

# Investigative Digital Ethnography: Methods for Environmental Modeling

Brian Friedberg

Version 1.0

Last updated: October 20, 2020



# Table of Contents

<b>Introduction</b>	<b>3</b>
<b>Research Practices and Perspective</b>	<b>3</b>
Actor-Focused	5
Information-Focused	6
Impact-Focused	6
<b>Elements of Investigative Digital Ethnography</b>	<b>7</b>
<b>Investigative Ethnography Stages</b>	<b>10</b>
1. Identify Topic and Artifacts	10
2. Identify Media Ecosystem and Influencers	12
3. Create Monitoring Environment	13
4. Develop Monitoring Strategy	14
5. Audit Key Assumptions	15
6. Archive and Analyze Findings	16
<b>Technical Setup</b>	<b>17</b>
Account Creation	17
Platforms	18

# Introduction

Investigative Digital Ethnography pairs journalistic and forensic techniques with ethnographic observation in a practice useful for academic researchers, policy makers, and the press. It is a particularly useful method for understanding disinformation campaigns. By taking the long-form approach of an investigation, this method may follow breaking news, or be used to analyze a specific case after the immediate event is over. The researcher is ideally tracking one central topic or case and may discover additional components as the investigation progresses. At some point, the gathering of information must end, and the ethnographer must move on to analysis.

The ethnographic method situates people in spaces marked by distinct rituals, beliefs, and cultural production. An ethnographer engages with the subjects to varying degrees, and in the case of digital ethnography, with the traces they leave behind. Observing online communities properly takes time, and the ethnographic process requires a commitment to observation during breaking news events and also during the downtime in between.

This investigative ethnographic method merges the pointed search for specific information that defines journalistic and legal investigation, with the long-term observation that defines ethnography. While an individual investigation may lead to one output in the form of an article, a long-term ethnography composed of many investigations can reveal valuable hidden details that may not have been significant to a single investigation. Investigative Digital Ethnography is an investment in future content and a means to study social media platforms even in the face of increased restrictions for researchers and journalists.

## Research Practices and Perspective

Researching and combatting the spread of disinformation is challenging, in large part because of the few barriers disinformers encounter in conducting operations, and the affordances offered by new communication technologies and platforms that spring up on what seems to be a daily basis. While certainly not a new phenomenon, disinformation has been exacerbated by networked communication technologies, specifically media consumption patterns and the rise of social media platforms. Furthermore, disinformation is a global issue found in both established democracies and authoritarian or illiberal regimes. When we talk of the study of disinformation, we are referring to multiple approaches to a sociotechnical phenomenon, particularly the dynamic back and forth between humans and the technology they engage with, and how this discursive relationship challenges social institutions.

Various terms and concepts have been used to describe problematic information, such as “fake news,” “disinformation,” “misinformation,” or “propaganda.”<sup>1</sup> Recent years have seen an increase of reports from journalists, research organizations, think tanks, and government agencies on the role of disinformation and media manipulation more broadly. Communication scholarship, Science and Technology Studies, Political Science, Computer Science and Sociology are a few of the formal disciplines that identify the types of information, how they spread, and, in some cases, measure their social impact.

This research may be quantitative or qualitative, with both types adding unique insights into how information systems become susceptible to media manipulation campaigns. Quantitative approaches measure disinformation at scale, using survey research to gauge people’s attitudes and exposure to disinformation, or data science to ascertain how widely false information spreads online. Qualitative approaches to the study of disinformation dissects the content itself, identifying common themes in media artifacts, how they are made, and the communities that benefit from their amplification.

Both approaches have their specific strengths and weaknesses. Quantitative approaches are necessary for empirical studies analyzing disinformation at scale, and are often utilized by social, political, and data scientists. These approaches can illustrate how social media users share or interact with information, estimate the aggregate impacts of disinformation, as well as capturing trends. This sort of empirical evidence can and should be repeatable, and the collection of it has allowed tech companies to make critical interventions in the circulation of disinformation on their platforms. However, quantitative work cannot alone account for the unique cultural and ideological positions held by communities in which disinformation rapidly spreads, and particularly those who reliably produce such disinformation. It is here that qualitative research, driven by small case studies examining the content of disinformation materials, can use social theory to help us understand the motivations of disinformation agents and media manipulators, particularly those adept at leveraging the features of major social media platforms and turning them into sociotechnical vulnerabilities. These qualitative investigations can reveal common techniques used in the creation of persuasive disinformation materials, the ideological motivations of the communities producing them, and suggest techniques to counter disinformation messaging. Qualitative techniques are often limited by scale and repeatability and can be marred by subjective “anecdotal” (cases interpreted through personal experience rather than systematic collection and analysis).

