Guide to Open Source Intelligence (OSINT)

Michael Edison Hayden

Funded by the John S. and James L. Knight Foundation.
# Table of Contents

*Introduction: Peering Behind the Mask*  

1. Open and Closed Networks  

2. Searching the Open Web  

3. Verifying the Authenticity of Who and What You Find on Social Media  

4. Verifying the Authenticity of the Images and Videos You Find Online  

5. Exploring Fringe Websites  

6. Using Archives, Saving Your Work  

7. Learning New Platforms and Interacting with Hostile Communities  

8. Ethics and Personal Security  

*Appendix: Tools*
Introduction: Peering Behind the Mask

Open source intelligence, which researchers and security services style OSINT, is one of the most valuable tools to a contemporary reporter, because of the vast amount of publicly available online information.

Reporters conducting OSINT-based research should aspire to use the information they gather online to peer behind the superficial mask of the internet—the anonymous avatars on Twitter, for example, or the filtered photographs on Instagram—and tell the story of the real, flesh-and-blood human beings on the other side of our screens.

Every time we go online, we give up part of our identity. Sometimes, it comes in the form of an email used to make a Twitter account. Other times, it’s a phone number for two-factor authentication, or days’ and weeks’ worth of timestamps suggesting when a user is awake and asleep. Journalists can piece together clues like this and use them to tell stories which are of interest to the public.

The following guide is written to provide a basic foundation not only for doing that work, but also for verifying the information, archiving findings, and interacting with hostile communities online.

The closer we get to understanding the people who make the influential and newsworthy aspects of the internet happen—and their motivations—the easier our work of discovery becomes.
1. Open and Closed Networks

Open-source reporting refers to the effort of retrieving information that is publicly available online, part of what we will call an open network.

When you communicate through devices like phones, laptops and tablets, you use both open and closed networks without even thinking. You probably talk to some of the same friends while dancing back and forth between open and closed networks, and do so within a span of only a few minutes. Reporters should be conscious of the difference.

When you think about the difference between open and closed networks, think about how you use Facebook: You want the public to find the link to the story you wrote, so you post it publicly on your timeline. But you may not want the public to read your comments making fun of your boss’s outfit, so you post that to a colleague over the website’s Messenger app. The public post is open and the snarky comment on Messenger is closed. These are two different types of communication, and how you report on the use of them will require different approaches.

Open Networks

Ask yourself: Does the conversation you are viewing require an invitation? If the answer is no, you are looking at what we call an open network. Twitter users, particularly journalists putting off a deadline, chat with one another quite a bit about what is “on the TL,” or their publicly visible timeline of tweets. What is “on the TL” is part of an open and searchable network. So is a conversation on imageboard forums like 4chan and 8chan. So is anything you can read by searching Google. When LeBron James signed with the Los Angeles Lakers in July 2018, for example, the comment section of a Bleacher Report blog post about the news lit up with chatter. The comment section
about James’ move is part of an open network. So is Yahoo Answers. So is a neo-Nazi forum like Iron March.

Examples of Open Networks:

Twitter, Facebook, YouTube, Instagram, Gab, Reddit, Voat, 4chan, 8chan, Parler, Minds.com, Steam, Kiwi Farms, the comment section of a New York Times article, public donation messages on a GoFundMe page.

Reporting on Public-Facing Content

Let’s say a 20-year-old woman named Jane Smith wanted to wish Joseph Stalin a happy birthday on Facebook because she belonged to the “tankie” community of former Soviet Union apologists. Smith posts, “Happy Birthday to my handsome hero Joseph Stalin!” on December 18. Two days later, Smith robs a bank and posts an Instagram video where she tells the camera that she did it in the name of “destroying Western civilization.”

Provided that both of Smith’s accounts are set to public, and you have verified that the accounts are authentic, you have the makings of an interesting story about a tankie bank robber: Jane Smith, the 20-year-old suspect in today’s heist, appeared to praise Joseph Stalin on Facebook, and told her audience of roughly 900 Instagram followers that she targeted the bank as part of an effort to destroy “Western civilization.”

When accounts are set to public, anyone with access to the internet can see what is posted. If anyone can see it, anyone can report on it. Most of what you will read about in this guide revolves around how to find, verify, and analyze content found on open networks.

Closed Networks
When Twitter users gossip in private “DMs,” that is a part of a closed network within Twitter’s platform. The users do not want strangers to see what they are writing to one another, and users need to be invited to read it. When you text someone, you’re doing it on a closed network. When you create a chat room that is only available to people through an invitation, you’re doing it on a closed network. Does the conversation you are viewing require an invitation? If the answer is yes, you are looking at a closed network. Encrypted messaging apps like WhatsApp, Signal and Wire are good examples of closed networks because they require an invitation from another person to view and participate in a conversation.

Keybase and Telegram are examples of encrypted messaging apps where a good amount of material about the users on it is public facing and part of an open network, but the conversations are not always presented that way.

There are websites like “Disboard” which purport to allow you to search for servers related to the gaming app Discord on the open web, but we will call that part of a closed network, because many of the servers worth writing about will require an invitation to join. Disboard and similar websites do not appear to pick up on a complete list of active Discord servers.

Examples of Closed Networks:

Discord, Riot, Signal, WhatsApp, Telegram, Wire, Kik, Threema, Keybase, messenger, text message, Facetime, Skype, Google Hangouts, the direct message feature on public-facing websites like Twitter.

Reporting on Other People’s Private Conversations

You may encounter screenshots posted on open networks of conversations that were captured on closed networks. These conversations provide challenges in terms of verification because they often cannot be authenticated by searching the internet.
Let’s say in a different version of the tankie bank robber story, you have searched for hours but cannot find Jane Smith’s Facebook or Instagram profile. The only thing you have to work with is a police issued-press release suggesting that this 20-year-old woman robbed a bank. In doing your due diligence, you call another 20-year-old woman named Cynthia Jackson who was close friends with Smith, according to local news reports. Jackson tells you, “Jane Smith is a tankie who robbed a bank in the name of communism and I have proof.”

Jackson shares with you screenshots of private text messages she claims to have had with Smith prior to the robbery. One screenshot appears to show Smith sharing a picture of young Joseph Stalin. The text says, “I’ll be robbing a bank in two days to show my love for him.”

Unfortunately, you cannot take this information and rush to publish your scoop without doing a little more work. Why?

- Screenshots can absolutely be faked.
- You were not there and do not know the full context of the remarks.
- You need more proof to be sure the remarks came from Smith and not someone else.

Will Jackson let you take a look on her phone to see the extended conversation with Smith? Can you in some way verify that the number used was Smith’s? Can you reach out to more people to help back up the claim that Jackson and Smith were friends? The closer you are to verifying the authenticity of what Jackson shared, the closer you are to reporting out your findings. Still, an important rule of thumb is to always verify everything, regardless of whether it was shared publicly or privately.
2. Searching the Open Web

Nine times out of ten, typing a word or phrase into your Chrome search bar will give you the answer you’re looking for as a reporter. You can find a band’s website, or a Congressman’s office line and email. A few clicks back, you can even find some guy’s 2007 Wordpress blog where he analyzes the psychosexual subtext of the alien suit from the film Spider-Man 3. It’s that tenth time that requires some effort.

Using Google’s Advanced Operators

Google offers some in-house tools to help with digging for hard to find material. These are called search operators and there are many of them at your disposal. I would list them all here, but only a few of them have been useful to me in my work. The ones that have been useful, however, have saved me time and turned up things I would not have otherwise found.

Google provides an extremely useful tool for searching within websites, for example: Site:[website address] “keyword”. The tool is particularly useful with fringe social media sites that offer ineffective in-house search tools, like the white nationalist-friendly Twitter knock off Gab.

Here’s one way I use that operator: Eric Striker is the pseudonym of a podcast host and Daily Stormer writer named Joseph Jordan who posted on Gab under the handles @estriker, @Eric_Striker, and @Eric_StrikerDS before dropping out of that site altogether in February of 2019. When reporting on Jordan for SPLC, I would put “Eric Striker” into Gab’s search tool. The results I’d get back would be chaotic, and of little use to my work.

To get around that I would type “site:Gab.com Eric Striker” into Google, and bring up the @Eric_Striker handle with the first result. The search would also produce scores of
his relevant posts. A more concentrated search, “site:gab.com/Eric_Striker whites”, for example, turns up specific posts wherein the user mentioned the word “white” or “whites.” The tool enabled me to bring up his Gab posts dating back to when he first joined that site.

