The building blocks of political disinformation campaigns

By Nick Biasini, Kendall McKay and Matt Valites

Talos
Cisco Security Research

TABLE OF CONTENTS

INTRODUCTION

The 2016 U.S. presidential election was a watershed moment for modern disinformation campaigns. Russia's desire to influence the election and efforts to undermine the American democratic process have since become a model for how a sophisticated, well-funded state actor can carry out an effective disinformation operation to achieve considerable gains with lasting effects. The tools and tactics used in such campaigns are as important as the mission's goals. Understanding this infrastructure – the major players, content creation, delivery mechanisms and more – is key to countering the ongoing global threat of disinformation.

As Cisco Talos discovered in "What to expect when you're electing," our four-year investigation into election security, securing elections is an extremely difficult, complex task. In this report, we continue our research into election security by focusing on the infrastructure behind disinformation campaigns to better understand and identify early signs of foreign influence and deceptive content. We discuss the types of actors and organizations involved, the content they create, and the tools actors use to share their messaging.

To help illustrate these infrastructure components and their importance, we delve deeper into several topics that are often overlooked by security researchers. Using a brief case study, we look at how actors exploit like-minded audiences in Facebook groups to quickly increase their reach and enlist large, unwitting audiences to promote their narratives. We cover social media platforms' special treatment of politicians – as outlined in their policy documents – that allow them to post content that would otherwise meet the criteria for removal under normal circumstances. We also share Talos' own experience of being blocked by security controls when attempting to promote election-related content on certain platforms.

Lastly, in our "Outlook" section, we make several key judgments about the disinformation threat landscape, including up-and-coming threats, how adversaries will attempt to avoid detection, and assessments on actor behavior for major players like Russia and China.

INFRASTRUCTURE

For the purposes of this blog, infrastructure involves the systems, tools, personnel, and technology required to conduct a disinformation campaign. This can include, but is not limited to, hardware, software, services and human capital. One of the starkest findings of this research is the low barrier to entry. It is relatively easy to leverage open-source tools and social media platforms to start a disinformation campaign. The success of such a campaign, however, can be affected by a number of different factors, many of which we identify in this report. Let's look at the components of a disinformation campaign, beginning with the actors.

PLAYERS

Who are the organizations and people involved in disinformation campaigns and what roles do they play in building or using the infrastructure? Figure 1 shows the relationships between the various entities involved, which are explained in greater detail below.

The building blocks of political disinformation campaigns

PROVOCATEUR

Behind every disinformation effort is a provocateur. This actor is the campaign's central figure, responsible for establishing the campaign's strategic goals and organizing its execution. As it relates to election security, state actors like Russia are typically the provocateurs. They leverage a range of resources to carry out their campaigns, including private and state-linked companies, social media platforms, and their own intelligence services.

Provocateurs often use third-party agencies to provide or supplement in-house disinformation services. These agencies come in two varieties: independent or state-linked. Independent entities are legitimate private digital marketing companies that engage in global influence operations. The Tunisian company UReputation attempted to influence elections in North Africa over the past year. British Columbia digital marketing firm AggregateIQ attempted to influence the Brexit vote. Israeli firm Archimedes Group targeted the 2019 Nigerian presidential elections and Newave in the United Arab Emirates (UAE) disseminated anti-Qatar and anti-Muslim Brotherhood narratives. In rare cases, there are known direct links between digital marketing companies and politicians or state actors. Most often though, while companies' messaging aligns with the political goals of a government or party, direct links between these marketing firms and state governments or individuals are often difficult to find.

As they identify and decommission fraudulent assets associated with coordinated inauthentic behavior, Facebook publishes information on whom is behind the abuse of their platform giving insight to the size of the disinformation campaigns. Figure 2 shows that ad spend, platform footprints and follower reach vary widely for external agencies and an Iranian state broadcaster.