As platform companies like Twitter and Facebook are rolling back API access for research and failing to meet the needs of social and computer scientists struggling to study the impact of disinformation, new techniques for data collection and analysis are becoming common. An application program interface (API) is a set of routines, protocols, and tools for building software applications and collecting analytic information from these services. While APIs are essential tools for software engineers and those engaged in commercial services online, the limitation of

---

<sup>1</sup> Caroline Jack, “Lexicon of Lies,” *Data & Society*, Data & Society Research Institute, August 9, 2017, <https://datasociety.net/library/lexicon-of-lies/>.

API access makes scientific analysis of social media data more difficult to collect and process. API access for researchers and advertisers at one time was far more open and are increasingly being rolled back to restrict free access. Additionally, the microtargeting of social media users by domestic and foreign political interests, such as the Cambridge Analytica scandal, has made platform transparency and data harvesting a contested international regulatory issue.<sup>2</sup> Today, platform companies are notoriously hesitant to grant access to researchers examining their systems. Instead, they may provide differential access to researchers in situations that are in the interest of platform companies. What this means is that it is now harder to critically study the role in which major platforms play in the distribution of disinformation. Lastly, collecting and storing large amounts of quantitative social media data is taxing on time and resources, leading to a need for new innovative techniques to define and measure discrete events and cases.

There are three different approaches to using investigative ethnography to understand misinformation: **Actor-Focused, Information-Focused, and Impact-Focused.**

## Actor-Focused

We will begin by examining the actor-focused method of studying misinformation and disinformation, which is primarily employed by qualitative social scientists, journalists, and activists. Here, you identify the individuals and communities that generate and disseminate media manipulation campaigns. After identifying the artifacts of disinformation of interest, you follow these objects across the media ecosystem by using investigative techniques, leading back to the first or most significant instances of circulation. By ascertaining *who* is creating and amplifying the disinformation in question, you can begin to understand *why* they are doing it.

Once the actors have been identified, you gather evidence of their tactics and collate it in a spreadsheet. Taking screenshots and downloading video content is crucial for the integrity of your dataset. Often automated or off-the-shelf analytics software will not store your data. In the case of mass takedowns, researchers will lose access to data, so programs like Hunchly can be useful for protecting the integrity of your dataset.

When manipulators breach platforms' terms of service policies, they may have their accounts removed, and certain artifacts may be banned from the platform. A clear example of actor-focused investigation is how journalists and activists deplatformed *Infowars* and Alex Jones.<sup>3</sup> After significant critical attention from journalists (and activists) to Jones' significant role in spreading science and political disinformation, particularly his claims that the Sandy Hook school shooting was a conspiracy propagated by "crisis actors," major platforms reached a

---

<sup>2</sup> Nicholas Confessore, "Cambridge Analytica and Facebook: The Scandal and the Fallout So Far," *The New York Times*, April 4, 2018, <https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html>.

<sup>3</sup> Jack Nicas, "Alex Jones Said Bans Would Strengthen Him. He Was Wrong," *The New York Times*, September 4, 2018, <https://www.nytimes.com/2018/09/04/technology/alex-jones-infowars-bans-traffic.html>.

consensus to remove Jones and his brand from their services. This significantly decreased his reach and mitigated the impact his for-profit content could have on the information ecosystem. While quantitative techniques may help bolster these kinds of investigations, the Jones example underscores that the work is primarily qualitative in nature. Journalists and activists established a pattern of dangerous behavior by Jones and a rationale for deplatforming that platforms relied on to make decisions.

## Information-Focused

Another method we will examine is information-focused research, wherein quantitative and qualitative methods are used to determine the spread of media artifacts on specific platforms. Information-focused investigations often begin when the researcher identifies a particularly unique piece of media, be it a fake story, imposter account, or a viral rumor. Measuring the spread of information requires you to systematically track the link sharing, gather publicly available metadata and analytic information, and analyze them in comparison to the spread of other kinds of genuine information.

While information-focused techniques do not account for motivation, they do help show the scale of a false story or conspiracy theory, which can aid qualitative researchers in knowing where to look for impact in the wider society. This data can be used in policy recommendations, both for government regulation or platform-specific standards.

An example of information-focused studies are the works of Harvard's Berkman Klein center, particularly the report entitled *Partisanship, Propaganda, and Disinformation: Online Media and the 2016 U.S. Presidential Election*.<sup>4</sup> By identifying how information moves in the partisan fringes online, researchers identified the 100 media sources that were in heaviest circulation on platforms during the 2016 election. Of these, seven highly partisan sources received substantially more attention on social media than links from other media outlets. These sites do not necessarily all engage in misleading or false reporting, but they are clearly highly partisan and motivated by ideological points of view. Based on the recommendations in studies like these, major platform companies may take significant actions to change search results to hide or downplay false or misleading information on their services and promote verified content from reliable sources, or in some cases remove certain malicious content all together.