Here are some other useful operators within Google that bring up material you may not get from dropping a string of keywords into a search bar:

- **Cache:** [website/page name], produces the latest cache result of a website or page.
- **Inurl:** [Keyword], produces results with specific words inside of the URL. For example, let’s say I were searching for old commencement documents related to Queens College’s class of 2013. I could search for “Queens College 2013 inurl:pdf,” and limit my search to only PDF documents.
- **filetype:** [Keyword], will also limit to type, as in PDF documents or Word documents.
- **Adding a “-” character** before a search operator will exclude the indicated operator from the results—for example, “-filetype:pdf” if you’d rather Google not return direct links to PDF files.
- **AROUND(X),** produces a proximity search between two words. If I wanted to search for “Goldman Sachs AROUND(5) fraud,” I could see all of the times Goldman Sachs, appeared within five words of “fraud” in a Google search.
- **inposttitle:** [Keyword], retrieves blog posts and articles with certain keywords in the title.
- **Date Range:** Type in your search with any of the above operators. Hit the box below the search bar called “Tools” and click on the drop down menu marked “Any time.” Scroll down until it says “custom date range” and enter the date.

For almost everything else, Google has an effective [advanced search tool] and largely self-explanatory fields for searching within it.
Searching Beyond Google

Google doesn’t, of course, own the web, even if it feels like that sometimes, and different search engines often produce different results. There are certainly privacy reasons not to use Google, considering the degree to which the company mines personal information.

Here are a few search engines that have given me alternative search results to what was found on Google (in no particular order):

- DuckDuckGo
- Yandex
- Bing
- Baidu
- Monstercrawler.com

“BvsG.org” is also a website that lines the results from Google and Bing up against one another side-by-side, if you want to compare results on a topic. It’s a neat trick but I’ve never found it to be particularly useful; your mileage may vary, especially if you’re actually writing about web search technology. Some of the above Google-specific operators will work across multiple search engines, though not all.

Boardreader.com is a search website that indexes forums. The advanced search tool will give you results from specific websites, including Reddit, Voat, 4chan, 8chan, and others.

Search engines are an industry, and searchenginewatch.com covers the arrival of new search-related websites that crop up. There are hundreds of them, but many produce complete duplicates of one another.
Searching Social Media

*Facebook*

Facebook presents unique challenges to reporters for a number of reasons. For one, the website is massive, boasting over two billion monthly active users. For another, every user’s timeline can drift seamlessly between public and private posts, and reaching deep into a person’s timeline for relevant information can be daunting. There are some shortcuts to searching Facebook, but they are far from perfect.

Every Facebook profile has a numerical ID, which can be found on several free sites if they are public facing. Here are three sites that pull Facebook accounts by number:

- Findmyfbid.com
- Commentpicker.com
- SmallSEOtools.com

The rapper Gucci Mane’s ID, for example, is 122044727596. Once you have obtained an ID, finding out specific details about Gucci Mane’s Facebook activity is a matter of plugging that number into search related URLs. For example:

- Stories tagged with Gucci Mane:  
  https://www.facebook.com/search/122044727596/stories-tagged
- Photos of Gucci Mane:  
  https://www.facebook.com/search/122044727596/photos-of
- Videos of Gucci Mane:  
  https://www.facebook.com/search/122044727596/videos-of

Here are some more specific search options to tag after the numerical ID:

- /photos-by/
The /reshare-stories-by/ tag is particularly helpful in determining what kind of user and how many—shared someone’s viral post.

Some websites provide direct tools for performing deep searches of Facebook profiles, and I prefer to use them directly rather than try to figure the vagaries of Facebook as they relate to a given search. Here are two with which you can play around at your leisure—type your own handle into them and what you find may compel you to delete your account. These tools make it easy to unearth a birthday party you attended in 2009, for example, and the pictures from it that you can no longer easily see by simply scrolling top to bottom on your own page.

- Stalkscan.com
You can imagine how websites like this could be particularly useful in producing information in a breaking situation. Let’s say a shooting suspect’s Facebook ID has been obtained and verified— websites like this can help you to unearth newsworthy details with quickness.

Using a Facebook String Search

Facebook’s in-house search sometimes produces underwhelming results, as I mentioned, and I have found this to be particularly true when searching for junk and so-called fake news articles. One way I work around this is by conducting a Facebook string search—simply writing a custom URL and asking Facebook to generate that page by entering it into the address bar. Let’s take a look at an article on the junk news site Newspunch.com, called “Deep State Panic After ‘Dark Overlord’ Leaks Thousands of 9/11 Papers.”

The URL for the article is:

Dropping the headline into Facebook’s in-house search bar produces one result, and a search for the URL there produces no results. A Facebook string search, however, demonstrates that scores of people have shared that story on Facebook, including several highly suspect accounts.


It appears a little confusing but let’s take a look what was plugged into that URL to produce these results. The words in bold were plugged in from the headline—just the first three words, initially. This brings up all the posts with these words in the title.
If I wanted to add an additional word to the search—in case there are too many stories about a deep, dark, state of panic, for example—I would do that by adding an additional “%20” to that URL followed by a keyword.

The second thing to plug into this template is the date range. The “story” was published on Newspunch.com on January 5, 2019. The date range I chose for the search was January 4, 2019 to January 6, 2019.

Those are the only parts that need to change to produce a more effective search. Don’t change anything else and plug your keywords into that template. The string search will also work for specific phrases, and can help you understand how pervasive a fake news story has become. Here’s a string search of posts in 2018 that combined the words “adrenal” and “Qanon,” referring to the false conspiracy that Democrat elites are harvesting people’s adrenal glands for use in a satanic ritual. Note which areas were swapped out in comparison to the previous URL (these are bolded).


Tracking Facebook Users Following a Change in Name

Facebook allows you to change your name once every 60 days. Some expert level trolls have been able to survive multiple suspensions from the social network by 1) buying off old handles from shady vendors and 2) Changing their usernames every 60 days to avert a ban. Here are some steps you can take to find them:

- Bookmark the Facebook account and the profile will remerge regardless of name change
- Search for the account by numerical ID, rather than name
- If you don’t have a numerical ID, search for past names on Facebook and keep an eye out for tagged material. If the name is tagged “Joe Smith” but the subject has changed his name to “Gary Brown,” the Gary Brown profile will appear when you click on Joe Smith.

Twitter

Twitter has one of the more useful in-house search tools in the industry. It allows users to search deep into its archive, and to do it by date and keyword. The tool has helped produce a culture within the site wherein users dig up old tweets for the purpose of demonstrating another user’s hypocrisy, sometimes to comic effect. The only catch is the search tool is slightly out-of-the-way and hard for casual users of the site to find.

Enter a keyword into Twitter’s search engine, and when the results have been retrieved, click on the link that says “show” in the box that says “search filters.” Then click on the link that reads “advanced search.” You can then build your search from words, phrases, hashtags, the exclusion of words, specific people, mentions of specific people, and from different date ranges, among other criteria.
Here’s a search for when @DennysDiner tweeted the word “dope” between January 2012 and September 2016:

Tracking Twitter Users by Numerical ID

Like Facebook, users have a numerical address as well as a handle. Websites like Tweeter ID, https://tweeterid.com/, help to pull that ID for you. Baseball player Jacob deGrom’s ID is 723517166 and his teammate Noah Syndergaard’s ID is 361351581. If you are in the business of tracking specific users, making note of the numerical ID will help you pinpoint their current whereabouts on that site.

Copy the following URL and plug the numerical ID after the equals sign:
Jacob deGrom: https://twitter.com/intent/user?user_id=723517166
Noah Syndergaard: https://twitter.com/intent/user?user_id=361351581

If those Mets starting pitchers change their names and handles on the site, their user IDs will not change.

Tracking Twitter Users Who Are Repeatedly Suspended
Chances are if you’re tracking someone down and meeting with resistance, it will be a person who is up to shadier business than pitching for the New York Mets. Here are some tips to finding users who are repeatedly banned from the site.