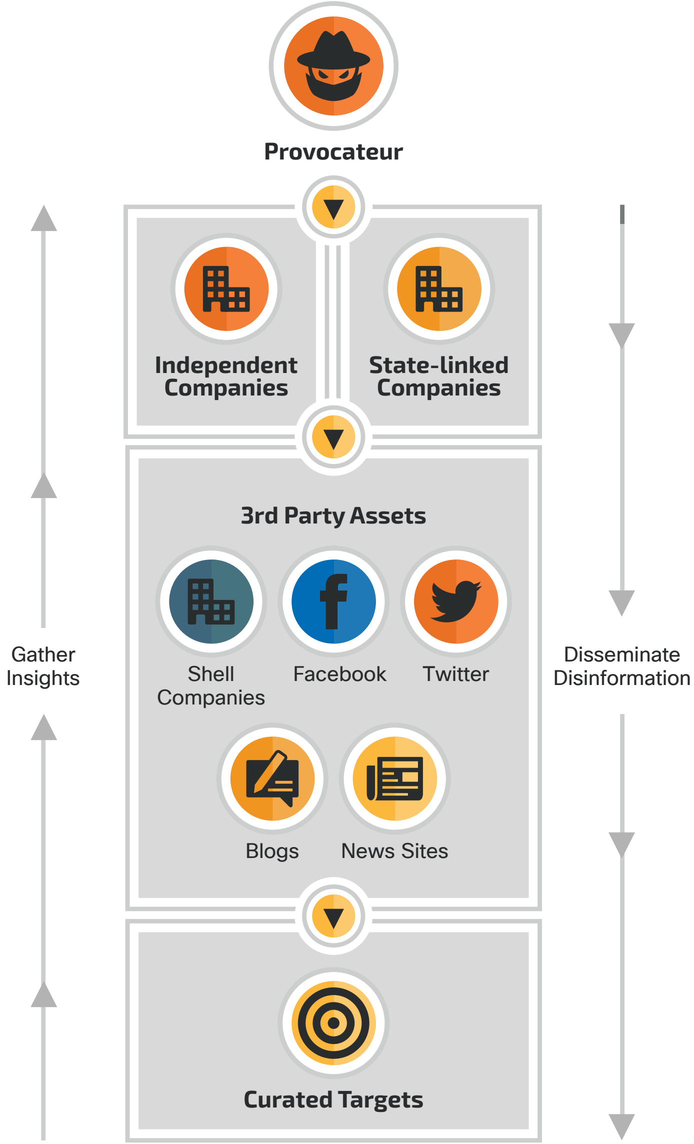
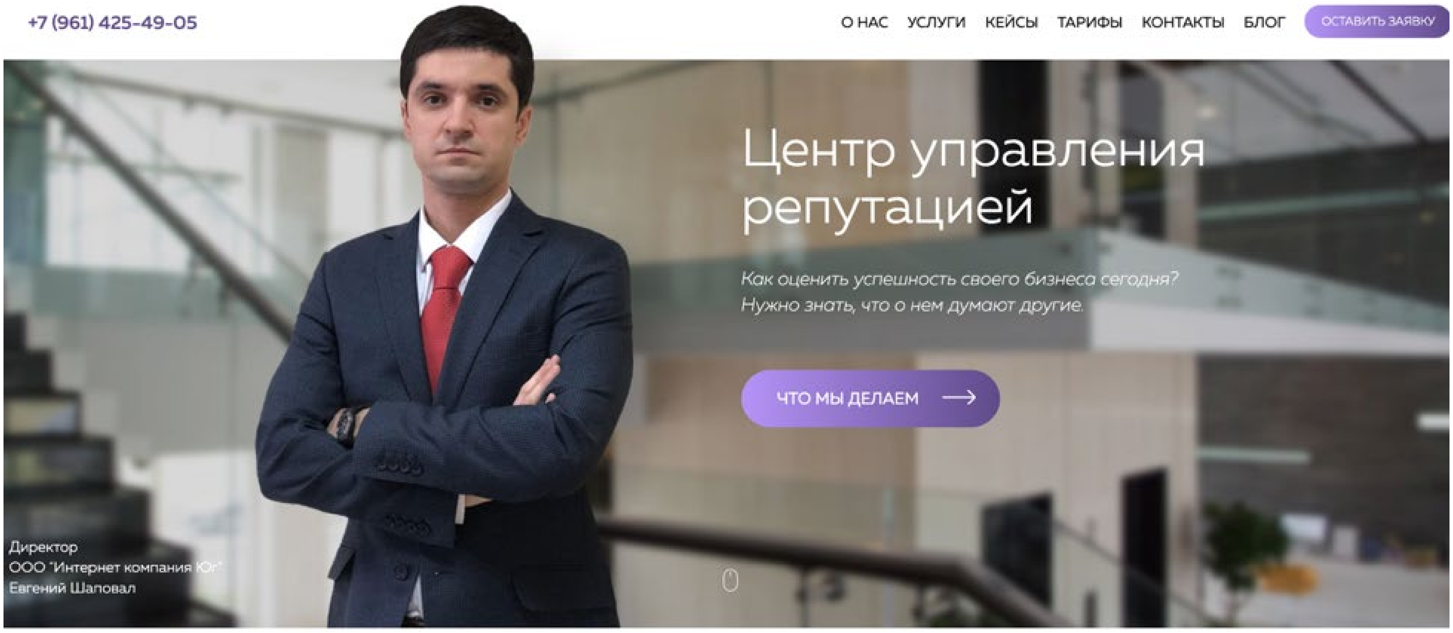