## Impact-Focused

A final method of disinformation studies is impact-focused, wherein researchers identify groups that are particularly vulnerable to mis- and disinformation and examine the integrity of their

---

<sup>4</sup> Yoichi Benkler, Robert Faris, and Hal Roberts, *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics* (Oxford, England: Oxford University Press, 2018), <https://oxford.universitypressscholarship.com/view/10.1093/oso/9780190923624.001.0001/oso-9780190923624>.

media ecosystems. This method is useful for directing harm reduction and counter messaging. Impact-focused research relies on both qualitative and quantitative research methods. An impact-focused approach is particularly useful for studying medical disinformation, especially how certain communities may be exposed repeatedly to spurious evidence designed to spread harmful information about health and wellness.

A few well-known examples of impact-focused research include work into anti-vaccination extremists targeting autism support communities to seed misinformation about the inaccurate and disproven linkage between vaccines and autism-spectrum disorders, and the study of unproven holistic remedies for chronic illnesses spreading in communities providing support to those living with cancer.<sup>5</sup> While it varies platform to platform, researchers using impact-focused methods can assess how misinformation spreads across different targeted groups, as well as analyze subsequent changes in opinion and behavior when scientifically accurate information is introduced to those groups to counter the misinformation.

## Elements of Investigative Digital Ethnography

Regardless of the focus chosen for an investigation, our focus here is online communities and the media artifacts that pass through them. An investigator's research question should dictate the methods employed as they trace the provenance and impact of an object and describe the sociotechnical conditions of the environments it was produced and distributed in. We suggest keeping these key questions in mind: where did an artifact come from, where did it spread, and why was it produced?

While we present three ways of studying disinformation, it's important to bear in mind that these methods can be messy. Often, as you pursue data to answer your research questions, you may switch from one approach to another, or borrow from parts of a different technique. That is OK. Investigating actors, information, and impact requires you to systematically approach your subject using methods that are both reliable and rigorous. A hallmark of any ethnographic approach is taking on a position simulating the online experience of a community or subject of inquiry.

---

<sup>5</sup> Julia Carrie Wong, "How Facebook and YouTube Help Spread Anti-Vaxxer Propaganda," *The Guardian*, February 1, 2019, <http://www.theguardian.com/media/2019/feb/01/facebook-youtube-anti-vaccination-misinformation-social-media>; Abby Ohlheiser, "They Turn to Facebook and YouTube to Find a Cure for Cancer — and Get Sucked into a World of Bogus Medicine," *The Washington Post*, June 25, 2019, [https://www.washingtonpost.com/lifestyle/style/they-turn-to-facebook-and-youtube-to-find-a-cure-for-cancer--and-get-sucked-into-a-world-of-bogus-medicine/2019/06/25/6df3ddae-7cdc-11e9-a5b3-34f3edf1351e\\_story.html](https://www.washingtonpost.com/lifestyle/style/they-turn-to-facebook-and-youtube-to-find-a-cure-for-cancer--and-get-sucked-into-a-world-of-bogus-medicine/2019/06/25/6df3ddae-7cdc-11e9-a5b3-34f3edf1351e_story.html).

We ground our ethnographic practice within the disciplines of anthropology, sociology, and communication studies. To gain a foundation in the formal practice of online anthropological research, we recommend the collection *Digital Ethnography: Principles and Practice* (2015).<sup>6</sup> Formative to the investigative digital ethnographic method is Biella Coleman’s work on the hacktivist collective Anonymous, in which she establishes research and representational techniques to document spaces where privacy and pseudonymity are prized.<sup>7</sup> Joan Donovan’s ethnographic work on the Occupy movement provides a grounding in the communication infrastructures of social movements.<sup>8</sup> Alice E. Marwick and Becca Lewis’ research details how the economics and infrastructure of social media has allowed motivated actors to manipulate popular conversation and normalize extremist positions.<sup>9</sup>

Other pertinent ethnographic examples include Jonathan Corpus Ong’s research, which highlights the need to understand social identities and moral justifications in troll industries, research which draws from media production studies and its tradition of talking with workers in creative industries. Crystal Abidin’s work on online influencers --celebrity figures who drive popular conversation on given topics or communities -- illustrates the global phenomenon of distributed cultural production and its social impact.<sup>10</sup> Sahana Udupa employs an ethnographic approach to describe political trolling cultures and social movements, exploring the social and economic incentives that drive people to do unsavory and abusive online work in India.<sup>11</sup>

Online investigations into manipulation campaigns often lead to spaces where violence is celebrated, or crime is orchestrated, and extreme prejudice normalized. In Jessie Daniel’s research on early white supremacist online organizing, she identifies these motivated actors as “innovation opportunists” who exploit emergent communication technologies to spread racist propaganda.<sup>12</sup> Teo Ballvé’s offline ethnography in violent conflict zones, and harm prevention therein, is also informed by the techniques and outcome-oriented frames of investigative

---

<sup>6</sup> Sarah Pink, Heather Horst, John Postill, Larissa Hjorth, Tania Lewis, and Jo Tacchi, *Digital Ethnography: Principles and Practice* (Los Angeles: SAGE Publications, 2015).