Look for signifiers like variations of names and avatars. Many users who are repeatedly suspended use signifiers like avatars and names to notify fans of their return to the site. An infamous troll who goes by the name “Spicci,” uses UFC fighter Gegard Mousasi as his avatar, for example. Spicci once bragged to me that he had been on and off Twitter with over 300 different handles. Spectre, a white nationalist troll and podcaster, used some variation of the same two or three photos of actor James Spader as his avatar for two years on the site. An analysis of 17 of his previous handles (he has likely used scores of others) helps to show what I mean by signifiers:

@AmericanSpectre
@JamesSpectreTM
@SpectreActual
@ActualSpectre
@BasedSpader
@SpectatorTM
@InSpectre_
@SpectreAmerican
@ThirdSpectre
@thedisrespectre
@TheSpectreTM
@JamesSpectreTM
@SpectreActual
@ActualSpectre
@BasedSpader

The user makes it clear to admirers of his racist and anti-Semitic posts who he is, both by using a specific type of picture and names with recognizable elements. Users like “Spectre” often operate in the same circles of different accounts. One way to pinpoint users like this is to see the first groups of accounts they follow, which are often similar.
Twitter users who repeatedly evade suspension also sometimes buy older handles from gray-market dealers. These accounts come with a cluster of pre-existing followers, so judging by the first accounts that follow them is often futile. Another way to find older handles is to locate older conversations through a search. When Twitter suspends an account, the timeline vanishes, but their handles remain tagged in conversations.

Once you find old handles, Google keeps old posts visible in cache for months even after the account has been taken down, and is usually your best bet for locating content that has been deleted from the site.

**YouTube**

YouTube has a fairly straightforward and effective search engine. Use the “+” symbol ahead of words for which you definitely want results and “-” for words you want omitted. If you’re looking for Kate Bush music videos, and a BBC-made documentary of her is clouding the results, type “Kate Bush -BBC” to get the material you want.

Google, Bing and other search engines offer a video option during a search, if you prefer to search off of YouTube for YouTube videos.

**Searching YouTube for Speech by Text**

One useful trick on
YouTube is the ability to search for transcripts of words spoken on videos posted to the site. When the video is opened up on your screen, click the icon with three dots on it in the lower right hand corner of the menu options. It’s immediately below the icon to expand to full screen. The ellipsis icon will give you three options: “Report,” “Open Transcript,” and “Show Translation.” Click “Open Transcript” and use it to search for specific keywords if you’re lost looking for a quote in a longer video. The tool is particularly helpful for scanning the content of dense podcasts or live streams posted to the site. If “Open Transcript” is unavailable for the video, click “Show Translation,” which opens the video to a page with subtitles.
**Instagram**

Instagram is arguably the least complicated of the major social networks because it mostly consists of images and hashtags. This also makes its content far more opaque to users without technical expertise than material on YouTube, Facebook, or Twitter. The site’s in-house search bar is utilitarian and gets the job done.

I frequently use Google’s site search tool to browse Instagram from my desktop and get a mix of both hashtags and account names in one place. Sometimes this shakes loose material you wouldn’t see from the in-house tool.

For example, here are the results for “India elections.”

You can also try Webstagram, an Instagram specific search engine, to see if it produces any leads you missed from the in-house search tool.

**Social Media Tools for Breaking Situations**

Every newsroom is different, but breaking situations are always the same. An event takes place that immediately overtakes the newscycle, and everyone snaps to attention, trying to unearth new information about what happened. Beyond the other suggestions in this section, I
would like to highlight two tools that are useful in terms of searching social media in specifically breaking news situations.

- TweetDeck provides results for multiple searches at one time and allows users to view multiple Twitter timelines on the same screen. Let’s say you create a feed that only follows the accounts of official police and fire stations around your home city, one that follows government officials, and another one that follows only local reporters -- TweetDeck enables you to watch all of those feeds at the same time, keeping you on top of news in your city as it breaks. It’s also particularly useful for picking up trends on a national or international breaking story before people begin tweeting about something under a unified hashtag, like “#MallAttack.” You can search “Mall attack,” “shooting,” and “Minneapolis” at the same time and track the results as they first come in. If something internationally significant happens, like “#Brexit,” you can track Brexit-related news across four different columns by pairing that hashtag with different keywords.

- Map.snapchat.com: Snapchat provides a “heat map” of activity by location. If there is a fire in Oklahoma City, for example, type Oklahoma City into the search engine and the map will show pockets of activity being generated from the app down to the block level.
3. Verifying the Authenticity of Social Media Accounts

Full disclosure: I was duped by what was apparently a fake Julian Assange Twitter account in February of 2018. You do not want this to happen to you.

There were two unverified accounts purporting to be Assange with high follower counts, and I apparently quote-tweeted the wrong one. My tweet, which used a tone that could easily be perceived as being critical of the dissident journalist, was up for only a few minutes before I took it down. But the allegedly real Assange account (it has since been suspended) mocked me with a comment, and I was instantly swarmed by hundreds of trolls. The not-so-subtly pro-Assange Russian news website RT even ran a 200 word story about this mix-up. Later, the neo-Nazi website Daily Stormer featured me in their headline story on February 11, 2018, savoring the error of someone they saw as an enemy:

“Michael Edison Hayden ... is a 90 IQ Arab who has attempted to make a career out of harassing members of the Alt-Right and their mothers. It is refreshing to see him so viciously publicly humiliated,” the website’s editor Andrew Anglin wrote.

I learned two lessons from this situation: 1) I try to verify everything I see online before I comment on it, and 2) I try to hold my social media posts to the same degree of fact-checking scrutiny I would the draft of a story, even though I’m often writing in a different voice. (I still sometimes fail at this!) This chapter is only about verification, but you might find the second point above to be useful advice.

Investigate the Source
It’s eerily easy to create a social media account that looks like it belongs to someone else and fool the internet with it. For example, President Donald Trump rose to power in no small part through the help of his verified and extremely active Twitter account, @RealDonaldTrump. The avatar Trump uses, which features a close-up photo of his face knotted into a sober-looking grimace, has been seen so many times on that website that it probably appears in people’s dreams.

And, because Twitter thrives on recognition, scores of users have snatched up some variation of the same avatar image and name, each hoping to exploit that feeling of familiarity to rack up retweets and likes. Here are 15 accounts that appear after President Trump’s handle in a search for the phrase “realDonald” that use some variation of the same avatar:

@RealRealDonaldT
@RealTrump2016
@RealDonaldTrumpCntxt
@RealDonaldTrumpTrmu
@RealPresDonald
@RealDonaldTrumpTRfan
@DonaldTrumpEgo
@TrumpbutTrue
@RealDonaldTrumpTrumpr
@RealDonaldTrumpTranp
@realDonaldTrumpNG
@RealDonaldTrumpTrmpu
@Fake_RealDonald
@AmDonaldTru

---

Donald Trump Fan Retweeted

Donald Trump Fan @realPresDonald · Feb 5

Replying to @WhiteHouse @realDonaldTrump

Retweet if you support President Trump! #SOTU #SOTU2019 #MAGA
Some of these accounts are designed to trick people into believing that they’re seeing the president appear on their timeline, and some likely aren’t. But it really doesn’t matter why these accounts exist. What matters is that each of them can deceive you, especially when you’re scrolling quickly through a busy timeline.

Before you report on a social media handle take the following steps:

- Slow down and don’t worry about being the first person to notice something. Worry about being accurate and thorough instead.
- Check the start dates and follower counts of profiles. This information is available on the user pages of major and minor social media platforms. Ideally, taking this step will rule out any accounts that were recently started for the primary purpose of deception.
- Read biographical information for disclaimers. Fan accounts and parody accounts often declare what they are in order to avert being suspended for impersonating someone else.
- Read back through the user’s timeline. Do the interactions seem consistent with the type of person you believe this account holder to be?
- Search for different variations of the same name on the same platform and compare them to determine which ones might be inauthentic.
- Search for other variations of the same name on other social media platforms to build a case that supports authenticity: Did the person post the same or similar material across different platforms? Does the person appear to congregate in similar social circles across different platforms?
- Always reach out to the person who issued a comment upon which you’re reporting for confirmation using traditional journalistic methods.
- Check for verification: Is there a blue checkmark next to the user name? Should there be?

If you do this every single time, you will run a much lower risk of the kinds of embarrassing mistakes that damage your credibility as a reporter.
Be Aware of Fake Verification Checks and Websites that Produce Fake Posts

Verified handles on Facebook, Twitter and Instagram certainly make it easier to determine whether or not public figures use certain accounts, but not everyone who is newsworthy is verified (as was the case with Assange). Adding further complication is the fact that verification checks can be faked in screenshots. Furthermore, people build websites wherein users can create fake posts.

FakeTrumpTweet.com, for example, is a website that produces screen caps of fake tweets from Trump’s @realDonaldTrump account. The site is popular with the alt-right movement, whose adherents sometimes use it to make Trump say things they want to hear. When Arizona Senator John McCain died on August 25, 2018, The Daily Stormer used that site to produce a screenshot of a fake Trump tweet saying, “I like senators who don’t die.” Following the publication of the fake tweet, the post was shared both by people who appeared to understand the trolling nature of the post, and also by people who believed the tweet to be authentic.