Figure 1: General evolution of a disinformation campaign.

Organization	Country of Operations	Ad Spend (in USD)	Social Assets	Followers
UReputation	Tunisia	$331,000	837	3,800,000
Archimedes Group	Israel	$812,000	256	2,800,000
New Waves / NewWave	Egypt / UAE	$167,000	387	13,700,000
Islamic Republic of Iran Broadcasting Group	Iran	$1,600	540	512,000

Figure 2: A breakdown of publicly known disinformation campaigns.



Центр управления репутацией

Как оценить успешность своего бизнеса сегодня? Нужно знать, что о нем думают другие.

Директор
ООО "Интернет компания Юг"
Евгений Шаповал

Figure 3: A screenshot of a now-defunct website belonging to Reputation Management Center.

The second type of agency that performs disinformation services are those with direct ties to well-funded governments. The most widely publicized example of this is the Russian Internet Research Agency (IRA), a state-sponsored private company that spearheaded Moscow's disinformation operation to influence the 2016 U.S. Presidential election. Another Russian example is a newer group used in a campaign dubbed "Secondary Infektion." These entities overcame disinformation problems such as account creation and content dissemination with wildly different solutions, some of which we'll discuss shortly.

Leaked documents give us an idea of the various budgetary aspects of disinformation campaigns. Data shows that by the summer of 2016 the IRA was spending $1.25 million a month targeting Americans with social media messaging. Some of this money paid for their staff. Junior analysts made $1,100 a month, bloggers $1,200, and senior management $4,200. Staffing numbers vary, with estimates ranging from 400 to 1000 employees. Despite its small size, the IRA created the illusion of a massive group of supporters by creating and managing countless social media accounts.

SHELL COMPANIES

Many of the agencies, including private commercial organizations and especially those that are state-linked, employ networks of shell companies either to provide some cover for the handling of financial transactions, establish a reputation, or both. An example of the latter is Concord Management and Consulting LLC, a Russian company listed in Robert Mueller's Indictment of 13 Russian companies for – among other things – conspiracy to defraud the U.S. as part of the 2016 presidential election. Concord, using a series of additional shell companies – some sharing the same mailing address as the IRA – was the primary funding source for the IRA's disinformation campaigns. Dzheykhun Aslanoz, a Russian national accused of having led the IRA's operations targeting the 2016 U.S. elections, is also listed as the general director of a company called Azimut, one of the shell companies used by Concord for transferring funds to the IRA. Mueller's report highlights numerous other connections between IRA officers and various additional shell companies.