<sup>7</sup> E. Gabriella Coleman, *Hacker, Hoaxer, Whistleblower, Spy: The Many Faces of Anonymous* (Verso, 2014).

<sup>8</sup> Joan Donovan, “Toward a Militant Ethnography of Infrastructure: Cybercartographies of Order, Scale, and Scope across the Occupy Movement.” *Journal of Contemporary Ethnography* 48, no. 4 (August 2019): 482–509, <https://doi.org/10.1177/0891241618792311>.

<sup>9</sup> Alice E. Marwick, *Status Update: Celebrity, Publicity, and Branding in the Social Media Age* (New Haven: Yale University Press, 2013), <https://www.jstor.org/stable/j.ctt5vkzxr>; Becca Lewis and Alice E. Marwick, “Media Manipulation and Disinformation Online,” *Data & Society*, Data & Society Research Institute, May 15, 2017, <https://datasociety.net/library/media-manipulation-and-disinfo-online>; Becca Lewis, “Alternative Influence,” *Data & Society*, Data & Society Research Institute, September 18, 2018, <https://datasociety.net/library/alternative-influence/>.

<sup>10</sup> Crystal Abidin, “Internet Celebrity: Understanding Fame Online,” *European Journal of Communication* 33, no. 6 (December 2018): 696–97, <https://journals.sagepub.com/doi/abs/10.1177/0267323118814646a>.

<sup>11</sup> Sahana Udupa, “Extreme Speech| Nationalism in the Digital Age: Fun as a Metapractice of Extreme Speech,” *International Journal of Communication* 13 (0): 22.

<sup>12</sup> Jessie Daniels, “The Algorithmic Rise of the ‘Alt-Right,’” *Contexts* 17, no. 1 (February 2018): 60–65, <https://doi.org/10.1177/1536504218766547>.



journalism.<sup>13</sup> We also refer to the online investigative techniques and guides of *Bellingcat*, a journalistic collective at the forefront of open-source investigations, as well as the methods of the Southern Poverty Law Center’s Michael Edison Hayden.<sup>14</sup>

To begin the Investigate Ethnographic process, whichever method or perspective you choose, you must first choose an entry point into the information ecosystem that is propagating a disinformation artifact. Start with simple questions: What are the research questions driving the investigation? Are you investigating a bad actor, an impacted group, or a trending topic? What platforms are relevant for your investigation? Who is driving these conversations? While the scope of the investigation may grow over time, the initial point of entry should focus around why a disinformation artifact matters, to whom, and where it is spreading.

After answering these initial questions and identifying the best entry point into an information ecosystem, your monitoring can commence. The monitoring process involves the creation of new accounts for tracking, algorithmic-assisted snowball sampling to establish accounts of interest, and establishing the multi-platform scope of communities in question. By establishing who is of interest, you can then trace topics and actors across platforms through the sharing of keywords, memes, and slogans associated with a community, establishing the scope of the monitoring environment. With properly maintained monitoring accounts, algorithmic suggestions of other topics or accounts to follow become an invaluable discovery tool. Let the algorithms do the work for you. At this point, you should decide on a monitoring strategy, and set a regular schedule for checking in on the observed environment.

After you create a monitoring environment and strategy, document your investigation. Field notes should include maps of individuals, topics, and non-human actors involved in a topic of inquiry, lists of important accounts or URLs in a database, and short descriptions of important events in this ecosystem so that changes may be tracked over time. A well-maintained database of accounts, URLs, hashtags, and keywords may then be prepared for follow-up quantitative analysis. The more comprehensive the ethnography, the more robust and accurate any quantitative analysis will be in the future.

This investigative ethnographic process requires time and deep observation. It involves becoming familiar with a community’s internal slang, code words, memes, and other cultural practices, as well as what media they consume, and who they interact with. Determining what is regular versus what is irregular behavior for a community is critical to identify which manipulation campaigns are important to track. While popular press and critical reporting may help inform the early stages of setting up an investigation, as the researcher you must spend your own time verifying and expanding on the previous work of journalists. To gain additional insight into the often purposefully obtuse online spaces in which disinformation campaigns form,

---

<sup>13</sup> “Guides,” Bellingcat, accessed October 16, 2020, <https://www.bellingcat.com/category/resources/how-tos/>.

