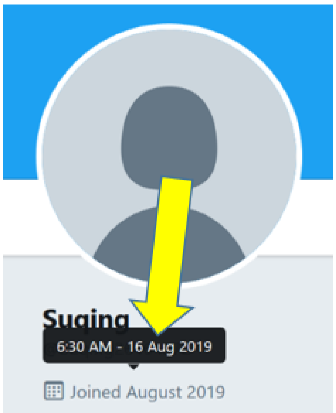
12 Retweets 111 Likes

We collected the account names that fit our pattern and then identified their creation dates. This revealed that most of the accounts were created on Aug. 16.



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

We gathered the exact creation date and time of the accounts by simply hovering over the profile's "joined" information, as shown below:

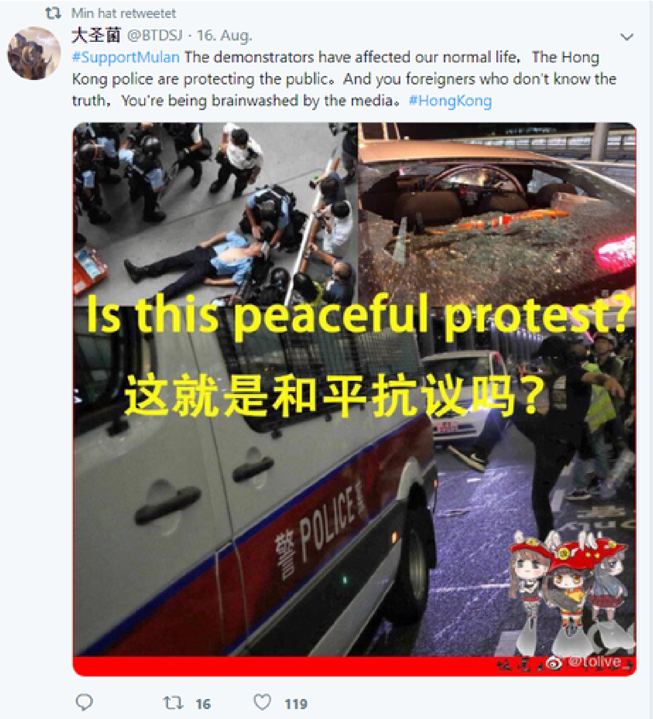


Suqing

6:30 AM - 16 Aug 2019

Joined August 2019

With the set of accounts in front of us, we began the manual analysis of the content they were sharing. It quickly became clear that the accounts in our list had all tweeted in favor of Yifei and against the Hong Kong protesters.



Min hat retweetet

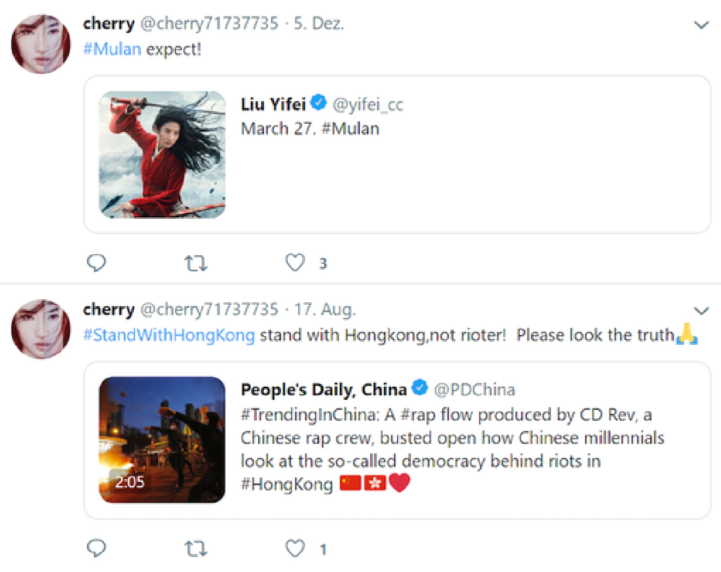
大圣菌 @BTDSJ · 16. Aug.

#SupportMulan The demonstrators have affected our normal life。 The Hong Kong police are protecting the public。 And you foreigners who don't know the truth，You're being brainwashed by the media。 #HongKong

16 119

Many of the accounts in our list became inactive after Aug. 17 or 18, which again showed an element of coordination. We don't know exactly why they went dormant, but it's possible that Twitter required additional verification steps for the creators to log in and they were unable to comply. Another option is that they simply stopped tweeting because the account creators did not want to raise further suspicion after Twitter started suspending pro-China accounts.

However, a few months later, we noticed that several of the accounts were active again. This time they spread positive messages about Yifei and her film, "Mulan."



cherry @cherry71737735 · 5. Dez.

#Mulan expect!

Liu Yifei @yifei_cc

March 27. #Mulan

3

cherry @cherry71737735 · 17. Aug.

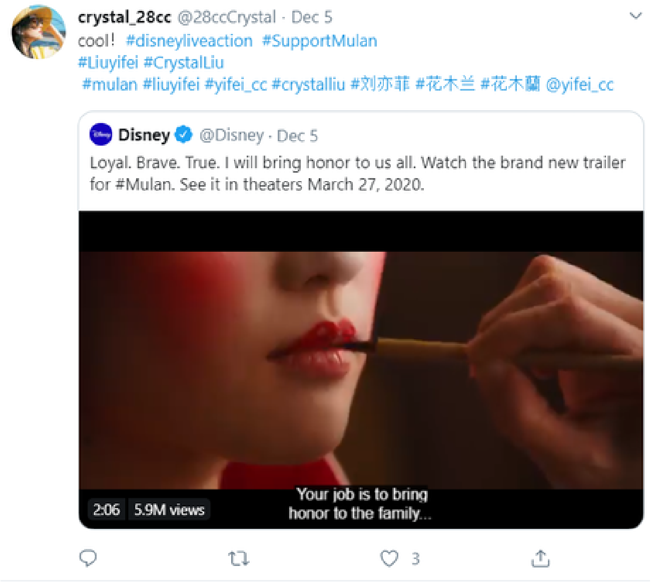
#StandWithHongKong stand with Hongkong,not rioter! Please look the truth🙏

People's Daily, China @PDChina

#TrendingInChina: A #rap flow produced by CD Rev, a Chinese rap crew, busted open how Chinese millennials look at the so-called democracy behind riots in #HongKong 🇨🇳🇭🇰❤️

1

We also found pro-"Mulan" accounts with other username patterns or creation dates that were continuously spreading messages in favor of Yifei. We did this by searching for tweets that included hashtags like #SupportMulan or #liuyifei

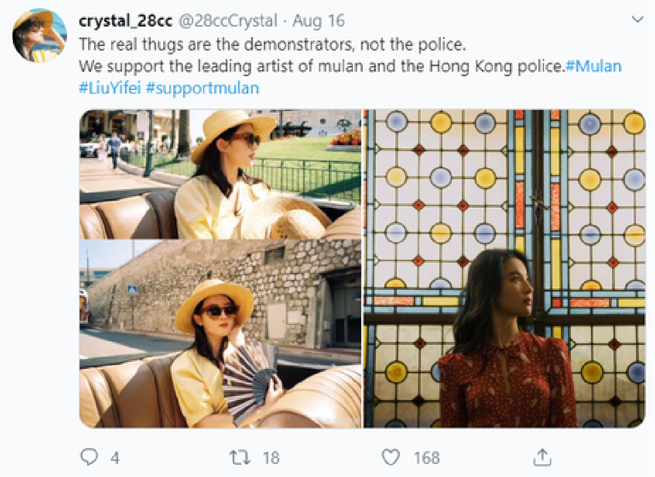


crystal_28cc @28ccCrystal · Dec 5

cool！ #disneyliveaction #SupportMulan
#Liuyifei #CrystalLiu
#mulan #liuyifei #yifei_cc #crystalliu #刘亦菲 #花木兰 #花木蘭 @yifei_cc

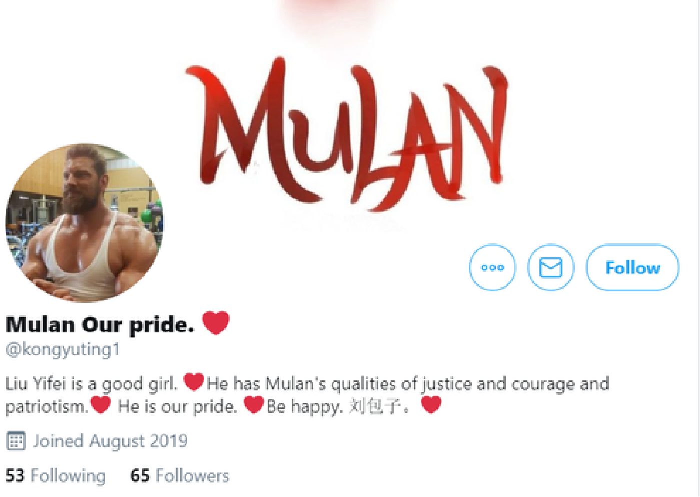
Disney @Disney · Dec 5

Loyal. Brave. True. I will bring honor to us all. Watch the brand new trailer for #Mulan. See it in theaters March 27, 2020.



crystal_28cc @28ccCrystal · Aug 16

The real thugs are the demonstrators, not the police.
We support the leading artist of mulan and the Hong Kong police.#Mulan
#LiuYifei #supportmulan



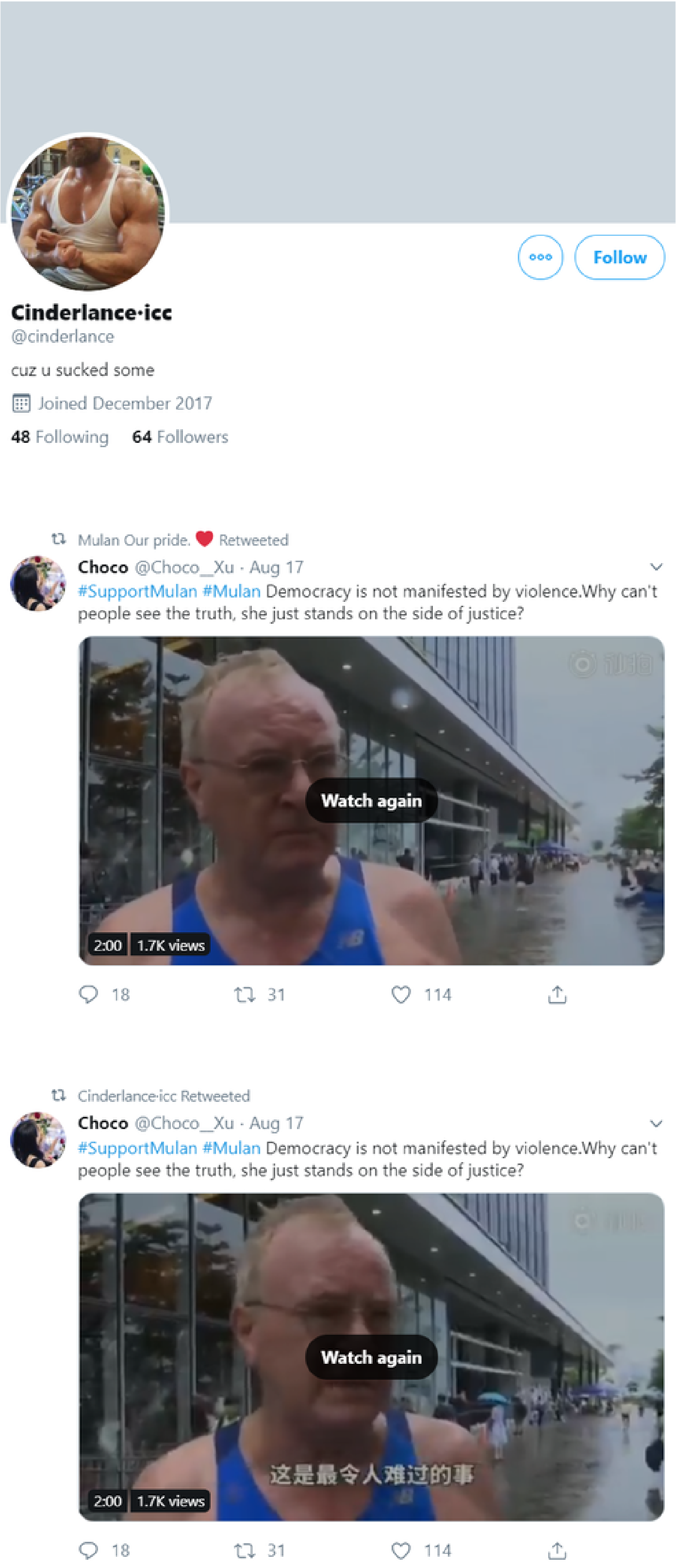
Mulan Our pride. ♥

@kongyuting1

Liu Yifei is a good girl. ♥He has Mulan's qualities of justice and courage and patriotism.♥ He is our pride. ♥Be happy. 刘包子。 ♥

Joined August 2019

53 Following 65 Followers



Cinderlance·icc
@cinderlance

cuz u sucked some

Joined December 2017

48 Following **64** Followers

Mulan Our pride. ❤ Retweeted

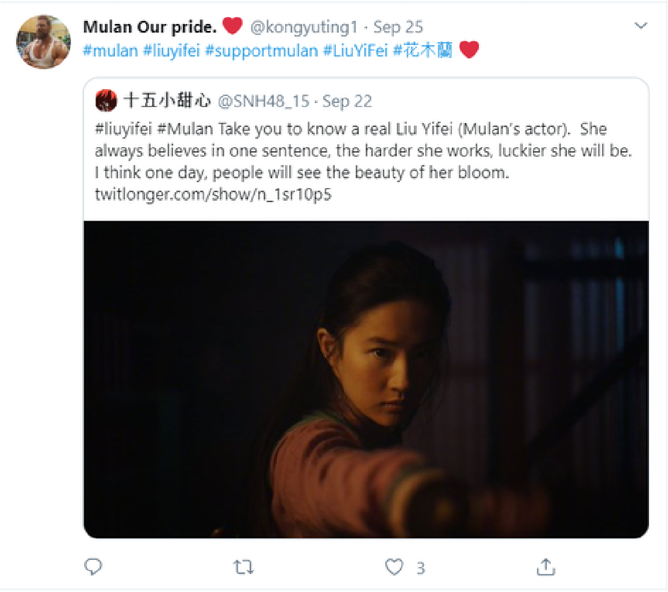
Choco @Choco_Xu · Aug 17
#SupportMulan #Mulan Democracy is not manifested by violence.Why can't people see the truth, she just stands on the side of justice?

18 31 114

Cinderlance·icc Retweeted

Choco @Choco_Xu · Aug 17
#SupportMulan #Mulan Democracy is not manifested by violence.Why can't people see the truth, she just stands on the side of justice?

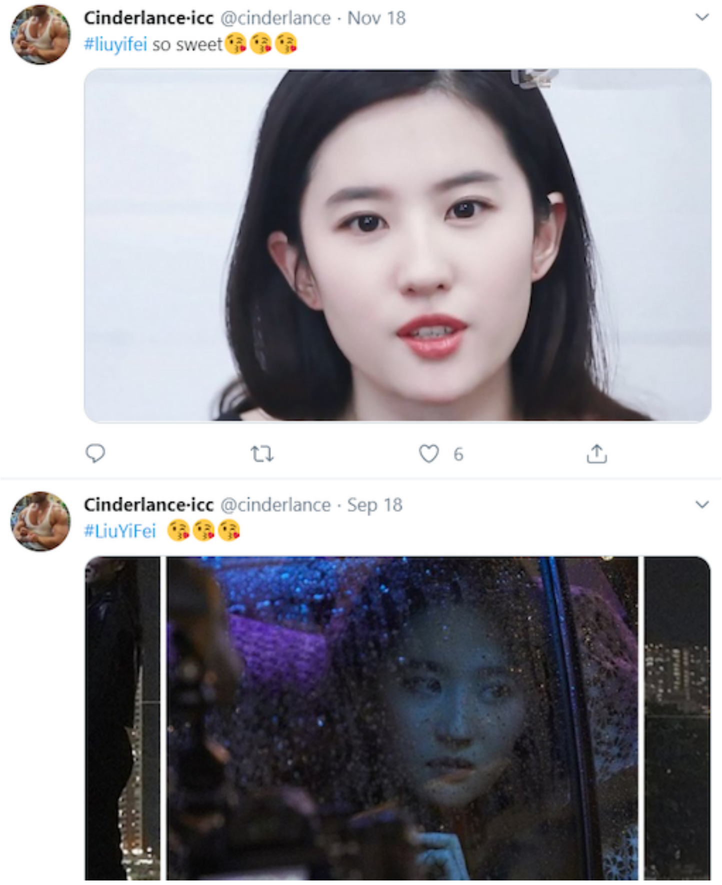
18 31 114



Mulan Our pride. ❤️ @kongyuting1 · Sep 25
#mulan #liuyifei #supportmulan #LiuYiFei #花木蘭 ❤️

十五小甜心 @SNH48_15 · Sep 22
#liuyifei #Mulan Take you to know a real Liu Yifei (Mulan's actor). She always believes in one sentence, the harder she works, luckier she will be. I think one day, people will see the beauty of her bloom.
twitlonger.com/show/n_1sr10p5

♡ 3



Cinderlance-icc @cinderlance · Nov 18
#liuyifei so sweet😘😘😘

♡ 6

Cinderlance-icc @cinderlance · Sep 18
#LiuYiFei 😘😘😘

It seems the accounts changed their strategy from criticizing the Hong Kong protesters to promoting the actress and her movie, perhaps to avoid being blocked from Twitter.

The case study shows how it's possible to combine manual and automated techniques to quickly uncover a network of suspicious Twitter accounts. It also illustrates that it's useful to look for additional accounts and activity even after a platform announces a takedown of accounts.

Here, we were able to use some simple search techniques and account details to identify a larger set of accounts that showed strong indicators of being engaged in coordinated inauthentic activity.

4. Monitoring for fakes and information operations during breaking news

Written by: Jane Lytvynenko

***Jane Lytvynenko** is a senior reporter at BuzzFeed News, where she focuses on disinformation, cyber security and online investigations. She has uncovered social media manipulation campaigns, cryptocurrency scammers and financially motivated bad actors spreading disinformation. Her work also brings accessible fact-checking to wide audiences during times of crisis. Jane is from Kyiv, Ukraine, and currently resides in Toronto, Canada.*

When news breaks, it can be hours or even days until reporters and officials are fully able to make sense of a situation. As early evidence and details begin to flow over social networks and other online platforms, bad actors can emerge to sow division or distrust, or make a quick buck off a worried news consumer's attention. Those same well-meaning consumers and other sources can also unintentionally spread false or misleading information. The mix of heightened emotions and slow trickle of news in the early minutes and hours of an event makes it necessary for journalists to be equipped to effectively monitor, verify and — when necessary — debunk breaking news. A fake tweet, image, social media account or article takes just a few minutes to create, while real information struggles to keep up.

The key to monitoring and debunking during breaking news is to lay a strong foundation before it happens. This means having a solid grounding in basic verification skills, such as those outlined in the first Verification Handbook, understanding how to monitor social networks and platforms, and knowing how to respond if you or your colleagues become targeted by bad actors. Reporters should never put online safety on the back burner.

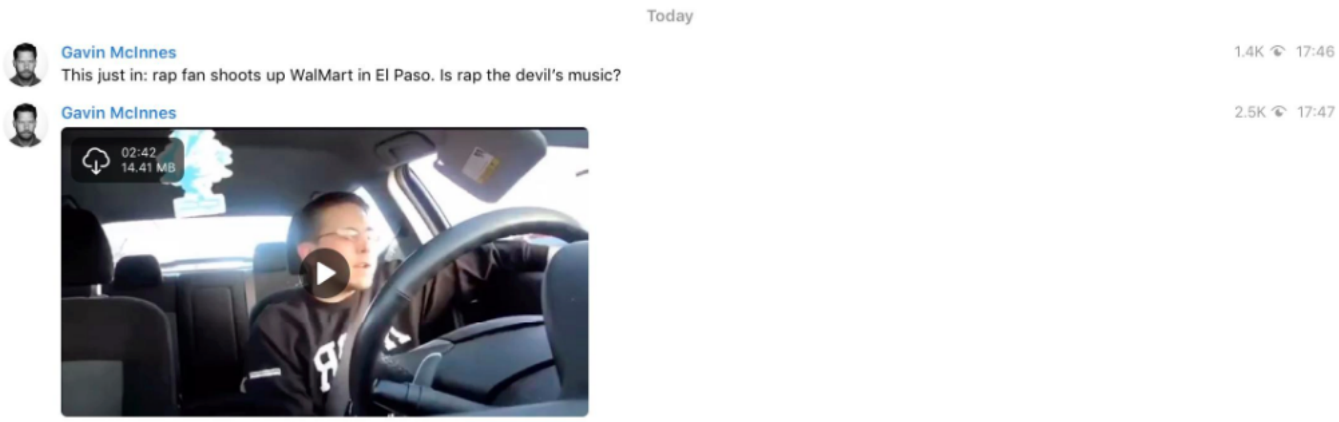
When news breaks, the first step is to identify key impacted communities. During the 2018 shooting at the high school in Parkland, Florida, reporters scoured the Snapchat map for videos of what was happening to the students trapped inside classrooms. In contrast, during Hurricane Irma in 2017, it was key to focus on Facebook, where those affected tried to find information. Understanding how each social network functions and how it intersects with a given event is essential.

This chapter will focus on tools a reporter can use for monitoring and debunking breaking news. Not every tool will be right for every situation, and understanding who has been affected can help you know which places to focus on most.

Three things to look for

As platforms and reporters work hard to fight disinformation, bad actors have evolved their tactics to avoid detection. Still, some consistent patterns of content and behavior emerge repeatedly.

1. Doctored or out-of-context imagery. The infamous image of a shark swimming on a flooded highway has been making rounds and continuing to trick people for years. (It was also the subject of a case study in the first Handbook.) Photos and videos from that have previously debunked are what fact-checkers and debunkers call zombie hoaxes and are important to watch for. Imagery spreads much faster across digital platforms than text, so focusing on them is often fruitful.



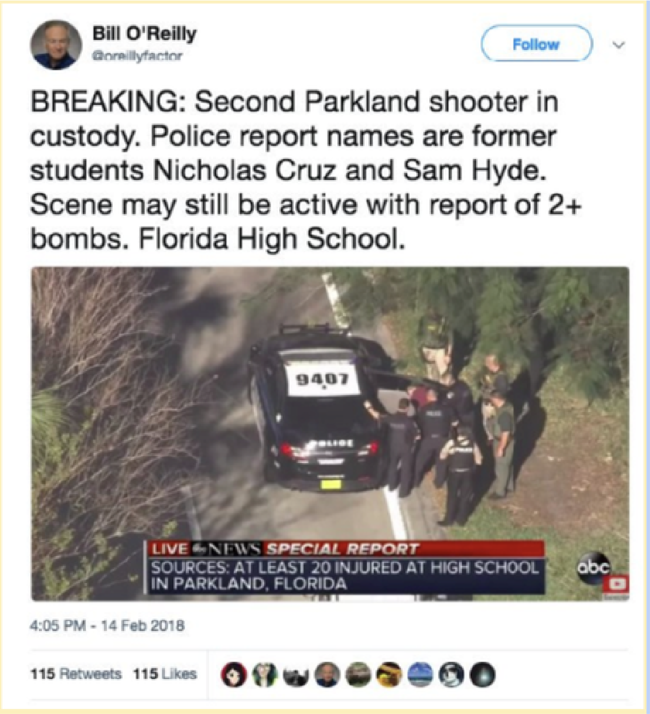
Today

Gavin McInnes
This just in: rap fan shoots up WalMart in El Paso. Is rap the devil's music?
1.4K 17:46

Gavin McInnes
2.5K 17:47

During the El Paso shooting at a Walmart in 2019, far-right figures tried to misrepresent an old YouTube video unrelated to the suspect.

2. False victims or perpetrators. During the YouTube headquarters shooting, social networks were littered with false claims of suspects. During the U.S. midterm elections in 2018, false rumors about ballots being cast by illegal immigrants were spread by the U.S. president. False perpetrators show up during most big breaking news events.



Bill O'Reilly
@oreillyfactor

Follow

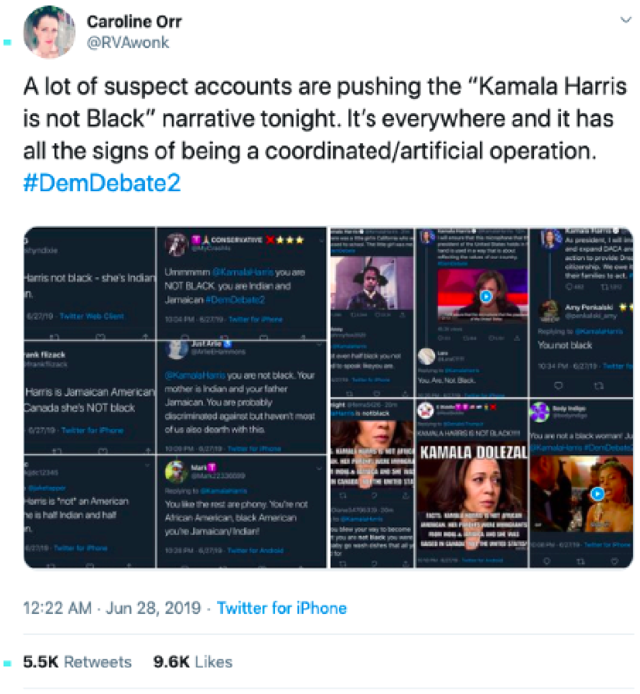
BREAKING: Second Parkland shooter in custody. Police report names are former students Nicholas Cruz and Sam Hyde. Scene may still be active with report of 2+ bombs. Florida High School.

4:05 PM - 14 Feb 2018

115 Retweets 115 Likes

During the 2018 Parkland shooting, a fake Bill O'Reilly account tried to spread a false name for the suspect.

3. Harassment and brigading. While not strictly disinformation, bad actors commonly try to harass people involved in a news event as a way of silencing them. It's also a sign that a group of people is paying attention to an event and may try different tactics down the road. “Brigading” is when a group of people work together to create the impression of a groundswell of engagement or reaction, by doing things such as up- or down voting content or flooding a user with comments.



Caroline Orr
@RVAwonk

A lot of suspect accounts are pushing the "Kamala Harris is not Black" narrative tonight. It's everywhere and it has all the signs of being a coordinated/artificial operation. #DemDebate2

12:22 AM · Jun 28, 2019 · Twitter for iPhone

5.5K Retweets 9.6K Likes

After a Democratic leadership debate in 2019, anonymous accounts spread the same message about Kamala Harris's race.

Best practices for archiving and publishing

Before looking for hoaxes, set up a folder for your documents and start a spreadsheet for what you find. Immediately take a screenshot of each hoax and relevant piece of content you discover, and archive the page. (The Archive.org web browser extension is a free, quick and effective tool for archiving content.) Be sure to record the original and archived URLs of the content in your spreadsheet. This enables you to come back to what you found and look for patterns after the dust settles.

To avoid helping spread pages associated with dis- or misinformation, be sure to link to the archived URL in any articles or social media posts instead of the original. It's also a best practice to watermark your images with a clear label such as "False" or "Misleading" to ensure they are spread and indexed with the proper context. If you do write an article, focus your headline and copy on what's true, rather than primarily saying what's false. Studies have shown that repeating falsehoods can cause people to retain the incorrect information

Your role is to minimize the repetition of falsehoods as much as possible, and to steer people toward accurate information.

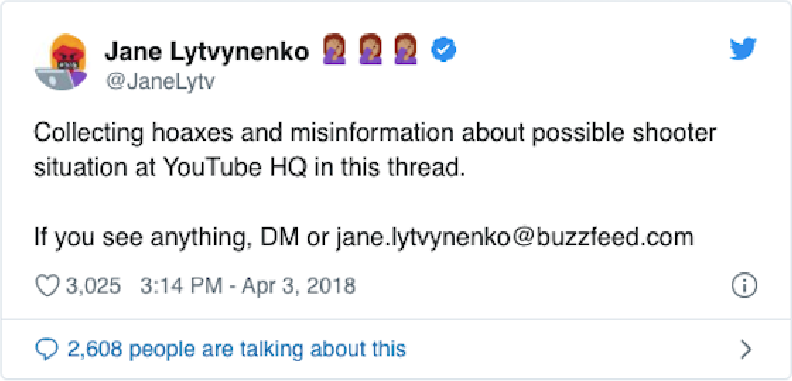
Identifying keywords and locations

As the event unfolds, come up with a list of locations and relevant keywords.

For location, take into account the city, state and country, and any relevant local terms such as the nickname for a city or affected neighborhood. During elections, you should also use the county or relevant electoral district name. This information is used to monitor geotagged posts and to search for mentions of the location. Also be sure to identify and begin monitoring the social accounts of any relevant local authorities, such as police and fire departments, politicians and local news outlets.

Next, identify key terms. This can include words like victim, suspect, shooter, shooting, flood, fire, the confirmed names of anyone involved and more general wording like "looking for" — think of the language people would use in the situation aside from key terms. If you find a credible account posting about being in the midst of the event you're monitoring, note their username and read their full feed. Looking through their friends or followers list is also a helpful way to find others in the area who might have been affected.

Note that during stressful situations, people may misspell locations or names. For example, during the 2019 Kincade fire in California, some tweeted #kinkaidfire because of autocorrect issues. Include common misspellings in your searches and try to identify possible autocorrect mistakes by typing key terms on your device and watching what suggestions pop up.



Jane Lytvynenko 🤦‍♀️🤦‍♀️🤦‍♀️ ✓
@JaneLytv

Collecting hoaxes and misinformation about possible shooter situation at YouTube HQ in this thread.

If you see anything, DM or jane.lytvynenko@buzzfeed.com

♡3,025 3:14 PM - Apr 3, 2018

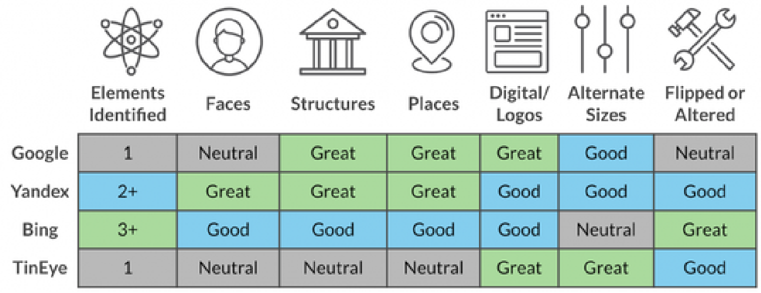
2,608 people are talking about this

This is also a good time to reach out to any sources you know in the relevant location or who are part of communities that might be targeted with harassment or disinformation, and ask what they've seen online. You can tell your audience that you're on the lookout for disinformation and other problematic content related to the event. Coordinate with your newsroom's social media team to help spread the word about your monitoring and to see if they have seen anything of note.

Key Image Tools

1. Image search

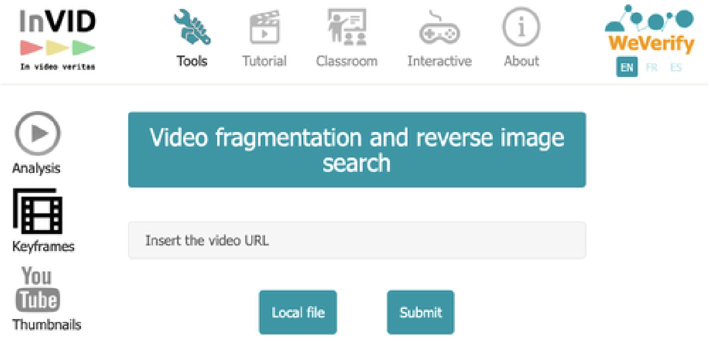
Reverse image search is an indispensable tool. It's easy to search Google for an image by right-clicking on an image and selecting "Search Google for Image" in the Chrome web browser. But it's always a good idea to search an image using different tools. If you install the InVID browser extension, you can right-click on an image and search it across different tools. This reverse image search comparison chart created by Domain Tools shows the relevant strengths and weaknesses of different reverse image products:



	Elements Identified	Faces	Structures	Places	Digital/ Logos	Alternate Sizes	Flipped or Altered
Google	1	Neutral	Great	Great	Great	Good	Neutral
Yandex	2+	Great	Great	Great	Good	Good	Good
Bing	3+	Good	Good	Good	Good	Neutral	Great
TinEye	1	Neutral	Neutral	Neutral	Great	Great	Good

InVID

InVID is a free browser extension and the best platform for helping you analyze and verify videos. It allows for users to paste a URL into its engine, which will then extract thumbnails from the video. You can run reverse image searches on these thumbnails to see where else this video has appeared on the web.



2. TweetDeck search

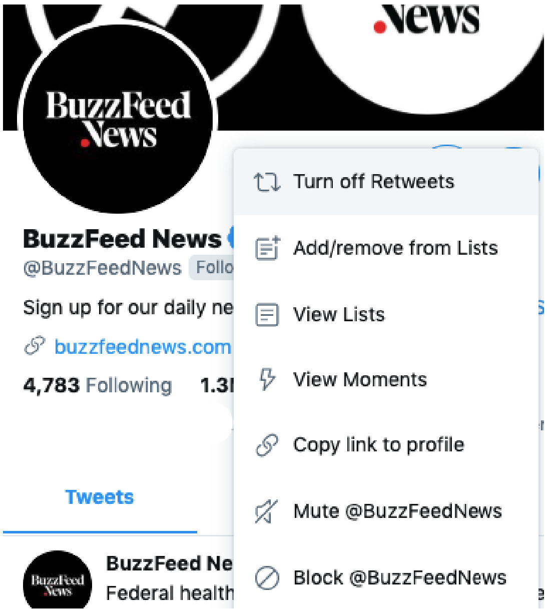
The best way to search Twitter is by using TweetDeck, which allows you to create unique columns for searches and lists.

Finding and duplicating relevant lists is key for staying abreast of a situation. You can use Google to search for Twitter lists using a simple formula. Type *site:twitter.com/*/lists* into the search engine and then add a key word in quotes, for example “Alabama reporters.” The final search string is therefore:

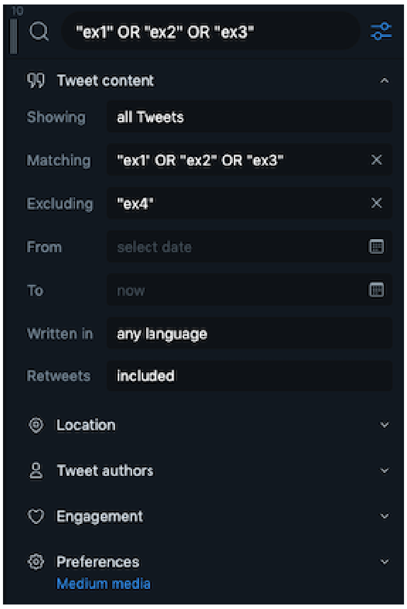
site:twitter.com//lists “Alabama reporters”*

This will bring up any lists that other Twitter users have created that include the phrase “Alabama reporters” in the title.

Once you’ve found a list that’s relevant for your needs, you need to duplicate it so you can add it to TweetDeck. Use this app: http://projects.noahliebman.net/listcopy/connect.php to duplicate as many as you like. It’s ideal to duplicate a list rather than to follow it because you can add or remove users as you like.



Along with finding and adding lists to TweetDeck columns, you want to create columns with specific search filters that enable you to quickly monitor for keywords as well as images and videos. To look for multiple keywords, wrap them in quotes and put “OR” between them, such as “Kincade” OR “Kinkade.” You can also exclude certain words if they produce irrelevant results. Most people no longer tag their tweets by location, so you can leave that field blank to cast a wider net.



"ex1" OR "ex2" OR "ex3"

Tweet content

Showing: all Tweets

Matching: "ex1' OR "ex2" OR "ex3"

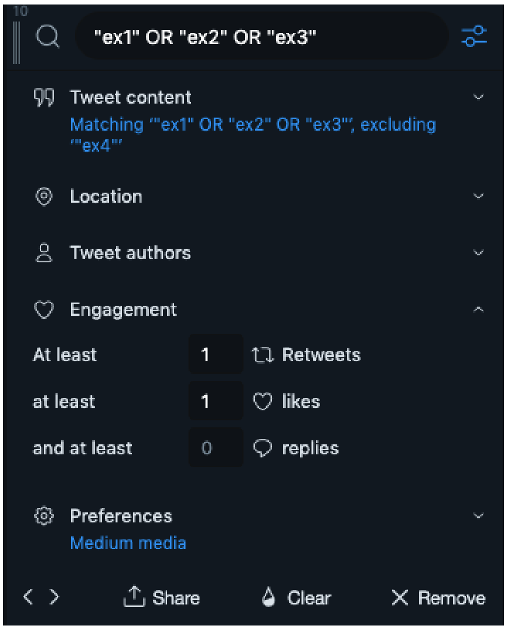
Excluding: "ex4"

From: select date

To: now

Written in: any language

Retweets: included



"ex1" OR "ex2" OR "ex3"

Tweet content

Matching '"ex1" OR "ex2" OR "ex3"', excluding '"ex4"'

Engagement

At least 1 Retweets

at least 1 likes

and at least 0 replies

If you want to narrow your results, set the “From” field to a day or two before the event took place, as this will make sure you don’t miss tweets because of possible time zone issues. If you’re still getting too many results, try filtering them by engagement to surface only the posts that others have liked or retweeted. You can also try breaking key terms into separate columns. For example, put locations in one column and other keywords in another. I usually break out a third column for possible names of suspects or victims and their misspellings.

Finally, if you’re seeing a very high volume of tweets, it’s a good idea to create a new column with your best keywords and to set the “Showing” option under the “Tweet content” filter to show only photos and videos. This will give you a feed that can help you spot viral or emerging visuals.

3. CrowdTangle

CrowdTangle is a web app and browser extension that’s free for newsrooms to use. (Contact the company if your newsroom is not set up with access.)

It’s a powerful tool that allows you to set up dashboards to monitor across Facebook, Instagram and Reddit. You can also search by keyword and set many filters, including time posted, language and engagement. CrowdTangle is especially useful for monitoring Facebook and checking where a URL may have been posted on social media.

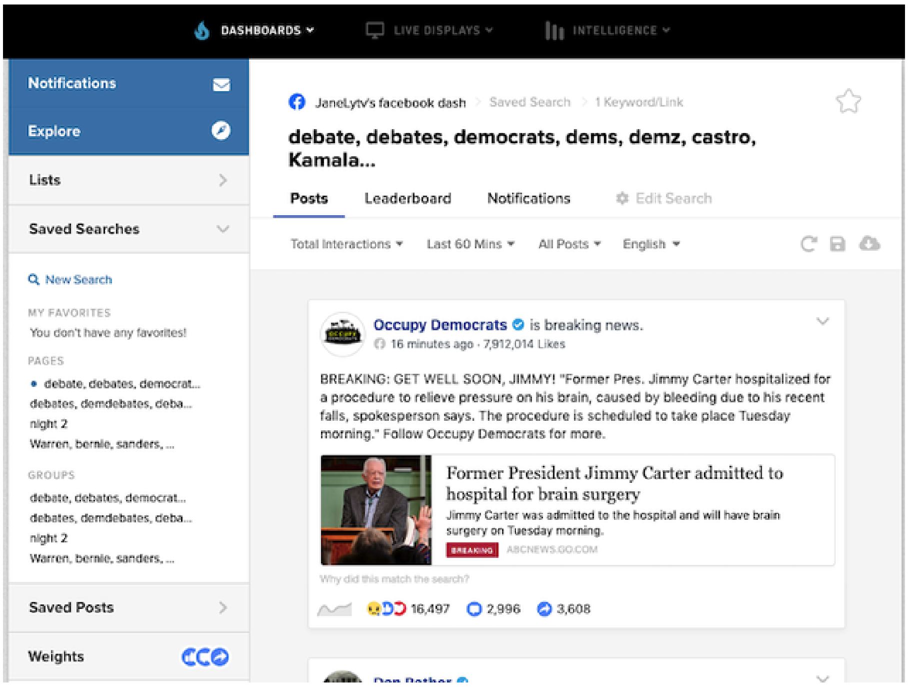
Once you have access, go to app.crowdtangle.com to get started and then click “Create New Dashboard.” Even if you don’t have access, the browser extension is free for anyone to use.