Not all shell companies are used to shield finances. Reputation Management Center, another Russian company cited in Mueller's indictment and owned by Aslanoz, claims on its now-defunct website, shown in Figure 3, it used "bots with a history" to "mimic live behavior" and form "a positive image of the company." Additionally, they provide services to “drown negative reviews in a sea of positive information” and to “write unique content that forms a positive image of the company.” We identified similar unsubtle language on other websites belonging to private commercial agencies performing disinformation, such as UReputation’s description of their cyber influence service which sends “targeted messages to specific categories of recipients to influence their perception of a brand or personality.”

Whether independent or state-linked, the modus operandi of both entities is similar in the assets they leverage to execute their campaigns. Most organizations use a combination of social media platforms, news sites – fake, legitimate and state-funded — blog platforms, and shell companies. Each asset works together to establish a reputation, generate content, disseminate disinformation, or shelter the funds that ultimately run these campaigns. The rest of this paper will discuss the ways and means these players use their assets.

CONTENT

One of the most important aspects of disinformation is the content one is trying to spread. This content can take on a variety of forms but most commonly is a social media post, blog post, or article. Mature agencies will often gather insights to tune their dissemination efforts. For instance, by collecting likes, friends, regions and even response times to posted content, agencies can better assess their targets and deliver custom content. In one instance, UReputation posted conflicting content for multiple sides of the 2019 Nigerian election (which they would later try to influence) possibly to gather data on supporters for future targeting. Intelligence gathering can also come from external sources. The NationBuilder software is one such tool that targets and manages communications with expectant voters. This software is known to have been used by agencies in political elections such as Todd Stone’s 2018 run to be leader of British Columbia’s Liberals.

While not all agencies – especially DIY-scale efforts – gather insights, all disseminate content. The goal is to make this content as visible as possible to influence as many opinions as possible, and there are multiple methods to seed the content.

Typically, content is posted in the form of “articles” that are supposed to look like traditional, more fact-based journalism. Sometimes these sites are fake and owned by the disinformation agencies, such as fakenewschecking[.]com, an ironically named site known to have posted biased information about candidates early in the 2019 Tunisian presidential election. Once published, this content is amplified by various social media posts. Other times, smaller “fake news” platforms are leveraged. These are typically low-profile web sites that purport to be against “fake news,” yet publish factually questionable information. Most of these sites can be created with simple servers hosting any one of a variety of Content Management Systems (CMS) such as WordPress, Drupal or Joomla. Most of these CMS are open-source and free to use. Any costs incurred are small when compared to the cost benefit the content provides. The campaigns that we analyzed all included these content sites in some form.

However, adversaries have begun changing their approach to leverage fake news sites. We now have evidence that some content is being created by fake personas and is successfully being pushed to high-profile, legitimate platforms. The Daily Beast published an article early in July 2020 that demonstrated how this behavior worked. The article highlighted a reporter that focused on Middle Eastern socio-politics who had articles published in numerous platforms, including some prominent ones. However, the reporter was not a real person. Instead, these detailed sock puppets included fake professional LinkedIn profiles, social media profiles, and contributions to multiple news sites. This was enough to establish credentials that would trick legitimate news sources into posting content by the personas. All told, there were more than 15 fake personas that published almost 100 opinion pieces on nearly 50 different platforms. The details show the extent to which the actors went to portray this disinformation.

When creating a fake persona on the internet, a crucial step to help establish legitimacy is to create a headshot that can be used for bios or profile pictures. These images can simply be stolen off the internet, but in recent years, it has become easier to identify such inauthentic behavior through tactics like reverse image searches. Adversaries have adapted, implementing slight changes to avoid detection. In the Tunisia campaign, for instance, the actors took the images found on the internet and modified them by mirroring the image or cropping the size. These simple tasks help the photos avoid forgery detection and are becoming common practice among adversaries.