<sup>14</sup> Michael Edison Hayden, “A Guide to Open Source Intelligence (OSINT),” *Columbia Journalism Review*, June 7, 2019, [https://www.cjr.org/tow\\_center\\_reports/guide-to-osint-and-hostile-communities.php/](https://www.cjr.org/tow_center_reports/guide-to-osint-and-hostile-communities.php/).

you must become accustomed to their manner of behaving and speaking to cut through the noise.

Often, groups of actors come together in campaigns who may not be aligned with each other normally. These groups, who we refer to as networked factions, may put aside preexisting conflict in order to collaborate on a campaign. By paying special attention to the lead influencers of these factions, and who they regularly come into conflict with, you can help identify what these groups are motivated by, and how they define themselves against each other. Deep understanding of where one group ends and another begins is essential for establishing provenance of narratives and disinformation artifacts. Below, we identify six unique stages of the Investigative Digital Ethnographic method, with detailed steps to identifying pertinent information, establishing a monitoring environment, and next steps for data collection.

## Investigative Ethnography Stages

- 1 - Identify Topic or Artifact
- 2 - Identify Media Ecosystem & Influencers
- 3 - Create Monitoring Environment
- 4 - Develop Monitoring Strategy
- 5 - Audit Key Assumptions
- 6 - Archive and Analyze Findings

### 1. Identify Topic and Artifacts

To begin an investigation, one must choose a topic and artifact to investigate. Some examples are below:

**Topic:** Vaccine disinformation

**Artifact:** An infographic linking vaccines to autism

**Topic:** Child-trafficking disinformation

**Artifact:** A viral hashtag used to spread false information about a child-trafficking case

**Topic:** Election disinformation

**Artifact:** A viral social media post spreading incorrect voter information

**Topic:** Active-shooter misidentification

**Artifact:** A URL from a junk news site shared during a breaking news event

**Topic:** Extremist propaganda

**Artifact:** A meme linked to hate groups being shared on social media

These examples are meant to link a general research question to a specific object of investigation. By first clarifying what is, and is not, of interest to an investigation, your scope can be narrowed to help create an efficient monitoring environment. As the lifecycle of these campaigns rise and fall rapidly, an investigator, with proper preparation, may be equipped to track these events in real time, or, establish provenance for a case study after the fact. Often, topics or artifacts of interest for investigation will emerge as wedge issues around controversial social topics and debates during election cycles, or unverified rumors that circulate during breaking news events.

Wedge issues are divisive political issues, especially ones that are raised by a candidate for public office in hopes of attracting or alienating an opponent's supporters. During an election cycle, wedge issues will be prominent topics of press and online conversation. The larger framing of these debated topics happens in partisan press coverage and through posts by online influencers. Wedge issues will usually split along partisan lines, and will be presented as binary positions – for or against. Political candidates running for office will amplify these wedges in popular discourse, in interviews, and on social media.

In popular press, journalists of any perceived bias will highlight these issues and provide additional context – social, legal, political or cultural – reinforcing the wedges with supporting or detracting arguments. Partisan and hyperpartisan media will likely display clear political orientations in their editorial agendas, and have less defined barriers between objective reporting and opinion. These clearer biases help show political division and fault lines along wedges. Partisan political media — be it print, television, radio, digital, or otherwise — represents a myriad of ideological communities and interest groups, and is the next place to look for extended conversations around particular wedge issues. Within this coverage, both in mainstream media and alternative publications, pundits and influencers of note will be identifiable. The identification of key topics, publications and influencers are the foundation of a monitoring environment.

Breaking news events are another opportunity for identifying topics and artifacts. Before social media, breaking news was interjected as news bulletins or updates into broadcast and cable news programming, keeping viewers up to date with important details on a developing story. With the advent of social media, and the mass communication opportunities afforded by these technologies, Twitter, Facebook and other social media platforms become more than disseminators of journalistically vetted information. These platforms are used in real time by law enforcement, journalists, and other parties interested in establishing facts and seeking to shape public perception.

During breaking news events, hashtags will emerge where social media users share verified and unverified information, express personal opinions, and often spread disinformation artifacts. These artifacts may be shared as hoaxes, or strategic attempts to misguide law enforcement

and journalists. Often, disinformation artifacts or false statements will quickly circulate on social media before any official statements of fact from pertinent social institutions. In this time between a story breaking and facts being established, manipulation campaigns may be workshopped and seeded. Many contemporary cases of disinformation stem from confusion during these time periods.

**Action Item: Select a topic, based upon a wedge issue or breaking news events, and identify 1-3 artifacts for further investigation.**

## 2. Identify Media Ecosystem and Influencers

To gain deep insight into a manipulation campaign, you must first identify where it is occurring, and who is amplifying it. Media ecosystems, and the influencers who drive conversation in these environments, are key to compiling evidence of a campaign's origins and impact. A small account posting a false piece of information with no uptake on a website with little traffic is insufficient to impact the mass media environment. However, the strategic amplification of a disinformation artifact by an influential social media user on a large platform, later reported on by a hyperpartisan news site, is often sufficient to rapidly scale a campaign in a specific media ecosystem.