Always Track Quotes to the Source and Be Wary of Citations

Another minefield is the preponderance of fake quotes online.
For years, trolls attacked tech writer Sarah Jeong on Twitter, and she sometimes fought with them. When the New York Times announced her hiring on August 1, 2018, Twitter users combed through Jeong’s extensive output, highlighting words from these fights that, among other things, appeared to compare white people to urinating dogs.

The hashtag #SarahJeong trended across on Twitter on August 2, 2018 and controversy-stirring Jeong quotes were spread far and wide, making her a household name among Trump supporters. At least one fake quote passed along in this collection of real ones however was apocryphal, and appeared in the form of a meme. It was most likely first posted on August 5, 2018, on the subreddit /r/MensRights, based upon research I performed at that time, presumably by someone who sought to further inflame the anger that was growing against Jeong and her new employer.

“We don’t need a military draft. We need a castration lottery for white men,” the fake quote declared. “Every month we pull a birthday, sort the excess and snip some sacks, preferably in a big public gathering. Imagine how serene public transportation would become.”

Jeong’s critics seized on it. The right wing Twitter account @The_Trump_Train posted the “castration lottery“ meme-quote with an ominous message: “This is why we need AR-15s.” A network of conservative Facebook groups also picked up on the meme. People in Facebook comment threads even threatened Jeong with sexual assault and murder over it.

But one difference between this quote and numerous real ones Jeong issued about white people (besides from the degree to which this one felt instinctively hyperbolic) is the way it was presented. The “castration lottery” quote was presented only as a meme and without any corresponding screenshot. The meme featured the quote, an image of Jeong that crops up when you search for her on Google images, and a watermark with the The New York Times logo. But a search for “Jeong” and “castration lottery” returned nothing from her published writing.
Some lessons from this incident apply broadly to anyone attempting this kind of reporting:

- Never take a quote you see online at face value without confirming it with the source.
- Be wary of any memes or social media posts promoting quoted text, especially involving public figures who are at the center of controversy. Trolls gravitate to news that is trending.
- View any posts on Twitter that claim to retweet someone without sharing the source with suspicion. (For example “RT @MichaelEHayden I am an extremely satisfied New York Knicks fan.”)
- Contact the person who is being quoted to confirm the authenticity of the quote. In the case of the Sarah Jeong story, Jeong acknowledged to me that the preponderance of quotes being shared online were authentic, but that the “castration lottery” quote and several others were falsified.

**Reaching Out to Your Subject**

If a public social media post is integral to your story, you must reach out to the person who posted it for more context.

- Confirm the identity of the person using traditional journalistic methods like tracking down a phone number, when possible. Public databases like telephone books are more useful than we often remember; public libraries often have access to large quantities of public information, both in print and via government access to private databases on computer terminals within the libraries themselves.
- Always introduce yourself by your real name, and the outlet for which you are working.
- Inquire about the subject’s safety, particularly in breaking situations like mass shootings, accidents or natural disasters.
- When approaching users publicly, ask to redirect the conversation to a closed network to avoid becoming part of the story for other reporters.
- Ask for confirmation of identity and report it out in your story when applicable.
• Ask for context behind the content about which you are writing. Was the post part of a larger discussion about which you are unaware? Do the words in a post mean what you think they mean? (For example: Do the words possibly refer to a meme, a song lyric, a line from a film or an inside joke of which you are unaware?)

• Always ask if the subject would be willing to expound upon what was posted. Social media posts are already publicly visible, and your readers can often read the same thing you’re reading. Providing added information of value is one reason why they will read your work rather than simply find it for themselves.
4. Verifying the Authenticity of the Images and Videos You Find Online

It doesn’t matter if a falsified image looks ridiculous to you—someone out there probably thinks it’s real. You may want to give people the benefit of the doubt. Don’t do it.

A photo went viral in September of 2018 that appeared to show Donald Trump riding a raft over flood waters. The Carolinas had been walloped by Hurricane Florence beginning the previous month, causing $17 billion in damages. In the photo, Trump reaches out to a man who is stranded to hand him a red Make America Great Again hat.

The photo was originally taken in 2017 during Hurricane Harvey. It captured rescue workers from the Austin Fire Department helping a flood victim. An unknown indoctored the version that went viral the following year. The faked image first appeared on Facebook in 2017 and again in 2018, when something about it abruptly clicked with the social media hivemind. The image racked up 275,000 shares on that site before they removed it. Many people saw the image and knew it to be fake. Some didn’t.

In hindsight, you may ask yourself: Why would the president of the United States sail into rising flood waters wearing a suit and tie? Where was his secret service detail while this was happening? You may give it even more thought and ask: What would you do with a red Make America Great Again hat when you’re facing a potential death by drowning?
From the perspective of the technology companies that helped to spread this disinformation, none of these things matter. Websites like Facebook and Twitter are designed to encourage people to act with speed and frequency when they are logged in. It’s our job as reporters to slow down social media and hold viral content up to scrutiny, regardless of how ridiculous it might be. And that means taking the time to verify or debunk images that we already instinctively know are completely fake with as much zeal as those we’re absolutely certain are genuine.

**Using Images**

Start With Reverse Image Search.

Reverse image search is a widely used, simple way to determine the origin of photographs, but it’s hardly perfect. Essentially, you grab the image in question and upload it into a reverse image search and see what descriptors arise—the opposite process of finding images by entering keywords. It fails to find a given image as often as it succeeds, but these searches can quickly rule out a mislabeled or misattributed image on occasion.

Google provides the most popular reverse image search on the web, but using other sites can help to increase your pool of potential answers. Here are a few sites I use for reverse image search to provide a diversity of results.
Analyze Every Image with the Assumption That It’s Fake

Let’s say you’re reporting on a fatal stabbing incident that took place in Times Square and someone on Instagram shares an image they suggest is a photo of the killer. The photo has newsmaking potential—you see that someone who looks a lot like the alleged killer can be seen in it, arguing with the victim, someone dressed as Elmo. The person who posted it claims that the conversation took place just moments before the attack, and suggests that it may offer up evidence of a motive.

The first question that should come through your mind is: Who took the picture and when? If the person sharing the image appears to be based in Cedar Rapids, Iowa, and not New York City, New York, for example, that might be a clue that he or she didn’t take it. And, if he or she didn’t take it, you need to find out who did by tracking down the image elsewhere. Search for terms related to the photo across different social media platforms to find the original, and ascertain whether or not it was truly captured during the breaking situation in question. Use a reverse image search to determine whether it came out of the archives instead or was manipulated.

The next question that should come through your mind is: Is this really Times Square? Match up available details in the photo with what you can find on Google Maps and other mapping websites that provide details about location. Let’s say there’s a Bubba Gump Shrimp Co. restaurant in the photo. Can you match it to what you find on Google Maps or other mapping sites? Can you find images of a nearby location for side-by-side comparison?

- Search different platforms for words you would use to describe the photo to find other versions of it across social media and compare them. It’s best to use natural, conversational language in your searches. Using the above example: Search for...
'"Times Square" elmo arguing', then '"Times Square" elmo argument', then '"Times Square" elmo fighting" to jar loose as many posts as you can.

- Read the timestamps of different posts, as well as the language used in written descriptions. Pay special attention to text that suggests a firsthand understanding of a situation, like a credible explanation of why the person posting happened to be in that place at that time.
- Look for details that corroborate authenticity: Are there landmarks? Can you cross reference other images from the scene to match up people, signs or license plates?
- Investigate the poster’s available internet history to rule out people who were unlikely to have been at the scene of an event.
- Investigate the poster’s timeline for an ideological bent that might incentivize a person to spread a misleading photo. (To be clear: Such an ideological bent doesn’t mean that you should ignore images shared by that person. It’s just an added reason to hold someone up to scrutiny.)
- Reach out to the person directly for details about how the image was obtained.
- Use a reverse image search to determine if an image has appeared elsewhere.
- Use a service such as WolframAlpha to find the weather for the location and time the user claims the picture was taken. If it’s January in Times Square and everyone in the Instagram photo is wearing shorts, but the historical weather data says it was 4 degrees fahrenheit, be suspicious. Analysis of shadows can be used to crosscheck the time of day the photo was taken, as well—if a landmark or a person is casting a shadow to the east, the photo wasn’t taken before noon.
- In Google Street View, there is a date slider near the top of the page that can be used to view historical photographs of a given address from Google’s archives. If the photo you’re looking at came from 2015, visit the address on Google Maps and scroll back to 2015, or as close as you can get. Were all the buildings in the background of the 2015 photograph really there when the Google photo truck drove by? Were the signs the same?