CrowdTangle: Searching for Facebook posts

Click on “Saved Searches” on the left sidebar and then “New Search.” You have two options with Facebook: search pages and search groups. I’d recommend doing both. Enter as many keywords as you like by separating them with commas. Then you can set how you see the posts, for example by most recent, most popular and overperforming, which is a measurement of posts receiving more engagement than is normal for a given page. I toggle among the three based on the situation to make sure I see viral content and new content.

You’re also able to sort posts by a specific time frame and type. CrowdTangle recently added the ability to search posts by the location of the page they were posted by. By clicking on “English” and then picking “Country,” you can select only posts that are coming from pages that have declared their location to be within the U.S., for example. You can also do the opposite and search for posts coming from pages based in Iran, Russia, Saudi Arabia, the Philippines or India, for example. Keep a special eye on image and video-based posts, which tend to spread further and be more engaging.

Once you’ve set up a search with relevant results be sure to save it so you can keep coming back to it.



debate, debates, democrats, dems, demz, castro, Kamala...

Occupy Democrats is breaking news.
16 minutes ago · 7,912,014 Likes

BREAKING: GET WELL SOON, JIMMY! "Former Pres. Jimmy Carter hospitalized for a procedure to relieve pressure on his brain, caused by bleeding due to his recent falls, spokesperson says. The procedure is scheduled to take place Tuesday morning." Follow Occupy Democrats for more.

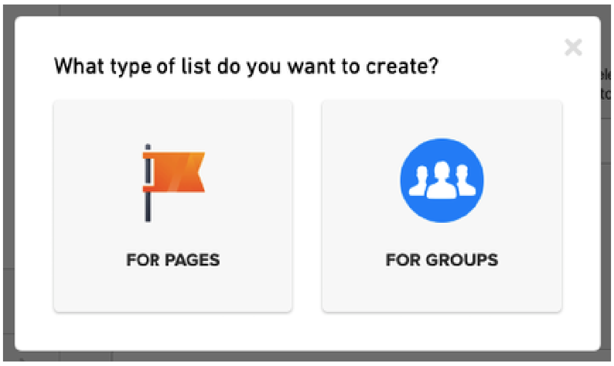
Former President Jimmy Carter admitted to hospital for brain surgery

Jimmy Carter was admitted to the hospital and will have brain surgery on Tuesday morning.

BREAKING ABCNEWS.GO.COM

CrowdTangle: Lists

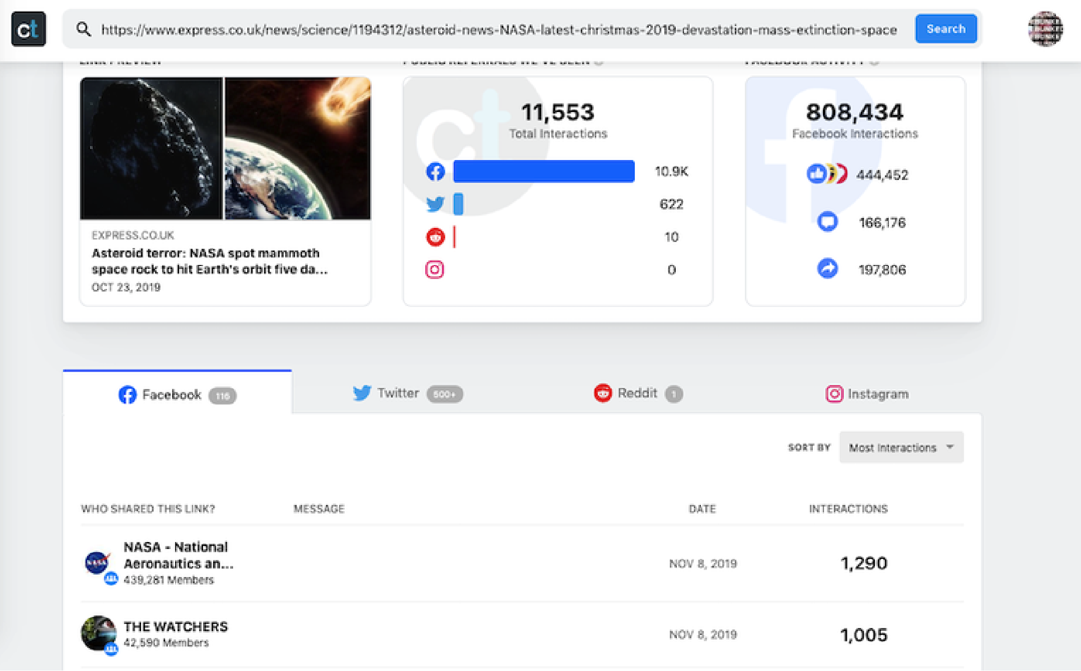
Like TweetDeck, CrowdTangle allows you to build lists of pages and public groups of interest. By clicking “Lists” on the left sidebar and then “Create List,” you can monitor pages or groups that match keywords you have chosen or pages whose URL you have. CrowdTangle also has a number of prebuilt lists you can view by clicking the “Explore” tab. As with Twitter, building lists of pages and groups talking about the event you’re covering is a good way to monitor the information environment.



What type of list do you want to create?

CrowdTangle: Link search

Another relevant CrowdTangle feature is link search. Go to https://apps.crowdtangle.com/search/ and paste in the URL or key terms of the content you're interested in. CrowdTangle will show you the top public sharers of the link across Facebook, Instagram, Reddit and Twitter. (Note that the Twitter results are restricted to the previous seven days.) This will help you understand how the content is spreading, whether there are any groups or individuals you should be investigating further, and whether the content has spread far enough to warrant a debunk. There are no simple rules on when to debunk a falsehood, but some good questions to ask are: Has it spread outside of its initial network of sharers? Has it been shared by figures of authority? Has it generated significant engagement? (The free browser extension delivers the same data as the link search tool, and both are free for anyone to use without a full CrowdTangle account.)



https://www.express.co.uk/news/science/1194312/asteroid-news-NASA-latest-christmas-2019-devastation-mass-extinction-space

EXPRESS.CO.UK
Asteroid terror: NASA spot mammoth space rock to hit Earth's orbit five da...
OCT 23, 2019

11,553 Total Interactions

Facebook	10.9K
Twitter	622
Reddit	10
Instagram	0

808,434 Facebook Interactions

👍	444,452
💬	166,176
↗	197,806

SORT BY Most Interactions

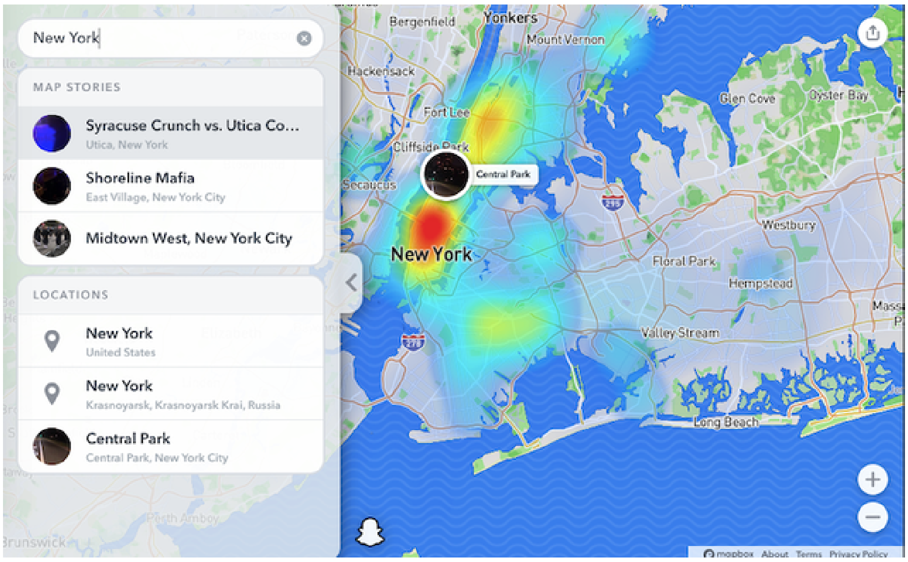
WHO SHARED THIS LINK?	MESSAGE	DATE	INTERACTIONS
NASA - National Aeronautics an... 439,281 Members		NOV 8, 2019	1,290
THE WATCHERS 42,590 Members		NOV 8, 2019	1,005

4. Instagram.com

Instagram is a useful place to monitor for hashtags and geotagged posts. Look up relevant locations where users may have tagged photos, and remember that location tags can also include neighborhoods and landmarks. Once you found someone who appears to have been involved in a news event, click through to their account and make sure you watch their stories — they're by far more popular than regular Instagram posts. Also look through the comments for other potential witnesses, and note any new hashtags that may have been used alongside their posts. If you want to archive someone's Instagram story for your files, you can use a site like storysaver.net to download it.

5. SnapMap

Disinformation on Snapchat is uncommon, but its public map feature is useful to help verify or debunk hoaxes. To get started, go to map.snapchat.com and enter a location of interest. This will show you a heat map of where content is being posted — the brighter the location, the more Snaps are coming from there. To save a useful Snap, click on three dots in the top right and select "Share." You'll be able to copy the URL of the Snap to look at later. (Be sure to screenshot it as well.)



Putting it all together

It's essential to practice using each tool before news breaks to avoid scrambling in the moment. Disinformation is meant to play on emotions and capitalize on gaps in news coverage. Keep that in mind as you search the web. You will also often come across accurate information that could help your colleagues. Write down everything that you know is true so you can recognize false things faster, and don't be afraid to ask any reporters your outlet has on the ground for help.

After the dust settles, it's helpful to look back at your saved images and posts. While in the moment you want to highlight individual falsehoods by way of public service journalism, in the aftermath you should take stock of any themes or patterns that can be seen. Were people targeted for their race or gender? Did hoaxes that originated on small, anonymous accounts become mainstream? Did any social media companies perform especially well or especially poorly? A wrap-up story can help your readers fully grasp the purpose and methods of the disinformation's spread. It will also serve as a research tool for you and your newsroom, showing you what might be useful to focus on the next time news breaks.

5. Verifying and questioning images

Written by: Hannah Guy , Farida Vis , Simon Faulkner

Farida Vis *is director of the Visual Social Media Lab and a professor of digital media at Manchester Metropolitan University. Her academic and data journalism work focuses on the spread of misinformation online. She has served on the World Economic Forum's Global Agenda Council on Social Media (2013-2016) and the Global Future Council for Information and Entertainment (2016-2019) and is a director at Open Data Manchester.*

Simon Faulkner *is a lecturer in art history and visual culture at Manchester Metropolitan University. His research is concerned with the political uses and meanings of images, with a particular focus on activism and protest movements. He is also the co-director of the Visual Social Media Lab and has a strong interest in the development of methods relevant to the analysis of images circulated on social media.*

Hannah Guy *is a Ph.D. student at Manchester Metropolitan University, examining the role of images in the spread of disinformation on social media. She is a member of the Visual Social Media Lab, where her current projects explore images shared on Twitter during the emergence of the Black Lives Matter movement, and Visual Media Literacy to combat misinformation in the context of Canadian schools.*

Communication on social media is now overwhelmingly visual. Photos and video are persuasive, compelling and easier to create than ever, and can trigger powerful emotional responses. As a result, they have become powerful vehicles of mis- and disinformation.

To date, the discussion of images within the context of mis- and disinformation has either focused on verification techniques or, more recently, been disproportionately focused on deepfake videos. Before considering deepfakes, as we do in the next chapter, it's essential to understand the more common low-tech use of misleading photos and videos, especially those shown out of context.

Given the widespread use of visuals in attempts to influence and manipulate public discourse, journalists must be equipped with fundamental image verification knowledge *and* with the ability to critically question and assess images to understand how and why they are being deployed. This chapter focuses on developing this second set of skills, and uses a framework we developed at the Visual Social Media Lab.

Building on verification

In the Visual Social Media Lab, we focus on understanding the roles online images play within society. While we mainly focus on still images, this also encompasses a range of different types of images: photos, composite images, memes, graphic images and screenshots, to name a few. Tackling visual mis- and disinformation requires its own set of strategies. To date, image verification by journalists has focused on establishing if the image is what they think it is. In the original "Verification Handbook," Trushar Barot outlined four core basic principles for image verification, which remain invaluable. The First Draft Visual Verification Guide is another useful resource that uses these principles by focusing on five questions for photos and videos:

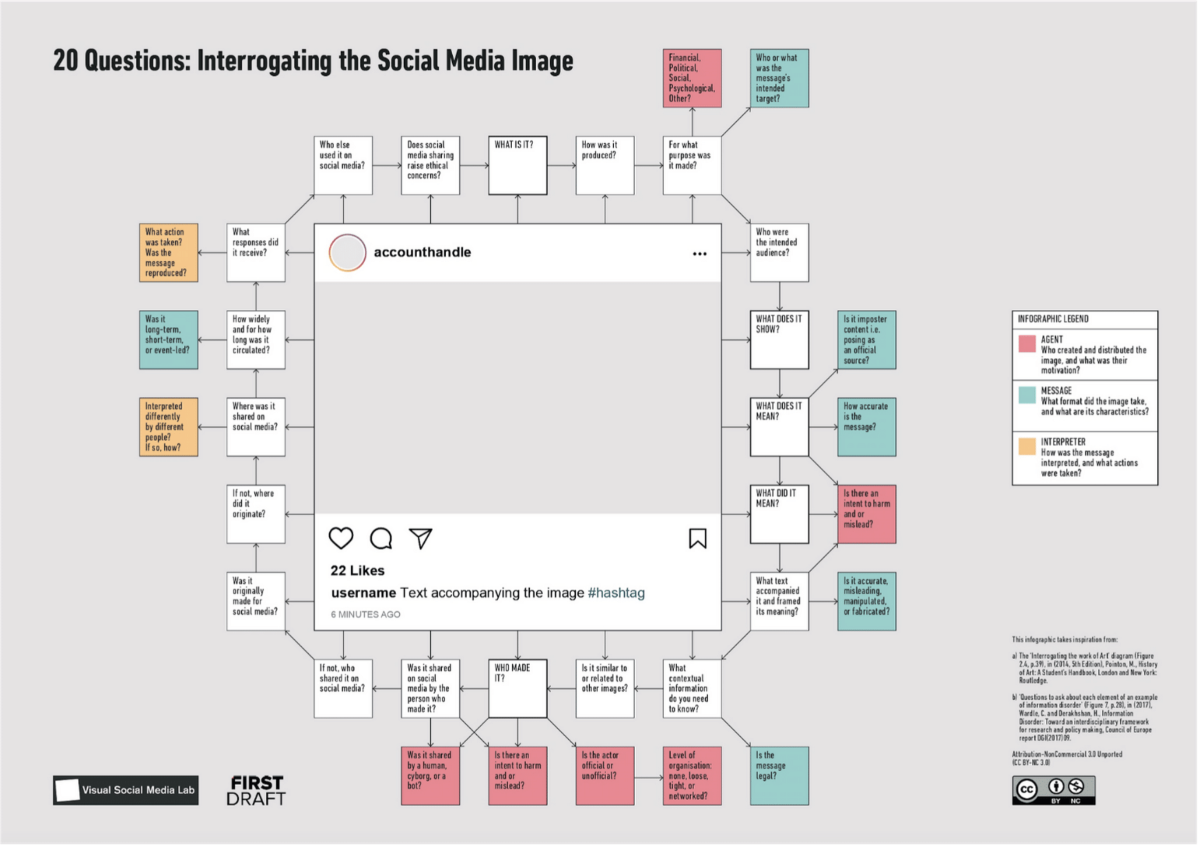
1. Are you looking at the original version?
2. Do you know *who* captured the photo?
3. Do you know *where* the photo was captured?
4. Do you know *when* the photo was captured?
5. Do you know *why* the photo was captured?

Standard tools that can help with investigating photos and video include InVID, Yandex Image Search, TinEye, Google Image Search and Forensically. These verification methods focus on the origin of the image.

While that method remains crucial, the strategies and techniques frequently used in mis- and disinformation, and in a range of forms of media manipulation, mean it is also important to consider how images are used and shared and by whom, and also what role journalists play in potentially further amplifying problematic images.

To go beyond standard forms of image verification, we have combined methods from art history with questions designed specifically for mis- and disinformation content. Our framework, "20 Questions for Interrogating Social Media Images," designed collaboratively with First Draft and journalists, is an additional tool journalists can use when investigating images.

Interrogating social media images



20 Questions: Interrogating the Social Media Image

Financial, Political, Social, Psychological, Other?

Who or what was the message's intended target?

Who else used it on social media?

Does social media sharing raise ethical concerns?

WHAT IS IT?

How was it produced?

For what purpose was it made?

What action was taken? Was the message reproduced?

What responses did it receive?

Who were the intended audience?

Was it long-term, short-term, or event-led?

How widely and for how long was it circulated?

WHAT DOES IT SHOW?

Is it imposter content i.e. posing as an official source?

Interpreted differently by different people? If so, how?

Where was it shared on social media?

WHAT DOES IT MEAN?

How accurate is the message?

If not, where did it originate?

WHAT DID IT MEAN?

Is there an intent to harm and or mislead?

Was it originally made for social media?

What text accompanied it and framed its meaning?

Is it accurate, misleading, manipulated, or fabricated?

If not, who shared it on social media?

Was it shared on social media by the person who made it?

WHO MADE IT?

Is it similar to or related to other images?

What contextual information do you need to know?

Was it shared by a human, cyborg, or a bot?

Is there an intent to harm and or mislead?

Is the actor official or unofficial?

Level of organisation: none, loose, tight, or networked?

Is the message legal?

INFOGRAPHIC LEGEND

🟥 AGENT
Who created and distributed the image, and what was their motivation?

🟦 MESSAGE
What format did the image take, and what are its characteristics?

🟧 INTERPRETER
How was the message interpreted, and what actions were taken?

This infographic takes inspiration from:

a) The 'Interrogating the work of Art' diagram (Figure 2.4, p.39), in (2014, 5th Edition), Pointon, M., History of Art: A Student's Handbook, London and New York: Routledge.

b) 'Questions to ask about each element of an example of information disorder' (Figure 7, p.28), in (2017), Wardle, C. and Derakhshan, H., Information Disorder: Toward an interdisciplinary framework for research and policy making, Council of Europe report DGI(2017)09.

As the title suggests, the framework consists of 20 questions that can be asked of any social media image (still image, video, gif, etc.), with an additional 14 questions aimed at digging deeper into different aspects of mis- and disinformation. The questions do not appear in a set order, but these five questions are useful to address first:

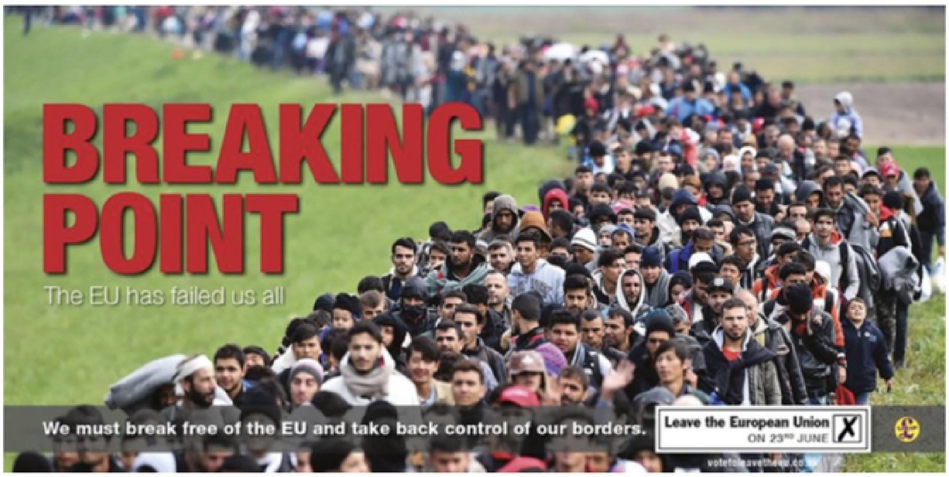
1. What is it?
2. What does it show?
3. Who made it?
4. What *did* it mean?
5. What does it mean?

Questions 1 to 3 are similar to established approaches to verification and are concerned with establishing what kind of image it is (a photograph, video, etc.), what it depicts and who made it. However, questions 4 and 5 take us somewhere else. They introduce considerations of meaning that encompass what the image shows, but also cover any meanings produced by the use of the image, including through its misidentification. When thought about together, questions 4 and 5 also allow us to focus on the changing nature of the meaning of images and on the ways that meanings ascribed to images through reuse can be significant in themselves. This doesn't simply concern what images are made to mean in a new context and how this misidentifies what they show, but also what the effects of such misidentifications are. This approach is no longer about verification, but more akin to the analysis of the meanings of images performed in disciplines such as art history and photo theory.

In the development and early deployment of this framework with journalists, we often heard that they had never thought about images in this much detail. Many said the framework helped them recognize that images are complex forms of communication, and that a clear method is required to question them and their meaning.

Most of the time, you will not need to answer all 20 questions in the framework to get a comprehensive understanding of what's going on with an image. The questions are there to fall back on. In our own work, we found them particularly useful when dealing with complex high-profile news images and videos that have received significant media attention and scrutiny. To show what this looks like in practice, here are three case studies with high-profile examples from the U.K. and U.S.

Case Study 1: Breaking Point, June 2016


BREAKING POINT
The EU has failed us all
We must break free of the EU and take back control of our borders.
Leave the European Union X
ON 23RD JUNE

What is it?

The "Breaking Point" image was a poster used by the UK Independence Party (UKIP) as part of its campaign during the EU referendum of 2016. It used a photograph taken by the photojournalist Jeff Mitchell in October 2015, focused on the refugee crisis.

What does it show?

A large queue of Syrian and Afghan refugees being escorted by Slovenian police from the border between Croatia and Slovenia to the Brezice refugee camp. The poster used a cropped version and added the text "BREAKING POINT: The EU has failed us all" and "We must break free of the EU and take back control of our borders." Because the refugees appear to move toward the viewer en masse, it has a strong visual impact.

Who made it?

The Edinburgh-based advertising firm Family Advertising Ltd., which was employed by UKIP for its Brexit campaign.

What did it mean?

UKIP did not try to misrepresent the content, but layered further meaning through adding slogans. Exploiting existing anti-immigrant and racist sentiment, this manipulation focused on generating further fear of immigration and refugees, on the basis of unsubstantiated claims and insinuations about EU border policy.

What does it mean?

In November 2019, in the run-up to the U.K. general election, the campaign organization Leave.EU also used a tightly cropped version of the photograph in an anti-immigration image uploaded to Twitter, making a clear reference back to UKIP's 2016 poster.

What other questions are useful to ask?

Is the actor official or unofficial? The key actor in creating and distributing the image, UKIP, is an official political party and not the type of actor usually associated with mis- and disinformation.

Is it similar to or related to other images? Some likened the poster to Nazi propaganda; it resonates both with previous anti-migrant imagery and a longer history of U.K. political posters involving queues, including one used by UKIP in May 2016 focused on immigration from the EU.

3 key takeaways:

- Official political parties and politicians can be actors in the spread of misinformation.
- Misinformation does not necessarily involve fake images or even the misidentification of what they show. Sometimes images can be used to support a message that misrepresents a wider situation.
- Some misinformation requires more than verification. There is a need to critically examine how real images are used to manipulate, and what such images do and mean.

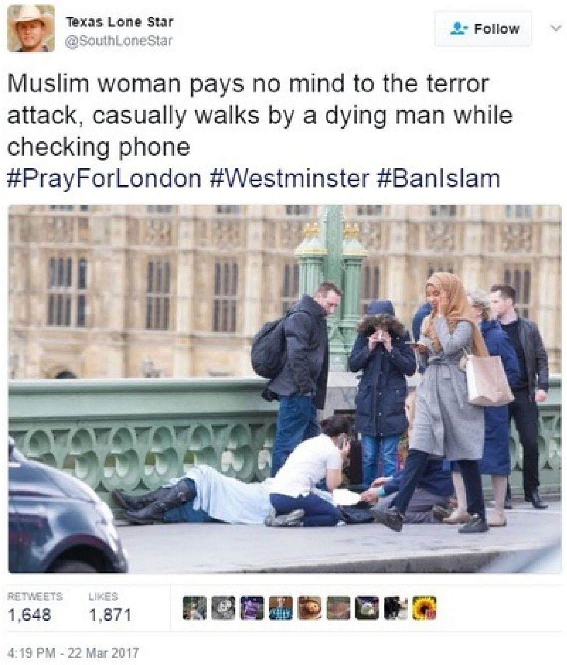
Examples of media coverage of this case:

Nigel Farage's anti-migrant poster reported to police — *The Guardian*

Brexit: UKIP's 'unethical' anti-immigration poster — *Al-Jazeera*

Nigel Farage accused of deploying Nazi-style propaganda as Remain crash poster unveiling with rival vans — *The Independent*

Case Study 2: The Westminster Bridge Photograph, March 2017



Texas Lone Star
@SouthLoneStar

Muslim woman pays no mind to the terror attack, casually walks by a dying man while checking phone
#PrayForLondon #Westminster #BanIslam

RETWEETS 1,648
LIKES 1,871

4:19 PM - 22 Mar 2017

What is it?

A tweet from a Twitter account that appears to be operated by a white Texan man, which received significant media attention. The account was later revealed to be operated by Russia's Internet Research Agency, and was used to spread mis- and disinformation. The tweet shared a photograph from the aftermath of the Westminster Bridge terrorist attack in London (March 22, 2017).

What does it show?

A Muslim woman walking past a group of people and a person on the ground, who has been injured in the terrorist attack. The text has Islamophobic connotations, claiming that the woman is purposefully ignoring the injured person, as well as an overtly anti-Islamic hashtag.

Who made it?

The Internet Research Agency worker who operated the @SouthLoneStar Twitter account, though it was not known to be an IRA account at the time of the tweet. The picture itself was taken by press photographer Jamie Lorriman.

What did it mean?

In March 2017, this appeared to be a tweet from a right-wing Texan Twitter user, who interpreted the photograph as showing that the Muslim woman did not care about the injured person. It suggested this example spoke to a larger truth about Muslims.

What does it mean?

As of today, the tweet is evidence of the Internet Research Agency's purposefully spreading Islamophobic disinformation in the aftermath of a terrorist attack.

What other questions are useful to ask?

What responses did it receive? This tweet received a significant response from the mainstream media. Dozens of U.K. newspapers reported on it, in some cases more than once. While most of these articles condemned @SouthLoneStar, it also moved the tweet from the confines of social media and opened it to a mainstream audience. After the image spread, the woman in the photo spoke out to say that she was distraught over the attacks at the time, and that “not only have I been devastated by witnessing the aftermath of a shocking and numbing terror attack, I’ve also had to deal with the shock of finding my picture plastered all over social media by those who could not look beyond my attire, who draw conclusions based on hate and xenophobia.”

Is it similar to or related to other images? The image that was circulated most of the time was one of seven images taken of the woman. Others showed clearly that she was distraught, something few publications picked up on.

How widely and for how long was it circulated? The added mainstream media attention means the tweet spread widely. However, within a few days, circulation slowed significantly. It was recirculated in November 2017, when it was discovered that @SouthLoneStar was operated by the Internet Research Agency. This later November circulation was notably smaller in the mainstream media compared to March.

3 key takeaways:

- Visual disinformation is not always wholly false and can involve elements that are based on truth. The photograph is real, but its context has been manipulated and falsified, and it relies on the reader/viewer not knowing what the woman was actually thinking in that moment.
- Journalists should think carefully about bringing further attention to such emotionally fueled, controversial and potentially harmful disinformation by reporting on it, even with positive intentions.
- More attention could be paid toward correcting disinformation-based news stories and ensuring that the true picture of events is most prominent. The limited coverage in November means that some readers may not have found out that the tweet was Russian disinformation.

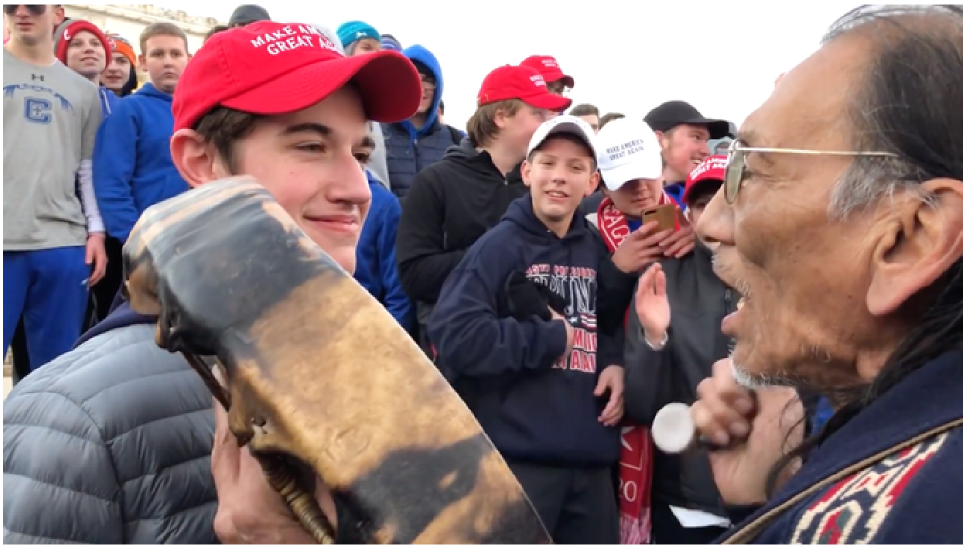
Examples of media coverage of this case:

People are making alarming assumptions about this photo of ‘woman in headscarf walking by dying man’ — *Mirror*

‘Who is the real monster?’ Internet turns on trolls who criticised ‘indifferent’ Muslim woman seen walking through terror attack — *Daily Mail*

British MP calls on Twitter to release Russian ‘troll factory’ tweets — *The Guardian*

Case Study 3: Lincoln Memorial confrontation, January 2019



What is it?

A video of a group of students from Covington Catholic High School who took part in the pro-life March for Life and an indigenous man, Nathan Phillips, who was accompanying other Native Americans in the Indigenous Peoples March.

What does it show?

A confrontation between one of the students from Covington Catholic High School and Phillips. The two demonstrations converged on the Plaza, with a large group of Covington students wearing MAGA hats supposedly facing off against Phillips. This paints a picture of a lone Native American facing off against a barrage of young alt-right bullies.

Who made it?

The video was first uploaded to Instagram by an Indigenous Peoples March participant. This received nearly 200,000 views. Hours later, the video was uploaded to Twitter, receiving 2.5 million views before being deleted by the original account. The video was then reposted across different social media sites, subsequently grabbing mainstream media attention. Within 24 hours, several articles about the video had been published.

What did it mean?

The initial narrative that spread online presented the video as a straightforward faceoff between Philips and the students in which the students were seen as intentionally taunting and ganging up on Phillips.

What does it mean?

A much longer video of the encounter, which emerged several days after the first video, painted a more complex picture. The memorial was also occupied by a group of Black Hebrew Israelites, who were taunting passersby, including the Covington students and Indigenous Peoples March participants. This led to a heated standoff between all three groups, with Phillips allegedly trying to pacify. This is where the first video begins.

What other questions are useful to ask?

What contextual information do you need to know?

Without the longer video, and the knowledge that the Black Hebrew Israelites were present and actively fueling conflict, all context is lost. While the students were recorded saying racist things, what led to this was more complicated than simply alt-right teens ganging up on an elderly indigenous man.

Where was it shared on social media?

While the video was originally shared on Instagram by someone who attended the Indigenous Peoples March, this received limited attention. It was subsequently reuploaded to Twitter and YouTube by other users, which greatly amplified awareness and secured mainstream media attention. Therefore, the attention came from these reuploads and not from the original video on Instagram.

3 key takeaways:

- When such emotion-ladened visuals spread so quickly online, it is easy to lose context and allow the superficial, reactionary online narrative to take control.
- In retrospect, some journalists argued that the initial articles served to fuel the controversy and further push the incorrect narrative. This suggests that, without proper investigation, mainstream media can unintentionally

continue the spread of misinformation.

- The speed with which the video spread online meant a lot of mainstream media outlets “fell for” the narrative pushed on social media and did not investigate further. Many news sites were forced to retract or correct their articles once true events emerged, and some were sued.

Examples of media coverage of this case:

Native American Vietnam Vet Mocked And Surrounded By MAGA Hat-Wearing Teens — *UNILAD*

Outcry after Kentucky students in Maga hats mock Native American veteran — *The Guardian*

Fuller video casts new light on Covington Catholic students' encounter with Native American elder — *USA Today*

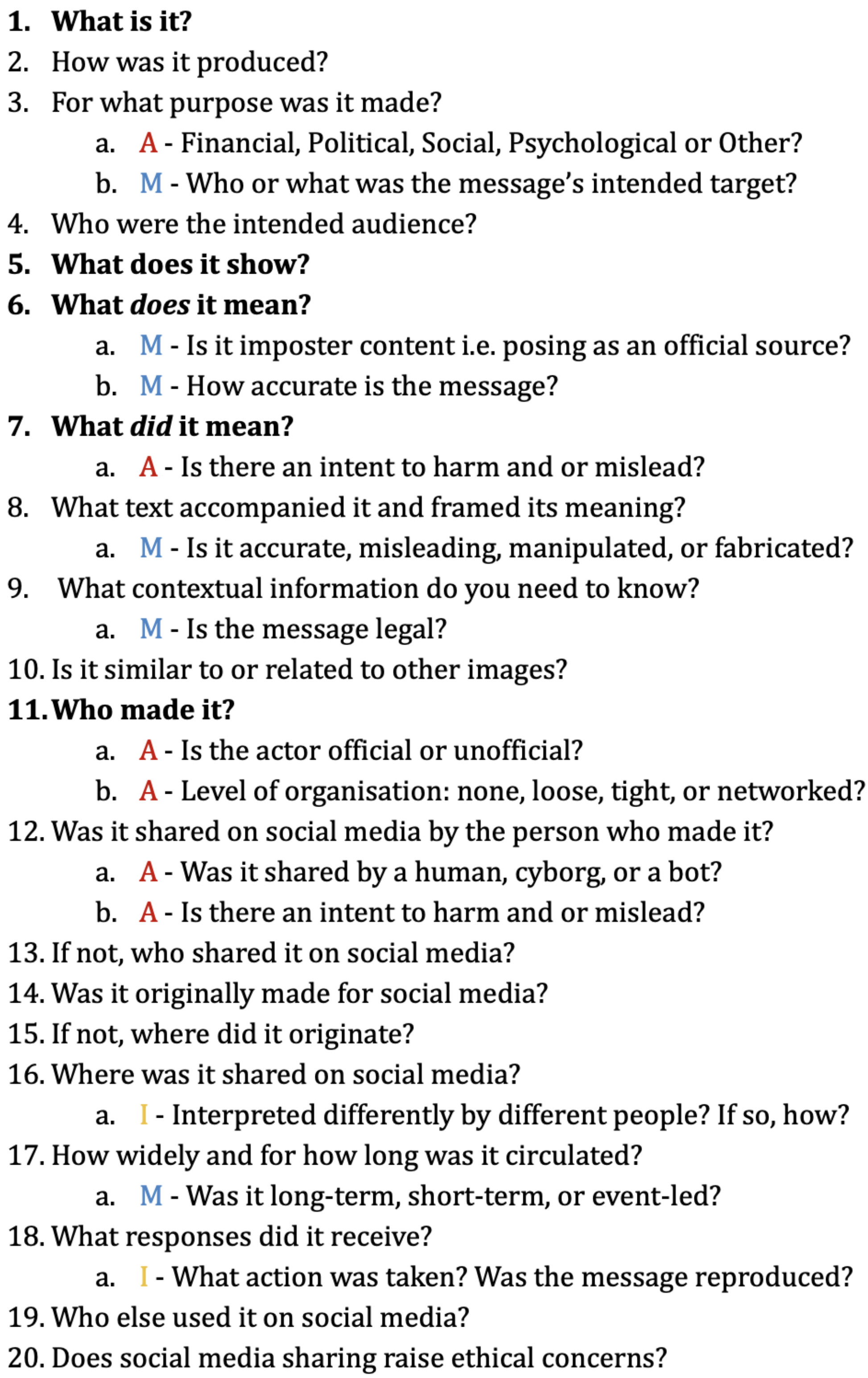
Conclusion

So much of what is shared on social media is visual. Journalists must be equipped with the ability to critically question and assess images to unearth important content and intent. The speed with which visual misinformation can spread further highlights the need for journalists to proceed with caution and make sure to investigate image-related stories fully before publishing. The “20 Questions for Interrogating Social Media Images” is an additional tool journalists can use when investigating images, especially when the story is primarily centered on something visual. Not every question in the framework is relevant to every image, but the five basic questions are a strong starting point and build on basic verification skills, with the aim of developing more accurate and more in-depth reporting.

APPENDIX

Below is the full list of questions from the 20 Questions framework, including 14 prompt questions specifically focused on mis- and disinformation. As we have noted in the chapter, there are five questions that are useful to address first (in bold). The prompt questions relate to either the agent, the message or the interpreter of the mis- and disinformation:

- AGENT (A) - Who created and distributed the image, and what was their motivation?
- MESSAGE (M) - What format did the image take, and what are its characteristics?
- INTERPRETER (I) - How was the message interpreted, and what actions were taken?



1. **What is it?**
2. How was it produced?
3. For what purpose was it made?
    a. A - Financial, Political, Social, Psychological or Other?
    b. M - Who or what was the message's intended target?
4. Who were the intended audience?
5. **What does it show?**
6. **What *does* it mean?**
    a. M - Is it imposter content i.e. posing as an official source?
    b. M - How accurate is the message?
7. **What *did* it mean?**
    a. A - Is there an intent to harm and or mislead?
8. What text accompanied it and framed its meaning?
    a. M - Is it accurate, misleading, manipulated, or fabricated?
9. What contextual information do you need to know?
    a. M - Is the message legal?
10. Is it similar to or related to other images?
11. **Who made it?**
    a. A - Is the actor official or unofficial?
    b. A - Level of organisation: none, loose, tight, or networked?
12. Was it shared on social media by the person who made it?
    a. A - Was it shared by a human, cyborg, or a bot?
    b. A - Is there an intent to harm and or mislead?
13. If not, who shared it on social media?
14. Was it originally made for social media?
15. If not, where did it originate?
16. Where was it shared on social media?
    a. I - Interpreted differently by different people? If so, how?
17. How widely and for how long was it circulated?
    a. M - Was it long-term, short-term, or event-led?
18. What responses did it receive?
    a. I - What action was taken? Was the message reproduced?
19. Who else used it on social media?
20. Does social media sharing raise ethical concerns?