Another way that actors, particularly state actors, disseminate disinformation is through state media. A common example of this is in Russia, where state-funded media outlets are mouthpieces for pro-Kremlin narratives. Sputnik and Russia Today (RT) are two such entities that have a huge global reach, operating in 100 countries and broadcasting in over 30 languages. Sputnik and RT are often where initial messages originate before they are amplified on other platforms. These types of national media stations have large budgets and significant global reach, making them highly effective tools for spreading the government's strategic narratives.

Russia is not the only country to leverage its state media to disseminate disinformation. In China, Chinese Central TV (CCTV), China Daily, and other media outlets operate under tight Chinese Communist Party (CCP) control. Beijing also pushes messages on social media that promote CCP viewpoints. A group of internet commentators known as the "50 centers" spreads pro-CCP messages on social media and publish fake news in content farms, among other activities. Iran also has several state-backed media entities pushing disinformation, including hardline outlets run by the Islamic Revolutionary Guards Corps (IRGC), Iran's elite military force, and English-language channels to appeal to sympathetic viewers in the West. Facebook has identified The Islamic Republic of Iran Broadcasting Corp. as being behind foreign interference against numerous countries. In the chart above, they achieved modest reach with a relatively small budget.

In addition to organized state-run media operations, recent campaigns detail a new adversary tactic for carrying out influence operations: compromising news sites. In a report published in July 2020, researchers uncovered evidence that actors may be actively working to compromise legitimate news websites to seed disinformation content. This includes replacing existing articles with new content or generating new articles. This is a dangerous escalation, as illegitimate content could now be published onto legitimate news websites directly. This again shows how these actors will always be working to try and get their content into the public's purview through any means necessary.

PLATFORMS

Once they've created the content, adversaries must deliver the information to users by making it available. Social media is the most popular form of disseminating this content, but there's many other options. Take the Secondary Infektion campaign. In this Russian campaign targeting Ukraine, the actors leveraged blog platforms and the online forum Reddit as methods of pushing content. This runs in stark contrast to the previous campaigns that heavily relied on social media for amplification. In Secondary Infektion, the Russian actors pushed fake "leaked documents" and other content onto these blog platforms with the hope of shaping opinion inside Ukraine. What is most notable about this campaign, however, is the lack of success when compared to some of the other campaigns run out of Russia. We'll share some of the reasons for this shortly.

DISSEMINATION AND AMPLIFICATION

Other platforms we commonly see abused are web pages and their associated platforms. This can include fake news sites or sites created by actors, or legitimate news sites posting content generated from these groups, like the Daily Beast article referenced above. Once the content is seeded on multiple platforms, the adversaries must then amplify that message by using social media and associated troll accounts.

The most effective way to amplify content is through the use of troll accounts. These accounts take several different forms in the various campaigns we have analyzed. In the case of the IRA, they used a hybrid approach of leveraging established accounts run by actual people and bot-based accounts that had all their content scripted. From a detection perspective, using manually operated accounts is more difficult to detect but will also require a significant human capital investment to create the account and establish a seemingly legitimate usage history. By contrast, automated or bot accounts are less expensive, produce content at a faster rate, and can be used in conjunction with many open-source projects but are easier to detect.

Assuming some of the actors have basic programming or scripting knowledge, they could also use the plethora of open-source libraries that facilitate interaction with social media platforms to create their own tools. Additionally, many of the software developers at these platforms have released their own set of libraries or tools to interact with the platform programmatically like the Facebook business Python SDK. Even so, like other social media platforms, the Facebook API does not let users automate account creation, requiring clever scripting using something like an automated WebDriver to create and verify accounts in an actual browser as opposed to programmatically via a service API.