First Draft News, a non-profit group which supports journalists, also offers a useful checklist for verifying photos.
Verifying Original Images Using Metadata

If you have successfully contacted the photographer, ask him or her to send you the original image. Simple viewers that allow you to view the EXIF metadata assigned to the photo by the camera are available online; a good one is Jeffrey’s Image Metadata Viewer. Most social media services strip out image metadata, though a few image-sharing services, such as Flickr, retain it; obtaining the original image may not always be possible, but it is ideal.

Image metadata will give you a wide variety of information about the photo: Whether it was taken with the front “selfie” camera or the rear camera of a cell phone, or the date and time the photo was taken using the camera’s clock. In breaking situations, authentic images will often be from phones, which are usually set by default to adjust their internal clocks automatically using information from local cell towers. On an iPhone with location services enabled for the camera, the EXIF data will also contain latitude and longitude.

Using Video

Authenticating video follows more or less the same processes as images. Older videos can be scraped, just as easily as images can be lifted. And it isn’t at all uncommon to find people repurposing older videos or distorted videos during breaking news stories to spread disinformation or go viral.

One primary difference to consider is that videos can be selectively edited to change their overall impact. Whereas a distorted image is inherently doctored, a video may be authentic in part while being false on the whole.

Speech can be clipped out seamlessly and scenes can be deliberately cut short and displayed out of context. Sound effects can be added to alter the atmosphere of what is being seen. Just as with images, the first step is to ascertain the source of the video.
- Track down the source of the video and rule out reuploads and scrapes, which are videos exported from another source. Scraped videos frequently lose quality because they re-record an original clip.
- Search for keywords used in the video’s description to compare them to other versions, when applicable. Be sure to search outside of YouTube to websites like Vimeo and also fringe platforms like Bitchute, Pewtube, and D Tube.
- Investigate the account that posted the video: Does the person have a history of posting dodgy-looking content? Does the poster have what appears to be a strong ideological bent?
- Confirm that the person who claims to have captured the video was in fact in the place from which the video appears to emerge.
- Gather as much information about the video as possible for purposes of verification: People, clothing, time, place.
- Always reach out to the person who posted the video to ask tough questions.
- Some websites offer reverse video search, but the tools available at the moment are a little wonky in my experience. Stills from videos, however, can sometimes produce results in a reverse image search, which can help you track down a clip’s origin. Amnesty International has a tool to help automate this process, called YouTube Dataviewer—if a video has already been uploaded to YouTube, but is, for example, fraudulently re-uploaded with a misleading description, Dataviewer will find the original, the date and time it was uploaded. It will also provide a thumbnail, while also showing any other copies of the video on the site.
- First Draft News also offers a useful checklist for verifying videos.

**Don’t Rely on Your Ability to Recognize a Face**

Seeing a particular human face isn’t reliable confirmation of a person’s identity in an image or video, and should never be used to prove verification. You may believe you are looking at a specific person. Your hunch may even be right. But faces are tricky and are much less reliable than the appearance of inanimate things like signs, buildings, or even clothing.
**Requesting Permission to Use an Image or Video**

Major publications typically provide a handbook for using images and video found on the web and often require reporters and editors use specific language in requesting permission to use them. The Tow Center also covered the matter of seeking permission in depth in 2014 [Global Study of User Generated Content](#).

Freelancers rarely have to worry about obtaining visual elements for a story themselves, but if you are working freelance and you do have to reach out to someone, always be courteous and clear about your intentions to use what you find. Specify exactly how the image will be used and make a record of the person giving you permission to use it.

When in doubt, embed a social media post into the body of a story instead of appropriating it directly. It’s important to remember though that not everyone is looking for the kind of widespread exposure they would get by being featured in a major publication. The best bet is always to make contact with uploader before using the post in your story.
5. Exploring Fringe Websites

For many Americans, it’s easy to ascribe a sense of permanence to the things they see and do online. Maybe they log onto a few major websites like Facebook, Twitter, and Instagram. Maybe they check their fantasy sports teams on corporate-backed websites like CBS or ESPN. Maybe they get a little information here and there about politics on Apple News. Maybe they do a little shopping, or take in a little pornography. Activists on both the left and the right would call people like this “normies,” implying that they aren’t clued into the minute details of 24/7 internet politics.

But on the fringes of the internet, non-“normies” are plotting to reshape normie opinions on urgent issues like immigration, health care, and war by pushing memes and other propaganda into their line of sight. Casual internet users may be unaware that these same people gather to hash out plans about them on obscure, often-clunky sites that have no sense of permanence. I’ll call them fringe websites here, and they span sites like 4chan, which get hundreds of thousands of page views per day, to obscure forums that get attract eyes than the average tweet.

And on the fringes, there is far less permanence. Influential communities with thousands of users disappear overnight without a trace: The white nationalist forum “The Right Stuff,” which had just under 10,000 users, was reduced to a single thread in May of 2018 after its proprietors ran into legal trouble. WrongThink.Net, a say-anything Facebook clone with a large constituency of racist and anti-Semitic users, found the service down for “maintenance” following the Pittsburgh Synagogue shooting in October of 2018. The website never returned.

From the rise of populist right-wing figures like Trump and Brazilian president Jair Bolsonaro to the Brexit vote, words whispered on tiny forums like this have found their way into the mouths of people with real power across the world.
What is a Fringe Website?

For our purposes, a fringe website is any esoteric community that resides off the grid of major websites like Facebook, Google, YouTube, Twitter, Snapchat, and Reddit, where people gather for public or at least quasi-public discussion. I’ll focus here on communities that have a disproportionately outsized impact on the culture at large—particularly when it comes to political discourse—but anything that works outside of the mainstream social media industry will cut it.

Alternatives to mainstream platforms tend to draw a crowd that sees itself as in conflict with the neoliberal cultural hegemony, which means that they naturally will court users who want to change our values, laws and language. In my experience reading and interviewing people who use these sites, one of the primary reasons they flee mainstream platforms is the desire to use language that may be deemed unacceptable on them. Language that may be acceptable on fringe websites that is not acceptable in the mainstream would include hate speech, the glorification of violent ideologies like Nazism, language praising ISIS, explicitly organizing harassment campaigns, and threats of violence. This clash between culture and anti-culture often creates news, which is why we cover the fringe with an interest that is disproportionate to its size.

Examples of Fringe Websites

Here are some examples of fringe websites. Please keep in mind that some of these websites exist in a tenuous reality and could be gone or marginalized overnight:

- For discussion: Voat, Gab, FreeSpeechExtremist, Parler, Minds, My Posting Career, Mastodon, VK, 4chan, 8chan, Endchan, Fascist Forge, Kiwi Farms
- For video hosting: BitChute, D Tube

Finding New Discussion Hubs
There is every likelihood that someone is posting something that is relevant to a story you will want to write on a website you and I have never used before. How do you find something when you don’t know what you’re looking for? While there is no right answer to this question, the easiest way to find these partly concealed hubs of discussion is to give close readings to similar or sympathetic discussion forums.

Communities on the internet are built through word of mouth, almost like restaurants that abruptly become popular in a town or neighborhood. You may be tracking a particular Twitter user and find links to his profile on Gab. Or, you may be surfing through Gab and encounter a link to a video on D Tube. Always bookmark links and make notes of your travels online to avoid losing track of something you found noteworthy—remember, much of this could be gone the next time you open your browser. I’ll elaborate more on archiving your work in Chapter six.

**Tracking the Flow of Users from Fringe Networks**

When “The Right Stuff” forum was reduced to one thread and its extensive history wiped in May of 2018, its white nationalist user base didn’t die—they migrated elsewhere. Similarly, when Gab was taken offline for a week between the end of October and beginning of November in 2018 following the Tree of Life synagogue shooting, users regrouped on several different sites.

Tracking the migration of users from a disabled platform to new places online requires either offline connections to users of the original platform (so you can ask them directly about where they’re headed), or that you find activity that points to a migration.

A poster from “The Right Stuff,” for example, linked to a room on “Riot” as that forum was closing down. Riot is a messaging and discussion app similar to Discord, and two hours after that user’s post, the room went from non-existence to having several hundred people in it, discussing the fate of the white race, just as they had on the forum. Before Gab was briefly taken offline, users saw its demise coming. Some took advantage of Gab’s waning hours before a shutdown to suggest they regroup on
WrongThink.Net. But when WrongThink.Net began to see a number of controversial new arrivals, its proprietors abruptly and somewhat mysteriously took the website down for maintenance. Users then quickly recommended the cryptocurrency website Minds.com, which is where high-profile Gab users like Christopher Cantwell, “Microchip,” and anti-Semitic politician Patrick Little ultimately reconvened in Gab’s absence.