The framework takes inspiration from:

1. The "Interrogating the work of Art" diagram (Figure 2.4, p.39), in (2014, 5th Edition), Pointon, M. History of Art: A Student's Handbook, London and New York: Routledge.
2. "Questions to ask about each element of an example of information disorder" (Figure 7, p. 28), in (2017), Wardle, C. and Derakshan, H., Information Disorder: Toward an interdisciplinary framework for research and policy making. Council of Europe report DGI(2017)09.

6. How to think about deepfakes and emerging manipulation technologies

Written by: Sam Gregory

Sam Gregory *is program director of WITNESS (www.witness.org), which helps people use video and technology to fight for human rights. An award-winning technologist and advocate, he is an expert on new forms of AI-driven mis/disinformation and leads work around emerging opportunities and threats to activism and journalism. He is also co-chair of the Partnership on AI's expert group focused on AI and the media.*

In the summer of 2018, Professor Siwei Lyi, a leading deepfakes researcher based at the University of Albany, released a paper showing that deepfake video personas did not blink at the same rate as real people. This claim was soon covered by Fast Company, New Scientist, Gizmodo, CBS News and others, causing many people to come away thinking they now had a robust way of spotting a deepfake.

Yet within weeks of publishing his paper, the researcher received videos showing a deepfake persona blinking like a human. As of today, this tip isn't useful or accurate. It was the Achilles' heel of a deepfakes creation algorithm at that moment, based on the training data being used. But within months it was no longer valid.

This illustrates a key truth about deepfake detection and verification: Technical approaches are useful until synthetic media techniques inevitably adapt to them. A perfect deepfake detection system will never exist.

So how should journalists verify deepfakes, and other forms of synthetic media?

The first step is to understand the cat-and-mouse nature of this work and be aware of how the technology is evolving. Second, journalists need to learn and apply fundamental verification techniques and tools to investigate whether a piece of content has been maliciously manipulated or synthetically generated. The approaches to image and video verification detailed in the first Verification Handbook, as well as in First Draft's resources related to visual verification all apply. Finally, journalists need to understand we're already in an environment where falsely claiming that something is a deepfake is increasingly common. That means the ability to verify a photo or video's authenticity is just as important as being able to prove it has been manipulated.

This chapter expands on these core approaches to verifying deepfakes, but it's first important to have a basic understanding of deepfakes and synthetic media.

What are deepfakes and synthetic media?

Deepfakes are new forms of audiovisual manipulation that allow people to create realistic simulations of someone's face, voice or actions. They enable people to make it seem like someone said or did something they didn't. They are getting easier to make, requiring fewer source images to build them, and they are increasingly being commercialized. Currently, deepfakes overwhelmingly impact women because they're used to create nonconsensual sexual images and videos with a specific person's face. But there are fears deepfakes will have a broader impact across society and in newsgathering and verification processes.

Deepfakes are just one development within a family of artificial intelligence (AI)-enabled techniques for synthetic media generation. This set of tools and techniques enable the creation of realistic representations of people doing or saying things they never did, realistic creation of people/objects that never existed, or of events that never happened.

Synthetic media technology currently enables these forms of manipulation:

- Add and remove objects within a video.
- Alter background conditions in a video. For example, changing the weather to make a video shot in summer appear as if it was shot in winter.
- Simulate and control a realistic video representation of the lips, facial expressions or body movement of a specific individual. Although the deepfakes discussion generally focuses on faces, similar techniques are being applied to full-body movement, or specific parts of the face.
- Generate a realistic simulation of a specific person's voice.
- Modify an existing voice with a "voice skin" of a different gender, or of a specific person.
- Create a realistic but totally fake photo of a person who does not exist. The same technique can also be applied less problematically to create fake hamburgers, cats, etc.
- Transfer a realistic face from one person to another, aka a deepfake.

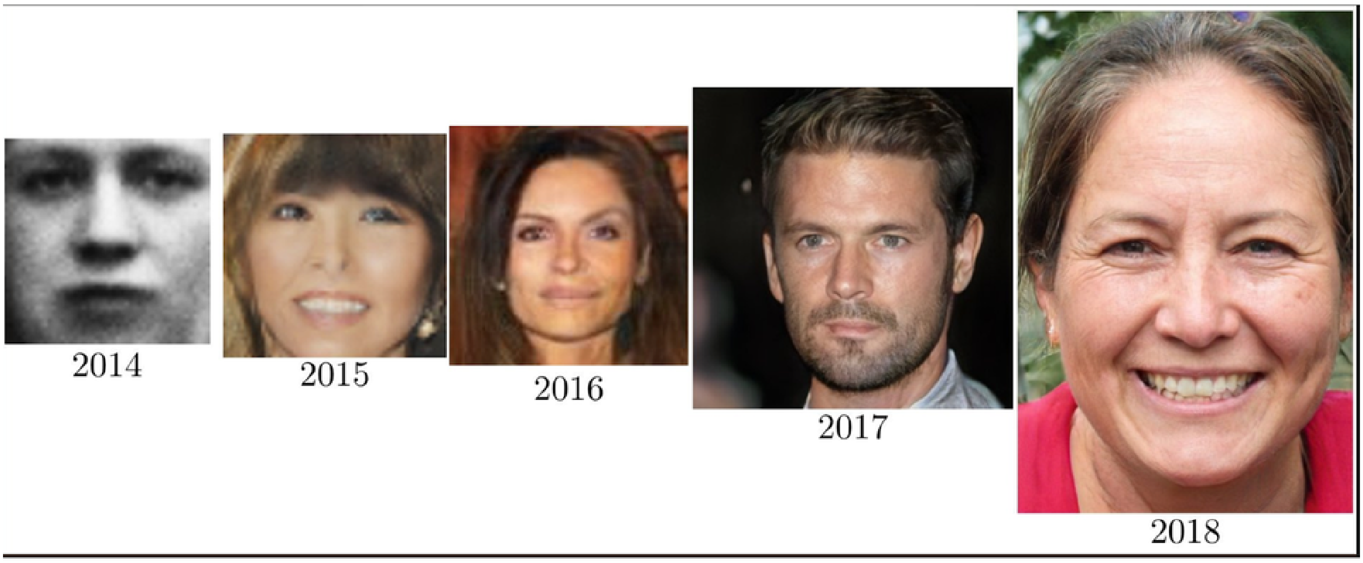
These techniques primarily but not exclusively rely on a form of artificial intelligence known as deep learning and what are called Generative Adversarial Networks, or GANs.

To generate an item of synthetic media content, you begin by collecting images or source video of the person or item you want to fake. A GAN develops the fake — be it video simulations of a real person or face-swaps — by using two networks. One network generates plausible re-creations of the source imagery, while the second network works to detect these forgeries. This detection data is fed back to the network engaged in the creation of forgeries, enabling it to improve.

As of late 2019, many of these techniques — particularly the creation of deepfakes — continue to require significant computational power, an understanding of how to tune your model, and often significant postproduction CGI to improve the final result. However, even with current limitations, humans are already being tricked by simulated media. As an example, research from the FaceForensics++ project showed that people could not reliably detect current forms of lip movement modification, which are used to match someone's mouth to a new audio track. This means humans are not inherently equipped to detect synthetic media manipulation.

It should also be noted that audio synthesis is advancing faster than expected and becoming commercially available. For example, the Google Cloud Text-to-Speech API enables you to take a piece of text and convert it to audio with a realistic sounding human voice. Recent research has also focused on the possibility of doing text to combined video/audio edits in an interview video.

On top of that, all the technical and commercialization trends indicate that it will continue to become easier and less expensive to make convincing synthetic media. For example, the below image shows how quickly face generation technology has advanced.



Credit: EFF

Because of the cat-and-mouse nature of these networks, they improve over time as data on successful forgeries and successful detection is fed through them. This requires strong caution about the effectiveness of detection methods.

The current deepfake and synthetic media landscape

Deepfakes and synthetic media are — as yet — not widespread outside of nonconsensual sexual imagery. DeepTrace Lab's report on their prevalence as of September 2019 indicates that over 95% of the deepfakes were of this type, either involving celebrities, porn actresses or ordinary people. Additionally, people have started to challenge real content, dismissing it as a deepfake.

In workshops led by WITNESS, we reviewed potential threat vectors with a range of civil society participants, including grassroots media, professional journalists and fact-checkers, as well as misinformation and disinformation researchers and OSINT specialists. They prioritized areas where new forms of manipulation might expand existing threats, introduce new threats, alter existing threats or reinforce other threats. They identified threats to journalists, fact-checkers and open-source investigators, and potential attacks on their processes. They also highlighted the challenges around "it's a deepfake" as a rhetorical cousin to "it's fake news."

In all contexts, they noted the importance of viewing deepfakes in the context of existing approaches to fact-checking and verification. Deepfakes and synthetic media will be integrated into existing conspiracy and disinformation campaigns, drawing on evolving tactics (and responses) in that area, they said.

Here are some specific threats they highlighted:

- **Journalists and civic activists will have their reputation and credibility attacked**, building on existing forms of online harassment and violence that predominantly target women and minorities. A number of attacks using modified videos have already been made on women journalists, as in the case of the prominent Indian journalist Rana Ayyub.
- **Public figures will face nonconsensual sexual imagery and gender-based violence as well as other uses of so-called credible doppelgangers.** Local politicians may be particularly vulnerable, as they have plentiful images but less of the institutional structure around them as national-level politicians to help defend against a

synthetic media attack. They also often are key sources in news coverage that bubbles up from local to national.

- **Appropriation of known brands** with falsified in-video editing or other ways in which a news, government, corporate or NGO brand might be falsely attached to a piece of content.
- **Attempts to plant manipulated user generated content into the news cycle,** combined with other techniques such as source-hacking or sharing manipulated content to journalists at key moments. Typically, the goal is to get journalists to propagate the content.
- **Utilization of newsgathering/reporting process weaknesses such as single-camera remote broadcasts** (as noted by the Reuters UGC team) and gathering material in hard-to-verify contexts such as war zones or other places.
- As deepfakes become more common and easier to make at volume, they will **contribute to a fire hose of falsehood** that floods media verification and fact-checking agencies with content they have to verify or debunk. This could overload and distract them.
- **Pressure will be on newsgathering and verification organizations to prove that something is true, as well as to prove that something is not falsified**. Those in power will have the opportunity to use plausible deniability on content by declaring it is deepfaked.

A starting point for verifying deepfakes

Given the nature of both media forensics and emerging deepfakes technologies, we have to accept that the absence of evidence that something was tampered with will not be conclusive proof that media has not been tampered with.

Journalists and investigators need to establish a mentality of measured skepticism around photos, videos and audio. They must assume that these forms of media will be challenged more frequently as knowledge and fear of deepfakes increases. It’s also essential to develop a strong familiarity with media forensics tools.

With that in mind, an approach to analyzing and verifying deepfakes and synthetic media manipulation should include:

1. Reviewing the content for synthetic media-derived telltale glitches or distortions.
2. Applying existing video verification and forensics approaches.
3. Utilizing emerging new AI-based approaches and emerging forensics approaches when available.

Reviewing for telltale glitches or distortions

This is the least robust approach to identifying deepfakes and other synthetic media modifications, particularly given the evolving nature of the technology. That said, poorly made deepfakes or synthetic content may present some evidence of visible errors. Things to look in a deepfake for include:

- Potential distortions at the forehead/hairline or as a face moves beyond a fixed field of motion.
- Lack of detail on the teeth.
- Excessively smooth skin.
- Absence of blinking.
- A static speaker without any real movement of head or range of expression.
- Glitches when a person turns from facing forward to sideways.

Some of these glitches are currently more likely to be visible on a frame-by-frame analysis, so extracting a series of frames to review individually can help. This will not be the case for the frontal-lateral movement glitches — these are best seen in a sequence, so you should do both approaches.

Applying existing video verification approaches

As with other forms of media manipulation and shallowfakes, such as miscontextualized or edited videos, you should ground your approach in well-established verification practices. Existing OSINT verification practices are still relevant, and a good starting point is the chapters and case studies in the first Handbook dedicated to image and video verification. Since most deepfakes or modifications are currently not fully synthesized but instead rely on making changes in a source video, you can use frames from a video to look for other versions using a reverse image search. You can also check the video to see if the landscape and landmarks are consistent with images of the same location in Google Street View.

Similarly, approaches based on understanding how content is shared, by who, and how may reveal information about whether to trust an image or video. The fundamentals of determining source, date, time and motivation of a piece of content are essential to determining whether it documents a real event or person. (For a basic grounding in this approach, see this First Draft guide.) And as always, it's essential to contact the person or people featured in the video to seek comment, and to see if they can provide concrete information to support or refute its authenticity.

New tools are also being developed by government, academics, platforms and journalistic innovation labs to assist with the detection of synthetic media, and to broaden the availability of media forensics tools. In most cases, these tools should be viewed as signals to complement your best-practices based verification approach.

Tools such as InVID and Forensically help with both provenance-based image verification and limited forensic analysis.

Free tools in this area of work include:

- FotoForensics: An image forensics tool that includes the capacity for Error Level Analysis to see where elements of an image might have been added.
- Forensically: A suite of tools for detecting cloning, error level analysis, image metadata and a number of other functions.
- InVID: A web browser extension that enables you to fragment videos into frames, perform reverse image search across multiple search engines, enhance and explore frames and images through a magnifying lens, and to apply forensic filters on still images.
- Reveal Image Verification Assistant: A tool with a range of image tampering detection algorithms, plus metadata analysis, GPS geolocation, EXIF thumbnail extraction and integration with reverse image search via Google.
- Ghiro: An open-source online digital forensics tool.

Note that almost all of these are designed for verification of images, not video. This is a weakness in the forensics space, so for videos it is still necessary to extract single images for analysis, which InVID can help with. These tools will be most effective with higher resolution, noncompressed videos that, for example, had video objects removed or added within them. Their utility will decrease the more a video has been compressed, resaved or shared across different social media and video-sharing platforms.

If you're looking for emerging forensics tools to deal with existing visual forensics issues as well as eventually deepfakes, one option is to look at the tools being shared by academics. One of the leading research centers at the University of Napoli provides online access to their code for, among other areas, detecting camera fingerprints

(Noiseprint), detecting image splices (Splicebuster) and detecting copy-move and removal detection in video.

As synthetic media advances, new forms of manual and automatic forensics will be refined and integrated into existing verification tools utilized by journalists and fact-finders as well as potentially into platform-based approaches. It's important that journalists work to stay up to date on the available tools, while also not becoming overly reliant upon them.

Emerging AI-based and media forensics approaches

As of early 2020, there are no tested, commercially available GAN-based detection tools. But we should anticipate that some will enter the market for journalists either as plug-ins or as tools on platforms in 2020. For a current survey of the state-of-field in media forensics including these tools you should read Luisa Verdoliva's 'Media Forensics and Deepfakes: An overview'.

These tools will generally rely on having training data (examples) of GAN-based synthetic media, and then being able to use this to detect other examples that are produced using the same or similar techniques. As an example, forensics programs such as FaceForensics++ generate fakes using existing consumer deepfakes tools and then use these large volumes of fake images as training data for algorithms to perform fake detection. This means they might not be effective on the latest forgery methods and techniques.

These tools will be much more suited to detection of GAN-generated media than current forensic techniques. They will also supplement new forms of media forensics tools that deal better with advances in synthesis. However, they will not be foolproof, given the adversarial nature of how synthetic media evolves. A key takeaway is that any indication of synthesis should be double-checked and corroborated with other verification approaches.

Deepfakes and synthetic media are evolving fast and the technologies are becoming more broadly available, commercialized and easy to use. They need less source content to create a forgery than you might expect. While new technologies for detection emerge and are integrated into platforms and into journalist/OSINT-facing tools, the best way to approach verification is using existing approaches to image/video, and complement these with forensics tools that can detect image manipulation. Trusting the human eye is not a robust strategy!

7. Monitoring and Reporting Inside Closed Groups and Messaging Apps

Written by: Claire Wardle

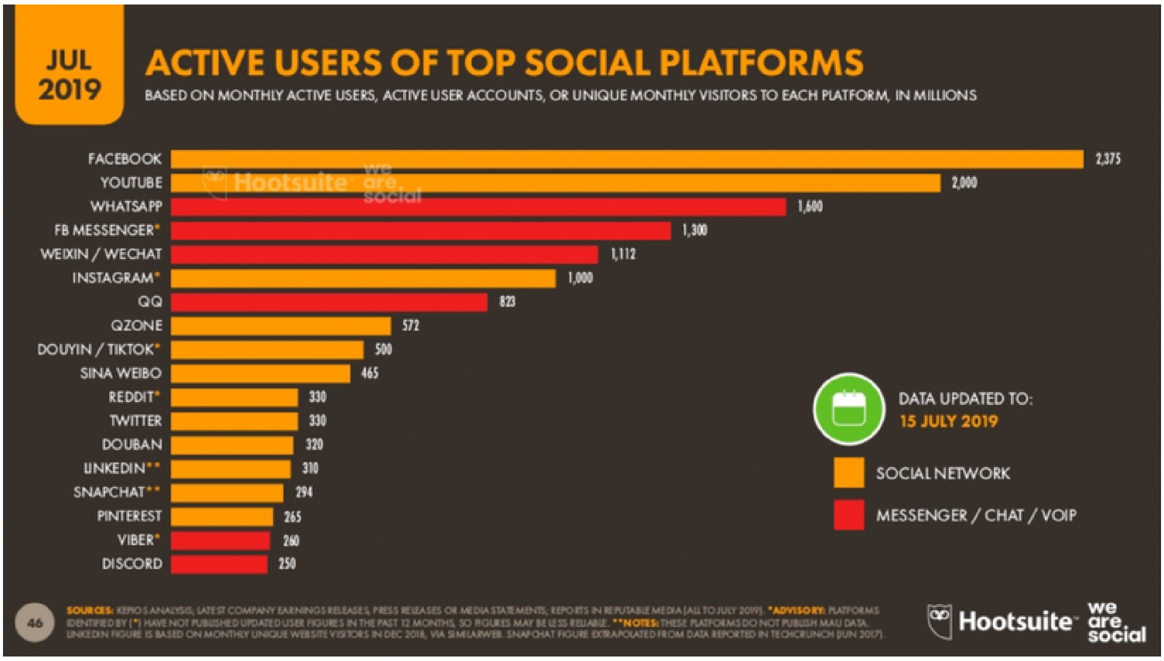
Claire Wardle leads the strategic direction and research for First Draft, a global non-profit that supports journalists, academics and technologists working to address challenges relating to trust and truth in the digital age. She has been a Fellow at the Shorenstein Center for Media, Politics and Public Policy at Harvard's Kennedy School, the Research Director at the Tow Center for Digital Journalism at Columbia University's Graduate School of Journalism and head of social media for UNHR, the United Nations Refugee Agency.

In March 2019, Mark Zuckerberg talked about Facebook's "pivot to privacy," which meant the company was going to emphasize Facebook groups, as a recognition that people were increasingly drawn to communicating with a smaller number of people in private spaces. Over the last few years, the importance of smaller groups for social communication has been clear to those of us working in this space.

In this chapter, I will explain the different platforms and applications, talk about the challenges of monitoring these spaces, and end with a discussion of the ethics of doing this type of work.

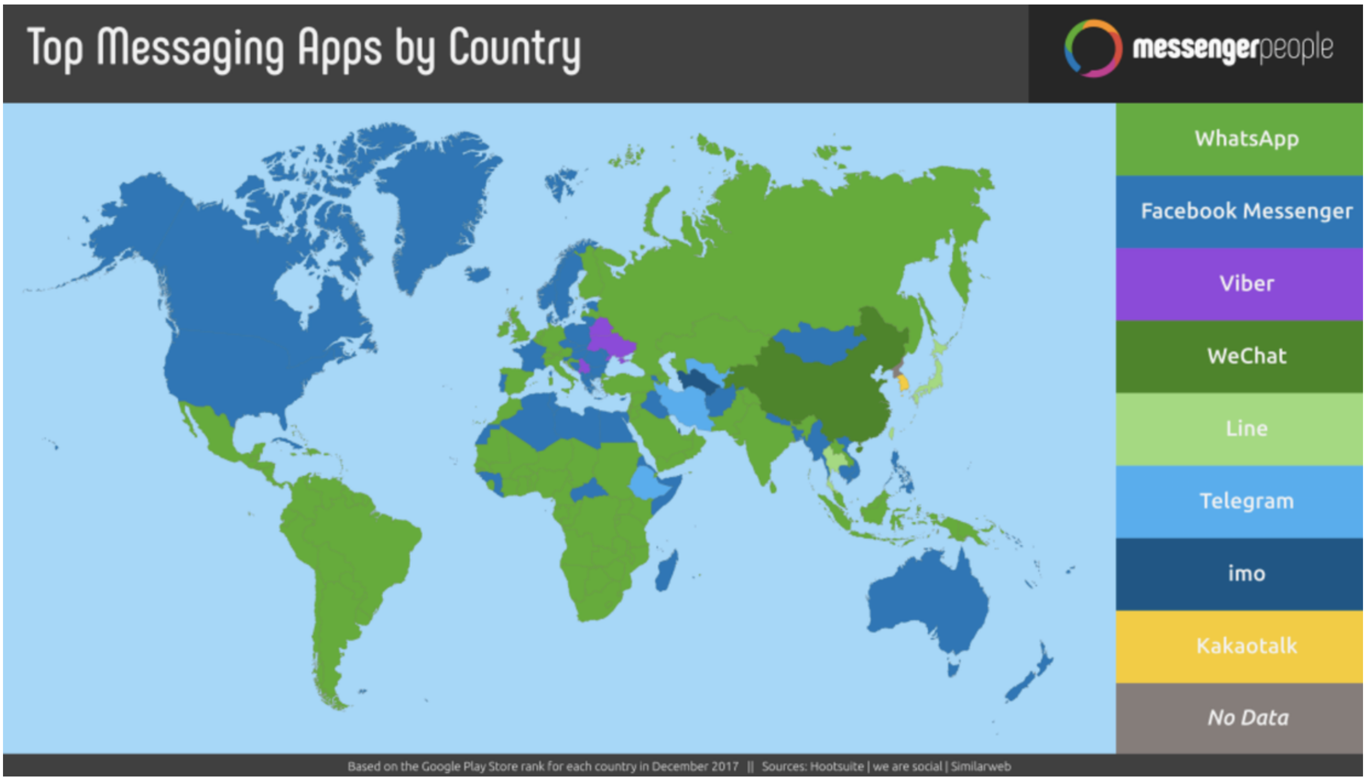
Different platforms and applications

Recent research by We Are Social shows the continuous dominance of Facebook and YouTube, but the next three most popular platforms are WhatsApp, FB Messenger and WeChat.



In many regions around the world, chat apps are now the dominant source of news for many consumers, particularly WhatsApp, for example, in Brazil, India and Spain.

Certainly, WhatsApp and FB Messenger are popular globally, but in certain countries, alternatives are dominant. For example, in Iran, it is Telegram. It's Line in Japan, KakaoTalk in South Korea and WeChat in China.


Top Messaging Apps by Country

Based on the Google Play Store rank for each country in December 2017 || Sources: Hootsuite | we are social | Similarweb

All these sites have slightly different functionality, in terms of encryption, group or broadcast features, and additional options such as in-app commerce opportunities.

Closed Facebook groups

There are three types of Facebook Groups: Open, Closed and Hidden.

- Open groups can be found in search and anyone can join.
- Closed groups can be found in search but you have to apply to join.
- Hidden groups cannot be found in search and you have to be invited to join.

Increasingly, people are congregating on Facebook groups, partly because they're being pushed by Facebook's algorithm but also because people are choosing to spend time with people they already know, or people who share their perspective or interest.

Discord

According to Statista, in July 2019, Discord had 250 million monthly active users (for comparison, Snap had 294 million, Viber had 260 million and Telegram had 200 million). Discord is popular with the gaming community, but in recent years, it has also become known as a site where people congregate in "servers" (a form of group in Discord) to coordinate disinformation campaigns.

One aspect of Discord and some closed Facebook groups is that you will be asked questions before you are accepted into that group. These questions might be about your profession, your religion, your political beliefs or your attitudes toward certain social issues.

Encryption, groups and channels

One reason these platforms and applications have become so popular is that they offer different levels of encryption. WhatsApp and Viber are currently the most secure, offering end to end encryption. Others, like Telegram, FB Messenger and Line, offer encryption if you turn it on.

Certain apps have groups or channels where information is shared to large numbers of people. The largest WhatsApp group can hold 256 people. FB Messenger groups hold 250. In Telegram, a group can be private or publicly searchable, and can hold 200. Once it hits that number it can be converted into a supergroup and up to 75,000 people can join. Telegram also has channels, a broadcast capability inside an app. You can subscribe to a channel and see what's being posted there, but you can't post your own content in response.

Ongoing monitoring

There is no doubt that misinformation circulates on closed messaging apps. It is difficult to independently assess whether there is more misinformation on these platforms than on social media sites, because there is no way of seeing what is being shared. But we know it is a problem, as high-profile cases from India, France and Indonesia have shown us. And in the U.S., during the shootings in El Paso and Dayton in August 2019, there were examples of rumors and falsehoods circulating on Telegram and FB Messenger.

The question is whether journalists, researchers, fact-checkers, health workers and humanitarians should be in these closed groups to monitor for misinformation. If they should be in these groups, how should they go about their work in a way that is ethical and keeps them safe?

While there are significant challenges to doing this work, it is possible. However, keep in mind that many people who use these apps are doing so specifically so they will not be monitored. They use them because they are encrypted. They expect a certain level of privacy. This should be central to anyone working in these spaces. Even though you can join and monitor these spaces, it's paramount to be aware of the responsibility you have to the participants in these groups, who often do not understand what is possible.

Techniques for searching

Searching for these groups can be difficult, as there are different protocols for each. For Facebook groups, you can search for topics within Facebook search and filter by groups. If you want to use more advanced Boolean search operators, search on Google using your keywords and then add site:facebook.com/groups.

For Telegram, you can search in the app if you have an Android phone, but not if you have an iPhone. There are desktop applications like https://www.telegram-group.com/. Similarly for Discord, there are sites such as https://disboard.org/search

Decisions around joining and participating

As mentioned, some of these groups will ask questions to secure entry. Before trying this, you should talk to your editor or manager about how to answer these questions. Will you be truthful about who you are and why you are in the group? Is there a way to join by staying deliberately vague? If not, how can you justify that decision to hide your identity (this might be necessary if you are joining a group that could jeopardize your safety if you identify yourself as a journalist). If you gain access will you contribute in any way, or just “lurk” to find information you can corroborate elsewhere?

Decisions about automatically collecting content from groups

It is possible to find “open” groups by searching for links that are posted to other sites. These then appear in search engines. It is then possible to use computational methods to automatically collect the content from these groups. Researchers monitoring elections in Brazil and India have done this, and I know anecdotally of other organizations doing similar work.

This technique allows organizations to monitor multiple groups simultaneously, which is often impossible otherwise. A key point is that only a small percentage of groups are findable this way, and they tend to be groups desperate for wide membership, so are not representative of all groups. It also raises ethical flags for me personally. However, there are guardrails that can be employed by securing the data, not sharing with others, and de-identifying messages. We need cross-industry protocols about doing this type of work.

Tiplines

The other technique is to set up a tipline, where you encourage the public to send you content. The key to a tipline is having a simple, clear call to action, and that you explain how you intend to use that content. Is it simply for monitoring trends, or are you going to reply to them with a debunk once you've investigated what they're sent you?

Returning to the ethical questions, which impact so much around working with closed messaging apps, it's important that you're not just "taking" content, or in other words being extractive. And putting ethics aside for one minute, all the research shows that if audiences don't know how their tips are being used, they are significantly less likely to keep sending them in. People are more willing to help if they feel like they're being treated like partners.

The other aspect, however, is how easy it is to game tiplines by sending in hoax content, or by one individual or small group sending in lots of the same content to make it appear to be a bigger problem that it is.

Ethics of reporting from closed messaging groups

Once you've found content, the question is how to report on it. Should you be transparent about how you found it? As part of their community guidelines, many groups ask that what is discussed in a group does not get shared more widely. If the group is full of disinformation, what will be the impact of your reporting on it? Can you corroborate what you have found in other groups or online spaces? If you report, might you put your own safety, or that of your colleagues or family at risk? Remember that doxxing journalists and researchers (or worse) is part of the playbook for some of the darker groups online.

Conclusions

Reporting from and about closed messaging apps and groups is full of challenges, yet those sources will become increasingly important as spaces where information is shared. As a first step, think about the questions outlined in this chapter, talk to your colleagues and editors, and if you don't have guidelines in your newsroom about this type of reporting, start working on some. There are no standard rules on how to do this. It depends on the story, the platform, the reporter and a newsroom's editorial guidelines. But it is important that all the details are considered before you start this kind of reporting.

7a. Case Study: Bolsonaro at the Hospital

Written by: Sérgio Lüdtke

***Sérgio Lüdtke** is a journalist and editor of Projeto Comprova, a coalition of 24 media organizations working collaboratively to investigate rumors about public policy in Brazil.*
In 2018, Comprova reviewed suspicious content shared on social media and messaging apps about the presidential elections in Brazil.

On Sept. 6, 2018, a month before Brazil's presidential election, far-right candidate Jair Bolsonaro held a campaign event in downtown Juiz de Fora, a city of 560,000 people 200 kilometers from Rio de Janeiro.

It had been a week since Bolsonaro became the leader in the first-round polls for the Brazilian presidential election. He took the first position after the candidacy of former President Luiz Inácio Lula da Silva, the previously isolated leader in the polls, was barred by the Superior Electoral Court.

Bolsonaro, however, was losing in the runoff simulations to three of the four closest candidates in the polls.

Bolsonaro’s situation was worrisome, since he had only two 9-second daily blocks in the free electoral broadcasts on TV. Brazilian electoral rules require radio and TV stations to give free time to political parties to publicize their proposals. This time is distributed according to the number of seats won by each party in the last election of the House of Representatives. Bolsonaro’s lack of seats meant very little free airtime. As a result, he had to rely on his supporters on social networks and make direct contact with voters on the streets.

In Juiz de Fora, as in other cities he visited before, Bolsonaro participated in a march by being carried on the shoulders by his supporters. He was trailed by a crowd of admirers when the march was suddenly interrupted. In the middle of the crowd, a man reached out and stabbed the candidate. The knife left a deep wound in Bolsonaro’s abdomen — and opened a Pandora’s Box on social networks.

Rumors and conspiracy theories spread, with some accusing Adélio Bispo de Oliveira, the man who stabbed Bolsonaro, of being linked to the party of former President Dilma Rousseff, who was removed from office in 2016. Fake photos showed the attacker standing next to Lula. That Bispo had been affiliated with the left-wing Partido Socialismo e Liberdade (PSOL), and his lawyers’ refusal to say who was paying their fees only served to feed the conspiratorial claims.

At the same time, videos and messages that tried to undercut Bolsonaro gained traction on social media platforms. Some of the malicious content claimed the stabbing was staged, that Bolsonaro had actually been in hospital to treat cancer, and that the photos published showing the surgery had been forged.

The stabbing gave Bolsonaro a reason to withdraw from campaign activities, but earned him a better position in the polls. (Eventually, of course, Bolsonaro won the election.)

On Sept. 19, nearly two weeks after the attack, Eleições sem Fake, a WhatsApp group monitoring program created by the University of Minas Gerais, identified an audio recording that was making the rounds. The audio was shared by 16 of nearly 300 groups monitored by the project; some were Bolsonaro supporters.

That same day our organization, Comprova, began to receive, also by WhatsApp, requests from readers to verify the integrity of the recording.

In the audio, which was just over one minute long, an angry man with a voice that resembles that of Bolsonaro argued with someone appearing to be his son, Eduardo, and complains about being kept in the hospital. On the recording, the man said he can no longer stand "this theater," suggesting that it was all an act.

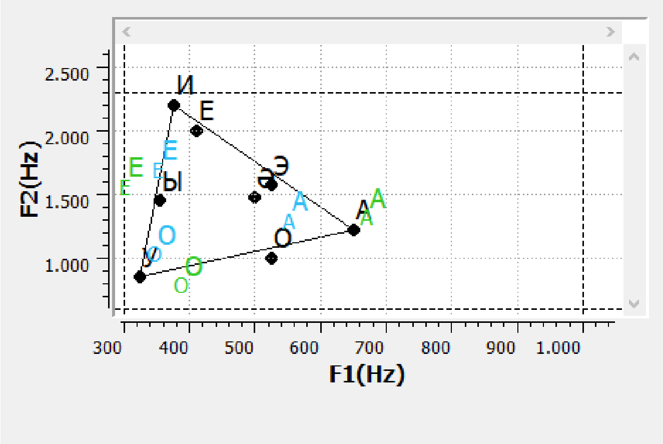
That day, Bolsonaro was still a patient to the Semi-Intensive Care Unit at Albert Einstein Hospital in São Paulo. The medical report said he had no fever, was receiving intravenous nutrition, and had recovered bowel function.

Comprova could not find the original source of the recording. The audio primarily spread through WhatsApp at a time when files could still be shared in up to 20 conversations. This enabled it to spread rapidly, and soon make its way to other social networks. It became impossible to track it back to the original source. (WhatsApp has since restricted the number of groups you can forward a message to.)

Unable to identify the author(s) of the recording, Comprova focused on a more conventional investigation, and requested the help of an expert report from Instituto Brasileiro de Perícia (the Brazilian Institute of Forensics). Experts compared the viral recording with Bolsonaro's voice in an April 2018 interview and concluded that the voice of the candidate was not the voice on the recording bring shared on social networks.

The experts made a qualitative analysis of the voice, speech and language markers of the man who spoke in the recording. Then they compared these parameters in each voice and speech sample. In this analysis, they investigated vowel and consonant patterns, speech rhythm and speed, intonation patterns, voice quality and habits presented by the speaker, as well as the use of specific words and grammatical rules.

For example, the below image shows a frequency analysis of "formants," the name of the pitches produced by vibrations of the vocal tract, the cavity where the sound produced at the larynx is filtered. The air inside the vocal tract vibrates at different pitches, depending on its size and shape of the opening. The image shows a frequency analysis of the formants using the vowels "a," "e" and "o." The green vowels correspond to the audio sample we obtained on WhatsApp, and the blue vowels correspond to a sample taken from an interview given by Bolsonaro a few days before the attack on him.



Additional analysis found that the speaker in the WhatsApp audio was found to have a typical accent from the countryside of the state of São Paulo. But this did not appear in Bolsonaro's speech patterns. Differences in resonance, articulation, speech rate and phonetic deviation were detected in the compared samples.

Comprova consulted a second expert. This professional also concluded that the voice in the recording differed from Bolsonaro's for several reasons. He said the tone of the voice appeared to be a little more acute than Bolsonaro's. He noted the pace of speech was also faster than another video recorded by the candidate at the hospital.

Another element that reinforced the conclusion that the audio was fake is the poor quality of the recording. According to experienced experts, this is a typical phony trick: Lowering the resolution of audios, videos and photos make analyzing them more difficult.

In terms of Bolsonaro's response, his sons, Flavio and Carlos, posted on social media to say the audio was "fake news."

If this audio went viral today, it would probably be harder to believe that the voice belonged to Bolsonaro. Before the election, with only 18 seconds a day on TV and his missing the campaign debates due to hospitalization and treatment, the current president's voice was not so well known. That created an opportunity for a faked audio recording to fool many.

More than a year later, however, it is still difficult to understand why groups in favor of Bolsonaro or campaigning for his candidacy shared this audio, which, if proved authentic, could have destroyed his candidacy. We will never fully know why these groups so eagerly shared this content. Even so, it's a powerful reinforcement of the fact that a piece of content that makes an explosive claim will spread rapidly across social media.

8. Investigating websites

Written by: Craig Silverman

Craig Silverman is the media editor of BuzzFeed News, where he leads a global beat covering platforms, online misinformation and media manipulation. He previously edited the "Verification Handbook" and the "Verification Handbook for Investigative Reporting," and is the author of "Lies, Damn Lies, and Viral Content: How News Websites Spread (and Debunk) Online Rumors, Unverified Claims and Misinformation."

Websites are used by those engaged in media manipulation to earn revenue, collect emails and other personal information, or otherwise establish an online beachhead. Journalists must understand how to investigate a web presence, and, when possible, connect it to a larger operation that may involve social media accounts, apps, companies or other entities.

Remember that text, images or the entire site itself may disappear over time — especially after you start contacting people and asking questions. A best practice is to use the Wayback Machine to save important pages on your target website as part of your workflow. If a page won't save properly there, use a tool such as archive.today. This ensures you can link to archived pages as proof of what you found, and avoid directly linking to a site spreading mis/disinformation. (Hunchly is a great paid tool for creating your own personal archive of webpages automatically while you work.) These archiving tools are also essential for investigating what a website has looked like over time. I also recommend installing the Wayback Machine browser extension so it's easy to archive pages and look at earlier versions.

Another useful browser extension is Ghostery, which will show you the trackers present on a webpage. This helps you quickly identify whether a site uses Google Analytics and/or Google AdSense IDs, which will help with one of the techniques outlined below.

This chapter will look at four categories to analyze when investigating a website: content, code, analytics, registration and connected elements.

Content

Most websites tell you at least a bit about what they are. Whether on a dedicated About page, a description in the footer or somewhere else, this is a good place to start. At the same time, a lack of clear information could be a hint the site was created in haste, or is trying to conceal details about its ownership and purpose.

Along with reading any basic "about" text, perform a thorough review of content on a website, with an eye toward determining who's running it, what the purpose is, and whether it's part of a larger network or initiative. Some things to look for:

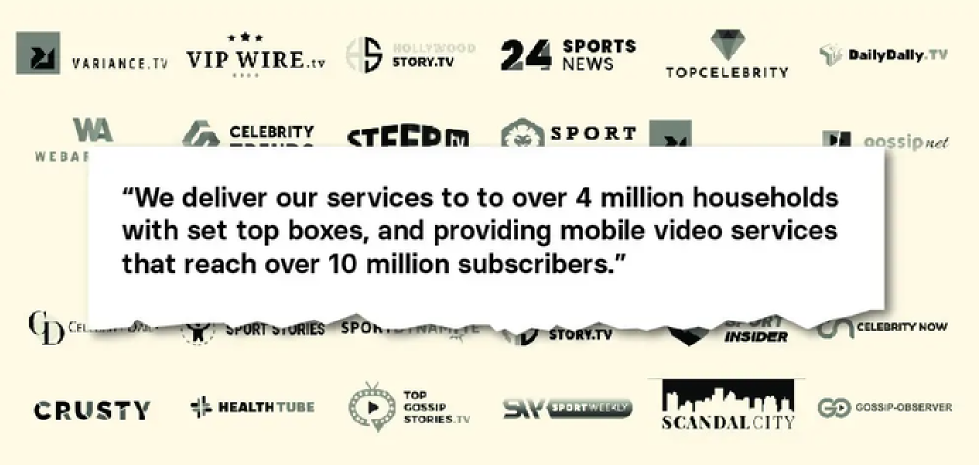
- Does it identify the owner or any corporate entity on its about page? Also note if it doesn't have an About page.
- Does it list a company or person in a copyright notice at the very bottom of the homepage or any other page?
- Does it list any names, addresses or corporate entities in the privacy policy or terms and conditions? Are those names or companies different from what's listed on the footer, about page or other places on the site?
- If the site publishes articles, note the bylines and if they are clickable links. If so, see if they lead to an author page with more information, such as a bio or links to the writer's social accounts.
- Does the site feature related social accounts? These could be in the form of small icons at the top, bottom or side of the homepage, or an embed inviting you to like its Facebook page, for example. If the page shows icons for platforms such as Facebook and Twitter, hover your mouse over them and look at the bottom left of your

browser window to see the URL they lead to. Often, a hastily created website will not bother to fill in the specific social profile IDs in a website's template. In that case, you'll just see the link show up as facebook.com/ with no username.