Media ecosystems are complex combinations of print, broadcast, digital, and social media that work together to create a self-referential information environment. Partisan and hyperpartisan media will be more likely to have a clear political orientation in editorial agenda, and have less defined barriers between objective reporting and opinion. Sites with low editorial control are usually key in developing alternative, conspiratorial, or disinformation narratives. Partisan political media, be it print, television, radio, digital or otherwise, represents a myriad of ideological communities and interest groups, and are useful for determining the affiliations of those who may be involved in a manipulation campaign.

Additionally, social media platforms themselves have specific affordances that make them useful for those propagating a disinformation campaign. The provenance of narratives or disinformation artifacts often have their first appearance on smaller websites or social media platforms before going viral online or garnering coverage in mass media. For example, a campaign participant may use Twitter in an attempt to bait journalists in taking an incorrect lead, and may use 4chan to organize collaborators. It is essential to find out what sites are of use for a particular topic, how each is used differently by participants and influencers, and the chronology of an artifact's appearance to establish provenance.

No disinformation artifact has power without social relevance granted by influencers. Influencers are visible pundits, journalists, or public figures who use media ecosystems as distribution networks. Influencers may be verified social media users, have large audiences, and their own personal websites. Some may use video streaming services, like YouTube, Periscope or Instagram Live to directly interact with their audiences and disseminate information during

breaking news events. For researching a disinformation campaign, influencers involved may be initially identified by their authorship of a relevant viral tweet, critical coverage from popular press, or popularity in partisan media. Once identified, these initially observed influencers help create the initial following list for new accounts used as the foundation for the creation of a monitoring environment.

**Action Item: Identify publications, social media platforms and influencers to defined media ecosystem**

### 3. Create Monitoring Environment

The establishment of an effective monitoring environment requires the creation of new social media accounts. Merely checking from your own account poses several problems. As social media platforms mine the data and behavior patterns of their users, search results are often customized to the user profile created during normal internet activity. To truly model the environment of a faction involved in a disinformation campaign, or that of a group targeted by disinformation, you must break out of your individual filter bubble to see the scope of an otherwise foreign environment.

First, you should create new social media accounts and browsing environments. Next, you want to “season” those accounts, i.e. influence recommendation algorithms to present relevant information. You do this by engaging in platform affordances, such as viewing individual posts or videos, liking or reposting them, and following recommended accounts. Using the influencers of note identified in stage 2, monitoring accounts should follow those influencers, and subsequently, other topics or individuals mentioned or reposted by this initial seed list.

Additional accounts of interest can be identified in three ways. First, look at who an influencer in question is following. Second, identify what key phrases or hashtags this influencer is using, and use in-platform search functionality to discover other individuals who are engaged in that conversation. Third, suggestion algorithms will offer new accounts to follow based on a user’s perceived interest, and will suggest other influential figures driving related conversations — in effect, algorithmic-assisted snowball sampling. By spending time reading what these accounts are posting, you will identify previously unforeseen topics or artifacts worthy of further investigation. Ultimately, these environmental modeling processes should replicate that of an individual who may be part of the community or faction under investigation.

The news these influencers share is also critical to modeling an environment. Some factions may primarily share mainstream news sources, often with additional commentary or framing. Others may share hyperpartisan sites with low editorial oversight. Finally, small publications, with little or no exposure outside these factions, are often shared internally and contain even lower-standard content than hyperpartisan news sites. Much of the news these communities share may not even be publications, but rather YouTube content. Identifying influencers who

have significant YouTube channels, and who those influencers are sharing, is essential to mapping out the information ecosystem of communities that do not value mainstream reporting.

Other than identifying influencers and news sources of import, there are three additional steps you should take to gain deep insight into the media intake of a community of interest. First, sign up for mailing lists. Many influencers and hyperpartisan publications maintain mailing lists that highlight narratives around wedge issues and breaking news events. Second, investigate podcasts or other audio streaming programs these influencers publish or appear as guests on. Finally, identify forums or other communication platforms where individuals may be communicating or coordinating.

This process will result in better results after several months of seasoning and training the recommendation algorithm. After an account has been well established, it will serve up content without the researcher having to do intensive manual searching. The recommendation algorithms will continue to surface users, hashtags, and URLs that you had not initially identified in Stage 1.