Fortunately, many pseudonymous participants in online forums enjoy the prominence they achieve in these communities and are loath to leave it behind when they migrate to a new service, so they use the same handles and avatars on different websites -- making them easy to find again. The neo-Nazi forum Iron March, The Right Stuff forum, and the Daily Stormer’s forum “The Goyim Know,” for example, featured many prolific posters whose pseudonyms and avatars also appear elsewhere. The use of the same name isn’t indicative of a lack of creativity, necessarily. The users want their friends and fans to be able to find them.

- Always make a note of conversations that indicate where users say they will move in the event of an emergency.
- Keep an eye on high profile names on forums and websites and try to search for their handles elsewhere.
- Check major websites like Twitter for conversations where people talk about reconvening on a different website by searching out the name of the original forum or website and scanning conversations about it.

Tracking Internet Campaigns from Fringe Websites into the Mainstream

Tracking and archiving discussions on fringe websites and being able to demonstrate their influence on mainstream culture is a big part of what open source intelligence analysts do. The work is time-consuming and requires that reporters sometimes learn
esoteric words, codes, and ways of seeing, but a lot of it is intuitive and merely requires some common-sense detective work.

I’ll use an obscure example of an obscure campaign that failed to take off: The hashtag #DisarmThem got a small and underwhelming amount of traction on Twitter and Facebook in October of 2018. The hashtag was usually shared along with an unconvincing flier that claimed to be from “Antifa.” It urged people to attempt to steal weapons from open carry enthusiasts in the name of performing some homemade rendition of gun control. (Never mind that the young anarchists behind the contemporary Antifa movement are mostly pro-gun.) The campaign was briefly amplified by a right-wing Facebook page which at the time had over 30,000 followers.

A reverse image search of the flier didn’t—for whatever reason—return any results on 4chan, where the troll campaign was created. But a site:4chan.org #DisarmThem search on Google did produce a result, which led me to a post discussing the planned troll. The discussion of the troll in turn referred me to the original 4chan post, with the flier. It’s important to note that the Facebook account was not the originator of the campaign, despite the degree to which it was likely the most visible share. The word amplification is used to describe a post that takes an internet campaign from obscurity and into the mainstream. #DisarmThem failed to become viral, but other 4chan-spawned hashtags like the white supremacist battle-cry #ItsOkaytoBeWhite (launched on October 31, 2017), were more successfully amplified, spreading across mainstream platforms.
#It'sOkaytoBeWhite even received sympathetic coverage from Tucker Carlson on FOX News.

When tracking down internet campaigns:

- Always do a reverse image search of fakes or memes to see where else they have appeared.
- Always search broadly across websites and forums for words, phrases and hashtags that have been passed along in an internet campaign.
- Always check timestamps on posts and compare them. Read language carefully to make sure that posts are not referring to a discussion that took place elsewhere.
- If you can’t prove with absolute certainty that a meme, hashtag or fake originated in a particular post, but evidence points in that direction, be honest with your reader. Use phrases like “most likely” or “appears to have” and then lay out the specific evidence that led you to form that conclusion.
- Always be sure to differentiate for your audience the origin of a campaign from the place from which it was amplified. The viral post that introduces something to a mainstream audience isn’t always the place where it spawned.

Using Tor and the Dark Web

The idea of the “dark web” is extremely overrated. Yes, there are apparently awful things on it. And there are also tall tales about things that don’t exist.

The only thing you need for our purposes is the Tor Browser, also known as an onion browser. It’s free and open source. The browser uses relays to conceal the IP address of the person behind it and allows you to access so-called onion websites, which are not visible on the so-called clear web.

The most famous example of a dark web site that I’ve had to use for work is the neo-Nazi site The Daily Stormer, which has been kicked off of over a dozen clear web domains like “.com,” “.red,” and “.hk.” When The Daily Stormer is down, editor Andrew Anglin keeps a homebase at an onion address:
https://dstormer6em3i4km.onion.to/ (Content warning: The website is vulgar, vicious and teeming with hate.)

If you are visiting a website with an .onion address, you will need to use the Tor browser to do it. It’s as simple as that, really. The only other reason to use Tor would be to obscure your IP address to the proprietors of a website. There are simpler ways to do this, which we will discuss in chapter seven.
6. Using Archives, Saving Your Work

Every reporter has lost track of something he or she found online, and the experience is frustrating. I’ve wasted the better part of a day looking for something I saw a week earlier but didn’t save. One way to overcome this is to get into the habit of taking screenshots and making archives of every interesting thing you see while you work.

Archiving Your Reporting

Chances are, the editors at your publication don’t want you linking to an extremist website in your story, and with good reason. Archives can be a way to refer to something online without potentially giving that site a bump in traffic. But archives also do more important things, like retain evidence of controversial material that social media companies are likely to pull down, or pages on fringe networks that are likely to disappear. Archived pages are sometimes your only evidence that you saw something online.

Here are sites that you can use to grab archives of pages:

- https://archive.org/web/
- http://archive.today/
- http://archive.is/
- http://archive.li/

These websites do lose details when copied. When in doubt, take screenshots, download video, and, in an absolute emergency, use your mobile phone to record what you are seeing by taking a picture or video of your computer screen. If nothing else, you can show the video to your editor as a way of backing up what you describe.
Using Archives to Unearth Unpublished Information

Searching archives of URLs can sometimes produce interesting results, even if the captures were ordered years ago by a different person. You can find someone’s evolving Twitter bio this way, or controversial posts they’ve since deleted. I find Archive.Org’s calendar layout, which shows a blue spot at a particular time when a capture was made, to be helpful, particularly when you’re reporting on events in someone’s background by date.

Using Website Scrapes

A “scrape” is a searchable reproduction of a website. Scrapes come in handy when you need to report on something that no longer exists online. Creating and storing entire scrapes of websites is admittedly quite a bit beyond my expertise, and are usually built by specialists who have a strong knowledge of website coding, among other things.

That said, I’m currently reporting on scrapes of forums in my work with Southern Poverty Law Center. They can be fun to use, almost like walking through a memory of a website, preserved in time. When reporting on a scrape, be aware of how the scrape may be limited in comparison to the original. Many scraped forums, for example, contain text without imagery. It’s useful to consider the possibility that one of the users may have been writing about a visual element that you can’t see in the scrape, which can potentially change the meaning of a post.

Here is an example of a story I wrote using a scrape to track down the digital history of a man living in Russia who aggressively advocated terrorism on an influential forum he founded before disappearing from the web in 2017: “Mysterious Neo-Nazi Advocated Terrorism for Six Years Before Disappearance.”

Cataloguing Screenshots

One question reporters ask each other quite frequently is: What do you do with all of your screenshots? By now you probably know that shift + command + 4 grabs a
screenshot on your Mac, and that you can use the snipping tool on Windows. But figuring out what to do with the dozens of screenshots you pull per day can be quite confusing. I found the suggestions listed here by using trial and error, and they’re far from perfect. By all means employ these suggestions, but use them to find a system that works for you, so you can find those old screenshots when they’re needed.

- Put all of your screenshots during the workday onto your desktop.
- Title every screenshot with a recognizable keyword while you work, like: “8chan.LGBTQ.troll.1” for something that came from 8chan, and deals with the LGBTQ community.
- Delete screenshots you know are unimportant as you work, as a way to reduce clutter.
- Email screenshots you find on your phone back to yourself with an easily searchable subject header during times when your find something on the go.
- Keep a series of named folders for screenshots on your desktop and drag your screens to the appropriate one every day before you log off.
- Consider making subfolders by date, and then dropping them into a larger folder by the end of the month. For example, “11.9.2020” for the daily folder, and then eventually moved into a folder named “11.2020.” Later, these can be dragged to a yearly folder, “2020.”
- Use storage sites like Google Drive to keep a database of your folders.

Using Hunchly

Hunchly is a browser extension created by Justin Seitz which records the pages you visit during OSINT sessions, as you visit them. It’s particularly useful for situations when pages get abruptly taken down.

Seitz produced a video for YouTube showing how Hunchly works, if you’re interested in checking out his product.

Cataloguing Telegram and Masking Your Phone Number
Telegram is surging in importance for people who cover extremist subcultures online. Extremists form groups or rooms on the app, and participants in one room can share content to their own room with a tap of the finger, sometimes making posts difficult to track.