- Does the site list any products, clients, testimonials or other people or companies that may have a connection and be worth looking into?
- Be sure to dig beyond the homepage. Click on all main menus and scroll down to the footer to find other pages worth visiting.

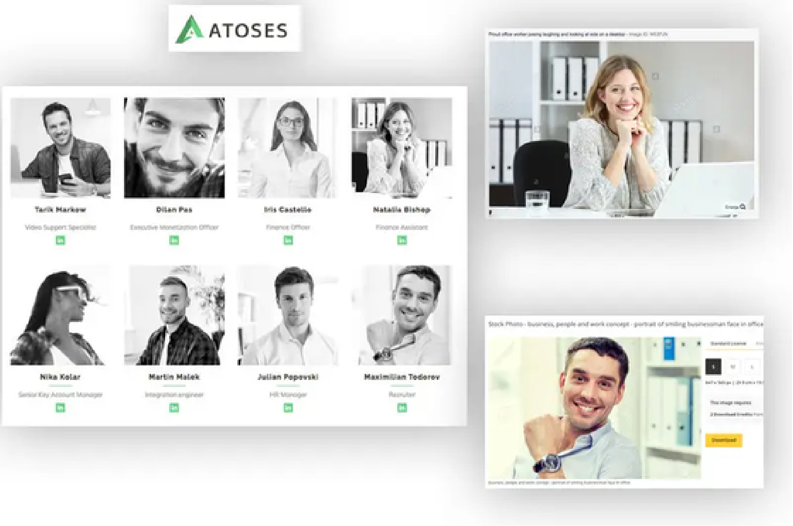
An important part of examining the content is to see if it's original. Has text from the site's About page or other general text been copied from elsewhere? Is the site spreading false or misleading information, or helping push a specific agenda?

In 2018 I investigated a large digital advertising fraud scheme that involved mobile apps and content websites, as well as shell companies, fake employees and fake companies. I ultimately found more than 35 websites connected to the scheme. One way I identified many of the sites was by copying the text on one site's About page and pasting it into the Google search box. I instantly found roughly 20 sites with the exact same text:



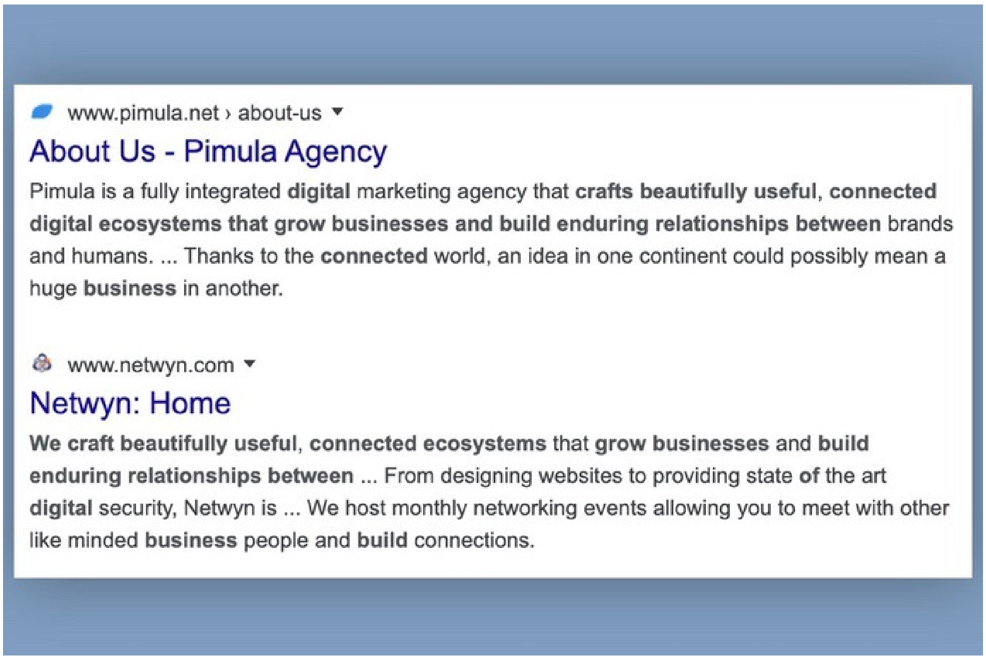
"We deliver our services to to over 4 million households with set top boxes, and providing mobile video services that reach over 10 million subscribers."

The fraudsters running the scheme also created websites for their front companies to help them appear legitimate when potential partners at ad networks visited to perform due diligence. One example was a company called Atoses. Its homepage listed several employees with headshots. Yandex's reverse image search (the best image search for faces) quickly revealed that several of them were stock images:



Atoses also had this text in the footer of its site: “We craft beautifully useful, connected ecosystems that grow businesses and build enduring relationships between online media and users.”

That same text appears on the sites of at least two marketing agencies:



www.pimula.net › about-us ▾

About Us - Pimula Agency

Pimula is a fully integrated **digital** marketing agency that **crafts beautifully useful**, **connected digital ecosystems that grow businesses and build enduring relationships between** brands and humans. ... Thanks to the **connected** world, an idea in one continent could possibly mean a huge **business** in another.

www.netwyn.com ▾

Netwyn: Home

We craft beautifully useful, **connected ecosystems** that **grow businesses** and **build enduring relationships between** ... From designing websites to providing state **of** the art **digital** security, Netwyn is ... We host monthly networking events allowing you to meet with other like minded **business** people and **build** connections.

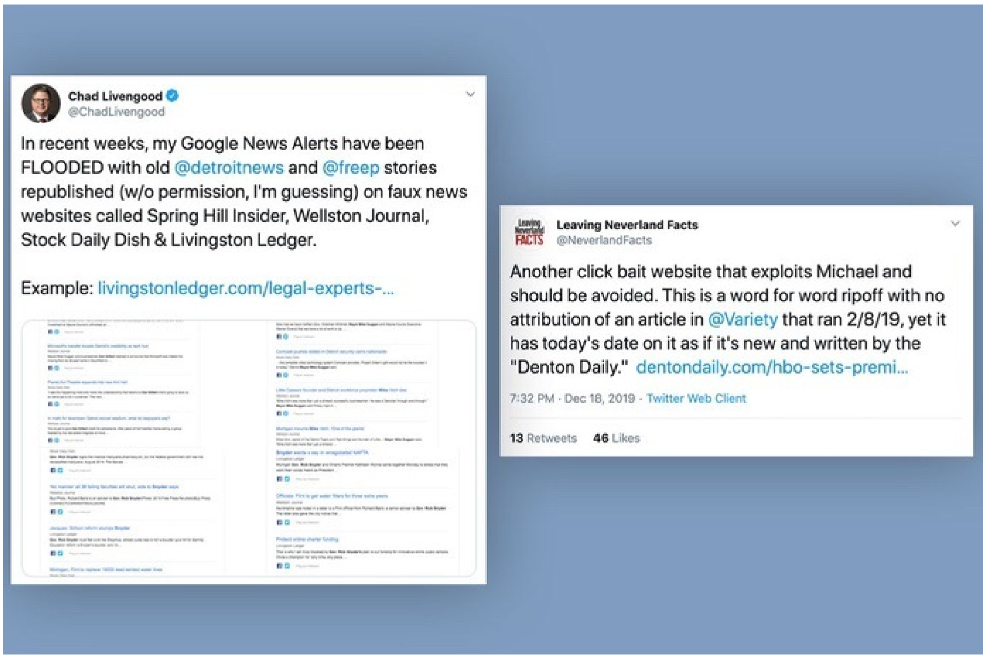
If a company is using stock images for employees and plagiarized text on its site, you know it’s not what it claims to be.

It’s also a good idea to copy and paste text from articles on a site and enter them into Google or another search engine. Sometimes, a site that claims to be a source of news is just plagiarizing real outlets.

In 2019, I came across a site called forbesbusinessinsider.com that appeared to be a news site covering the tech industry. In reality it was mass plagiarizing articles from a wide variety of outlets, including, hilariously, an article I wrote about fake local websites.

Another basic step is to take the URL of a site and search it in Google. For example, “forbesbusinessinsider.com.” This will give you a sense of how many of the site’s pages have been indexed, and may also bring up examples of other people reporting on or otherwise talking about the site. You can also check if the site is listed in Google News by loading the main page of Google News and entering “forbesbusinessinsider.com” in the search box.

Another tip is to take the site URL and paste it into search bars at Twitter.com or Facebook.com. This will show you if people are linking to the site. During one investigation, I came across a site, dentondaily.com. Its homepage showed only a few articles from early 2020, but when I searched the domain on Twitter, I saw that it had previously pumped out plagiarized content, which had caused people to notice and complain. These older stories were deleted from the site, but the tweets provided evidence of its previous behavior.



Chad Livengood
@ChadLivengood

In recent weeks, my Google News Alerts have been FLOODED with old @detroitnews and @freep stories republished (w/o permission, I'm guessing) on faux news websites called Spring Hill Insider, Wellston Journal, Stock Daily Dish & Livingston Ledger.

Example: livingstonledger.com/legal-experts-...

Leaving Neverland Facts
@NeverlandFacts

Another click bait website that exploits Michael and should be avoided. This is a word for word ripoff with no attribution of an article in @Variety that ran 2/8/19, yet it has today's date on it as if it's new and written by the "Denton Daily." dentondaily.com/hbo-sets-premi...

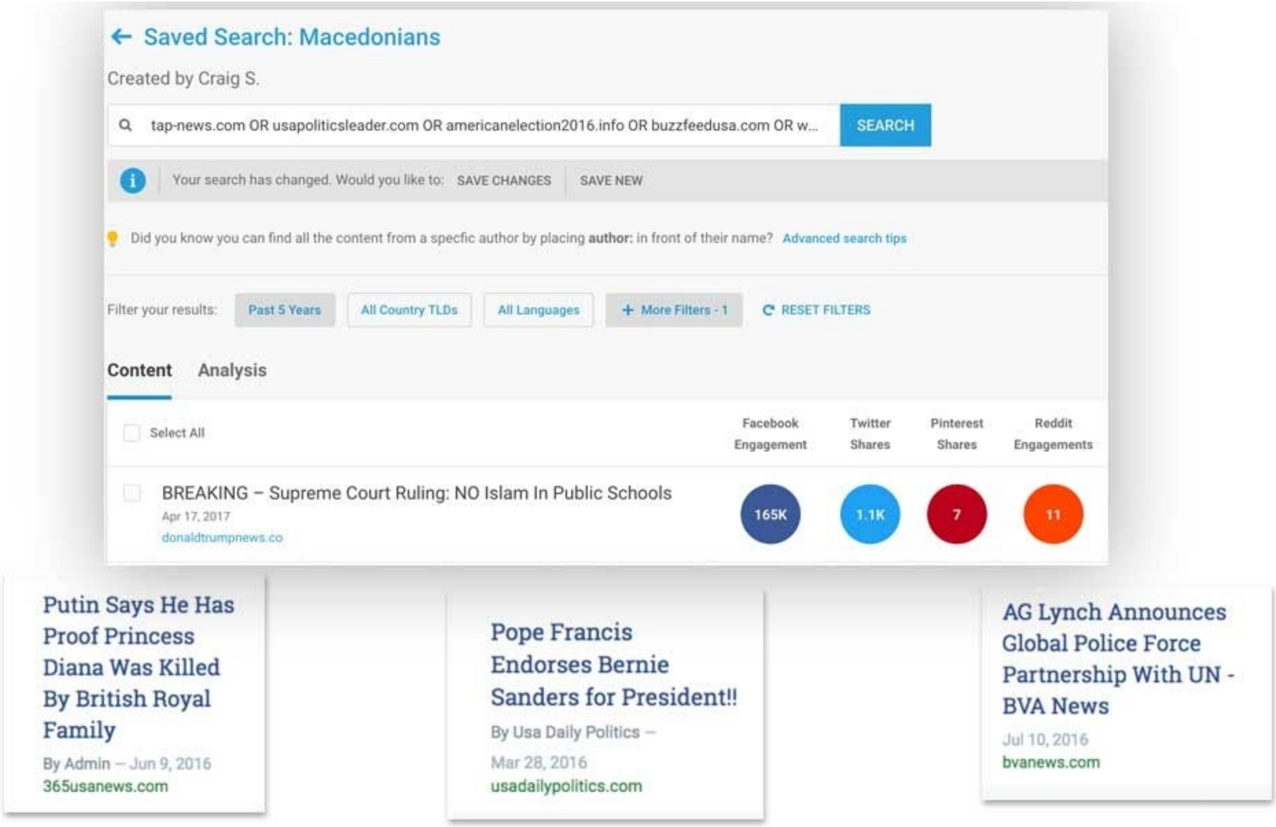
7:32 PM · Dec 18, 2019 · Twitter Web Client

13 Retweets 46 Likes

Once you've dug into the content of a website, it's time to understand how it spreads. We'll look at two tools for this: BuzzSumo and CrowdTangle.

In 2016, I worked with researcher Lawrence Alexander to look at American political news sites being run from overseas. We soon zeroed in on sites run out of Veles, a town in North Macedonia. We used domain registration details (more on that below) to identify more than 100 U.S. political sites run from that town. I wanted to get a sense of how popular their content was, and what kind of stories they were publishing. I took the URLs of several sites that seemed to be the most active and created a search for them in BuzzSumo, a tool that can show a list of a website's content ranked by how much engagement it received on Facebook, Twitter, Pinterest and Reddit. (It has a free version, though the paid product offers far more results.)

I immediately saw that the articles from these sites with the most engagement on Facebook were completely false. This provided us with key information and an angle that was different from previous reporting. The below image shows the basic BuzzSumo search results screen, which lists the Facebook, Twitter, Pinterest and Reddit engagements for a specific site, as well as some sample false stories from 2016:



← Saved Search: Macedonians

Created by Craig S.

tap-news.com OR usapoliticsleader.com OR americanelection2016.info OR buzzfeedusa.com OR w... SEARCH

Your search has changed. Would you like to: SAVE CHANGES SAVE NEW

Did you know you can find all the content from a specfic author by placing **author:** in front of their name? Advanced search tips

Filter your results: Past 5 Years | All Country TLDs | All Languages | + More Filters - 1 | RESET FILTERS

Content Analysis

□ Select All	Facebook Engagement	Twitter Shares	Pinterest Shares	Reddit Engagements
□ BREAKING – Supreme Court Ruling: NO Islam In Public Schools Apr 17, 2017 donaldtrumpnews.co	165K	1.1K	7	11

Putin Says He Has Proof Princess Diana Was Killed By British Royal Family
By Admin – Jun 9, 2016
365usanews.com

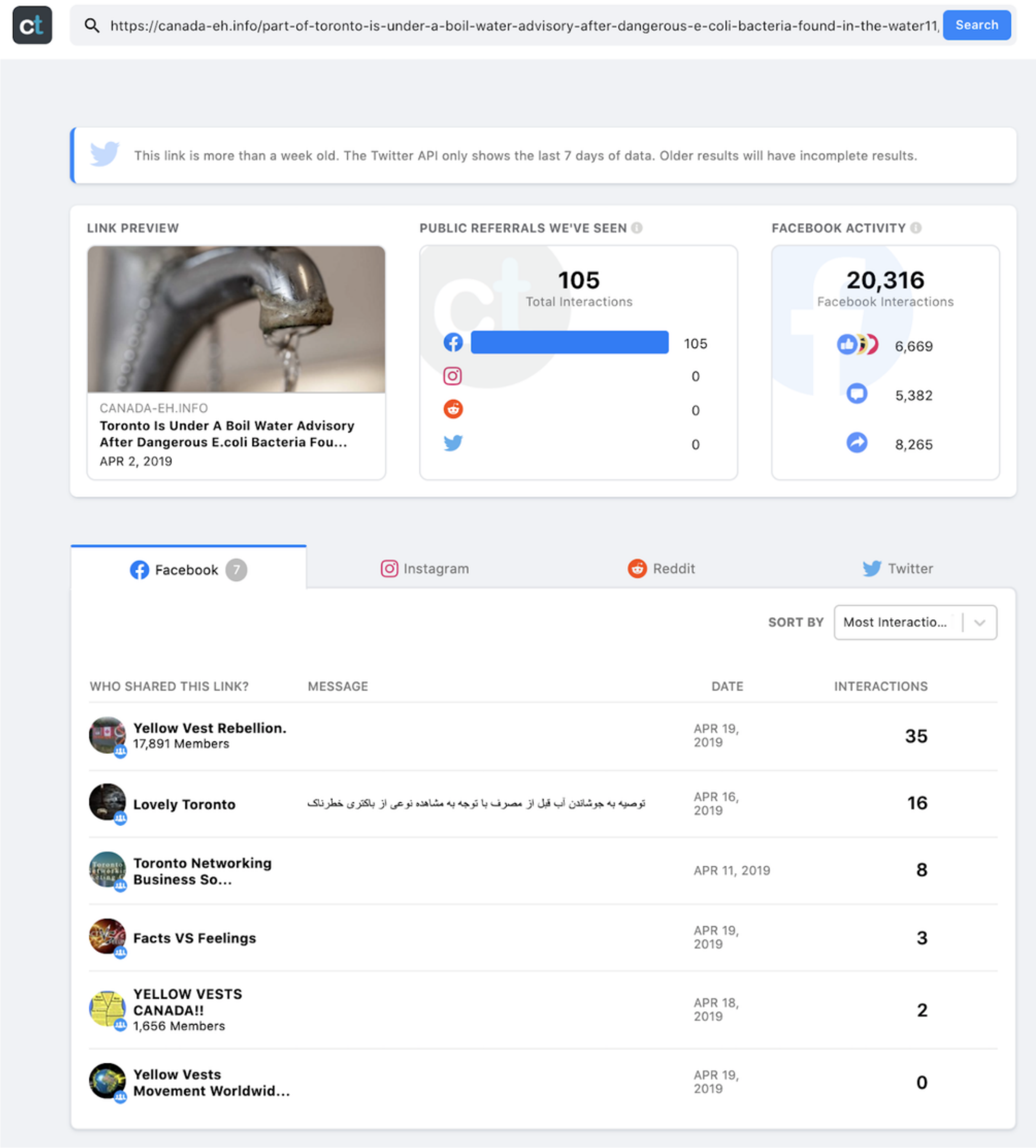
Pope Francis Endorses Bernie Sanders for President!!
By Usa Daily Politics –
Mar 28, 2016
usadailypolitics.com

AG Lynch Announces Global Police Force Partnership With UN - BVA News
Jul 10, 2016
bvanews.com

Another way to identify how a website's content is spreading on Facebook, Twitter, Instagram and Reddit is to install the free CrowdTangle browser extension, or use its web-based link search tool. Both offer the same functionality, but let's work with the web version. (These tools are free, but you need a Facebook account for access.)

The key difference between BuzzSumo and CrowdTangle is that you can enter the URL of a site in BuzzSumo and it will automatically bring up the most-engaged content on that site. CrowdTangle is used to check a specific URL on a site. So if you enter buzzfeednews.com, in CrowdTangle, it's going to show you engagement stats only for that homepage, whereas BuzzSumo will scan across the entire domain for its top content. Another difference is that CrowdTangle's link search tool and extension will show Twitter engagements only from the past seven days. BuzzSumo provides a count of all-time shares on Twitter for articles on the site.

As an example, I entered the URL of an old, false story about a boil water advisory in Toronto into CrowdTangle Link Search. (The site later deleted the story but the URL is still active as of this writing.) CrowdTangle shows that this URL received more than 20,000 reactions, comments and shares on Facebook since being published. It also shows some of the pages and public groups that shared the link, and offers the option to view similar data for Instagram, Reddit and Twitter. Remember: The Twitter tab will show tweets only from the past seven days.



https://canada-eh.info/part-of-toronto-is-under-a-boil-water-advisory-after-dangerous-e-coli-bacteria-found-in-the-water11, Search

This link is more than a week old. The Twitter API only shows the last 7 days of data. Older results will have incomplete results.

LINK PREVIEW

CANADA-EH.INFO
Toronto Is Under A Boil Water Advisory After Dangerous E.coli Bacteria Fou...
APR 2, 2019

PUBLIC REFERRALS WE'VE SEEN

105
Total Interactions

Facebook	105
Instagram	0
Reddit	0
Twitter	0

FACEBOOK ACTIVITY

20,316
Facebook Interactions

Reactions	6,669
Comments	5,382
Shares	8,265

Facebook 7 | Instagram | Reddit | Twitter

SORT BY Most Interactio...

WHO SHARED THIS LINK?	MESSAGE	DATE	INTERACTIONS
Yellow Vest Rebellion. 17,891 Members		APR 19, 2019	35
Lovely Toronto	توصیه به جوشاندن آب قبل از مصرف با توجه به مشاهده نوعی از باکتری خطرناک	APR 16, 2019	16
Toronto Networking Business So...		APR 11, 2019	8
Facts VS Feelings		APR 19, 2019	3
YELLOW VESTS CANADA!! 1,656 Members		APR 18, 2019	2
Yellow Vests Movement Worldwid...		APR 19, 2019	0

Note that the high number of total Facebook interactions is not really reflected in the small list of pages and groups we see. This is at least partly because some of the key pages that spread the link when it was first published were later removed by Facebook. This is a useful reminder that CrowdTangle shows data only from active accounts, and it won't show you every public account that shared a given URL. It's a selection, but is still incredibly useful because it often reveals a clear connection between specific social media accounts and a website. If the same Facebook page is consistently — or exclusively — sharing content from a site, that may signal they're run by the same people. Now you can dig into the page to compare information with the site and potentially identify the people involved and their motivations. Some of the Facebook link share results listed in CrowdTangle may also be of people sharing the article in a Facebook group. Note the account that shared the link, and see if they've spread other content from the site. Again, there could be a connection.

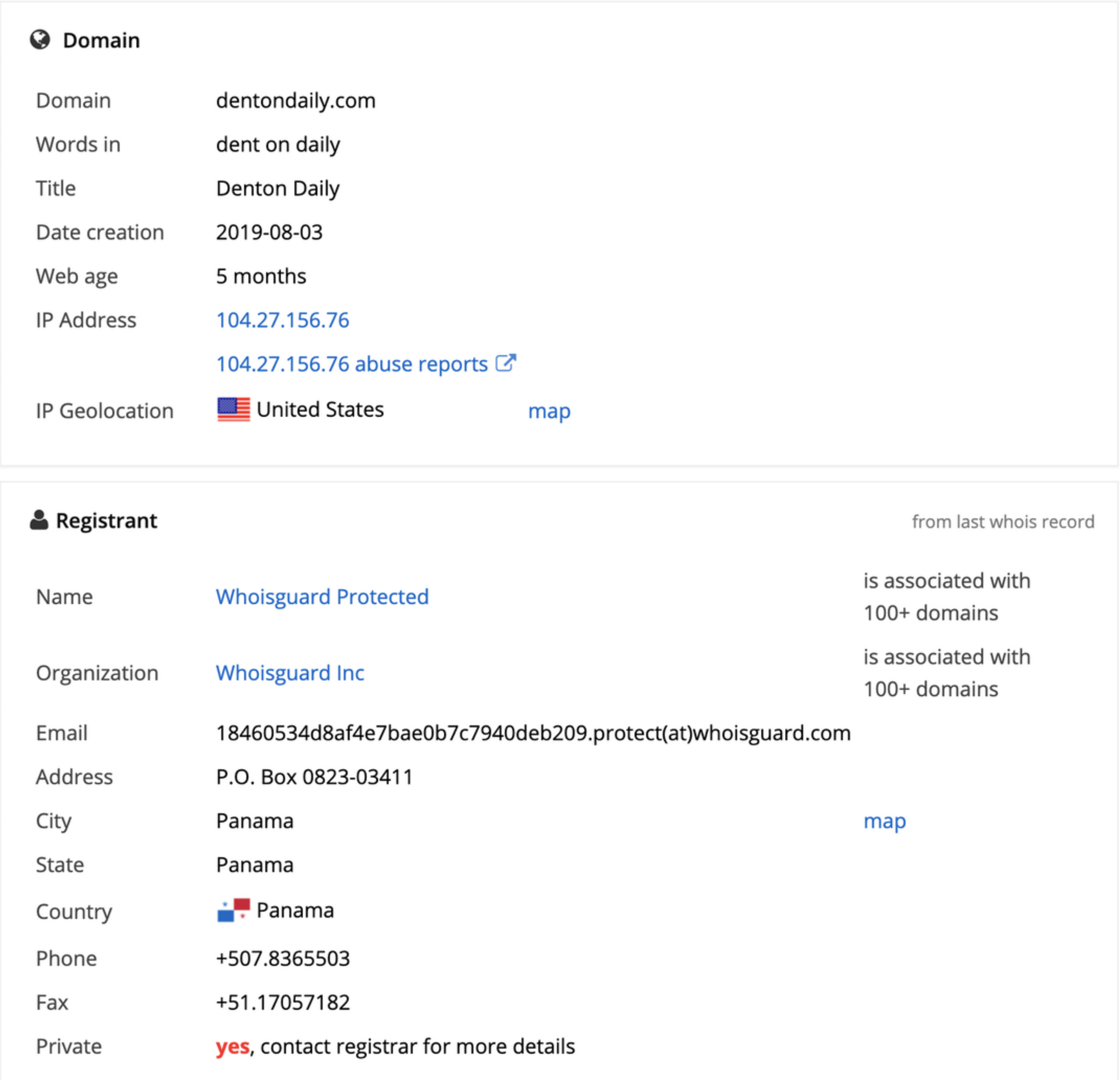
Registration

Every domain name on the web is part of a central database that stores basic information about its creation and history. In some cases, we also get lucky and find information about the person or entity that paid to register a domain. We can pull up this information with a whois search, which is offered by many free tools. There are also a handful of great free and low-priced tools that can bring up additional information, such as who has owned a domain over time, the servers it's been hosted on, and other useful details.

One caveat is that it's relatively inexpensive to pay to have your personal information privacy protected when you register a domain. If you do a whois search on a domain and the result lists something such as “Registration Private,” “WhoisGuard Protected,” or “Perfect Privacy LLC” as the registrant, that means it's privacy protected. Even in those cases, a whois search will still tell us the date the domain was most recently registered, when it will expire and the IP address on the internet where the site is hosted.

DomainBigData is one of the best free tools for investigating a domain name and its history. You can also enter in an email or person or company name to search by that data instead of a URL. Other affordable services you may want to bookmark are DNSlytics, Security Trails and Whoisology. A great but more expensive option is the Iris investigations product from DomainTools.

For example, if we enter dentondaily.com into DomainBigData, we can see it's been privacy protected. It lists the registrant name as “Whoisguard Protected.” Fortunately, we can still see that it was most recently registered in August 2019.



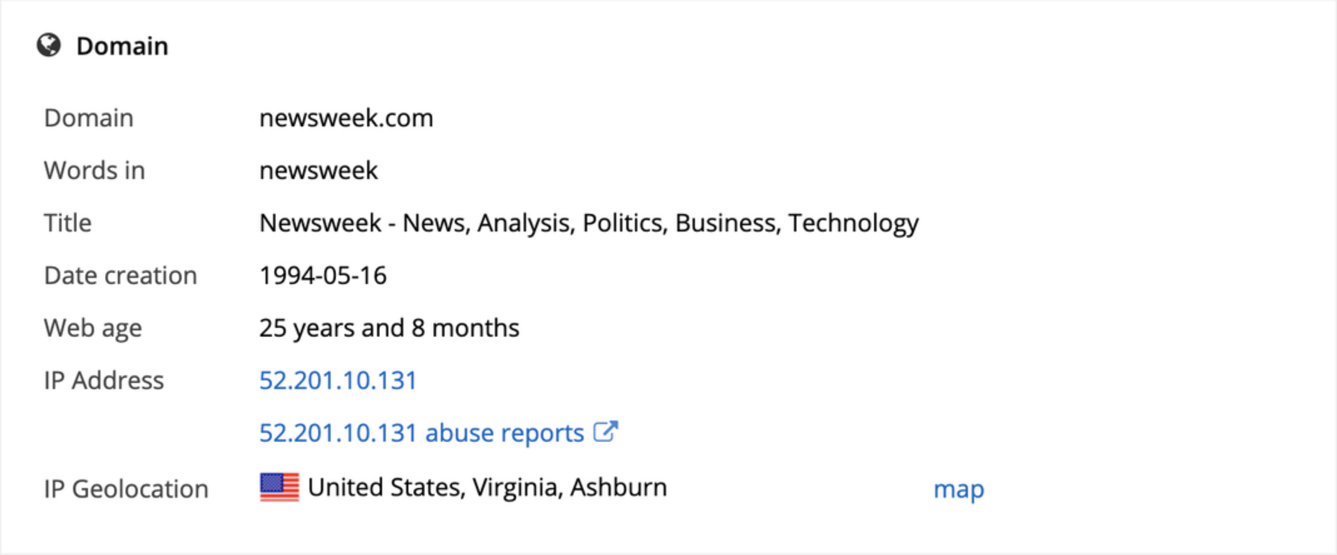
Domain

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

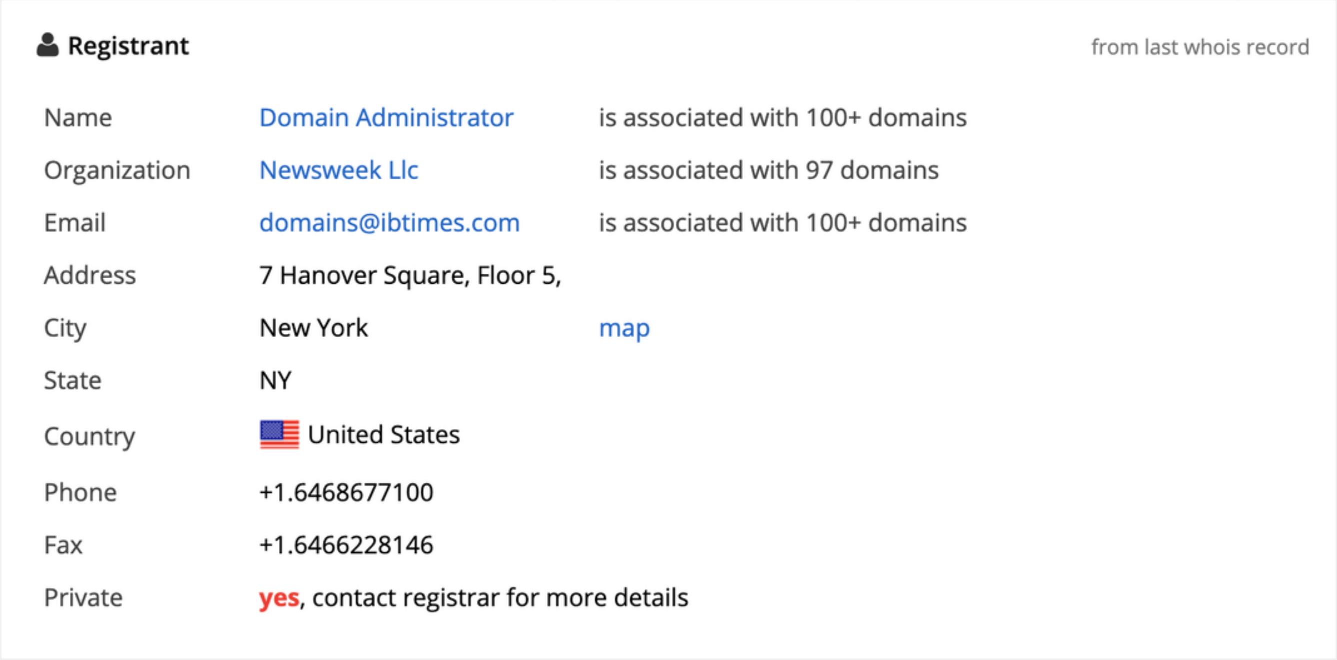
Name	Whoisguard Protected	is associated with 100+ domains
Organization	Whoisguard Inc	is associated with 100+ domains
Email	18460534d8af4e7bae0b7c7940deb209.protect(at)whoisguard.com	
Address	P.O. Box 0823-03411	
City	Panama	map
State	Panama	
Country	Panama	
Phone	+507.8365503	
Fax	+51.17057182	
Private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

For another example, let's search newsweek.com in DomainBigData. We immediately see that the owner has not paid for privacy protection. There's the name of a company, an email address, phone and fax numbers.



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	

We also see that this entity has owned the domain since May 1994, and that the site is currently hosted at the IP address 52.201.10.13. The next thing to note is that the name of the company, the email and the IP address are each highlighted as links. That means they could lead us to other domains that belong to Newsweek LLC, domains@ibtimes.com and other websites hosted at that same IP address. These connections are incredibly important in an investigation, so it's always important to look at other domains owned by the same person or entity.

As for IP addresses, beware that completely unconnected websites can be hosted on the same server. This is usually because people are using the same hosting company for their websites. A general rule is that the fewer the websites hosted on the same server, the more likely they may be connected. But it's not for sure.

If you see hundreds of sites hosted on a server, they may have no ownership connection. But if you see there are only nine, for example, and the one you're interested in has private registration information, it's worth running a whois search on the eight other domains to see if they might have a common owner, and if it's possible that person also owns the site you're investigating. People may pay for privacy protection on some web domains but neglect to do it for others.

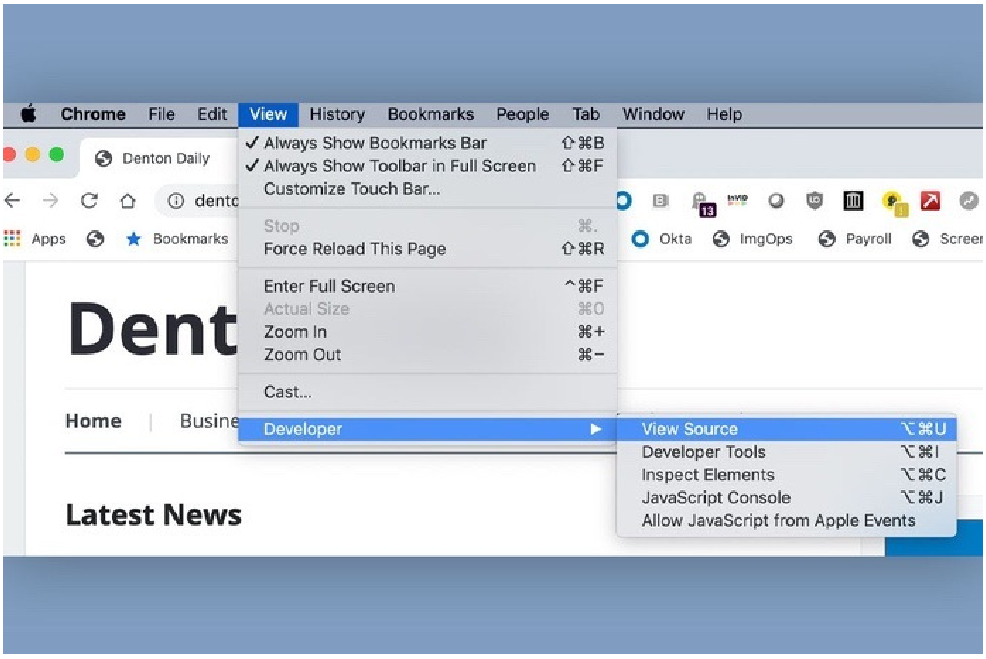
Connecting sites using IP, content and/or registration information is a fundamental way to identify networks and the actors behind them.

Now let’s look at another way to link sites using the code of a webpage.

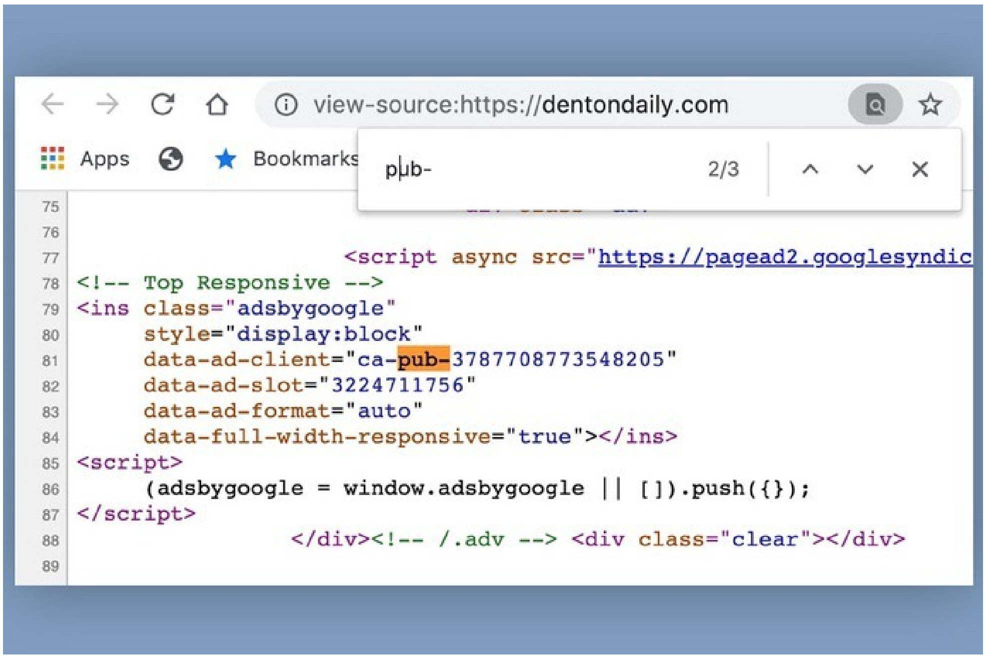
Code and analytics

This approach, first discovered by Lawrence Alexander, begins with viewing the source code of a webpage and then searching within it to see if you can locate a Google Analytics and/or Google AdSense code. These are hugely popular products from Google that, respectively, enable a site owner to track the stats of a website or earn money from ads. Once integrated into a site, every webpage will have a unique ID linked to the owner’s Analytics or AdSense account. If someone is running multiple sites, they often use the same Analytics or AdSense account to manage them. This provides an investigator with the opportunity to connect seemingly separate sites by finding the same ID in the source code. Fortunately, it’s easy to do.

First, go to your target website. Let’s use dentondaily.com. In Chrome for Mac, select the “View” menu then “Developer” and “View Source.” This opens a new tab with the page’s source code. (On Chrome for PC, press ctrl-U.)