**Action Items: Create social media accounts to use for monitoring, and identify publications and multimedia content of interest.**

## 4. Develop Monitoring Strategy

Once you establish a monitoring environment for your research, you should develop a strategy for investigation with careful consideration to time. Often, disinformation artifacts, and accounts that spread them, are removed for Terms of Service violation upon discovery. This creates a challenge for researchers attempting to establish definitive provenance for narratives or artifacts. For provenance research of past events, this may be less pressing than for tracking information during a breaking news event or long-term ethnographic research. Past events may not be able to be adequately mapped, due to account or artifact deletion. In this case, press and other investigations may help fill in information gaps. For active or long-term investigations, there are four factors to consider in developing a monitoring strategy.

First, create a check-in schedule based upon needs. Are you willing to commit time weekly, daily, or hourly? For investigations that track campaigns in real time, timely monitoring is necessary as is a regular schedule of checking on monitoring accounts. For more long-term and immersive ethnographic investigations, the monitoring environment may supplement or replace an individual's regular media intake. The degree to which the monitoring environment replaces a normal environment depends upon a researcher's needs, the toxicity of the monitoring environment, and the research support available.

Second, decide if this is a solo or group investigation. While this may be accomplished alone, working in a group setting may yield more comprehensive results, and allows researchers to

take shifts when tracking campaigns in real time. Group work also helps offset the more toxic elements of certain online communities, as it allows for peer support.

Third, choose your monitoring device: Will you be using a personal computer, phone, tablet or virtual machine? For toxic or extremist communities, mobile devices and tablets are not recommended, as your browsing history may begin to influence your device's user profile, which informs all manner of advertising tech.

Fourth, set your boundaries. How deep do you want to go? Public web, closed groups, or private messaging? For closed online environments, some interaction with moderators may be required to gain access to closed groups or forums. Any interaction opens up potential vulnerabilities for a researcher, and may require you to self-identify upon entry to satisfy institutional review or journalistic ethics standards. For long-term investigations, the possibility of parasocial relationships with subjects may develop. These must be critically examined as they arise, as these one-sided relationships with influencers and content creators may cloud a researcher's judgement or lead to unwanted empathic experiences. Identify how far is too far for your own needs.

**Action Items: Develop a strategy and schedule for monitoring, based upon research needs and personal boundaries.**

## 5. Audit Key Assumptions

After a monitoring schedule has been established, and knowledge of that community increases, audit the key assumptions you entered with. Revisit your original research questions and topics of interest to see how they compare with current findings and ongoing discovery. You may reassess partisan designations you previously applied to publications or influencers based upon new observations or instances of social change. New wedge issues may emerge as topics of conversation within observed communities, and breaking news events may provide opportunities for campaigns, previously nonexistent, to emerge. If your initial assumptions and research questions do not hold up to changes in the environment, amend them to match the new observations in the field.

To best audit assumptions in either short-term or long-term investigations, compare your findings with reporting and other researchers in the field. Do your observations hold up against journalists or other researchers? Are your initial research questions still relevant based upon the assertions of experts? Are you learning anything you can't already find in the popular press? If your findings are continually lagging behind critical reporting of a particular topic or community, revisit your media ecosystem and monitoring environments to add new topics and influencers that may have been missing.

**Action Items: Compare the findings of your monitoring environment to those of journalists and other researchers, and adjust as needed.**

## 6. Archive and Analyze Findings

Once the integrity of a monitoring environment is audited, begin to record and transcribe relevant findings. First, archive social media posts, images, and URLs so they may be used for later research. This may involve using an archiving service like archive.is or perma.cc, saving files natively on a personal computer, and/or taking screenshots. As disinformation and accounts associated with it are regularly removed for violating Terms of Service, building a body of evidence requires utilizing archival tools for capture, and is useful for surfacing evidence archived by other researchers or individuals once a campaign has concluded.

These findings should be organized. Create a database containing URLs of accounts of interest or online publishing. This database should be coded according to your needs. Some investigation tools, like Hunchly, allows researchers to record their web history to be easily exported into other database formats. Entry-level tools like Excel, Google Sheets or Airtable may be sufficient for most researchers' needs. Databases of findings should be designed to suit the project's needs, though there are several consistent categories that should be included:

- Influencers driving online conversation
- Hashtags, memes and slogans used by the community
- Popular and alternative press on a topic

Once archived and stored, the collected data should be used to create a relational map of associations between topics, artifacts, and human and non-human actors. These maps not only help document the ethnographic observations of the research process, but may yield additional insights that may form the foundation of future investigations. These maps may be done in word-processing software, or in more specialized mapping tools like Xminds.

These organized findings may then be translated into analysis, be it a narrative case study tracing the provenance of a disinformation artifact, observations of the organized behavior of the participants in a disinformation campaign, or mapping the influence networks that helped a viral rumor spread online. These findings may form the foundation of analysis in multiple academic or journalistic disciplines, and the accounts created to form the modeling environment will remain useful for future investigations. Coded databases of URLs and social media accounts may later be prepared for quantitative analysis so long as the data is stored in a CSV-compatible format.