For example, self-described “Islamophobe” Laura Loomer formed a room on Telegram after being banned from conventional social media in May of 2019. She used her room to post the addresses and telephone numbers of reporters who cover the far-right. Hardline neo-Nazis then spread Loomer’s posts into their own, less popular rooms, creating an echo effect across the app. When cataloguing images and text shared in rooms like these, always take a moment to determine the original source, which is marked with a link.

Some Telegram groups leave membership in a room open to anyone. Other groups are locked down, and require an invitation. Any user can reach out to an individual profile on the app, and invite them into a private conversation.

If you want to use an app on your phone without exposing your personal cell number, MySudo provides alternative numbers to use, which will help protect your identity.

**Masking Your Email Address**

When signing up for shady websites, you don’t have to give them your personal or professional email address. One way to get around it is by using a service like 33mail. 33mail will give you an email address and forward mail sent to that address to your personal or professional account without anyone knowing. A 33mail user will be able to see those emails in your inbox.

**Using a Virtual Machine**
I’m going to declare the subject of virtual machines to be a bit too complicated for this guide because the subject would require an entirely new chapter to explain in a credible way, and there are many guides available for your use, should you want to set one up.

That said, it’s worth learning what a virtual machine is and why you would use one for doing OSINT work, so I’ll define it here briefly.

A virtual machine runs on your computer and approximates another system altogether. When set up, it functions like a fresh new computer made entirely of math that operates inside of the one you’re using right now.

Why would you want to use a virtual machine? Let’s say you wanted to download a browser extension and you don’t trust it on your computer. Or let’s say you want to test a program but you feel fairly certain it will be infested with malware. These would be reasons you would start looking into the use of a virtual machine, like Virtual Box. Think of it as the containment unit from Ghostbusters for sketchy internet things.

When Gab CEO Andrew Torba created a browser extension called Dissenter, I used a virtual machine to look at it, because I did not trust it to behave itself in my Chrome bar. Installing Virtualbox and using it is not a simple process, so be sure to seek out additional guides for help.
7. Learning New Platforms and Interacting with Hostile Communities

A number of writers who have reported on 4chan have never posted anything on that website. 4chan, of course, is an imageboard website that has spawned countless memes and convincing fakes. People of high moral character have good reasons to want to keep their distance: A fair share of the users of /pol/, a subsection of 4chan focused on politics, are notoriously misogynistic, cruel and bigoted. Many of them would likely consider that description a badge of honor.

For my part, I’m in favor of reporters posting on 4chan. Not every hour, or every day, or even more than once per year. But using the site enough to know how it functions from the point of view of my subjects has helped me. While you needn’t be a baseball player to report on baseball, it stands to reason that someone who has thrown a baseball at least once before has an advantage in describing the motion over someone who hasn’t. Likewise, using a website will give you an advantage in reporting on that website.

Let’s take Twitter, a website nearly every reporter has used, as an example. When looking at a tweet: Is that a reply? Is it a retweet with comment? People who use Twitter everyday will be able to answer those questions without thinking about them, while someone who has never had an account may struggle.

To be clear: I’m not advocating that reporters spread hate speech and disinformation to understand what that experience is like. I’m merely suggesting that you try out websites upon which you’re building your stories to know how they function. Post an article that interests you on 4chan, for example, and ask the users what they think of it.

Breaking Situations
The internet is incredibly vast and no amount of preparation will ensure that you are already logged into every website or app that arises in a news story. Still, you should never make an assumption about the workings of an unfamiliar service, site, or tool that is suddenly vital to breaking news. If you have time to find out by testing out a strange website or an app for yourself, I recommend taking that extra ten or fifteen minutes to do it.

If you don’t have that kind of time on a tight deadline, tread carefully and write only what you know is provable. If you are reporting out a breaking situation and encounter a screenshot of a conversation on a platform you have never seen before, tell your editor about it, and make a plan for how you will describe what you’ve seen accurately. Be sure to determine if the site is part of an open or closed network.

**Gaining Unfettered Access**

While some forums and websites grant strangers total access, many others aren’t so generous, and require that you sign up for an account. Twitter, for example, requires that users log in to view another user’s history of replies (though the search operators described in chapter two will function whether or not a user is logged in). Fringe websites often require a login to use their in-house search functions.

Use your judgment about how you enter particular communities. For example, some websites require a vetting system. Fascist Forge, the neo-Nazi forum, requires essay-length submissions to the site to unlock part of it. Rooms on Riot and Discord may require vetting from a moderator. If you lie about who you are to gain access, your story may be ethically compromised, depending on the standards set by your publisher. Doing this can also create legal complications that simply telling the truth won’t.

Unicorn Riot, an activist group, has published personal information contained in Discord logs from white supremacist groups after gaining trust and entry, but Unicorn Riot’s writers don’t have the same legal requirements that are company-wide policy at a large media conglomerate like News Corp or ABC News/Disney. Check with your
editor about how to proceed in situations where permission is required to read certain material published online. The last thing you want is to lose your hard work over regulations you weren’t aware of and could easily have followed.

Building Relationships with Online Communities

Whether or not I’m getting juicy quotes is irrelevant to me in my work with the alt-right community, and it’s rare that I even quote those subjects in a story. Most of the conversations I have are off the record, in fact. And yet, despite the fact these chats don’t help to fill in the details of a story, I chat with those posters—people with whom I’m not friends, mind you—all day long on different apps. Why do this? (My therapist may have a different reply to this question but my professional answer follows.)

The reason reporters and researchers speak with members of any community daily is to be able to report with confidence about what motivates people within it. Even if it’s racist, sexist, homophobic or if it offends me, it’s useful for me to treat it as activism in my reporting whenever people are trying to use their computers to change the world in some way, big or small. To the extent that they succeed, that’s newsworthy.

People who operate in this world daily can point you in the direction of new stories, and sharpen your understanding of a beat. Conversations like this can help you navigate breaking news, especially when someone with whom you’ve corresponded can direct you to something happening online that you might otherwise miss.

Here’s what I recommend to help build relationships with subjects in hostile communities:

- Use humor and language that your subjects will understand without using offensive words that could get you fired.
- Try not argue or lose your temper. (This is not always easy. Talented trolls know how to make people angry.)
● Never agree to debate someone on a subject or speak too personally about politics.
● When you do offer an opinion, be reasonable, and make it clear that you are speaking for yourself and not your employer.
● Be sincere about your desire to build a relationship. If you feel confident that you don’t need a conversation on the record, say as much.
● Avoid situations where the feelings of unpleasantness outweigh the potential gain of obtaining a new source.
8. Ethics and Personal Security

Treat Every Private Conversation as If It Will Be Shown to Your Boss

I’ve had at least three of my private phone conversations recorded and used as podcasts by white nationalists without my permission. I’ve had my private direct message conversations screenshot and posted to social media by the CEO of Gab and others. Whenever you’re working, be keenly aware that in this era of microblogging and self-documentation, anyone you speak to may decide to publicize your conversation without asking permission to do it. If someone is hostile to you or your publication, this note is doubly important. It’s conceivable that the subject’s only motive for speaking to you is to produce material that will damage your reputation or your publication’s.

- Try to avoid expressing strong political opinions in a conversation.
- Avoid following someone’s lead into graphic sex talk, or anything else that could look ugly when taken out of context.
- Avoid disparaging your employer or colleagues, when asked about them.
- Never get into an ideological debate with your subject.
- Avoid appearing to approve of someone’s bigoted comments and ignore statements intended to shock or inflame anger.
- Always try to bring the conversation back to topics you can use for your story.

Ethical Use of Sock Accounts

The use of so-called sock accounts, or accounts one creates in order to impersonate a generic user of a particular website or member of a particular online community, come with strong ethical concerns for reporters. While use of sock accounts may enable you to blend in and monitor happenings of an online community, the act of
straightforwardly deceiving your subjects would be considered unethical by many publications.
Still, there are instances in which creating a sock account might be necessary. If you have to monitor Twitter users who block you to keep you from doing your job, for example, a sock account should be fine, provided that you’re only using it to look at their tweets. Likewise, if you’re entering a room on Discord or Riot where there is no vetting process, you don’t need to identify yourself by your name unless you’re there to ask questions.