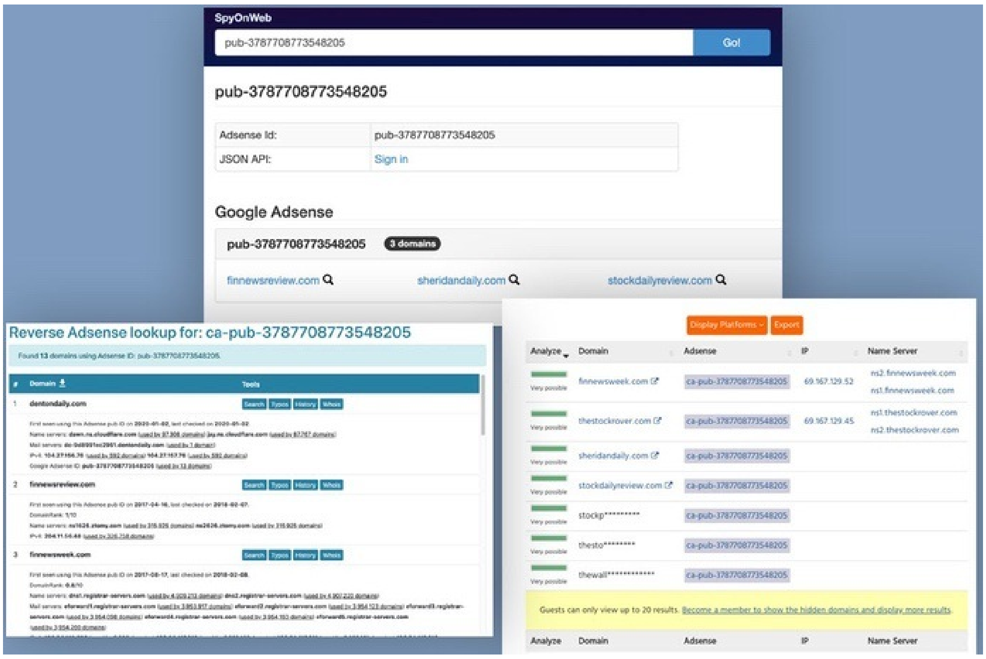


All Google Analytics IDs begin with “ua-” and then have a string of numbers. AdSense IDs have “pub-” and a string of numbers. You can locate then in the source code by simply doing a “find” on the page. On a Mac, type command-F; on a PC it’s ctrl-F. This brings up a small search box. Enter “ua-” or “pub-” and then you’ll see any IDs within the page.



```
                    <script async src="https://pagead2.googlesyndic
<!-- Top Responsive -->
<ins class="adsbygoogle"
     style="display:block"
     data-ad-client="ca-pub-3787708773548205"
     data-ad-slot="3224711756"
     data-ad-format="auto"
     data-full-width-responsive="true"></ins>
<script>
     (adsbygoogle = window.adsbygoogle || []).push({});
</script>
                </div><!-- /.adv --> <div class="clear"></div>
```

If you find an ID, copy it and paste it into the search box in services such as SpyOnWeb, DNSlytics, NerdyData or AnalyzeID. Note that you often receive different results from each service, so it's important to test an ID and compare the results. In the below image, you can see SpyOnWeb found three domains with the same AdSense ID, but DNSlytics and AnalyzeID found several more.



pub-3787708773548205

Adsense Id:	pub-3787708773548205
JSON API:	Sign in

Google Adsense

pub-3787708773548205 3 domains

finnewsreview.com sheridandaily.com stockdailyreview.com

Reverse Adsense lookup for: ca-pub-3787708773548205

Found 13 domains using Adsense ID: pub-3787708773548205

#	Domain	Tools
1	dentondaily.com	Search Typos History Whois
2	finnewsreview.com	Search Typos History Whois
3	finnewsweek.com	Search Typos History Whois

1 dentondaily.com
First seen using this Adsense pub ID on 2020-01-02, last checked on 2020-01-02
Name servers: dawn.ns.cloudflare.com (used by 37,308 domains) jay.ns.cloudflare.com (used by 87,767 domains)
Mail servers: dc-0d8891ec2861.dentondaily.com (used by 1 domain)
IPv4: 104.27.166.76 (used by 592 domains) 104.27.167.76 (used by 592 domains)
Google Adsense ID: pub-3787708773548205 (used by 13 domains)

2 finnewsreview.com
First seen using this Adsense pub ID on 2017-04-16, last checked on 2018-02-07.
DomainRank: 1/10
Name servers: ns1626.ztomy.com (used by 315,926 domains) ns2626.ztomy.com (used by 315,926 domains)
IPv4: 204.11.56.48 (used by 326,758 domains)

3 finnewsweek.com
First seen using this Adsense pub ID on 2017-08-17, last checked on 2018-02-08.
DomainRank: 0.8/10
Name servers: dns1.registrar-servers.com (used by 4,029,213 domains) dns2.registrar-servers.com (used by 4,007,220 domains)
Mail servers: eforward1.registrar-servers.com (used by 3,853,917 domains) eforward2.registrar-servers.com (used by 3,954,123 domains) eforward3.registrar-servers.com (used by 3,954,098 domains) eforward4.registrar-servers.com (used by 3,954,193 domains) eforward5.registrar-servers.com (used by 3,954,200 domains)

Analyze	Domain	Adsense	IP	Name Server
Very possible	finnewsweek.com	ca-pub-3787708773548205	69.167.129.52	ns2.finnewsweek.com ns1.finnewsweek.com
Very possible	thestockrover.com	ca-pub-3787708773548205	69.167.129.45	ns1.thestockrover.com ns2.thestockrover.com
Very possible	sheridandaily.com	ca-pub-3787708773548205		
Very possible	stockdailyreview.com	ca-pub-3787708773548205		
Very possible	stockp*********	ca-pub-3787708773548205		
Very possible	thesto********	ca-pub-3787708773548205		
Very possible	thewall*************	ca-pub-3787708773548205		

Guests can only view up to 20 results. Become a member to show the hidden domains and display more results.

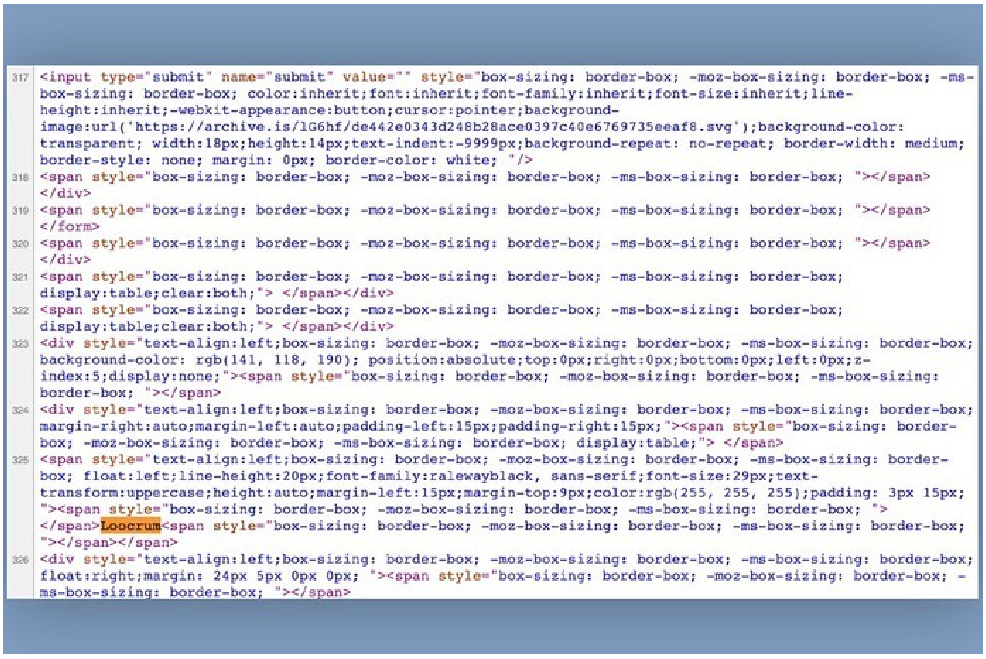
Analyze	Domain	Adsense	IP	Name Server

Sometimes a site had an ID in the past but it's no longer present. That's why it's essential to use the same view source approach on any other sites that allegedly have these IDs listed to confirm they're present. Note that AdSense and Analytics IDs are still present in the archived version of a site in the Wayback Machine. So if you don't find an ID on a live site, be sure to check the Wayback Machine.

All of these services deliver some results for free. But it's often necessary to pay to receive the full results, particularly if your ID is present on a high number of other sites.

A final note on inspecting source code: It's worth scanning the full page even if you don't understand HTML, JavaScript, PHP or other common web programming languages. For example, people sometimes forget to change the title of a page or website if they reuse the same design template. This simple error can offer a point of connection.

While investigating the ad fraud scheme with front companies like Atoses, I was interested in a company called FLY Apps. I looked at the source code of its one-page website and near the top of the site's code I saw the word "Loocrum" in plain text (emphasis added):



```
<input type="submit" name="submit" value="" style="box-sizing: border-box; -moz-box-sizing: border-box; -ms-
box-sizing: border-box; color:inherit;font:inherit;font-family:inherit;font-size:inherit;line-
height:inherit;-webkit-appearance:button;cursor:pointer;background-
image:url('https://archive.is/lG6hf/de442e0343d248b28ace0397c40e6769735eeaf8.svg');background-color:
transparent; width:18px;height:14px;text-indent:-9999px;background-repeat: no-repeat; border-width: medium;
border-style: none; margin: 0px; border-color: white; "/>
<span style="box-sizing: border-box; -moz-box-sizing: border-box; -ms-box-sizing: border-box; "></span>
</div>
<span style="box-sizing: border-box; -moz-box-sizing: border-box; -ms-box-sizing: border-box; "></span>
</form>
<span style="box-sizing: border-box; -moz-box-sizing: border-box; -ms-box-sizing: border-box; "></span>
</div>
<span style="box-sizing: border-box; -moz-box-sizing: border-box; -ms-box-sizing: border-box;
display:table;clear:both;"> </span></div>
<span style="box-sizing: border-box; -moz-box-sizing: border-box; -ms-box-sizing: border-box;
display:table;clear:both;"> </span></div>
<div style="text-align:left;box-sizing: border-box; -moz-box-sizing: border-box; -ms-box-sizing: border-box;
background-color: rgb(141, 118, 190); position:absolute;top:0px;right:0px;bottom:0px;left:0px;z-
index:5;display:none;"><span style="box-sizing: border-box; -moz-box-sizing: border-box; -ms-box-sizing:
border-box; "></span>
<div style="text-align:left;box-sizing: border-box; -moz-box-sizing: border-box; -ms-box-sizing: border-box;
margin-right:auto;margin-left:auto;padding-left:15px;padding-right:15px;"><span style="box-sizing: border-
box; -moz-box-sizing: border-box; -ms-box-sizing: border-box; display:table;"> </span>
<span style="text-align:left;box-sizing: border-box; -moz-box-sizing: border-box; -ms-box-sizing: border-
box; float:left;line-height:20px;font-family:ralewayblack, sans-serif;font-size:29px;text-
transform:uppercase;height:auto;margin-left:15px;margin-top:9px;color:rgb(255, 255, 255);padding: 3px 15px;
"><span style="box-sizing: border-box; -moz-box-sizing: border-box; -ms-box-sizing: border-box; ">
</span>Loocrum<span style="box-sizing: border-box; -moz-box-sizing: border-box; -ms-box-sizing: border-box;
"></span></span>
<div style="text-align:left;box-sizing: border-box; -moz-box-sizing: border-box; -ms-box-sizing: border-box;
float:right;margin: 24px 5px 0px 0px; "><span style="box-sizing: border-box; -moz-box-sizing: border-box; -
ms-box-sizing: border-box; "></span>
```

Googling that word brought up a company called Loocrum that used the exact same website design as FLY Apps, and had some of the same content. A whois search revealed that the email address used to register loocrum.com had also been used to register other shell companies I previously identified in the scheme. This connection between FLY Apps and Loocrum provided important additional evidence that the four men running FLY Apps were linked to this overall scheme. And it was revealed by simply scrolling through the source code looking for plain text words that seemed out of place.

Conclusion

Even with all of the above approaches and tools under your belt, you might sometimes feel as though you've hit a dead end. But there's often another way to find connections or avenues for further investigation on a website. Click every link, study the content, read the source code, see who's credited the site, see who's sharing it, and examine anything else you can think of to reveal what's really going on.

9. Analyzing ads on social networks

Written by: Johanna Wild

__Johanna Wild__ is an open-source investigator at Bellingcat, where she also focuses on tech and tool development for digital investigations. She has an online journalism background and previously worked with journalists in (post-) conflict regions. One of her roles was to support journalists in Eastern Africa to produce broadcasts for the Voice of America.

The ads you see on your social media timeline are not the same that the people sitting next to you on public transportation see on theirs. Based on factors like your location, gender, age and the things you liked or shared on the network, you might be shown ads for luxurious holiday suites in Málaga while your neighbor sees ads for Japanese mobile games.

Microtargeting, categorizing users into target groups to show them ads that fit their life circumstances and interests, has become a major concern during elections. The worry is that campaigns could target very small slices of the population with ads that stoke fear or hatred, or that spread false information. In general, ads from politicians placed on social networks are not subject to fact-checking. Facebook, for instance, reaffirmed in January 2020 that it will continue to allow any political ad as long as it abides by Facebook's community standards. This means specific user groups could be targeted with ads that contain disinformation on crucial political or social topics.

Until recently, it was nearly impossible for journalists and researchers to gain insights into the ads targeted to different users. In response to public criticism about the lack of transparency, several social networks created ad libraries that allow anyone to review information about ads published on their platforms.

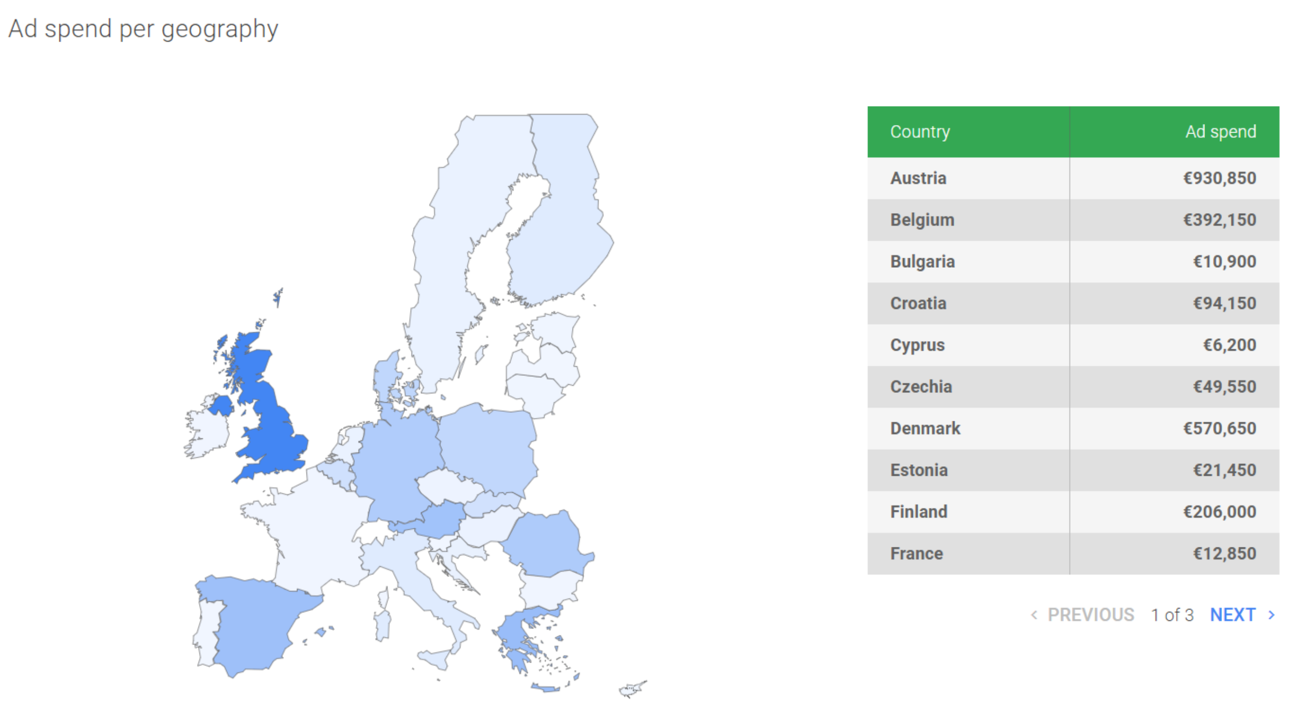
In particular, Facebook's library has been accused of not reliably showing all available ads. So whenever you use these libraries, take some time to check whether all the ads that you see on your timeline can also be found there.

Ad libraries are nevertheless an important step toward more transparency and provide journalists and others with exciting new ways of investigating digital advertisements. The following techniques will help you get started on investigating ads placed on major platforms like Google, Twitter and Facebook.

Google

Google's ads center is well hidden within its Transparency Report. Use this link to access the political advertising section, which provides information on Google and YouTube ads from the European Union, India and the United States.

The page for each region shows a list of countries and the total ad spend since the launch of the report.

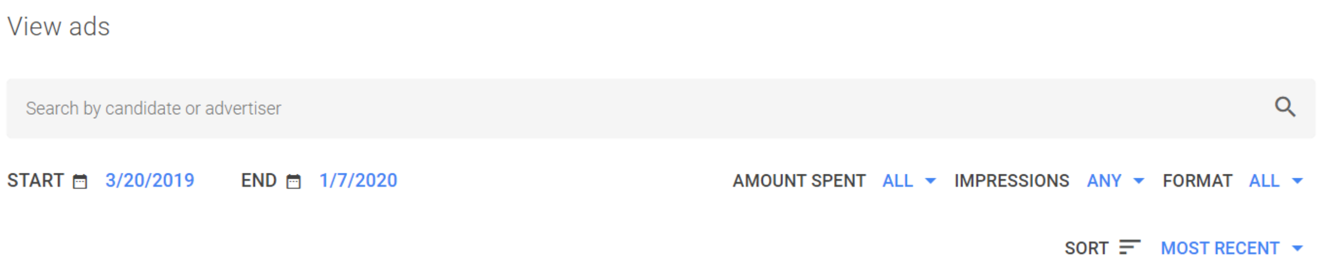


Ad spend per geography

Country	Ad spend
Austria	€930,850
Belgium	€392,150
Bulgaria	€10,900
Croatia	€94,150
Cyprus	€6,200
Czechia	€49,550
Denmark	€570,650
Estonia	€21,450
Finland	€206,000
France	€12,850

‹ PREVIOUS 1 of 3 NEXT ›

Click on a country and you will be led to a page containing its ads database:



View ads

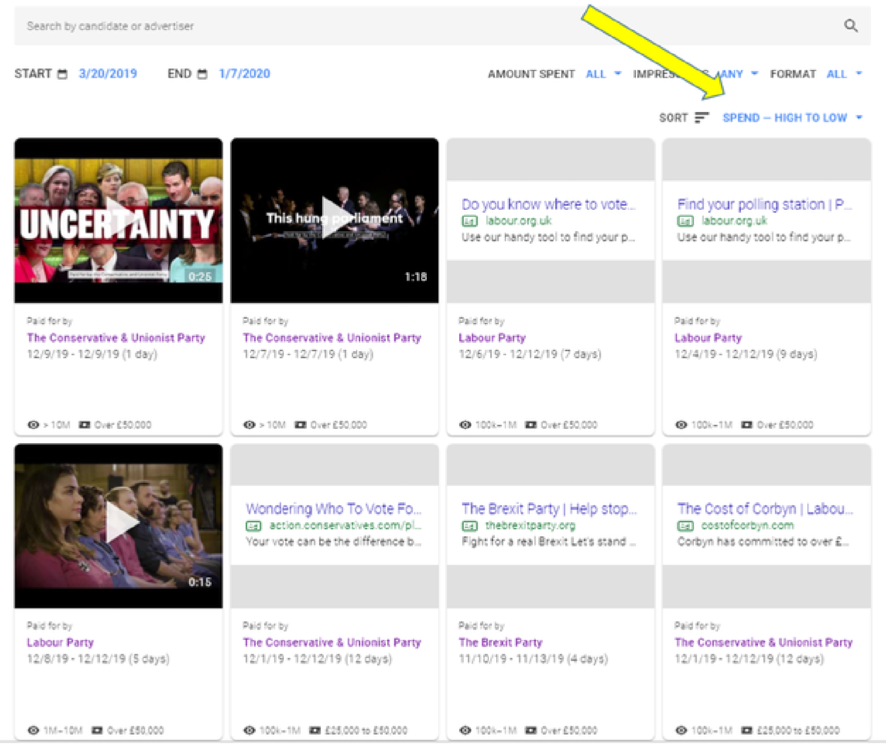
Search by candidate or advertiser

START 3/20/2019 END 1/7/2020 AMOUNT SPENT ALL IMPRESSIONS ANY FORMAT ALL

SORT MOST RECENT

You can filter the results by date, the amount of money spent and the number of times an ad is shown to users (impressions). You can filter by the format of the ad if you want to view results for video, image or text-based ads.

It’s also easy to find the biggest spenders. For example, if you want to view the biggest political ad campaigns placed in the U.K. since the launch of the report until January 2020, simply change the “sort” category to “spend – high to low,” as shown below.



START 3/20/2019 END 1/7/2020 AMOUNT SPENT ALL IMPRES ANY FORMAT ALL

SORT SPEND – HIGH TO LOW

Paid for by
The Conservative & Unionist Party
12/9/19 - 12/9/19 (1 day)
> 10M Over £50,000

Paid for by
The Conservative & Unionist Party
12/7/19 - 12/7/19 (1 day)
> 10M Over £50,000

Do you know where to vote...
Ad labour.org.uk
Use our handy tool to find your p...
Paid for by
Labour Party
12/6/19 - 12/12/19 (7 days)
100k–1M Over £50,000

Find your polling station | P...
Ad labour.org.uk
Use our handy tool to find your p...
Paid for by
Labour Party
12/4/19 - 12/12/19 (9 days)
100k–1M Over £50,000

Paid for by
Labour Party
12/8/19 - 12/12/19 (5 days)
1M–10M Over £50,000

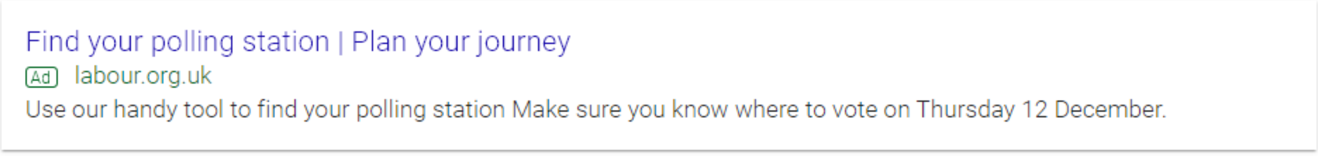
Wondering Who To Vote Fo...
Ad action.conservatives.com/pl...
Your vote can be the difference b...
Paid for by
The Conservative & Unionist Party
12/1/19 - 12/12/19 (12 days)
100k–1M £25,000 to £50,000

The Brexit Party | Help stop...
Ad thebrexitparty.org
Fight for a real Brexit Let's stand ...
Paid for by
The Brexit Party
11/10/19 - 11/13/19 (4 days)
100k–1M Over £50,000

The Cost of Corbyn | Labou...
Ad costofcorbyn.com
Corbyn has committed to over £...
Paid for by
The Conservative & Unionist Party
12/1/19 - 12/12/19 (12 days)
100k–1M £25,000 to £50,000

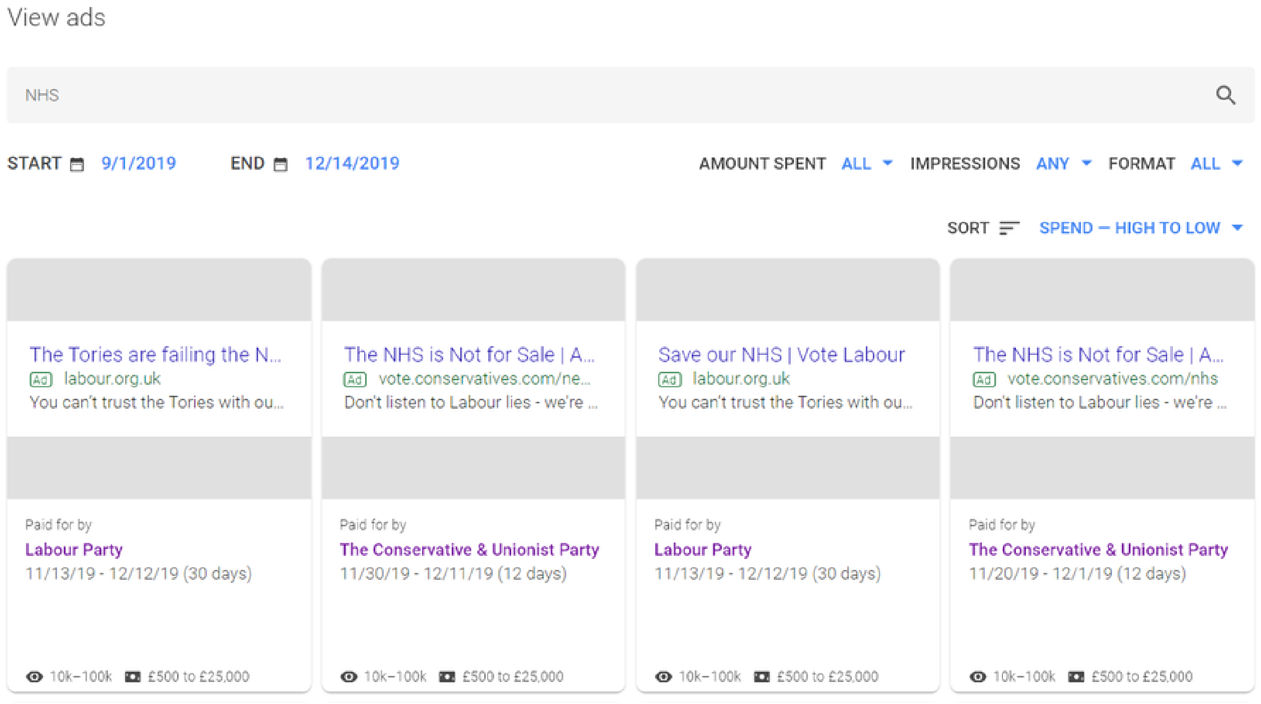
Unsurprisingly, the biggest ad buys came just before and on the day of the General Election, Dec. 12, 2019. You can also see that the Conservative & Unionist Party invested more than £50,000 each on two YouTube ads that ran for just one day.

The Labour Party, in contrast, spent more than £50,000 for an ad on Google's search results pages for a tool it said could help voters find their polling station.



Find your polling station | Plan your journey
Ad labour.org.uk
Use our handy tool to find your polling station Make sure you know where to vote on Thursday 12 December.

You can also search by keyword. Type in NHS (for National Health Service) and you will see that in November and December 2019 the Labour Party and the Conservatives purchased Google search ads to criticize each other's plans for the NHS.



View ads

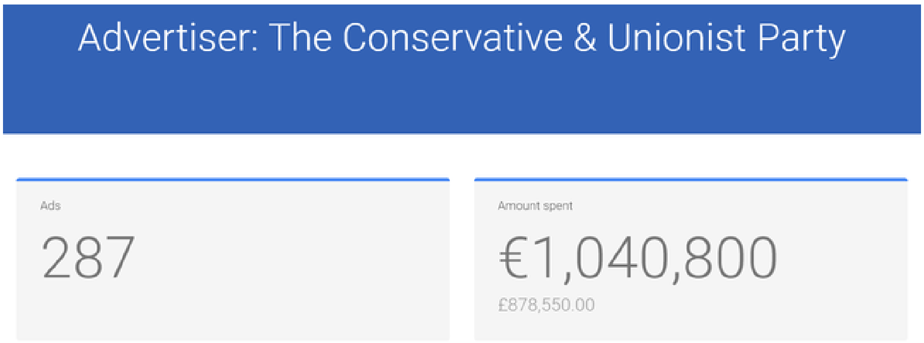
NHS

START 9/1/2019 END 12/14/2019 AMOUNT SPENT ALL IMPRESSIONS ANY FORMAT ALL

SORT SPEND — HIGH TO LOW

The Tories are failing the N... Ad labour.org.uk You can't trust the Tories with ou...	The NHS is Not for Sale | A... Ad vote.conservatives.com/ne... Don't listen to Labour lies - we're ...	Save our NHS | Vote Labour Ad labour.org.uk You can't trust the Tories with ou...	The NHS is Not for Sale | A... Ad vote.conservatives.com/nhs Don't listen to Labour lies - we're ...
Paid for by Labour Party 11/13/19 - 12/12/19 (30 days)	Paid for by The Conservative & Unionist Party 11/30/19 - 12/11/19 (12 days)	Paid for by Labour Party 11/13/19 - 12/12/19 (30 days)	Paid for by The Conservative & Unionist Party 11/20/19 - 12/1/19 (12 days)
10k–100k £500 to £25,000	10k–100k £500 to £25,000	10k–100k £500 to £25,000	10k–100k £500 to £25,000

By clicking on the name of the advertiser, you can also check the total amount of money they've spent on Google ads since the launch of the Transparency Report. Here's what that looked like for the two leading U.K. political parties as of January 2020:



Advertiser: The Conservative & Unionist Party

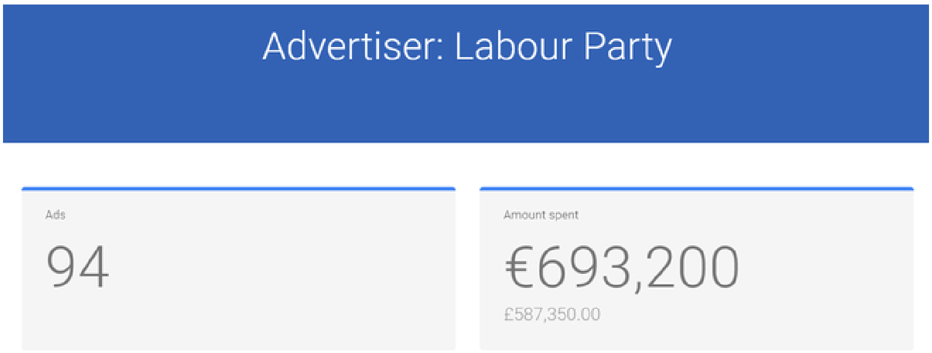
Ads

287

Amount spent

€1,040,800

£878,550.00



Advertiser: Labour Party

Ads

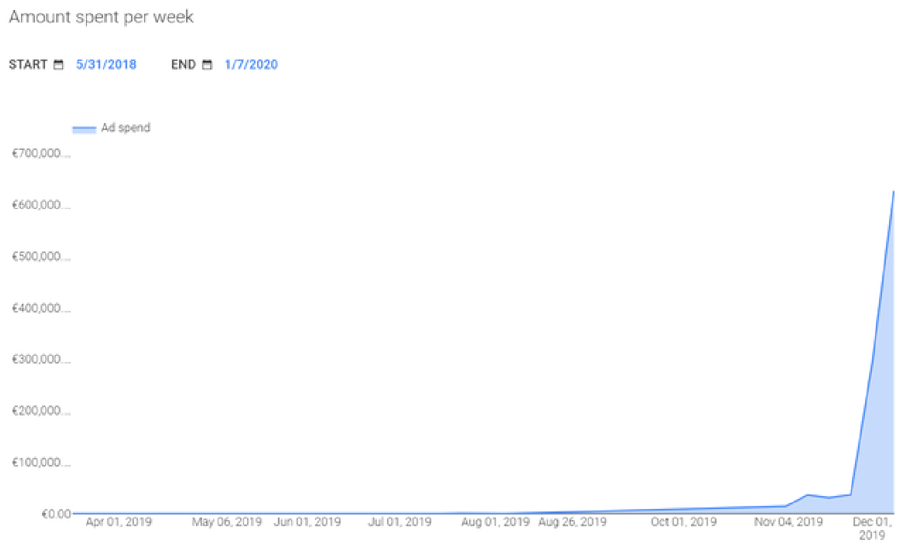
94

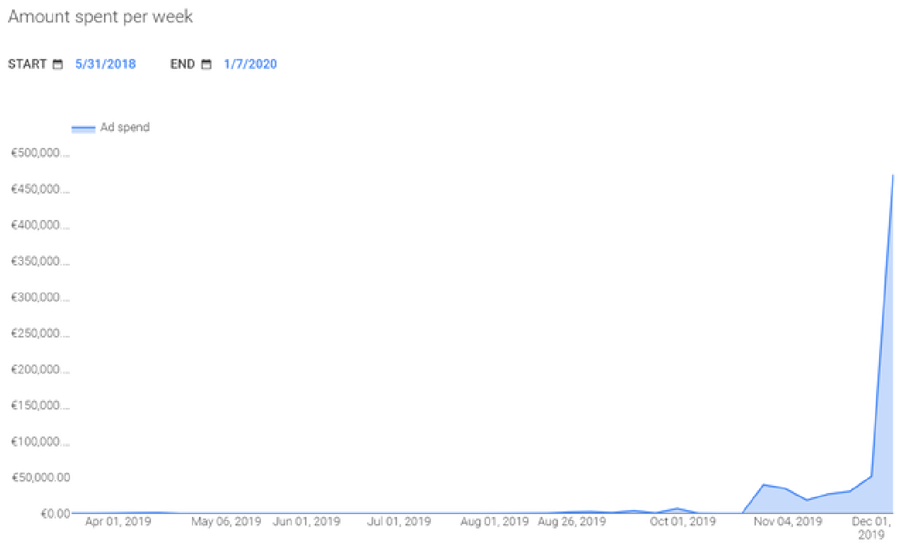
Amount spent

€693,200

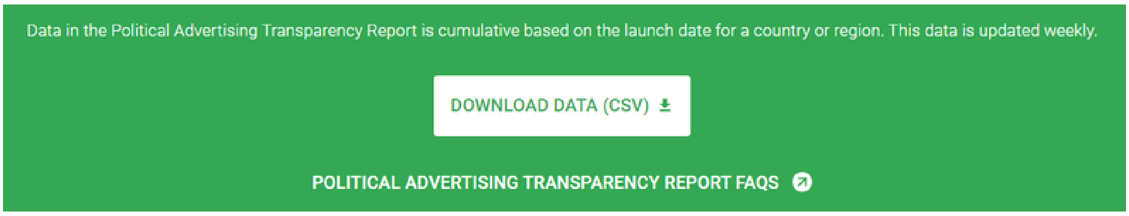
£587,350.00

You can also view a timeline of their spend. The reports on the left shows the spending pattern for the Conservative & Unionist Party, and the one on the right is for the Labour Party:





If you want to further analyze the ads database, scroll down until you see a green section called “download data,” which allows you to download the data in CSV format.



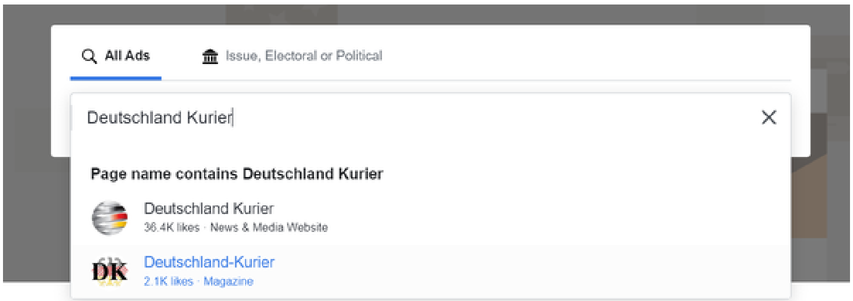
Data in the Political Advertising Transparency Report is cumulative based on the launch date for a country or region. This data is updated weekly.

This enables you to import the data into a spreadsheet program like Google Sheets or Excel so you can perform additional filtering and analysis.

Facebook

The Facebook ad library is divided into two parts: “All Ads” and “Issue, Electoral or Political.” If you click on “All Ads,” you can search for specific advertisers by name only, instead of also using keywords.

For example, if I want to see ads from Deutschland Kurier, a publication that often publishes content in support of German far-right party AfD, I can type its name and Facebook will recommend pages with that text:



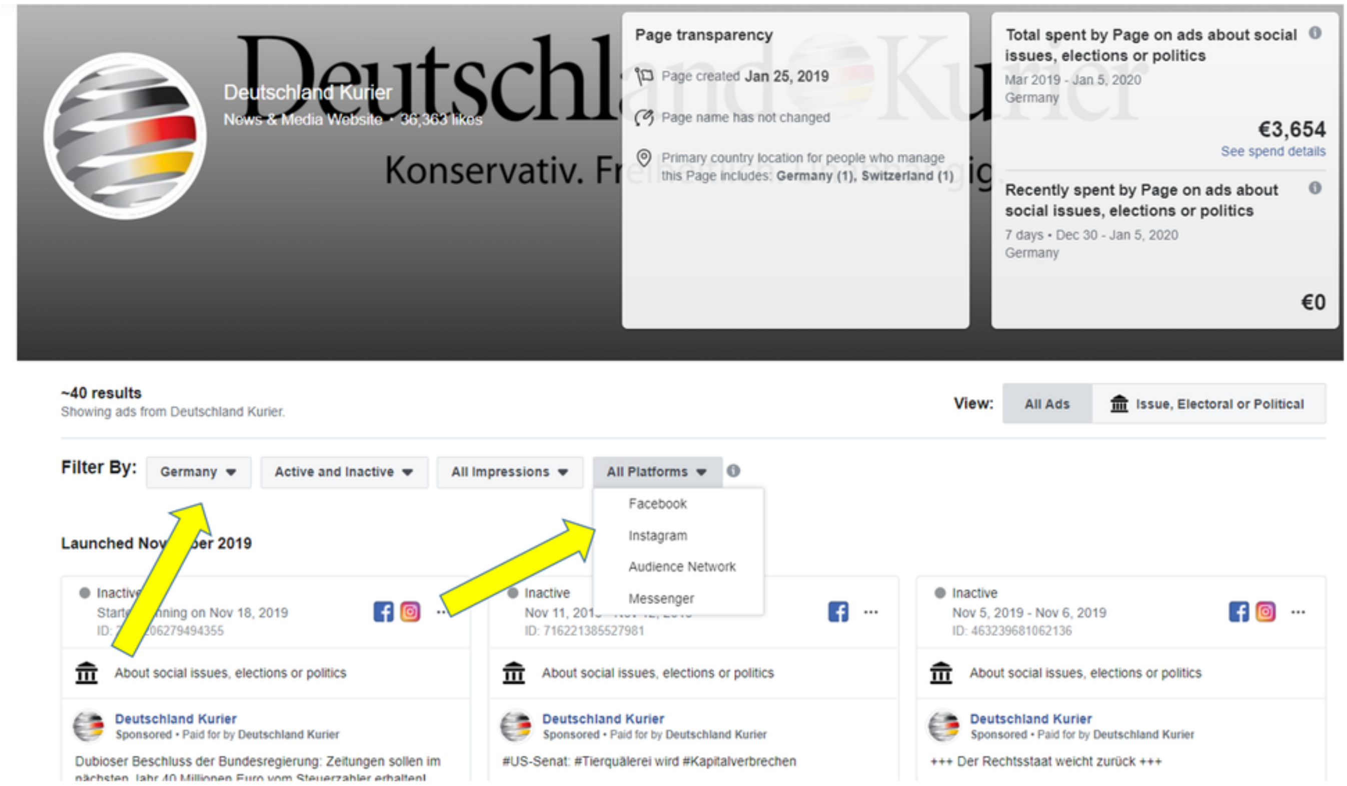
Deutschland Kurier

Page name contains Deutschland Kurier

Deutschland Kurier
36.4K likes · News & Media Website

Deutschland-Kurier
2.1K likes · Magazine

The results page shows that Deutschland Kurier placed ads worth 3,654 euros in Germany between March 2019 and January 2020.



Deutschland Kurier
News & Media Website · 36,363 likes

Page transparency
Page created Jan 25, 2019
Page name has not changed
Primary country location for people who manage this Page includes: Germany (1), Switzerland (1)

Total spent by Page on ads about social issues, elections or politics
Mar 2019 - Jan 5, 2020
Germany
€3,654
See spend details

Recently spent by Page on ads about social issues, elections or politics
7 days • Dec 30 - Jan 5, 2020
Germany
€0

~40 results
Showing ads from Deutschland Kurier.