Finally, the publications of findings from an ethnographic investigation must be decided based on the urgency of the topic. For example, when researching extremist communities, researchers may uncover time-sensitive information of hostile actions against a vulnerable or marginalized group. Evidence of such campaigns, or evidence of doxing or harassment, should be properly shared with the affected individuals, journalists, social media platforms, or law enforcement to reduce harm when possible. While data being collected may be destined for an eventual article, peer-reviewed journal, or public report, immediate evidence of harm should not be withheld from those targeted by harassment or manipulation campaigns.



**Action Items: Create a database and association map that forms the foundation of academic or journalistic analysis. Decide appropriate outputs based on harm reduction when applicable.**

# Technical Setup

## Account Creation

### 1. Create a new email account for registration

*Recommended:* Gmail, Protonmail, TutaNota

### 2. Choose a browser for research

*Recommended:* Chrome, Firefox, Brave

*For consideration:* TOR

*Note:* Ad-block and privacy plugins may be used for security, but they will disrupt the advertisements that may be of interest for an investigator modeling an otherwise foreign information ecosystem.

### 3. Create new phone number for registration

*Recommended:* Google Voice, Burner App, prepaid cell phone

### 4. Create new social accounts

*Recommended:* Facebook, Twitter, TikTok, YouTube, Reddit, Pinterest

*For consideration:* Snapchat, Discord, Telegram, WhatsApp, Mailing lists and forums

*Note:* Facebook is challenging to use for this research. *Slowly* like pages and join groups. This is important. If you move too quickly your account may be locked for appearing to be automated or otherwise inauthentic. Like, post and interact with subjects at your own discretion if appropriate for investigation. For some researchers, using their personal Facebook may be necessary, but not advisable for sensitive materials.

### 5. Curate recommendations

*Recommended:* Follow identified influencers and recommended social media accounts. Subscribe to related YouTube channels and subsequent recommended content. To season

your account and accurately model the experience of a user whose community you are investigating, let the recommendation algorithms do some work for you.  
*For consideration:* Like or repost related content to influence recommendation algorithms.

## Platforms

### 1. Search engines

[Google](#)

[YouTube](#)

[Yandex](#)

[Baidu](#)

[Bing](#)

[DuckDuckGo](#)

[Tineye](#)

[4plebs](#)

[Redditsearch](#)

### 2. Social media

[Facebook](#)

[Twitter](#)

[Instagram](#)

[Reddit](#)

[LinkedIn](#)

[Pinterest](#)

[Mastodon](#)

[Marco Polo](#)

[TikTok](#)

[VK](#)

[Tumblr](#)

[Parler](#)

[Nextdoor](#)

### 3. Chat and messaging

[Discord](#)

[Telegram](#)

[Snapchat](#)

[Slack](#)

[WhatsApp](#)

[Skype](#)

[Zoom](#)

[Kik](#)

[Wickr](#)

[Facebook Messenger](#)

[WeChat](#)

[Google Voice](#)

### 4. Blogs, self-Publishing, and Q&A

[Wordpress](#)

[Substack](#)

[Medium](#)

[Blogger](#)

[Quora](#)

[Stack Exchange](#)

[Disqus](#)

[Pastebin](#)

[Yahoo Answers](#)

[CuriousCat](#)

[Ask.fm](#)

### 5. Streaming platforms

[Facebook Live](#)

[Periscope](#)

[YouTube Live](#)

[Dlive](#)

[Streamlabs](#)

[Twitch](#)

### 6. Fundraising platforms

[Patreon](#)

[Gofundme](#)

[Kickstarter](#)

[Indiegogo](#)

[Subscribestar](#)

[Gumroad](#)

## 7. Alt-tech and the fringe web

[4chan](#)

[8kun](#)

[Minds](#)

[Bitchute](#)

[Kiwifarms](#)

[Entropy](#)

[Gab](#)

## 8. Paid archiving and research tools

[Hootsuite](#)

[Netlytic](#)

[Buzzsumo](#)

[Brandwatch](#)

[Lexus-Nexus](#)

[Critical Mention](#)

[Hunchley](#)

[Perma.cc](#)

[Xmind](#)

## 10. Free archiving and research tools

[Crowdtangle](#)

[Zotero](#)

[Wayback Machine](#)

[Archive.is](#)

[Social Mention](#)

[Tweetbeaver](#)

[Board Reader](#)

[Tweetdeck](#)

[Google Alerts](#)

[Mediacloud](#)

[Google Trends](#)

## 11. Database and notetaking tools (free and paid)

[Airtable](#)

[Google Sheets](#)

[Microsoft Excel](#)

[Evernote](#)

[Microsoft One Note](#)

[Google Keep](#)