Always Identify Yourself Truthfully Online and Be Clear about What You’re Doing

If you’re asking questions, remember that you’re not the story. They are. And, while some people may clam up and refuse to talk to a reporter, or kick you out of a room, or threaten to kill you, there’s nothing you can do about that besides from meditate on more creative ways to open up a conversation with them (and, of course, fill out a police report if you believe a threat to the physical safety of you or your family is credible). Tell the subject you find something they wrote was interesting, if it was interesting. Tell someone you don’t understand why they wrote a certain thing they wrote, if you don’t.
Encourage a person to elaborate on the behavior they exhibit online. Most subjects respond well to honesty from reporters, even if they don’t like reporters on principle.

Make sure the subject knows who you are, where you work, and what your purpose is before he or she is quoted in a story. If you’re interested in using or even referencing a piece of their content in the story, always make an attempt to speak to that person out of courtesy—even if that’s in excess of the rules your publication has for using images and video.

- State your full name to the person and disclose the publication for which you work.
- Tell the person the subject about which you’re writing.
- When a subject agrees to speak, but asks to be kept off the record and you agree, always honor your agreement.
- Once any person understands you’re a reporter, everything they tell you is on the record. People can’t say something and then make it off the record after the fact.
Be Aware of the Potential Risks

At least eight reporters received a postcard in the mail from the white supremacist group Patriot Front in June 2018. The card had a red, white, and blue design. The words on them read, “Patriotism with Teeth.” These mailings contained no threat, but the implication was obvious: Patriot Front wanted reporters to know they could connect with them in the physical world, if they wanted to. A friend of mine described them as “wish you weren’t here” messages.

I was one of the reporters to receive a card. It came to my parents’ house. Receiving the mail on any day would be unpleasant, but the timing made it worse. The mail—postmarked on June 25, 2018—arrived at the house on June 28, 2018—the same day that five people were murdered at the newsroom of Maryland’s Annapolis Gazette by a man who held a grudge against that paper. After receiving the card, I spent the next few days explaining to local police about the type of work I did, and why I might be at risk of retaliatory violence for it. This wasn’t a fun experience, but the incident was only one of three separate instances in which I was forced to file a police report over threats I’ve received simply for doing my job.

Building contacts with trolls and engaging in dialogue with people who are difficult or dangerous is an integral part of writing with authority about the internet. Avoiding potentially unpleasant interactions like this will hinder your ability to report, but interacting with hostile communities like right-wing or Islamist extremists can be dangerous. The responsible way to approach the job of engaging in these interactions is to be aware that they can lead to harassment, harassment of loved ones, and even acts of violence.

Protect Your Information Online

Here are some recommendations for protecting yourself online:
Use two-factor identification on every major website and app without exception.

I would also recommend using a password manager to enable you to use and store more complex passwords. LastPass gets the job done.

Delete your Facebook account or at minimum lock it down with the tightest possible security options.

If you choose to lock Facebook down rather than delete your account, trim down your friend circle to a small, manageable group who you can trust not to leak.

If you choose to lock Facebook down, rather than delete your account, remove all public facing posts and change your name to something only your closest friends recognize. Search your name and untag any public facing content that includes your birth name.

Have your immediate family members and other people who can be linked to your through an internet search lockdown their Facebook accounts to the tightest level of security.

Never accept friendship requests from people you don’t know.

Avoid using the geolocation tag on Twitter and don’t list your city.

Don’t post any personal information or family images to Twitter.

Don’t divulge any personal information on Twitter and try to keep it to only work.

Scrub old tweets that reveal details about your current whereabouts.

Avoid using image based social media like Instagram, especially if you’re fond of posting selfies. Selfies can contain visual clues to your whereabouts.

Use products like Delete Me to remove yourself from personal information websites like Pipl.com and Spokeo: https://abine.com/deleteME/

Delete Me doesn’t cover LexisNexis, so send them a separate contact here: https://optout.lexisnexis.com/

Use a second mobile device for work texts and calls, if possible.

Never tell your subjects your exact whereabouts.

Using VPNs
VPN is short for Virtual Private Network. These are services that strengthen anonymity and security by obscuring your location, and I recommend that everyone use one while doing this sort of work.

The one caveat is that I wouldn’t recommend going cheap on your VPN service by downloading a free one, which can’t be trusted. Chances are, the service is making money from you in another way. Here are three I would recommend:

- Proton VPN
- Tor Guard
- Nord VPN

What to Do When You Receive a Threat

Reporters receive threats all of the time. The overwhelming majority of them turn out to be nothing. The Capital Gazette shooting in Annapolis, Maryland on June 28, 2018, however, shows that sometimes threats to reporters turn into real-world violence. Don’t assume that you’ll be able to tell which threats will lead to real world action. Respond to what’s said proactively, and with seriousness.

- Always document every correspondence that makes you feel strange or uneasy. Do so by archiving social media posts and taking screenshots. Record everything, even if the situation is stressful and you’d prefer not to look at it.
- If someone threatens to kill you, and the threat is specific and serious, contact the police. The police aren’t likely to be much help to you as far as investigating who sent an anonymous threat online, but they can help you create formal documentation in case the situation escalates, which will be valuable if you need a restraining order or other form of legal safeguard.
- If you work in a newsroom, pass on news of any threats to your office or building’s security personnel.
- Make sure to let loved ones know that you have been threatened. It’s not a pleasant topic of conversation, but they should know about it because it could affect their lives.
• Don’t blame yourself for threats of violence you receive online. Nothing you do deserves this treatment.

• Contact the Committee to Protect Journalists (CPJ) if you feel unsafe. They have resources to help. This is especially true for freelancers, who may not have the security and legal infrastructure of a newsroom behind them: https://cpj.org/emergency-response/how-to-get-help.php

• PEN America published an Online Harassment Field Manual for writers who are targeted with abuse. It is thorough and provides a lot of useful advice.

**Log Off**

Being online from morning until night is not healthy for your mental state, and people who report on the internet may not realize the degree to which the job is negatively affecting their happiness. Periodically, make sure that you’re spending time without any devices in your hand. Go to the gym without your phone, for example, or take a walk without it. The time away from your screen will likely help to improve your work, as well as they way you feel.
Appendix: Tools

The following is a brief overview of the tools mentioned in this guide. The guide is meant to be an introduction, and a discussion of process. As a result, there are many more tools listed on websites like Michael Bazzell’s Inteltechniques.com, Bellingcat, and in Craig Silverman’s Verification Handbook. I recommend that you check those out and try them at your own pace.

Search Engines

- Google.com
- DuckDuckGo.com
- Yandex.com
- Bing.com
- Baidu.com
- Monstercrawler.com
- Boardreader.com

Google Operators

- Site:[website address]
- Cache:[website/page name]
- Inurl:[Keyword]
- AROUND(X)
- inposttitle:[Keyword]
- Advanced Search
- List of all Google Operators

Facebook Search Sites:

- Findmyfbid.com
- Commentpicker.com
- SmallSEOtools.com
- Stalkscan.com
- Stalkface.com

Facebook Search Operators:
• From the template: https://www.facebook.com/search/[FACEBOOK ID NUMBER]/[Operator]
  • /photos-by/
  • /photos-of/
  • /photos-uploaded/
  • /photos-tagged/
  • /photos-in/
  • /photos-keyword/
  • /photos-liked/
  • /photos-commented/
  • /photos-interested/
  • /photos-interacted/
  • /photos-recommended-for/
  • /recent-photos/
  • /photos-liked/
  • /photos-commented/
  • /stories-liked/
  • /stories-publishers/
  • /stories-media-tagged/
  • /stories-in/
  • /stories-tagged/
  • /stories-topic/
  • /stories-news/
  • /stories-recent/
  • /reshare-stories-by/
  • /videos-of/
  • Facebook String Search Template

Maps
  • Google Street View
  • Google Earth
  • Snapchat Map
  • Wikimapia
  • Yandex Maps

Reverse Image Search:
  • Google Image
- Bing
- Jeffrey’s Image Metadata Viewer
- Tin Eye
- Yandex

Other Social Search Tools:

- Twitter Advanced Search
- TweetDeck.com
- TweeterID
- Snapchat Map
- Discord Search
- Webstagram (Instagram Search)
- YouTube Dataviewer

Archiving:

- https://archive.org/web/
- http://archive.today/
- http://archive.is/
- http://archive.li/
- Hunchly

Security and Miscellaneous:

- Proton VPN
- Tor Guard
- Nord VPN
- LastPass
- Tor Browser
- VirtualBox
- MySudo
- 33mail
- LexisNexis Opt-Out
- Delete Me
- CPJ: Help
- Online Harassment Field Manual
- WolframAlpha (historical weather)
Conversation Apps:

- Wire
- Signal
- Telegram
- Keybase
- Threema
- Whatsapp
- Skype
- Kik