Filter By: Germany, Active and Inactive, All Impressions, All Platforms

Launched Nov ber 2019

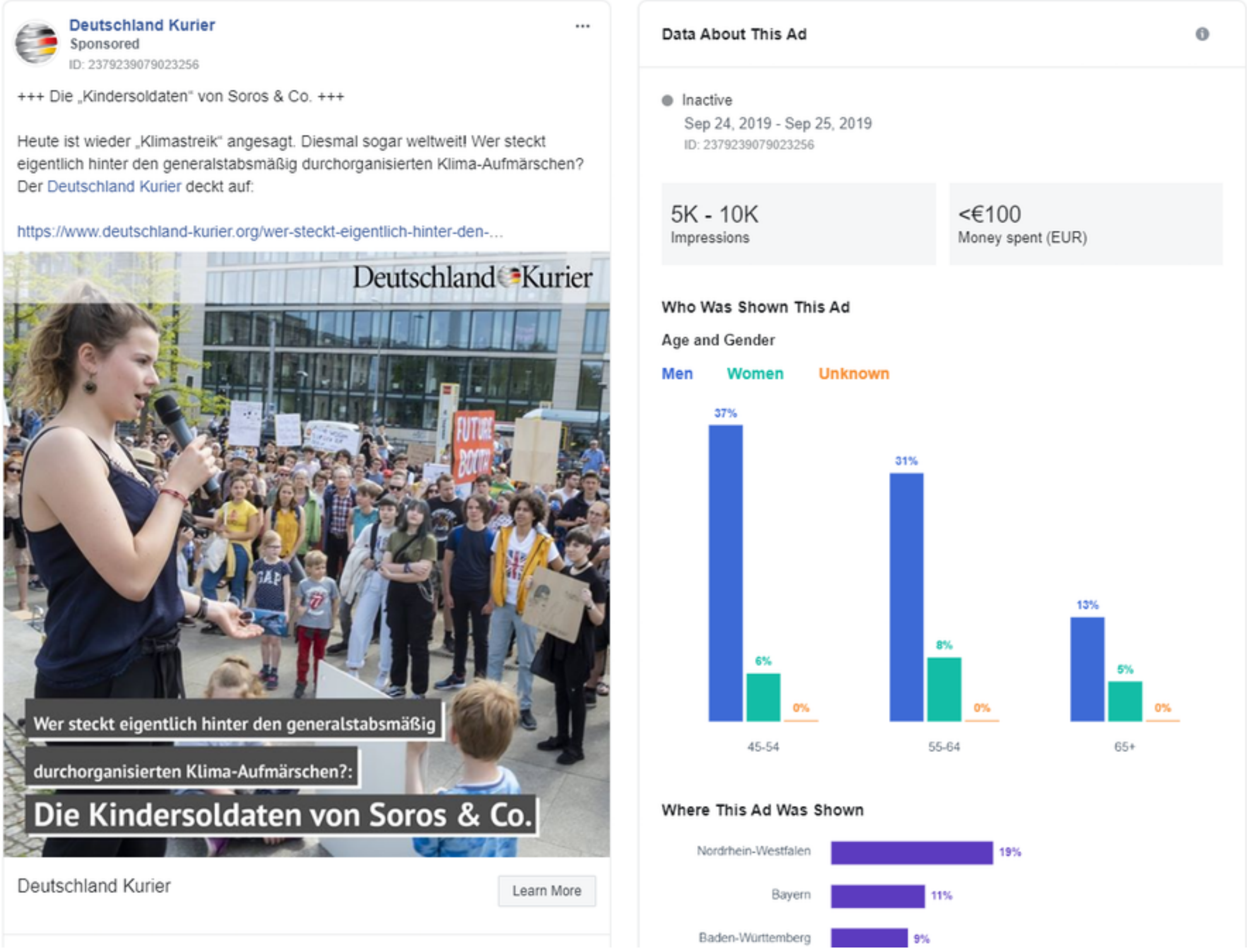
Inactive
Start nning on Nov 18, 2019
ID: 7 06279494355
About social issues, elections or politics
Deutschland Kurier
Sponsored • Paid for by Deutschland Kurier
Dubioser Beschluss der Bundesregierung: Zeitungen sollen im nächsten Jahr 40 Millionen Euro vom Steuerzahler erhalten!

Inactive
Nov 11, 201
ID: 716221385527981
About social issues, elections or politics
Deutschland Kurier
Sponsored • Paid for by Deutschland Kurier
#US-Senat: #Tierquälerei wird #Kapitalverbrechen

Inactive
Nov 5, 2019 - Nov 6, 2019
ID: 463239681062136
About social issues, elections or politics
Deutschland Kurier
Sponsored • Paid for by Deutschland Kurier
+++ Der Rechtsstaat weicht zurück +++

Once on the results page, make sure to select the correct country for your search (or "all"), and to choose whether you want to see ads from Facebook, Instagram, Messenger or Facebook Audience Network. Audience Network is an ad network operated by Facebook that places ads on mobile apps and websites other than Facebook's own properties. In most cases, the best choice will be to search across all platforms to get a full picture of an organization's ads.

On an individual ad you can click the "See ad details" button to view additional information.



Deutschland Kurier
Sponsored
ID: 2379239079023256

+++ Die „Kindersoldaten" von Soros & Co. +++

Heute ist wieder „Klimastreik" angesagt. Diesmal sogar weltweit! Wer steckt eigentlich hinter den generalstabsmäßig durchorganisierten Klima-Aufmärschen? Der Deutschland Kurier deckt auf:

https://www.deutschland-kurier.org/wer-steckt-eigentlich-hinter-den-...

Deutschland Kurier
Wer steckt eigentlich hinter den generalstabsmäßig
durchorganisierten Klima-Aufmärschen?:
Die Kindersoldaten von Soros & Co.

Deutschland Kurier

Data About This Ad

● Inactive
Sep 24, 2019 - Sep 25, 2019
ID: 2379239079023256

5K - 10K
Impressions

<€100
Money spent (EUR)

Who Was Shown This Ad

Age and Gender

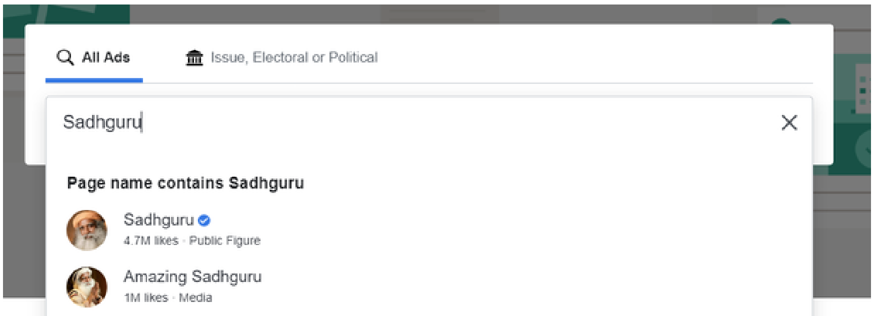
Where This Ad Was Shown

In this case, Deutschland Kurier spent less than €100 for this ad that calls climate change protesters "child soldiers of Soros & Co.," and it had between 5,000 and 10,000 impressions, mostly displayed to men aged 45 and older.

The second option for searching the ads library is to choose "Issue, Electoral or Political" database, which is an archive of ads about "social issues, elections or politics." The big advantage of this option is that you can search for any keyword you like, and these kinds of ads are archived by Facebook.

Let's look at an example.

Sadhguru is the name of a well-known Indian spiritual figure who says he's not associated with any political party. He has said he sees it as his duty to support any current government "to do their best." If you type in his name in the "All Ads" section, Facebook suggests Sadhguru's personal Facebook page.



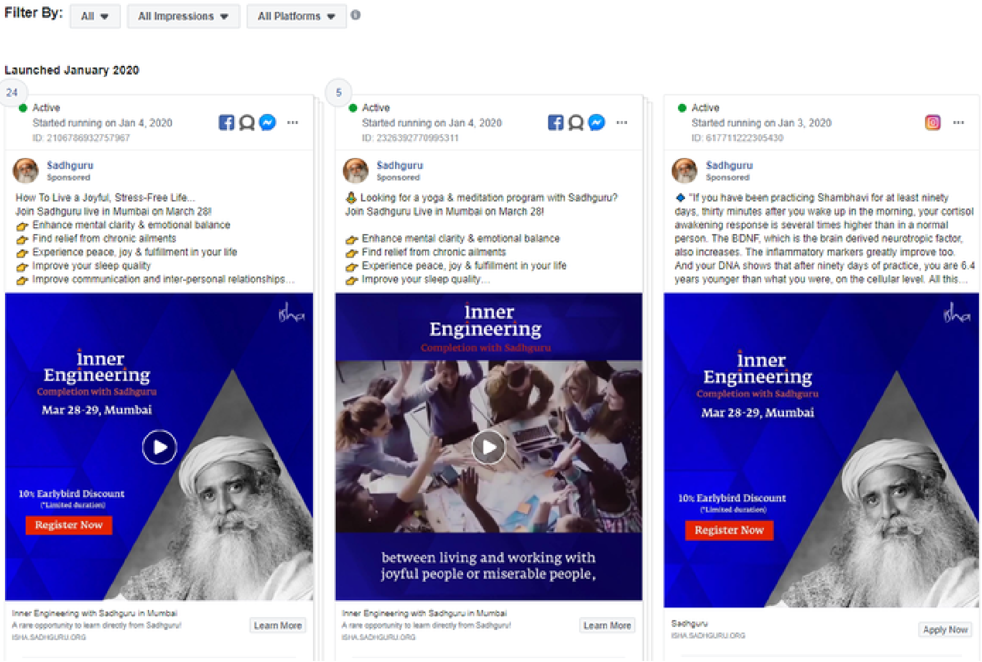
Sadhguru

Page name contains Sadhguru

Sadhguru
4.7M likes · Public Figure

Amazing Sadhguru
1M likes · Media

This shows us a selection of apolitical ads published by Sadhguru in which he promotes his yoga and meditation courses.



Launched January 2020

Active
Started running on Jan 4, 2020
ID: 2106786932757967

Sadhguru
Sponsored

How To Live a Joyful, Stress-Free Life...
Join Sadhguru live in Mumbai on March 28!
👉 Enhance mental clarity & emotional balance
👉 Find relief from chronic ailments
👉 Experience peace, joy & fulfillment in your life
👉 Improve your sleep quality
👉 Improve communication and inter-personal relationships...

Inner Engineering with Sadhguru in Mumbai
A rare opportunity to learn directly from Sadhguru!
ISHA.SADHGURU.ORG

Active
Started running on Jan 4, 2020
ID: 2326392770995311

Sadhguru
Sponsored

Looking for a yoga & meditation program with Sadhguru?
Join Sadhguru Live in Mumbai on March 28!

👉 Enhance mental clarity & emotional balance
👉 Find relief from chronic ailments
👉 Experience peace, joy & fulfillment in your life
👉 Improve your sleep quality...

Inner Engineering
Completion with Sadhguru
between living and working with
joyful people or miserable people,

Inner Engineering with Sadhguru in Mumbai
A rare opportunity to learn directly from Sadhguru!
ISHA.SADHGURU.ORG

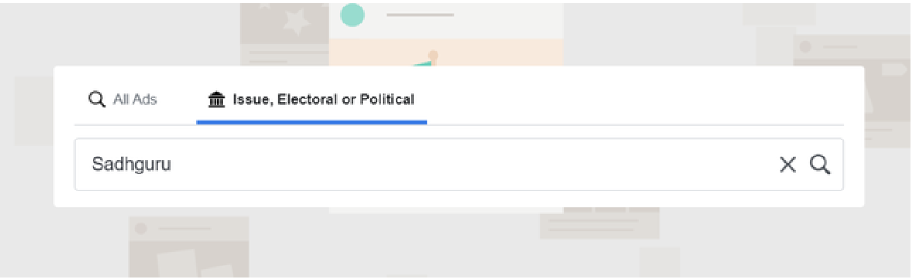
Active
Started running on Jan 3, 2020
ID: 617711222305430

Sadhguru
Sponsored

🔷 "If you have been practicing Shambhavi for at least ninety days, thirty minutes after you wake up in the morning, your cortisol awakening response is several times higher than in a normal person. The BDNF, which is the brain derived neurotropic factor, also increases. The inflammatory markers greatly improve too. And your DNA shows that after ninety days of practice, you are 6.4 years younger than what you were, on the cellular level. All this...

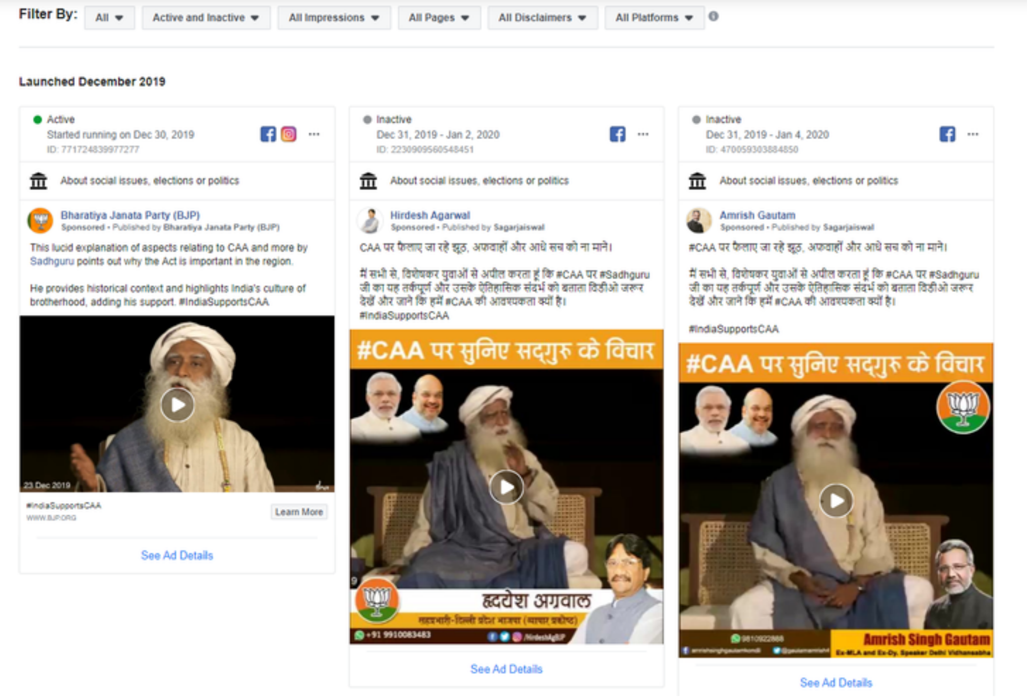
Sadhguru
ISHA.SADHGURU.ORG

Now let's type in his name into the "Issue, Electoral or Political" search bar without accepting the Facebook page suggestions that come up:



Sadhguru

The results change drastically. You can now see a collection of ads mentioning Sadhguru's name published by other accounts.



Launched December 2019

Active
Started running on Dec 30, 2019
ID: 771724839977277

About social issues, elections or politics

Bharatiya Janata Party (BJP)
Sponsored • Published by Bharatiya Janata Party (BJP)

This lucid explanation of aspects relating to CAA and more by Sadhguru points out why the Act is important in the region.

He provides historical context and highlights India's culture of brotherhood, adding his support. #IndiaSupportsCAA

#IndiaSupportsCAA
WWW.BJP.ORG

Inactive
Dec 31, 2019 - Jan 2, 2020
ID: 2230909560548451

About social issues, elections or politics

Hirdesh Agarwal
Sponsored • Published by Sagarjaiswal

CAA पर फैलाए जा रहे झूठ, अफवाहों और आधे सच को ना माने।

मैं सभी से, विशेषकर युवाओं से अपील करता हूं कि #CAA पर #Sadhguru जी का यह तर्कपूर्ण और उसके ऐतिहासिक संदर्भ को बताता विडीओ जरूर देखें और जाने कि हमें #CAA की आवश्यकता क्यों है।
#IndiaSupportsCAA

Inactive
Dec 31, 2019 - Jan 4, 2020
ID: 470059303884850

About social issues, elections or politics

Amrish Gautam
Sponsored • Published by Sagarjaiswal

#CAA पर फैलाए जा रहे झूठ, अफवाहों और आधे सच को ना माने।

मैं सभी से, विशेषकर युवाओं से अपील करता हूं कि #CAA पर #Sadhguru जी का यह तर्कपूर्ण और उसके ऐतिहासिक संदर्भ को बताता विडीओ जरूर देखें और जाने कि हमें #CAA की आवश्यकता क्यों है।

#IndiaSupportsCAA

#CAA पर सुनिए सद्गुरु के विचार
Amrish Singh Gautam
Ex-MLA and Ex-Dy. Speaker Delhi Vidhansabha
See Ad Details

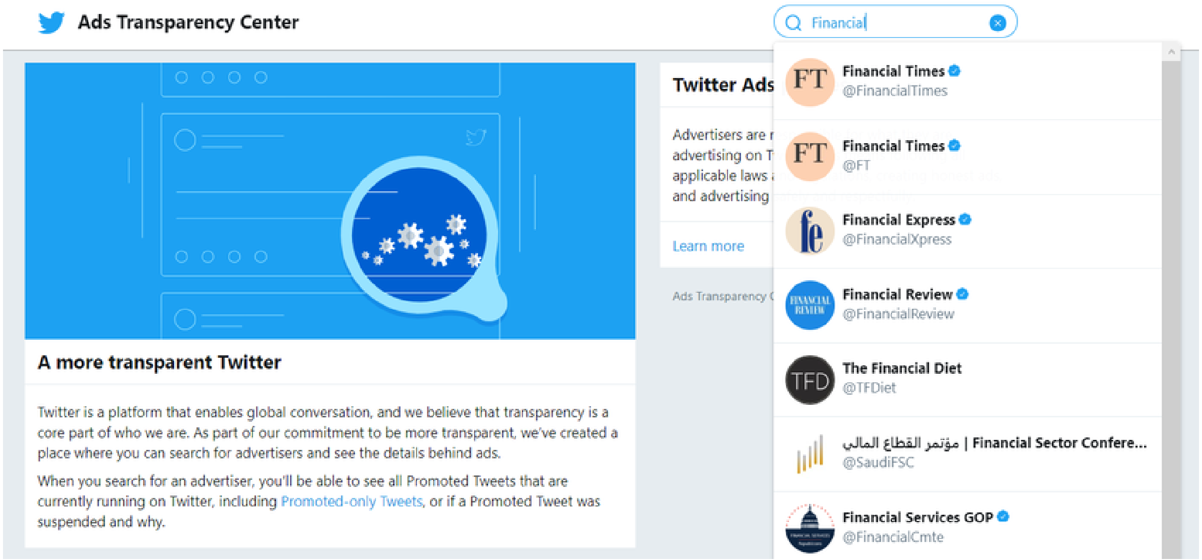
One ad from the ruling Indian nationalist party BJP shows a video in which Sadhguru pronounces his support for the party's controversial Citizenship Amendment bill. The bill allows unregistered immigrants from some of India's neighboring countries to attain Indian citizenship more easily but does not grant the same opportunity to Muslims. The ad provides one hint to the possible relationship between Sadhguru and BJP, a topic that is widely discussed in India.

This example shows how to use Facebook's ad library to add key information to your investigations. You may also want to have a look at the Facebook Ad library report, which extracts key insights from political ads in different countries.

Twitter

In late 2019, Twitter decided to ban political advertising from its platform. However, it's still possible to use the social network's ads transparency center to gain information about nonpolitical ads from the past seven days.

Finding ads is cumbersome because there's no keyword search functionality. To start a search, go to the box in the upper right corner and type in a specific username or handle.

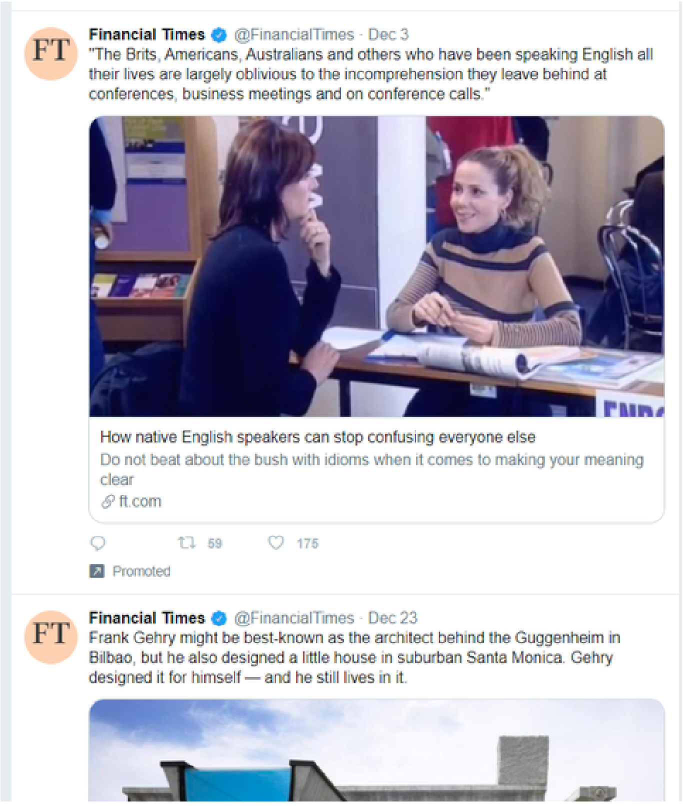

Ads Transparency Center
Financial
Twitter Ads
Advertisers are r
advertising on T
applicable laws a
and advertising
Learn more
Ads Transparency C
Financial Times @FinancialTimes
Financial Times @FT
Financial Express @FinancialXpress
Financial Review @FinancialReview
The Financial Diet @TFDiet
مؤتمر القطاع المالي | Financial Sector Confere... @SaudiFSC
Financial Services GOP @FinancialCmte

A more transparent Twitter

Twitter is a platform that enables global conversation, and we believe that transparency is a core part of who we are. As part of our commitment to be more transparent, we've created a place where you can search for advertisers and see the details behind ads.

When you search for an advertiser, you'll be able to see all Promoted Tweets that are currently running on Twitter, including Promoted-only Tweets, or if a Promoted Tweet was suspended and why.

If there were ads in the last seven days, you will now see them listed.



Financial Times @FinancialTimes · Dec 3
"The Brits, Americans, Australians and others who have been speaking English all their lives are largely oblivious to the incomprehension they leave behind at conferences, business meetings and on conference calls."

How native English speakers can stop confusing everyone else
Do not beat about the bush with idioms when it comes to making your meaning clear
ft.com

Financial Times @FinancialTimes · Dec 23
Frank Gehry might be best-known as the architect behind the Guggenheim in Bilbao, but he also designed a little house in suburban Santa Monica. Gehry designed it for himself — and he still lives in it.

Searching for The Financial Times, we can see it paid to try to generate more interest in its story "How native speakers can stop confusing everyone else." The tweet was sent on Dec. 3, 2019, but the ad information from Twitter doesn't detail when this paid promotion exactly ran.

To speed up your searches, you can use a small trick. Once you have conducted a search, take a look at the URL in your browser:

```
ads.twitter.com/transparency/FinancialTimes
```

The URL always uses the same structure, with a Twitter handle at the end. Simply delete the last part and replace it with another handle:

```
ads.twitter.com/transparency/Bellingcat
```

Refresh the page and you will now see the ads information for Bellingcat. If that account hasn't run any ads in the past seven days, you'll see the message "This account hasn't promoted any ads in the last seven days." Since you can only see ads from the previous seven days, the best thing you can do is to check back frequently to see if an account of note has run ads, and to take screenshots each time you see new ads.

Snapchat

The "Snap political ads library" offers insights into political, "issue related" or advocacy ads. The latter are defined as "ads concerning issues or organisations that are the subject of debate on a local, national or global level, or are of public importance." For instance, topics such as immigration, education or guns.

If you go to the library, you will see a list of years.

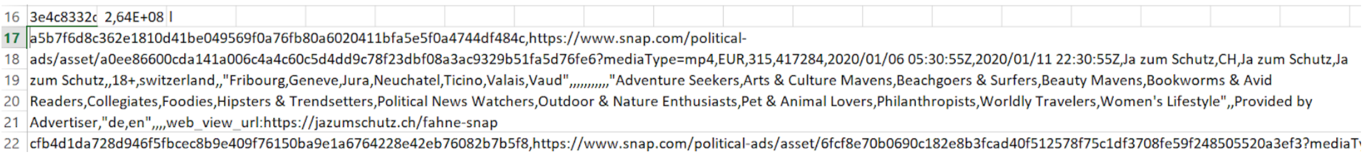
Archives

2018

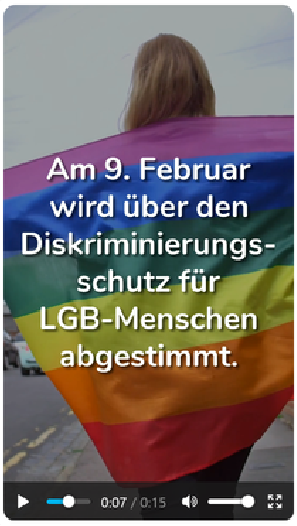
2019

2020

Click on one of the years and you can download a spreadsheet with all available information about ads from that year. The content of the spreadsheet doesn't look very exciting at first sight but it actually is! Each line represents an ad and it shows you who placed the ad, the amount of money spent on it, and even which characteristics were chosen to micro target users.



```
16  3e4c8332c  2,64E+08 l
17  a5b7f6d8c362e1810d41be049569f0a76fb80a6020411bfa5e5f0a4744df484c,https://www.snap.com/political-
18  ads/asset/a0ee86600cda141a006c4a4c60c5d4dd9c78f23dbf08a3ac9329b51fa5d76fe6?mediaType=mp4,EUR,315,417284,2020/01/06 05:30:55Z,2020/01/11 22:30:55Z,Ja zum Schutz,CH,Ja zum Schutz,Ja
19  zum Schutz,,18+,switzerland,,"Fribourg,Geneve,Jura,Neuchatel,Ticino,Valais,Vaud",,,,,,,,,,,,"Adventure Seekers,Arts & Culture Mavens,Beachgoers & Surfers,Beauty Mavens,Bookworms & Avid
20  Readers,Collegiates,Foodies,Hipsters & Trendsetters,Political News Watchers,Outdoor & Nature Enthusiasts,Pet & Animal Lovers,Philanthropists,Worldly Travelers,Women's Lifestyle",,Provided by
21  Advertiser,"de,en",,,,web_view_url:https://jazumschutz.ch/fahne-snap
22  cfb4d1da728d946f5fbcec8b9e409f76150ba9e1a6764228e42eb76082b7b5f8,https://www.snap.com/political-ads/asset/6fcf8e70b0690c182e8b3fcad40f512578f75c1df3708fe59f248505520a3ef3?mediaT
```


Am 9. Februar wird über den Diskriminierungs-schutz für LGB-Menschen abgestimmt.
0:07 / 0:15

In the example above, the advertiser wanted to target "Adventure Seekers,Arts & Culture Mavens,Beachgoers & Surfers,Beauty Mavens,Bookworms & Avid Readers,Collegiates,Foodies,Hipsters & Trendsetters,Political News Watchers,Outdoor & Nature Enthusiasts,Pet & Animal Lovers,Philanthropists,Worldly Travelers,Women's Lifestyle."

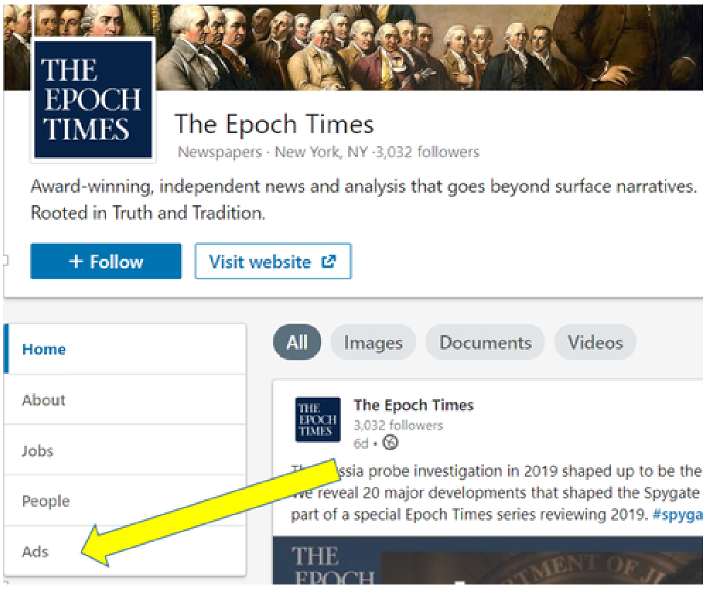
Other platforms do not offer this kind of targeting information in their ad libraries.

You also find a URL in the spreadsheet that allows you to see the actual ad. In this example, I found a message that encouraged people to order free rainbow flags in support of an upcoming vote in Switzerland related to the protection against discrimination of LGBT people.

LinkedIn

LinkedIn does not allow political ads on its platform and it does not have an ads library. Luckily, there's another way to get insights into a specific company's advertising on the platform.

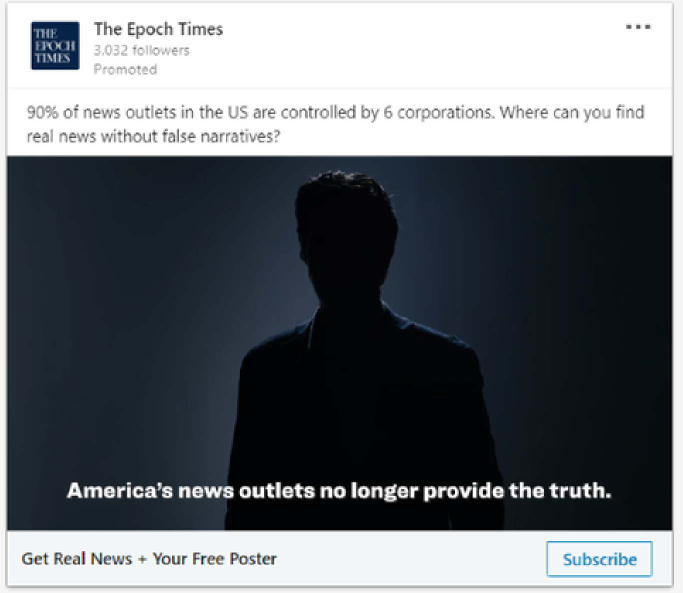
If you go to the company's LinkedIn page, you will see a tab called "Ads" at the bottom of the left column.



Award-winning, independent news and analysis that goes beyond surface narratives. Rooted in Truth and Tradition.

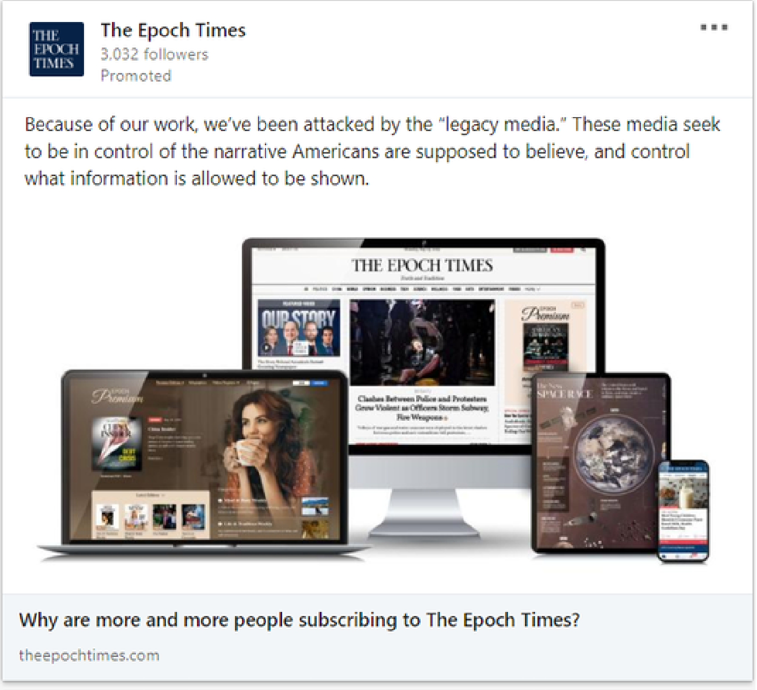
Th ssia probe investigation in 2019 shaped up to be the we reveal 20 major developments that shaped the Spygate part of a special Epoch Times series reviewing 2019. #spyga

Click on that tab and LinkedIn will show you a list of all ads published by that company in the previous six months. Using this feature, it was possible to see that the Epoch Times was still publishing ads on LinkedIn after it had been banned from doing the same on Facebook. The company's two ads claimed that "America's news outlets no longer provide the truth" and contrasted that claim by presenting The Epoch Times as "independent" and "non-partisan media."



THE EPOCH TIMES
The Epoch Times
3,032 followers
Promoted
America's news outlets no longer provide the truth.
Get Real News + Your Free Poster
Subscribe

90% of news outlets in the US are controlled by 6 corporations. Where can you find real news without false narratives?



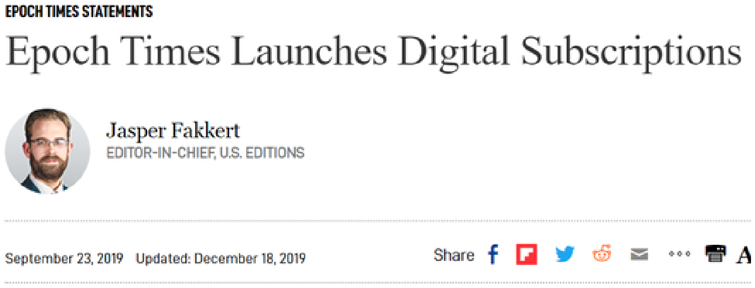
The Epoch Times
3,032 followers
Promoted

Because of our work, we've been attacked by the "legacy media." These media seek to be in control of the narrative Americans are supposed to believe, and control what information is allowed to be shown.

THE EPOCH TIMES
OUR STORY
Premium
Clashes Between Police and Protesters Grow Violent as Officers Storm Subway, Fire Weapons
The New SPACE RACE

Why are more and more people subscribing to The Epoch Times?
theepochtimes.com

The exact publishing dates are not visible, but you can click on the ad (this will work even if it is not active on LinkedIn) and sometimes the destination site provides a more concrete date. The first Epoch Times ad led to a text dated as "September 23, 2019" and "Updated: December 18, 2019," which helped estimate when it might have been online.



EPOCH TIMES STATEMENTS

Epoch Times Launches Digital Subscriptions

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

September 23, 2019 Updated: December 18, 2019

Once you get to know their hidden features, ad libraries are an easy and powerful addition to your digital investigation arsenal, and an important element to check when investigating a person or entity with a social media presence.

10. Tracking actors across platforms

Written by **Ben Collins**

***Ben Collins** is an NBC News reporter covering disinformation, extremism and the internet. For the past five years, he's reported on the rise of conspiracy theories, hate communities, foreign manipulation campaigns and platform failures. He previously worked at The Daily Beast, where his team discovered the accounts, groups and real-life events created by Russia's Internet Research Agency troll farm during the 2016 U.S. election.*

On August 3, 2019, Patrick Crusius walked into an El Paso Walmart and killed 22 people in a white nationalist-motivated shooting. But before he entered the store, he posted a manifesto to the /pol/ political discussion board on 8chan.net, an anonymous message board that has in recent years become a gathering place for white nationalists. The /pol/ boards on 4chan and 8chan are almost entirely unmoderated, and by the summer of 2019, 8chan had become a gathering place of violent white nationalist content and discussion.

Partly because of this, 8chan users would sometimes alert authorities and journalists when a new, violent manifesto was posted. This was done by adding comments beneath the manifesto itself and through online tip submissions to media or law enforcement. When the El Paso shooter first submitted his manifesto — which initially went up with the wrong attachment — one user replied "Hello FBI." The correct manifesto was then posted directly underneath the comment flagging the FBI.

This sort of self-reporting can be critical information for journalists in the wake of these tragedies. In some cases, slightly benevolent users will take to more open and mainstream, civilian parts of the web like Reddit and Twitter to call out manifestos or suspicious posts made before shootings. This is essential because it's easy to miss a relevant post or comment on 4chan and 8chan.

Anonymous platforms like 4chan and 8chan play an important role in the online mis- and disinformation and trolling ecosystem because they're where people often work together to hatch and coordinate campaigns. Reddit, another popular place where users are largely anonymous, hosts a diverse array of online communities. Some are heavily moderated subreddits that can help users trade stories about hobbies or discus news and events; others are basically free-for-alls where hate can breed unabated. It's essential for journalists to know how to monitor and report on all of these communities, and know the intricacies of how they operate.

With that in mind, here are five rules to abide when events require you to use 4chan or 8chan (or its newer iteration 8kun) to inform your reporting:

1. Don't trust anything on 4chan/8chan.
2. Don't trust anything on 4chan/8chan.
3. Don't trust anything on 4chan/8chan.
4. Some useful information pertaining to (or even evidence of) a crime, trolling campaign or disinformation might be found on 4chan/8chan.
5. Don't trust anything on 4chan/8chan.

I can't stress how important it is for reporters to follow rules 1, 2, 3 and 5, even if it prevents them from getting some of the important juice that could be garnered from number 4. These websites are literally built to troll, spread innuendo and falsehoods about perceived enemies, push lies about marginalized people, and, occasionally, post quasi-funny lies framed as true stories about what it's like to be a teenager.

This is evidenced by the fact they have been used as dumping grounds for manifestos by white nationalist, incel and other aggrieved young male shooters.

Let's say it one more time: If it's on 4chan or 8chan (which we'll continue to refer to as 8chan from here on out, despite its merely nominal name change to 8kun), there's a very good chance it's a lie meant to sow chaos and mess with reporters. Don't go into a thread asking for more details. Don't post anything, actually. You will be targeted by people with too much time on their hands.

Confirming the manifesto

This is why it's so helpful when members of these communities make an effort to call out manifestos or other newsworthy content. The "Hello FBI" comment on 8chan is how I found out about the El Paso manifesto's existence. Shortly after reports of the shooting, I searched Twitter with the keywords "El Paso 4chan" and "El Paso 8chan." Searching for "[city name] + [8chan or 4chan or incels.co] or other extremist sites provides a useful template for any similar event.

My Twitter search revealed that a few users had shared screenshots of the shooter's 8chan posts, though most had falsely attributed the post to someone on 4chan. So I needed to look for the post.

What's the fastest way to search for an 8chan post? Google. In the aftermath of the shooting, I searched for "site:8ch.net" then added a part of a sentence from the alleged 8chan post from the shooter. (Note: 4chan automatically deletes posts from its servers after a certain period of time, but there are automatic 4chan archive sites. The most comprehensive one is called 4plebs.org. Archived 4chan posts can be found by simply replacing 4chan in the URL with 4plebs, and removing the "boards" prefix. For example: boards.4chan.org/pol/13561062.html could be found at 4plebs.org/pol/13561062.html.)

During some shootings, it might be beneficial to try searching for "site:4chan.net + 'manifesto' or 'fbi'" and use Google's search options to restrict your time frame to the past 24 hours. Chan users might have already attempted to rat out the shooter in replies to their post.

My initial search strategy didn't turn up the relevant 8chan post, which led me to believe this was a quickly created hoax. But something didn't sit right. The post shown in the screenshot on Twitter did, in fact, have a user ID and post number. These details led me to think it was real, and not a simple fake. On 8chan, each post comes from a unique user ID, which is algorithmically generated and displayed next to the post date. This system allows users to have a static ID so they can identify themselves within a thread.

This user ID system, by the way, is how people know "Q" from the QAnon conspiracy theory is actually him. Users can create de facto permanent usernames and passwords by entering a username in the ID field while making a post, followed by a #, followed by a password.

This user ID is how I knew the same person who mistakenly posted the PDF with the name of the shooter on it was the same user as the one who posted the actual manifesto two minutes later. Both posts shared the randomly created same user ID: 58820b.

Next to a user ID is a post number, which is a somewhat permanent artifact that creates a unique URL for each post. The screenshot of the El Paso manifesto shared on Twitter included a post ID of "No.13561062." This creates the url 8ch.net/pol/res/13561062.html. You can use this URL convention across both 4chan and 8chan.

But in this case, the post didn't exist. I thought maybe it had been deleted. (I later learned that 8chan owner Jim Watkins removed it once he was alerted to its content.)

With the post gone, my last best hope was that it had been archived by someone who recognized its importance. Thankfully, a quick-thinking 8chan user saved the post on the archive site archive.is. Pasting the URL into the "I want to search the archive for saved snapshots" box of archive.is revealed that the manifesto post was real, and now I could view it.

But there was a new problem: When was it first posted on 8chan? I needed an accurate timestamp to confirm that the manifesto was posted before the El Paso shooter began his rampage.

Both 4chan and 8chan localize their timestamps, making it a complicated task to derive the real time from archiving sites. Fortunately, there's a foolproof way around this. Right-clicking the timestamp and clicking "inspect element" will bring up the site's source code, and it will highlight a section that starts with '<time unixtime='[number].'"

Copy and paste that number into an Epoch/Unix timestamp converter, like unixtimestamp.com, and you'll get a to-the-second post timestamp in UTC time. Converting from UTC time to El Paso time revealed that the manifesto was posted at 10:15 a.m. Central Time — minutes before the shooting began.

This work helped me confirm that the manifesto posted on 8chan manifesto was, in fact, a legitimate piece of evidence in a case of racist domestic terrorism.

Tracking actors across platforms

In 2017, Lane Davis, a former "Gamergate researcher" (read: professional internet stalker) for disgraced alt-right figure Milo Yiannopoulos, killed his father in his own home.

Davis had gotten into an argument with his parents, and a 911 call revealed he was spouting far-right internet extremist jargon shortly before the attack. He referred to his parents as "leftist pedophiles" before his father called the police to help him kick Davis out of their home, where his son still lived.

Davis was known as "Seattle4Truth" online, and in YouTube videos he frequently referred to fictitious secret pedophile rings he believed were the driving force behind liberalism. One video on YouTube under his name was titled, "Progressive ideology's deep ties to pedophilia."

A reporter's dream scenario in online extremism investigations is a perpetrator using a static username across platforms, and that was the case with Davis. He identified himself as Seattle4Truth on YouTube and on Reddit, where his posts revealed an even more conspiracy-addled brain.

How was that discovered? By simply putting seattle4truth into the Reddit username URL convention: reddit.com/u/[username]. Once there, you can sort by newest posts, most popular posts, and most "controversial," which ranks posts by how a combination of how many times they were upvoted and downvoted.

One way to quickly research a username is to use Namechk, which searches for a username across close to 100 internet services. As I detail below, that doesn't mean the same person is running these accounts, but it's an efficient way to see where the username is being used so you can dig in and research. You can also Google any username you're interested in.

It's also important to be aware of the kind of super-niche internet communities where your target could be active. A 2017 school shooter in New Mexico, William Edward Atchison, was identified by users on KiwiFarms, a site primarily devoted to anti-trans bullying, as @satanicdruggie. Users said he was active on Encyclopedia Dramatica, an anything-goes meme site that can sometimes host extremist rhetoric.

Not only was Atchison active on Encyclopedia Dramatica, he was a SysOp there, which means he was an administrator and power user. (We confirmed with users on the site who developed real-life, Skype-centric relationships with Atchison that the accounts were his. Atchison would voluntarily point users to other accounts of his own, in case of a ban.) A Google search of his username using the string "site:encyclopediadramatica.rs + [username]" revealed he went by Satanic Druggie, but also names like "Future School Shooter" and "Adam Lanza," the name of the Sandy Hook shooter.

His posting history across the web revealed an obsession with school shootings that even the police didn't discover in the wake of the shooting.

It's again important to emphasize that the presence of a username across platforms does *not* guarantee the accounts were created by one person. In one famous example, notorious far-right disinformation agents Ian Miles Cheong, Mike Cernovich, InfoWars and GatewayPundit all claimed a man who killed two people and injured 10 others at a Jacksonville video game tournament was anti-Trump.

Their reason? The shooter, David Katz, used the username "Ravens2012Champs" in online video game tournaments, and an anti-Trump user on Reddit had a similar username: "RavenChamps."

The coverage was as breathless as it was incorrect. The InfoWars headline read "Jacksonville Madden Shooter Criticized 'Trumptards' on Reddit," and the story claimed he "hated Trump supporters."

RavenChamps, it turns out, was an entirely different person, a Minnesota factory worker named Pavel.

"I'm alive you know?" he wrote on Reddit hours after the shooting. (The real shooter killed himself after committing the massacre.)

You need a lot more than just a username, but it can be a key starting point to further your reporting as you contact law enforcement, dig in public records and make phone calls.

Tracking campaigns in close to real time

Disinformation and media manipulation campaigns often spread across on Reddit and 4chan, and some are traceable in real time.

For example, 4chan has been in the business of rigging online polls to boost preferred candidates for years. In 2016, 4chan posters repeatedly posted links to both national and hyperlocal news sites running polls in the wake of debates featuring the userbase's preferred candidate, Donald Trump.

Changing Google's search parameters to filter by posts in the "last hour," then searching "site:4chan.org 'polls'" will give you a pretty good window into the polls 4chan users are trying to manipulate in real time.

This has continued well into the next election cycle. 4chan polls boosted Tulsi Gabbard, whom they referred to as "Mommy," in polls on The Drudge Report and NJ.com. Using that simple Google search, anyone could see poll results shifted in real time after one channer told users to "GIVE HER YOUR POWER."

It's even easier to see active trolling operations on sites like Reddit's r/The_Donald community because of Reddit's useful "rising" feature.

Using the convention "reddit.com/r/[subreddit-name]/rising" shows results that are gaining steam at an unusual clip on a subreddit at any given hour.

You can also look at the posts that are overperforming posts across all of reddit.com/r/all/rising. This indexes every post across most Reddit communities. It does not search in quarantined subreddits, which are toxic communities with a habit for deeply offensive content and targeting other communities with trolling campaigns. Quarantined subreddits also don't index on Google, but the "reddit.com/r/[subreddit-name]/rising" will work for them. Quarantining works great for limiting the reach of trolling campaigns outside of centralized audiences, but it makes it harder to track how bad actors are organizing in the moment.

Overall, it's a good idea to keep tabs on the rising section of communities known for trolling campaigns, like r/the_donald, during big political news events, tragedies and elections.

The reality is that sometimes the things these platforms do to thwart bad actors can also make it more difficult for reporters to do important work. Tools can help, but so much of this is manual and requires approaches to verification that algorithms and computers can't reproduce.

At the end of the day, a computer can't replace this kind of work. It's up to us.

11. Network analysis and attribution

Written by: Ben Nimmo

Ben Nimmo *is director of investigations at Graphika and a nonresident senior fellow at the Atlantic Council's Digital Forensic Research Lab. He specializes in studying large-scale cross-platform information and influence operations. He spends his leisure time underwater, where he cannot be reached by phone.*

When dealing with any suspected information operation, one key question for a researcher is how large the operation is and how far it spreads. This is separate from measuring an operation's impact, which is also important: It's all about finding the accounts and sites run by the operation itself.

For an investigator, the goal is to find *as much of the operation as possible* before reporting it, because once the operation is reported, the operators can be expected to hide — potentially by deleting or abandoning other assets.

The first link in the chain

In any investigation, the first clue is the hardest one to find. Often, an investigation will begin with a tipoff from a concerned user or (more rarely) a social media platform. The Digital Forensic Research Lab's work to expose the suspected Russian intelligence operation "Secondary Infektion" began with a tipoff from Facebook, which had found 21 suspect accounts on its platform. The work culminated six months later when Graphika, Reuters and Reddit exposed the same operation's attempt to interfere in the British election. An investigation into disinformation targeting U.S. veterans began with the discovery by a Vietnam Veterans of America employee that its group was being impersonated by a Facebook page with twice as many followers as their real presence on the platform.

There is no single rule for identifying the first link in the chain by your own resources. The most effective strategy is to *look for the incongruous*. It could be a Twitter account apparently based in Tennessee but registered to a Russian mobile phone number; it could be a Facebook page that claims to be based in Niger, but is managed from Senegal and Portugal. It could be a YouTube account with a million views that posts vast quantities of pro-Chinese content in 2019, but almost all its views came from episodes of British sitcoms that were uploaded in 2016.

It could be an anonymous website that focuses on American foreign policy, but is registered to the Finance Department of the Far Eastern Military District of the Russian Federation. It could be an alleged interview with an "MI6 agent" couched in stilted, almost Shakespearean English. It could even be a Twitter account that intersperses invitations to a pornography site with incomplete quotations from Jane Austen's "Sense and Sensibility."

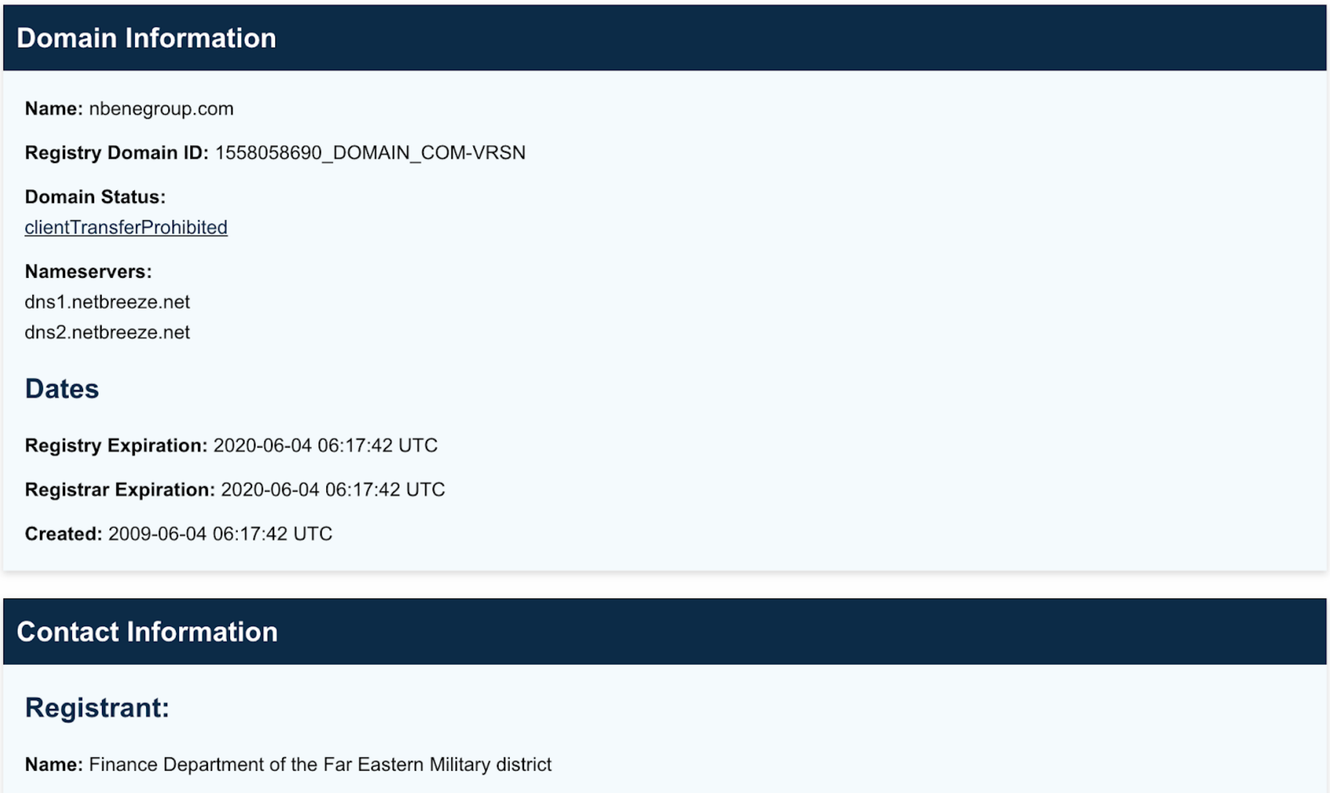
The trick with all such signals is to take the time to think them through. Investigators and journalists are so often pressured for time that it is easy to dismiss signals by thinking "that's just weird," and moving on. Often, if something is weird, it is weird for a reason. Taking the time to say "That's weird: Why is it like that?" can be the first step in exposing a new operation.

Assets, behavior, content

Once the initial asset — such as an account or website — is identified, the challenge is to work out *where* it leads. Three questions are crucial here, modeled on Camille François' Disinformation ABC:

- What information about the initial asset is available?
- How did the asset behave?
- What content did it post?

The first step is to glean as much information as possible about the initial asset. If it is a website, when was it registered, and by whom? Does it have any identifiable features, such as a Google Analytics code or an AdSense number, a registration email address or phone number? These questions can be checked by reference to historical WhoIs records, provided by services such as lookup.icann.com, domaintools.com, domainbigdata. com or the unnervingly named spyonweb.com.



Domain Information

Name: nbenegroup.com

Registry Domain ID: 1558058690_DOMAIN_COM-VRSN

Domain Status:
clientTransferProhibited

Nameservers:
dns1.netbreeze.net
dns2.netbreeze.net

Dates

Registry Expiration: 2020-06-04 06:17:42 UTC

Registrar Expiration: 2020-06-04 06:17:42 UTC

Created: 2009-06-04 06:17:42 UTC

Contact Information

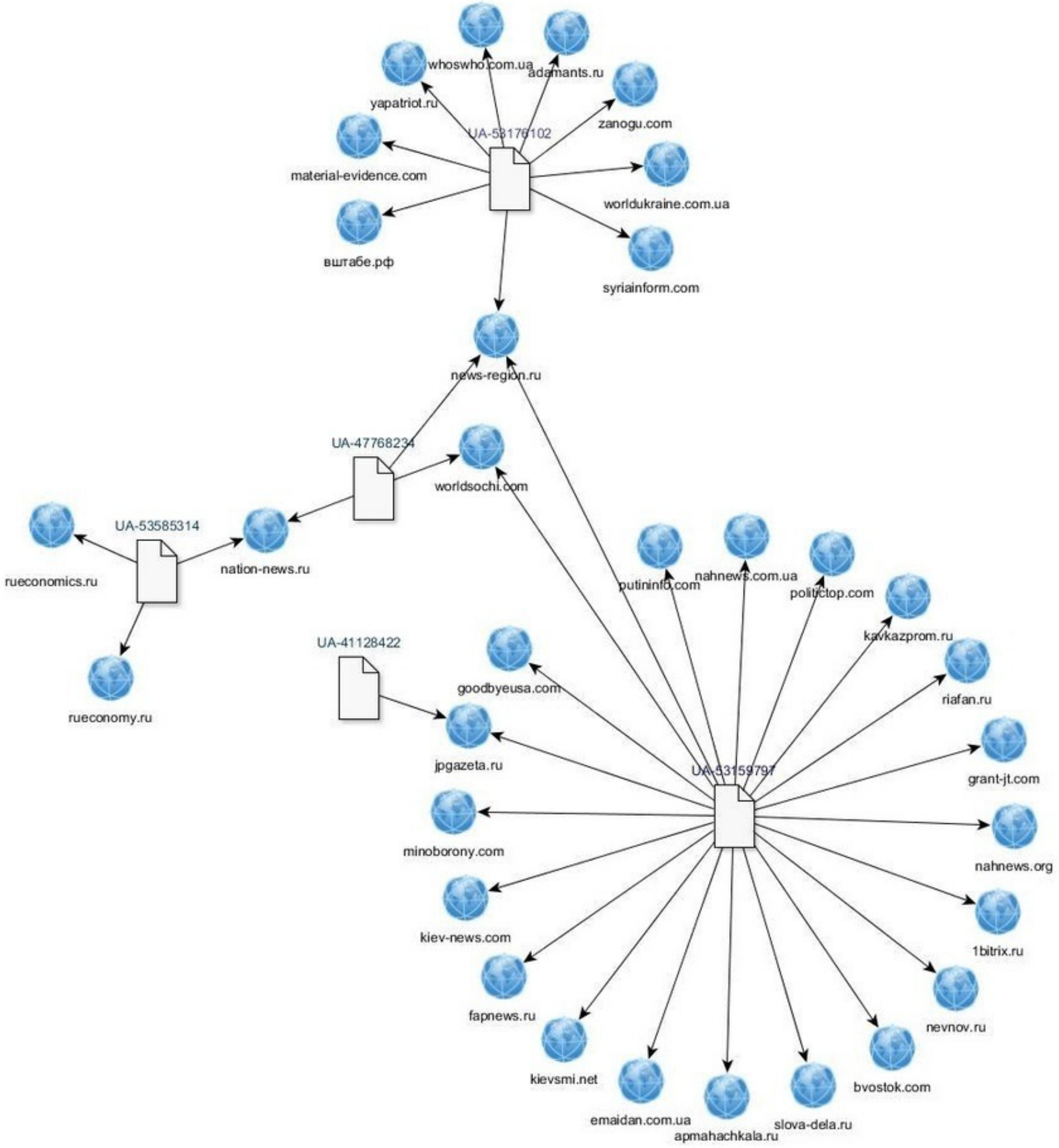
Registrant:

Name: Finance Department of the Far Eastern Military district

Web registration details for the website NBeneGroup.com, which claimed to be a "Youth Analysis Group," showing its registration to the Finance Department of the Far Eastern Military District of the Russian Federation, from lookup.icann.org.

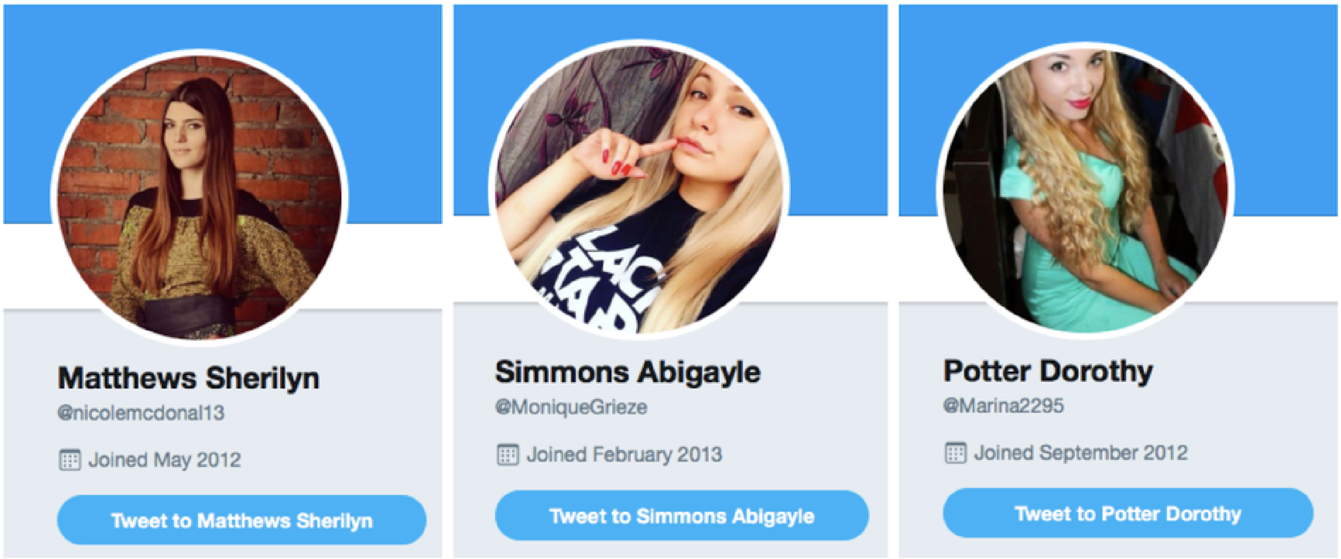
Website information can be used to search for more assets. Both domaintools.com and spyonweb.com allow users to search by indicators such as IP address and Google Analytics code, potentially leading to associated websites — although the savvier information operations now typically hide their registration behind commercial entities or privacy services, making this more difficult.

An early piece of analysis by British researcher Lawrence Alexander identified 19 websites run by the Russian Internet Research Agency by following their Google Analytics numbers. In August 2018, security firm FireEye exposed a large-scale Iranian influence operation by using registration information, including emails, to connect ostensibly unconnected websites.



Network of related websites connected by their Google Analytics codes (eight-digit numbers prefixed with the letters UA), identified by British researcher Lawrence Alexander

If the initial asset is a social media account, the guidance offered in the previous two chapters about bots and inauthentic activity, and investigating social accounts, applies. When was it created? Does its screen name match the name given in its handle? (If the handle is “@moniquegrieze” and the screen name is “Simmons Abigayle,” it’s possible the account was hijacked or part of a mass account creation effort.)



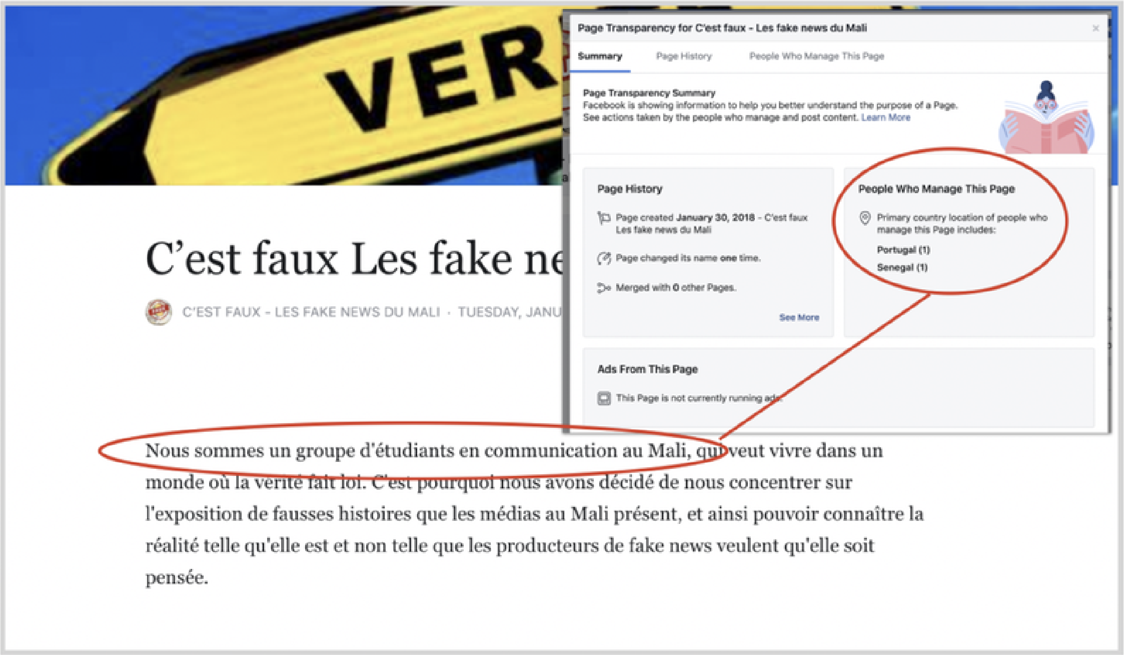
Matthews Sherilyn
@nicolemcdonal13
Joined May 2012

Simmons Abigayle
@MoniqueGrieze
Joined February 2013

Potter Dorothy
@Marina2295
Joined September 2012

Three Twitter accounts involved in a major bot operation *in August 2017. Compare the screen names with the handles, indicating that these were most probably accounts that had been hijacked, renamed and repurposed by the bot herder.*

Does it provide any verifiable biographical detail, or links to other assets on the same or other platforms? If it's a Facebook page or group, who manages it, and where are they located? Whom does it follow, and who follows it? Facebook's "Page transparency" and "group members" settings can often provide valuable clues, as can Twitter profile features such as the date joined and the overall number of tweets and likes. (On Facebook and Instagram, it's not possible to see the date the account was created, but the date of its first profile picture upload provides a reasonable proxy.)



C'est faux Les fake ne

C'EST FAUX - LES FAKE NEWS DU MALI · TUESDAY, JANU

Page Transparency for C'est faux - Les fake news du Mali

Page Transparency Summary
Facebook is showing information to help you better understand the purpose of a Page. See actions taken by the people who manage and post content. Learn More

Page History
Page created **January 30, 2018** - C'est faux Les fake news du Mali
Page changed its name **one** time.
Merged with **0** other Pages.

People Who Manage This Page
Primary country location of people who manage this Page includes:
Portugal (1)
Senegal (1)

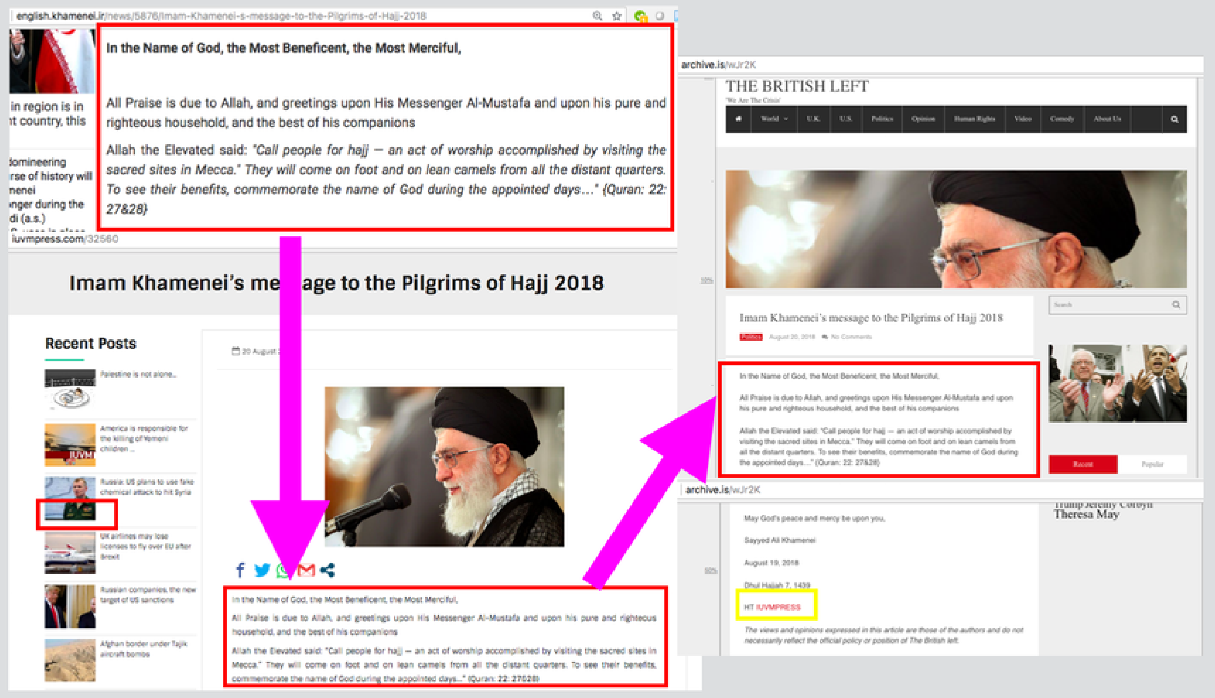
Ads From This Page
This Page is not currently running ads.

Nous sommes un groupe d'étudiants en communication au Mali, qui veut vivre dans un monde où la verité fait loi. C'est pourquoi nous avons décidé de nous concentrer sur l'exposition de fausses histoires que les médias au Mali présent, et ainsi pouvoir connaître la réalité telle qu'elle est et non telle que les producteurs de fake news veulent qu'elle soit pensée.

Website and Facebook Page transparency for ostensible fact-checking site "C'est faux — Les fake news du Mali" (It's false — fake news from Mali), showing that it claimed to be run by a student group in Mali, but was actually managed from Portugal and Senegal. Image from DFRLab*.*

Once the details of the asset have been recorded, the next step is to characterize its behavior. The test question here is, “What behavioral traits are most typical of this asset, and might be useful to identify other assets in the same operation?”

This is a wide-ranging question, and can have many answers, some of which may emerge only in the later stages of an investigation. It could include, for example, YouTube channels that have Western names and profile pictures, but post Chinese-language political videos interspersed with large quantities of short TikTok videos. It could include networks of Facebook or Twitter accounts that always share links to the same website, or the same collection of websites. It could include accounts that use the same wording, or close variations on the same wording, in their bios. It could include “journalist” personas that have no verifiable biographical details, or that give details which can be identified as false. It could include websites that plagiarize most of their content from other sites, and insert only the occasional partisan, polemic or deceptive article. It could include many such factors: The challenge for the researcher is to identify a combination of features that allows them to say, “This asset is part of this operation.”



english.khamenei.ir/news/5876/Imam-Khamenei-s-message-to-the-Pilgrims-of-Hajj-2018

In the Name of God, the Most Beneficent, the Most Merciful,

All Praise is due to Allah, and greetings upon His Messenger Al-Mustafa and upon his pure and righteous household, and the best of his companions

Allah the Elevated said: *"Call people for hajj — an act of worship accomplished by visiting the sacred sites in Mecca." They will come on foot and on lean camels from all the distant quarters. To see their benefits, commemorate the name of God during the appointed days…" {Quran: 22: 27&28}*

iuvmpress.com/32560

Imam Khamenei's me age to the Pilgrims of Hajj 2018

Recent Posts

Palestine is not alone..

America is responsible for the killing of Yemeni children ..

Russia: US plans to use fake chemical attack to hit Syria

UK airlines may lose licenses to fly over EU after Brexit

Russian companies, the new target of US sanctions

Afghan border under Tajik aircraft bombs

20 August

In the Name of God, the Most Beneficent, the Most Merciful,

All Praise is due to Allah, and greetings upon His Messenger Al-Mustafa and upon his pure and righteous household, and the best of his companions

Allah the Elevated said: "Call people for hajj — an act of worship accomplished by visiting the sacred sites in Mecca." They will come on foot and on lean camels from all the distant quarters. To see their benefits, commemorate the name of God during the appointed days…" (Quran: 22: 27&28)

archive.is/wJr2K

THE BRITISH LEFT

'We Are The Crisis'

Imam Khamenei's message to the Pilgrims of Hajj 2018

Politics August 20, 2018 No Comments

In the Name of God, the Most Beneficent, the Most Merciful,

All Praise is due to Allah, and greetings upon His Messenger Al-Mustafa and upon his pure and righteous household, and the best of his companions

Allah the Elevated said: "Call people for hajj — an act of worship accomplished by visiting the sacred sites in Mecca." They will come on foot and on lean camels from all the distant quarters. To see their benefits, commemorate the name of God during the appointed days…" (Quran: 22: 27&28)

archive.is/wJr2K

May God's peace and mercy be upon you.

Sayyed Ali Khamenei

August 19, 2018

Dhul Hajjah 7, 1439

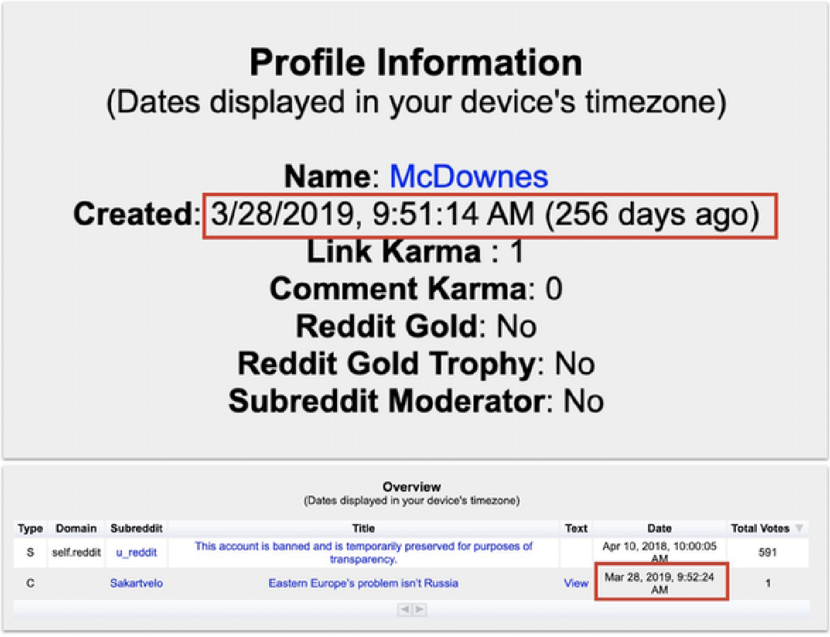
HT IUVMPRESS

The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of The British left.

Trump Jeremy Corbyn
Theresa May

Behavior patterns: An article originally posted to the website of Iran’s Ayatollah Khamenei, and then reproduced without attribution by IUVMpress.com and britishleft.com, two websites in an Iranian propaganda network. Image from DFRLab.

Sometimes, the lack of identifying features can itself be an identifying feature. This was the case with the “Secondary Infektion” campaign run from Russia. It used hundreds of accounts on different blogging platforms, all of which included minimal biographical detail, posted one article on the day they were created, and were then abandoned, never to be used again. This behavior pattern was so consistent across so many accounts that it became clear during the investigation that it was the operation’s signature. When anonymous accounts began circulating leaked US-UK trade documents just before the British general election of December 2019, Graphika and Reuters showed that they exactly matched that signature. Reddit confirmed the analysis.



Profile Information
(Dates displayed in your device's timezone)

Name: McDownes
Created: 3/28/2019, 9:51:14 AM (256 days ago)
Link Karma : 1
Comment Karma: 0
Reddit Gold: No
Reddit Gold Trophy: No
Subreddit Moderator: No

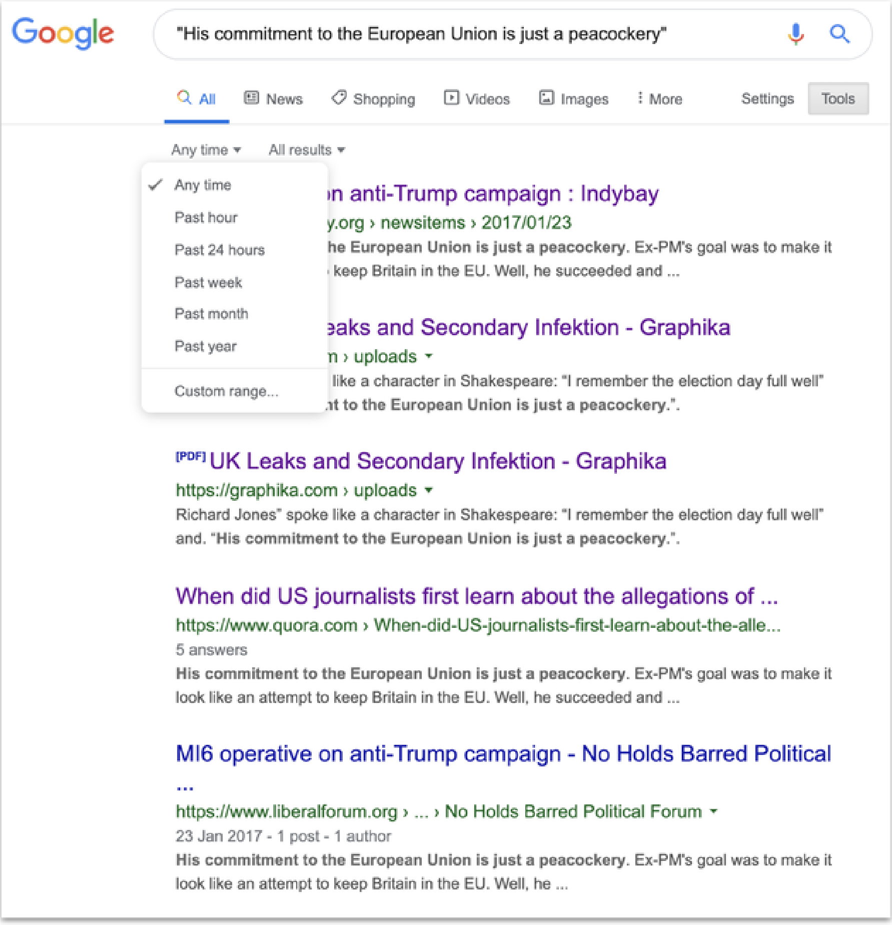
Overview
(Dates displayed in your device's timezone)

Type	Domain	Subreddit	Title	Text	Date	Total Votes
S	self.reddit	u_reddit	This account is banned and is temporarily preserved for purposes of transparency.		Apr 10, 2018, 10:00:05 AM	591
C		Sakartvelo	Eastern Europe's problem isn't Russia	View	Mar 28, 2019, 9:52:24 AM	1

Reddit profile for an account called "McDownes," attributed by Reddit to Russian operation "Secondary Infektion." The account was created on March 28, 2019, posted one article just over one minute after it was created, and then fell silent. Image from Graphika, data from redective.com.

Content clues can also help to identify assets that are part of the same network. If a known asset shares a photo or meme, it's worth reverse-searching the image to see where else it has been used. The RevEye plug-in for web browsers is a particularly useful tool, as it allows investigators to reverse search via Google, Yandex, TinEye, Baidu and Bing. It's always worth using multiple search engines, as they often provide different results.

If an asset shares a text, it's worth searching where else that text appeared. Especially with longer texts, it's advisable to select a sentence or two from the third or fourth paragraphs, or lower, as deceptive operations have been known to edit the headlines and ledes of articles they have copied, but are less likely to take the time to edit the body of the text. Inserting the chosen section in quotation marks in a Google search will return exact matches. The "tools" menu can also sort any results by date.



"His commitment to the European Union is just a peacockery"

ɪn anti-Trump campaign : Indybay
y.org › newsitems › 2017/01/23
he European Union is just a peacockery. Ex-PM's goal was to make it
keep Britain in the EU. Well, he succeeded and ...

eaks and Secondary Infektion - Graphika
n › uploads
like a character in Shakespeare: "I remember the election day full well"
nt to the European Union is just a peacockery.".

[PDF] UK Leaks and Secondary Infektion - Graphika
https://graphika.com › uploads
Richard Jones" spoke like a character in Shakespeare: "I remember the election day full well" and. "His commitment to the European Union is just a peacockery.".

When did US journalists first learn about the allegations of ...
https://www.quora.com › When-did-US-journalists-first-learn-about-the-alle...
5 answers
His commitment to the European Union is just a peacockery. Ex-PM's goal was to make it look like an attempt to keep Britain in the EU. Well, he succeeded and ...

MI6 operative on anti-Trump campaign - No Holds Barred Political ...
https://www.liberalforum.org › ... › No Holds Barred Political Forum
23 Jan 2017 - 1 post - 1 author
His commitment to the European Union is just a peacockery. Ex-PM's goal was to make it look like an attempt to keep Britain in the EU. Well, he ...

Results of a Google search for a phrase posted by a suspected Russian operation, showing the Google tools functionality to date-limit the search.

Assets that post text with mistakes have particular value, as errors are, by their nature, more unusual than correctly spelled words. For example, an article by a suspected Russian intelligence operation referred to Salisbury, the British city where former Russian agent Sergei Skripal was poisoned, as "Solsbury." This made for a much more targeted Google search with far fewer results than a search for "Skripal" and "Salisbury." It therefore produced a far higher proportion of significant finds.

With content clues, it's especially important to look to other indicators, such as behavior patterns, to confirm whether an asset belongs to an operation. There are many legitimate reasons for unwitting users to share content from information operations. That means the sharing of content from an operation is a weak signal. For example, many users have shared memes from the Russian Internet Research Agency because those memes had genuine viral qualities. Simple content sharing is not enough on its own to mark out an operational asset.

Gathering the evidence

Information and influence operations are complex and fast moving. One of the more frustrating experiences for an open-source researcher is seeing a collection of assets taken offline halfway through an investigation. A key rule of analysis is therefore to *record them when you find them*, because you may not get a second chance.

Different researchers have different preferences for recording the assets they find, and the needs change from operation to operation. Spreadsheets are useful for recording basic information about large numbers of assets; shared cloud-based folders are useful for storing large numbers of screenshots. (If screenshots are required, it is *vital* to give the file an identifiable name immediately: few things are more annoying than trying to work out which of 100 files called “Screenshot” is the one you need.) Text documents are good for recording a mixture of information, but rapidly become cluttered and unwieldy if the operation is large.

Whatever the format, some pieces of information should always be recorded. These include how the asset was found (an essential point), its name and URL, the date it was created (if known), and the number of followers, follows, likes and/or views. They also include a basic description of the asset (for example, “Arabic-language pro-Saudi account with Emma Watson profile picture”), to remind you what it was after looking at 500 other assets. If working in a team, it is worth recording which team member looked at which asset.

Links can be preserved by using an archive service such as the Wayback Machine or archive.is, but take care that the archives do not expose genuine users who may have interacted unwittingly with suspect assets, and make sure that the archive link preserves visuals, or take a screenshot as backup. Make sure that all assets are stored in protected locations, such as password-protected files or encrypted vaults. Keep track of who has access, and review the access regularly.

Finally, it’s worth giving the asset a confidence score. Influence operations often find unwitting users to amplify their content: indeed, that is often the point. How sure are you that the latest asset is part of this operation, and why? The level of confidence (high, moderate or low) should be marked as a separate entry, and the reasons (discussed below) should be added to the notes.

Attribution and confidence

The greatest challenge in identifying an information operation lies in attributing it to a specific actor. In many cases, precise attribution will lie beyond the reach of open-source investigators. The best that can be achieved is a degree of confidence that an operation is *probably* run by a particular actor, or that various assets belong to a specific operation, but establishing who is behind the operation is seldom possible with open sources.

Information such as web registrations, IP addresses and phone numbers can provide a firm attribution, but they are often masked to all but the social media platforms. That’s why contacting the relevant platforms is a vital part of investigative work. As the platforms have scaled up their internal investigative teams, they've become more willing to offer public attribution for information operations. The firmest attribution in recent cases has come directly from the platforms, such as Twitter’s exposure of Chinese state-backed information operations targeting Hong Kong, and Facebook’s exposure of operations linked to the Saudi government.

Content clues can play a role. For example, an operation exposed on Instagram in October 2019 posted memes that were almost identical with memes posted by the Russian Internet Research Agency, but stripped out the IRA’s watermarks. The only way they could have made these memes was to source the original images that were the basis for the IRA’s posts and then rebuild the memes on top of them. Ironically, this attempt to mask the origins of the IRA posts suggested that the originators were, in fact, the IRA.

Similarly, a large network of apparently independent websites repeatedly posted articles that had been copied, without attribution, from Iranian government sources. This pattern was so repetitive that it turned out to be the websites’ main activity. As such, it was possible to attribute this operation to pro-Iranian actors, but it was *not* possible to further attribute it to the Iranian government itself.

Ultimately, attribution is a question of self-restraint. The researcher has to imagine the question, “How can you *prove* that this operation was run by the person you’re accusing?” If they cannot answer that question with confidence to themselves, they should steer clear of making the accusation. Identifying and exposing an information operation is difficult and important work, and reaching to make an unsupported or inaccurate attribution can undermine everything that came before it.

11a. Case study: Attributing Endless Mayfly

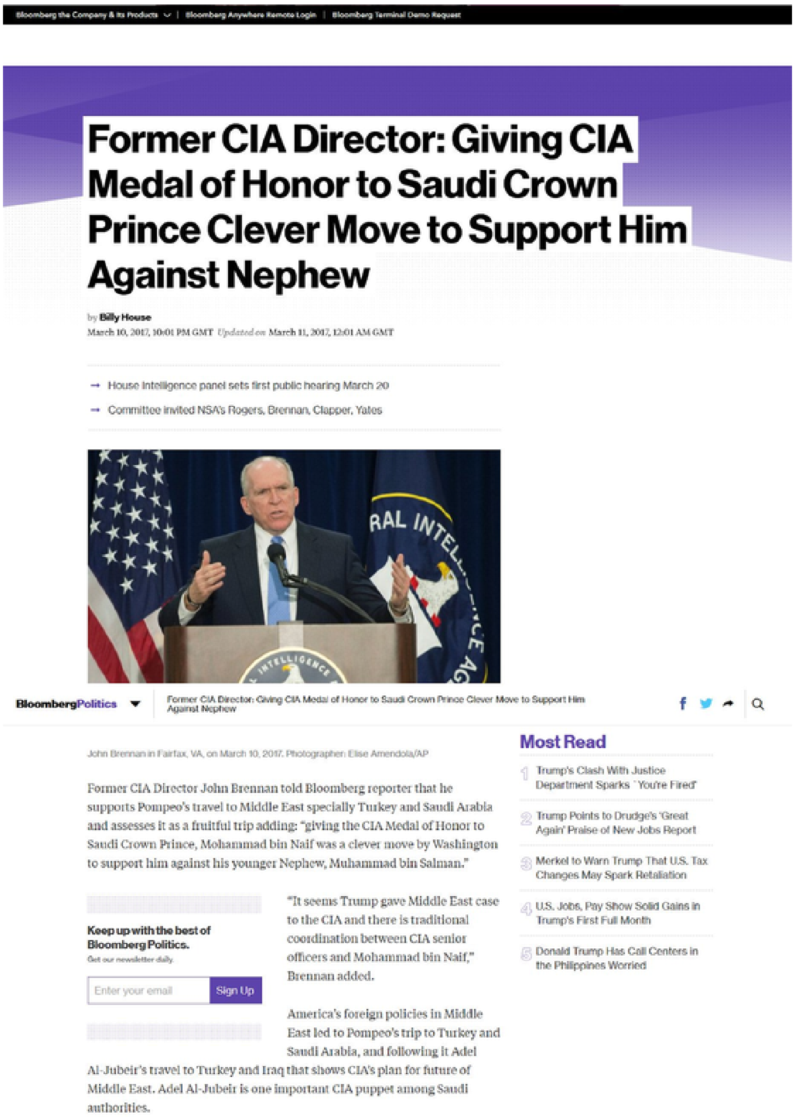
Written by: Gabrielle Lim

Gabrielle Lim *is a researcher at the Technology and Social Change Research Project at Harvard Kennedy School's Shorenstein Center and a fellow with Citizen Lab. She studies the implications of censorship and media manipulation on security and human rights.*

In April 2017, an inauthentic article spoofing British news outlet The Independent was posted to Reddit. This article falsely quoted former U.K. Deputy Prime Minister Nick Clegg as saying that then-Prime Minister Theresa May was "kissing up to Arab regimes." Savvy Redditors were quick to call out the post as dubious and false. Not only was it hosted on indep**n**edent.**co** as opposed to www.independent.co.uk, but the original poster was a shallow persona who had also posted several other inauthentic articles on Reddit.

From that initial inauthentic article, domain and persona, researchers at Citizen Lab spent the next 22 months tracking and investigating the network behind this multifaceted online information operation. Called Endless Mayfly, the goal of the operation was to target journalists and activists with inauthentic websites by spoofing established outlets's websites, and disseminating false and divisive information.

Broadly speaking, the network would spoof a reputable news outlet with an inauthentic article, amplify it through a network of websites and fake Twitter personas, and then either delete or redirect the inauthentic article once some online buzz was created. Below is an example of a spoofed article that masqueraded as Bloomberg.com by typosquatting on bloomber**q**.com:



Bloomberg the Company & Its Products | Bloomberg Anywhere Remote Login | Bloomberg Terminal Demo Request

Former CIA Director: Giving CIA Medal of Honor to Saudi Crown Prince Clever Move to Support Him Against Nephew

by Billy House
March 10, 2017, 10:01 PM GMT *Updated on* March 11, 2017, 12:01 AM GMT

→ House Intelligence panel sets first public hearing March 20
→ Committee invited NSA's Rogers, Brennan, Clapper, Yates

BloombergPolitics
Former CIA Director: Giving CIA Medal of Honor to Saudi Crown Prince Clever Move to Support Him Against Nephew

John Brennan in Fairfax, VA, on March 10, 2017. Photographer: Elise Amendola/AP

Former CIA Director John Brennan told Bloomberg reporter that he supports Pompeo's travel to Middle East specially Turkey and Saudi Arabia and assesses it as a fruitful trip adding: "giving the CIA Medal of Honor to Saudi Crown Prince, Mohammad bin Naif was a clever move by Washington to support him against his younger Nephew, Muhammad bin Salman."

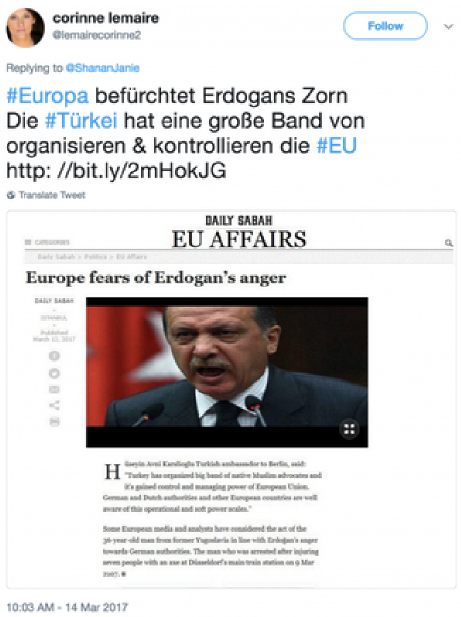
"It seems Trump gave Middle East case to the CIA and there is traditional coordination between CIA senior officers and Mohammad bin Naif," Brennan added.

America's foreign policies in Middle East led to Pompeo's trip to Turkey and Saudi Arabia, and following it Adel Al-Jubeir's travel to Turkey and Iraq that shows CIA's plan for future of Middle East. Adel Al-Jubeir is one important CIA puppet among Saudi authorities.

Most Read

1 Trump's Clash With Justice Department Sparks 'You're Fired'
2 Trump Points to Drudge's 'Great Again' Praise of New Jobs Report
3 Merkel to Warn Trump That U.S. Tax Changes May Spark Retaliation
4 U.S. Jobs, Pay Show Solid Gains in Trump's First Full Month
5 Donald Trump Has Call Centers in the Philippines Worried

This image shows two fake online personas affiliated with Endless Mayfly tweeting a link to a copycat version of the Daily Sabah, a Turkish news outlet. Note that the persona on the right, "jolie prevoit," is using a photo of actor Elisha Cuthbert as its profile photo.



corinne lemaire
@lemairecorinne2

Replying to @ShananJanie

#Europa befürchtet Erdogans Zorn
Die #Türkei hat eine große Band von organisieren & kontrollieren die #EU
http: //bit.ly/2mHokJG

EU AFFAIRS

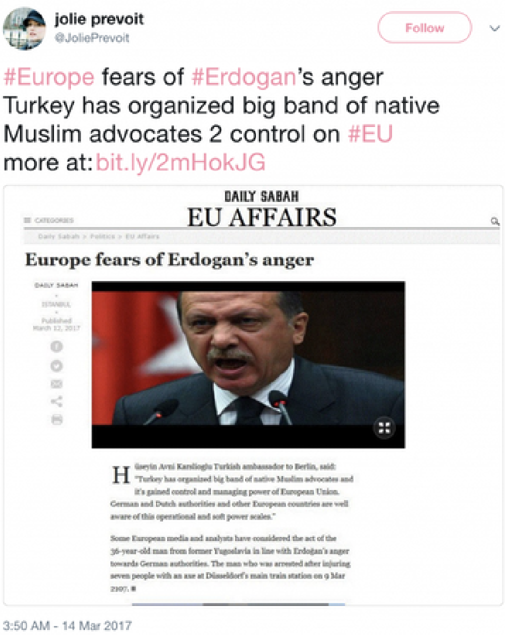
Europe fears of Erdogan's anger

DAILY SABAH · ISTANBUL · Published March 12, 2017

Hüseyin Avni Karslioglu Turkish ambassador to Berlin, said: "Turkey has organized big band of native Muslim advocates and it's gained control and managing power of European Union. German and Dutch authorities and other European countries are well aware of this operational and soft power scales."

Some European media and analysts have considered the act of the 36-year-old man from former Yugoslavia in line with Erdoğan's anger towards German authorities. The man who was arrested after injuring seven people with an axe at Düsseldorf's main train station on 9 Mar 2017. ■

10:03 AM - 14 Mar 2017



jolie prevoit
@JoliePrevoit

#Europe fears of #Erdogan's anger
Turkey has organized big band of native Muslim advocates 2 control on #EU
more at: bit.ly/2mHokJG

DAILY SABAH
EU AFFAIRS

Europe fears of Erdogan's anger

DAILY SABAH · ISTANBUL · Published March 12, 2017

Hüseyin Avni Karslioglu Turkish ambassador to Berlin, said: "Turkey has organized big band of native Muslim advocates and it's gained control and managing power of European Union. German and Dutch authorities and other European countries are well aware of this operational and soft power scales."

Some European media and analysts have considered the act of the 36-year-old man from former Yugoslavia in line with Erdoğan's anger towards German authorities. The man who was arrested after injuring seven people with an axe at Düsseldorf's main train station on 9 Mar 2017. ■

3:50 AM - 14 Mar 2017

By the time we published our report in May 2019, our dataset included 135 inauthentic articles, 72 domains, 11 personas, one fake organization and a pro-Iran publishing network that amplified the falsehoods found in the inauthentic articles. In the end, we concluded with moderate confidence that Endless Mayfly was an Iran-aligned information operation.

Endless Mayfly illustrates how you can combine network and narrative analysis with external reporting to arrive at attribution. It also highlights the difficulty involved in attributing information operations to a specific actor, why multiple indicators are required, and how to use a confidence level to indicate your level of certitude for the attribution.

Ultimately, attribution is a difficult task often constrained by imperfect information, unless you're able to elicit a confession or secure definitive proof. This is why attribution is often expressed as a probabilistic estimate in many media manipulation cases.

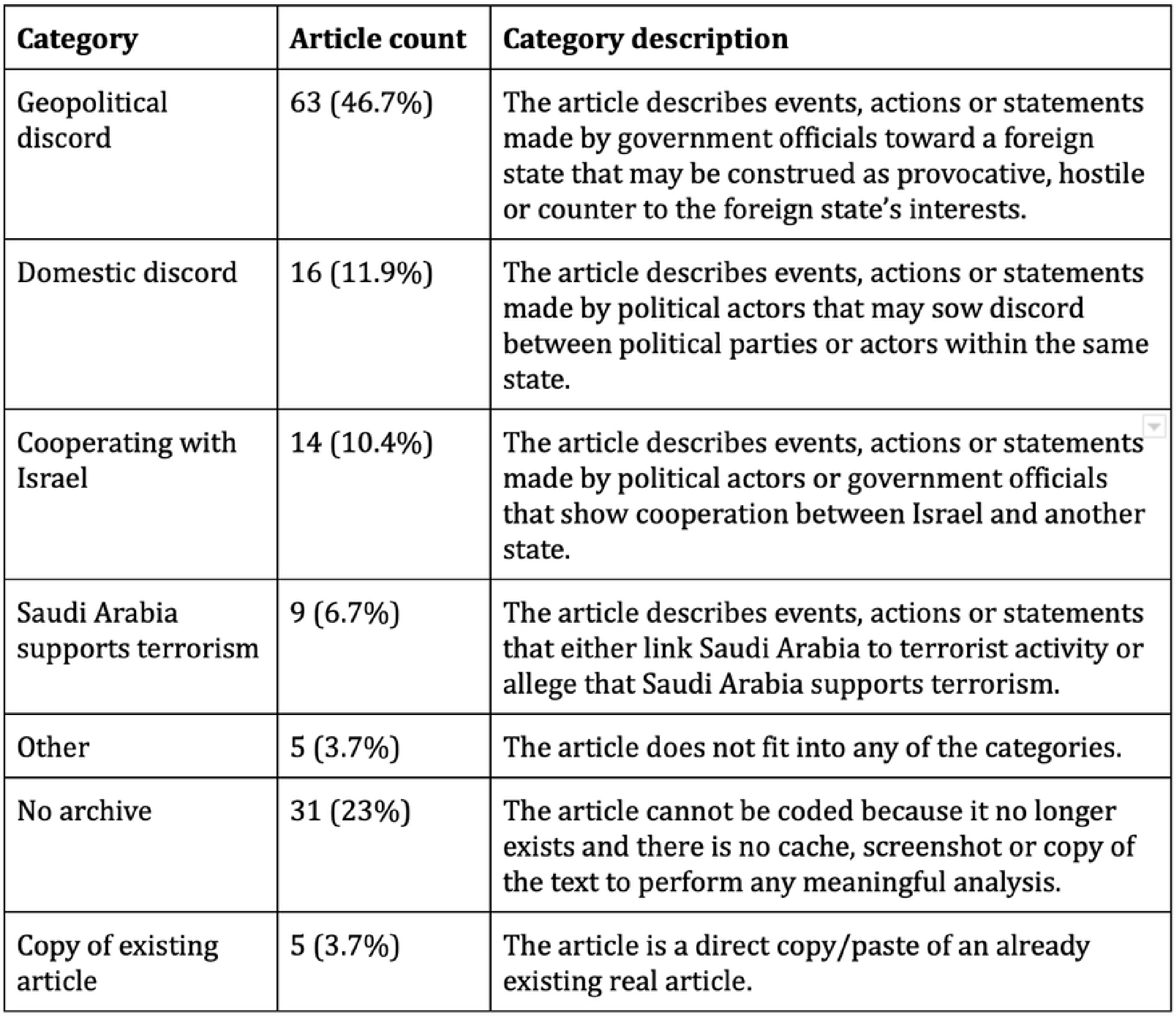
Triangulating multiple data points and analyses

Due to the clandestine nature of information operations, the ability for actors to engage in "false flag" campaigns, and the ephemeral nature of evidence, attribution should be the result of a combination of analysis and evidence. With Endless Mayfly, we concluded with moderate confidence that it was an Iran-aligned operation because of indicators derived from three types of analysis:

1. Narrative analysis
2. Network analysis
3. External reporting and analysis

1. Narrative analysis

Using content and discourse analysis on the 135 inauthentic articles collected in our investigation, we determined that the narratives being propagated were aligned with Iran's interests. Each article was coded into categories that were determined after an initial reading of all the articles. Two rounds of coding were conducted: The first round was executed independently by two researchers, and a second round was conducted together by the same researchers to resolve any discrepancies. This table represents the results of our coding process.



Category	Article count	Category description
Geopolitical discord	63 (46.7%)	The article describes events, actions or statements made by government officials toward a foreign state that may be construed as provocative, hostile or counter to the foreign state's interests.
Domestic discord	16 (11.9%)	The article describes events, actions or statements made by political actors that may sow discord between political parties or actors within the same state.
Cooperating with Israel	14 (10.4%)	The article describes events, actions or statements made by political actors or government officials that show cooperation between Israel and another state.
Saudi Arabia supports terrorism	9 (6.7%)	The article describes events, actions or statements that either link Saudi Arabia to terrorist activity or allege that Saudi Arabia supports terrorism.
Other	5 (3.7%)	The article does not fit into any of the categories.
No archive	31 (23%)	The article cannot be coded because it no longer exists and there is no cache, screenshot or copy of the text to perform any meaningful analysis.
Copy of existing article	5 (3.7%)	The article is a direct copy/paste of an already existing real article.

After all the articles were coded, we were able to determine the most common narratives propagated by Endless Mayfly. We compared these with our preliminary research on the region. This involved extensive research to understand the region's rivalries and alliances, geopolitical interests and threats, and history of information controls. This was necessary for us to contextualize the evidence and situate the narratives in the broader political context. With the results of the coding in hand, we determined that these narratives were most likely serving the interests of Iran.

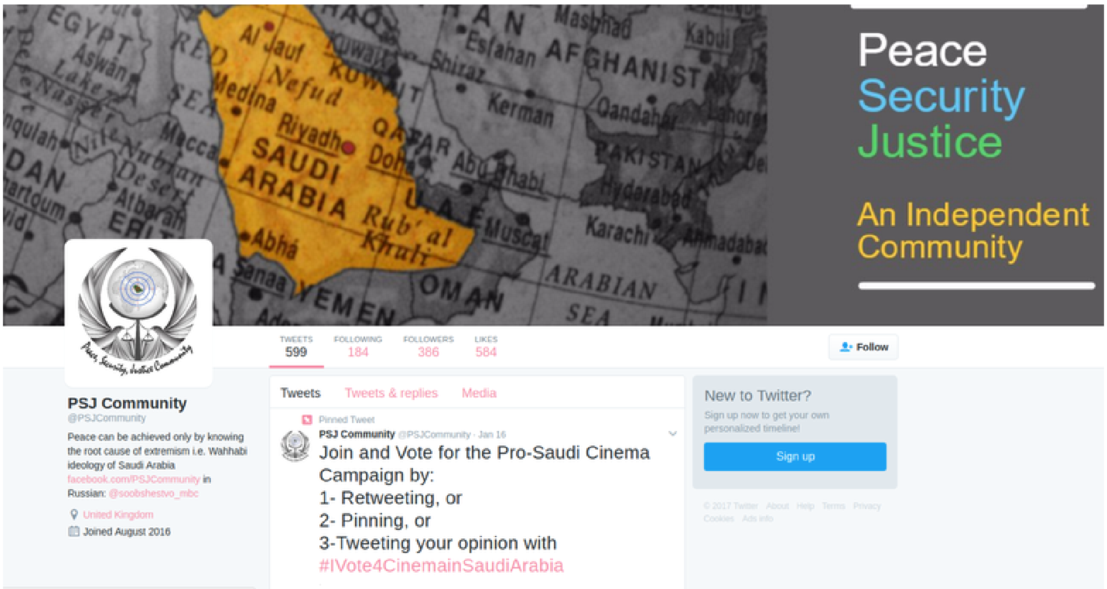
2. Network analysis

Network analysis was carried out to determine which domains or platforms were responsible for amplifying the content. For Endless Mayfly, two networks were involved in disseminating the inauthentic articles and their falsehoods: a network of pro-Iran websites, and a cluster of pro-Iran personas on Twitter. Both factored into Endless Mayfly's attribution because they consistently pushed stories that were in line with official Iranian policies, public statements and positions with regards to Saudi Arabia, Israel and the United States.

The publishing network — The publishing network consisted of a number of seemingly pro-Iran websites portraying themselves as independent news outlets. In total, we found 353 webpages across 132 domains that referenced or linked back to Endless Mayfly's inauthentic articles. This process involved a Google search of all the inauthentic articles' URLs and their headlines. In addition, we scanned the links tweeted by the personas in our network, identifying webpages that contained references or links to the articles.

Following this process, we identified the top 10 domains that most frequently referenced the inauthentic articles. Of these 10 domains, eight shared the same IP address or registration details, indicating they may be controlled by the same actor. The content of these sites was also skewed toward promoting Iranian interests. For example, IUVM Press, which linked to or referenced Endless Mayfly's inauthentic articles 57 times, hosted a PDF document titled "Statute" that explicitly stated they are against "the activities and projects of global arrogance states, the imperialism and Zionism," and that "The headquarters of the Union is located in the Tehran — capital of Islamic Republic of Iran."

The persona network — Similar to the inauthentic articles and the publishing network, the personas affiliated with Endless Mayfly on Twitter were decidedly critical of Saudi Arabia, Israel and Western nations in general. An analysis of their Twitter activity found these accounts pushed a combination of credible and inauthentic articles that were highly critical of Iran's political rivals. Take, for example, the Twitter account for the "Peace, Security, Justice Community," a fake organization identified by our investigation, shown below. Not only did it propagate content that was against Saudi Arabia, Israel and the U.S., the profile photo and header image also targeted Saudi Arabia. Note the cross hairs over Saudi Arabia in the profile photo, and the map used in the header image. The account's bio also explicitly calls out Saudi Arabia and Wahhabi ideology as the cause of extremism.



TWEETS 599 FOLLOWING 184 FOLLOWERS 386 LIKES 584

PSJ Community
@PSJCommunity

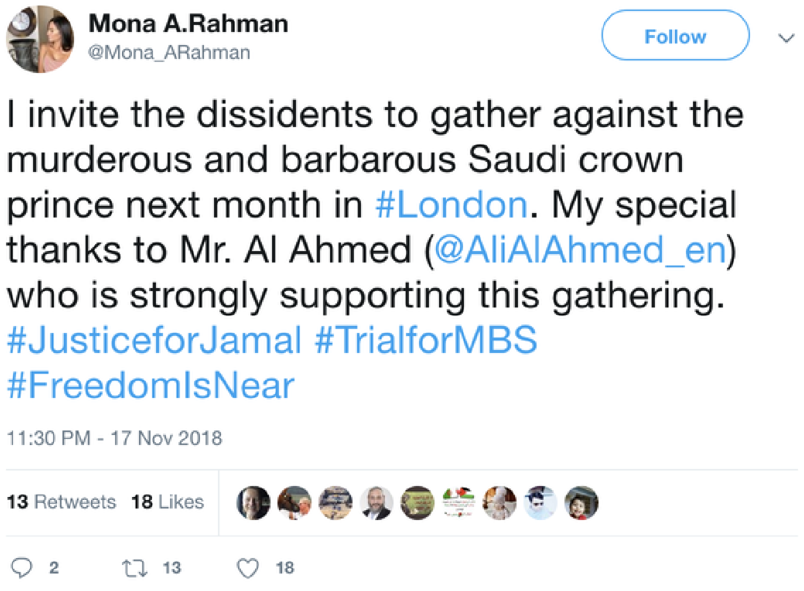
Peace can be achieved only by knowing the root cause of extremism i.e. Wahhabi ideology of Saudi Arabia
facebook.com/PSJCommunity in Russian: @soobshestvo_mbc

United Kingdom
Joined August 2016

Pinned Tweet
PSJ Community @PSJCommunity · Jan 16
Join and Vote for the Pro-Saudi Cinema Campaign by:
1- Retweeting, or
2- Pinning, or
3-Tweeting your opinion with #IVote4CinemainSaudiArabia

New to Twitter?
Sign up now to get your own personalized timeline!

Similarly, this tweet from another Endless Mayfly persona, "Mona A. Rahman," mentions journalist and Saudi critic Ali al-Ahmed while criticizing Saudi Arabia's crown prince, Mohammad bin Salman.



Mona A.Rahman
@Mona_ARahman

Follow

I invite the dissidents to gather against the murderous and barbarous Saudi crown prince next month in #London. My special thanks to Mr. Al Ahmed (@AliAlAhmed_en) who is strongly supporting this gathering. #JusticeforJamal #TrialforMBS #FreedomIsNear

11:30 PM - 17 Nov 2018

13 Retweets 18 Likes

2 13 18

3. External reporting and analysis

We also compared our findings and data with external reporting. Following a tip from FireEye in August 2018, for example, Facebook deactivated some accounts and pages linked to the publishing network used by Endless Mayfly. In its analysis, FireEye identified several domains that were part of the publishing network we had identified, like institutomanquehue.org and RPFront.com. Like us, they also concluded with moderate confidence that the "suspected influence operation" appears to originate from Iran. Facebook, in its announcement, similarly noted the operations most likely originated from Iran.

In addition, Twitter released a dataset of Iran-linked accounts that had been suspended for "coordinated manipulation." Although accounts with fewer than 5,000 followers at the time of suspension were anonymized, we were able to identify one Endless Mayfly persona (@Shammari_Tariq) in Twitter's dataset.

The assessments by Twitter, Facebook and FireEye were useful in corroborating our hypothesis because they surfaced evidence that was not part of our data collection efforts, and overlapped with Endless Mayfly assets we identified. For example, FireEye's analysis identified phone numbers and registration information connected to Twitter accounts and domains associated with Endless Mayfly — evidence that was not part of our dataset. Likewise, Facebook and Twitter presumably had account registration information, such as IP addresses, that we don't have access to. The additional data points identified by these external reports therefore helped expand the body of evidence.

Arriving at moderate confidence

In Endless Mayfly's case, the evidence we collected — the pro-Iran narratives, personas and publishing network — pointed to Iran as a plausible source of the information operation. This body of evidence was then compared to credible external reporting and research from FireEye, Facebook and Twitter, which corroborated our findings. Each individual piece of evidence, while insufficient on its own for attribution, helped confirm and strengthen our hypothesis when assessed holistically, and when compared to the totality of the evidence our investigation surfaced.

Despite the multiple indicators pointing to Iran, we still did not have definitive evidence. As such, we used a framework of cyber-attribution that's common within the intelligence community. It makes use of multiple indicators and probabilistic confidence (low, moderate, high), allowing researchers to convey their findings while qualifying their level of uncertainty.

Ultimately, we concluded that Endless Mayfly is an Iran-aligned operation with moderate confidence, which the U.S. Office of the Director of National Intelligence defines as meaning “the information is credibly sourced and plausible but not of sufficient quality or corroborated sufficiently to warrant a higher level of confidence.” We did not opt for a higher level of confidence because we felt that there was insufficient evidence to completely rule out a false flag operation — meaning someone trying to make it *look* like Iran was behind this operation — or a third party sympathetic with Iranian interests.

Attributing information operations like Endless Mayfly will almost always rely on incomplete and imperfect information. Attaching confidence levels to findings is therefore an important component of attribution — as it operates with an abundance of caution. Incorrect attribution or an inflated confidence level can have dire consequences, especially if government policies and retaliatory measures result from the faulty assessment. To avoid hasty and poor attribution practices, it’s important to consider multiple indicators, types of evidence and analyses, and to make use of a confidence level that considers alternative hypotheses and missing data.

11b. Case Study: Investigating an Information Operation in West Papua

Written by: Elise Thomas , Benjamin Strick

***Benjamin Strick** is an open-source investigator for the BBC, a Bellingcat contributor and an instructor in open-source techniques, geospatial intelligence and network analysis. He has a background in law and the military, and focuses on using OSINT/GEOINT, geolocation and intelligence methods for good, through human rights, conflict and privacy.*

***Elise Thomas** is a freelance journalist and a researcher working with the International Cyber Policy Centre at the Australian Strategic Policy Institute. Her writing has appeared in Wired, Foreign Policy, The Daily Beast, The Guardian and others. She also previously worked as an editorial assistant for the U.N. Office for the Coordination of Humanitarian Affairs, and as a podcast writer and researcher.*

In August 2019, separatist tensions flared up yet again in West Papua, a province that became part of Indonesia in a controversial decision in the 1960s. Since then, the region has suffered from widespread allegations of human rights abuses committed by Indonesian authorities to quash dissent.

Access to the region is heavily restricted, and foreign journalists have been banned from reporting in the province. All of this makes social media a crucial resource for monitoring and reporting on West Papua.

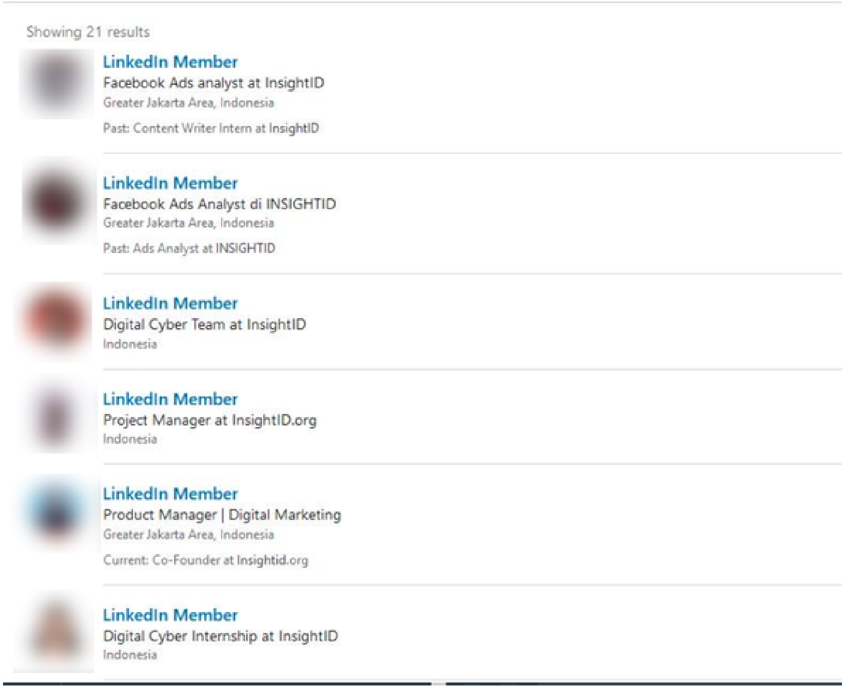
While trying to geolocate some of the footage that was coming out of the violence in FakFak, one of us identified two hashtags spreading on Twitter, #WestPapua and #FreeWestPapua.

Searches under those hashtags revealed a wave of fake accounts autoposting the same videos and same text using these same hashtags. The accounts also retweeted and liked one another's content, helping further amplify it and increase engagement on the hashtags.

The process for analyzing these automated accounts was detailed in Chapter 3. Building on that work, we expanded our investigation by working to identify the people or groups behind the operations. In the process, we uncovered a similar, smaller and apparently unrelated campaign, and were also able to identify the individual responsible. Operators of both campaigns eventually admitted their involvement after being approached by the BBC.

The size of the first campaign and the fact that it was operating across multiple platforms gave us a range of opportunities to find clues we could use to pivot on to find more information about the campaign's operators.

The first useful piece of information was the websites being shared by the network of Twitter and Facebook accounts. Whois searches revealed that four of the domains were registered using a fake name and a dummy email address, but with a real phone number. We entered the number into WhatsApp to see if it was connected to an account. It was, and that account also had a profile photo. Using Yandex reverse image search on that profile photo, we were able to connect the profile photo to Facebook, LinkedIn and Freelancer.com accounts. Through that associated LinkedIn account, we were able to find the person's current workplace, and see their colleagues.



Showing 21 results

LinkedIn Member
Facebook Ads analyst at InsightID
Greater Jakarta Area, Indonesia
Past: Content Writer Intern at InsightID

LinkedIn Member
Facebook Ads Analyst di INSIGHTID
Greater Jakarta Area, Indonesia
Past: Ads Analyst at INSIGHTID

LinkedIn Member
Digital Cyber Team at InsightID
Indonesia

LinkedIn Member
Project Manager at InsightID.org
Indonesia

LinkedIn Member
Product Manager | Digital Marketing
Greater Jakarta Area, Indonesia
Current: Co-Founder at Insightid.org

LinkedIn Member
Digital Cyber Internship at InsightID
Indonesia

The individual was an employee of a Jakarta-based company called InsightID, whose website said it offered “integrated PR and digital marketing program[s].”

We also gathered additional data points that InsightID was responsible for the information operation. On its website, InsightID referred to its work on the “Papua Program Development Initiative,” which “examines Papua rapid socio-economic development and explores its challenges.” Former InsightID employees and interns described producing video content, writing copy and translating content as part of their work on the Papua Development Project.

One former employee stated on their LinkedIn profile that their work could be seen on “West Papuan (Instagram, Facebook, Website).” West Papuan was one of five news websites involved in the campaign. Another InsightID employee created a YouTube account in their own name to host a video as part of the campaign. This video was then embedded on westpapuan.org.

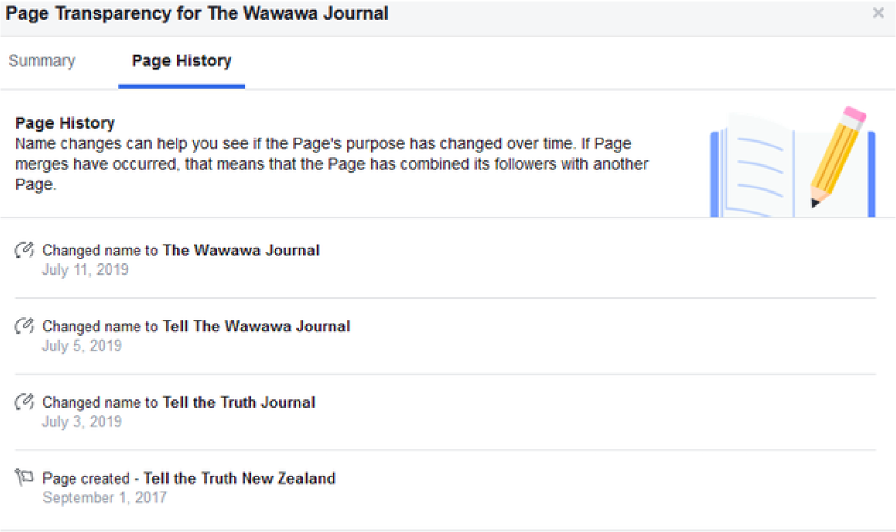
Further domain record searches revealed that InsightID’s co-founder used his company email address to register 14 domains on the same day, most of which clearly related directly to West Papua. These included westpapuafreedom.com, westpapuagenocide.com and westpapuafact.com. Each additional piece of information added to the evidence that InsightID was responsible for the operation.

At that point, BBC journalists attempted to contact InsightID for comment. Although the company didn’t respond, InsightID ultimately acknowledged its responsibility, saying in a social media post that “our content defends Indonesia against the hoax narrative of the Free Papua separatist groups.”

We were not able to identify the client who hired InsightID to conduct the information campaign.

While uncovering this larger operation, we also investigated a smaller network of three websites that masqueraded as independent news sources and had associated social media profiles. Although apparently not connected to the first campaign, these sites targeted international perceptions of the situation in West Papua, focusing on audiences in New Zealand and Australia.

The key to identifying the individual responsible was that the Facebook page for one brand, the Wawawa Journal, was originally called Tell the Truth NZ. We were able to see this by looking at the page’s naming history. This allowed us to link it back to the domain tellthetruthnz.com, which was registered to Muhamad Rosyid Jazuli.



Page Transparency for The Wawawa Journal

Page History
Name changes can help you see if the Page's purpose has changed over time. If Page merges have occurred, that means that the Page has combined its followers with another Page.

Changed name to **The Wawawa Journal**
July 11, 2019

Changed name to **Tell The Wawawa Journal**
July 5, 2019

Changed name to **Tell the Truth Journal**
July 3, 2019

Page created - **Tell the Truth New Zealand**
September 1, 2017

When approached by BBC journalists, Jazuli admitted to being the operator of the campaign. He works with the Jenggala Center, an organization created by Indonesia's vice president, Jusuf Kalla. It was created in 2014 to promote his reelection and support President Jokowi's administration.

What this investigation demonstrates is that identifying information campaigns and attributing them to the individuals and groups responsible does not necessarily require complicated techniques or tools — but it does require both patience and a certain amount of luck. This investigation relied on open-source resources such as Whois records, reverse image search, social media profiles and analysis of website source codes. The fact that the campaign was in operation across multiple platforms, in combination with the social media and LinkedIn profiles of InsightID's employees, was crucial in allowing us to piece together many small clues to build the bigger picture.

If there is a key lesson to take away from this example, it is to think about how you can use details or clues from one platform to pivot to another.

Credits

Editor: Craig Silverman

Contributing Editor: Claire Wardle

Copy Editor: Merrill Perlman

Contributors: 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

Production Manager: Arne Grauls

This book is published by the European Journalism Centre and was made possible thanks to funding from the Craig Newmark Philanthropies.

Created by European Journalism Centre

Supported by Google News Initiative