There are three primary types of accounts these groups leverage: aged, amplifiers and paid/stolen. Aged accounts are those that have existed for longer periods of time and have a history of established activity that make them appear more legitimate. They are typically the most valuable type of account for disinformation actors, largely because they've so far evaded detection while gaining as many followers as possible. These accounts are also the hardest to come by since they require a large amount of planning. The next type of account is primarily concerned with amplification. Often, this is the primary function of bots. One of the hallmarks of bot accounts is a lack of historical content or a recent account creation. The goal of these accounts is to amplify the original content that is being posted by the aged accounts. Secondary Infektion eschewed the concept of aged accounts and instead implemented a high-operational security methodology of almost entirely single-use accounts. Without an activity history and established group of followers associated with their accounts, the disseminated information lacked the sound board on which subsequent amplification relies. As a result, Secondary Infektion's campaigns exhibited no measurable effect. The final type of account we've seen leveraged is paid/stolen accounts. These are typically already existing accounts with a follower count close to aged accounts, but their access has either been bought through a marketplace or they have directly stolen credentials to gain access. In some cases, an actor has stolen access to a bunch of accounts and then sells that access to a third party. Regardless of how the access is obtained, adversaries consider these types of accounts highly valuable to adversaries.

DETECTION

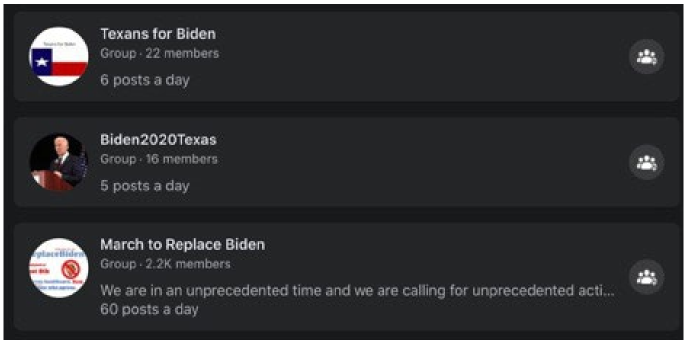
At a high level, the content is created by the actor and published to a variety of sites, after which aged, or paid/stolen, accounts start disseminating this content on various social media platforms. Finally, this content created by the aged accounts is spread further by the amplifiers whose primary purpose is to broadcast the message with the goal of it going viral and getting picked up by a larger audience. All of this is done by one group controlling all the levers behind the scenes. This allows for more coordinated behavior but is also something that most social media platforms are looking for and actively shutting down. Technology companies have worked on multiple responses to detect this type of behavior. One of the most common identifiers is around timing. For example, if an account responds to a tweet in a matter of seconds, that is likely inauthentic behavior. While people need time to read, process and respond, computers and more specifically bots, do not. Another way to identify malicious behavior is to take note of the way(s) in which the platforms are being accessed. It is exceedingly rare for a normal user to only connect to social media from a web browser or desktop application. If a user never connects from a cell phone, tablet, or other device, particularly without changing geographic locations periodically, it can be an indicator that the activity is possibly nefarious. There is a plethora of other techniques that can be used to identify bot behavior. [Our colleagues at Duo demonstrated some of those techniques in 2018.]()

While these measures help detect and block inauthentic content, bad actors are constantly updating their tactics to incorporate novel approaches to the problem.

A main problem with these types of detections is that they are largely reactionary, meaning that the controls have been implemented in response to something that has already happened. While these measures help detect and block inauthentic content, bad actors are constantly updating their tactics to incorporate novel approaches to the problem.

INFORMATION SILOS

Social media platforms foster information silos – places where users only interact with accounts, groups, or content exclusively aligned with their currently held beliefs. By using algorithms and machine learning, platforms like Facebook, Instagram and others suggest "friends" to connect with or groups to join based on similar interests or networks. By accessing groups of people with the same beliefs and interests, disinformation actors can quickly increase their reach and enlist large, unwitting audiences to promote their narratives.

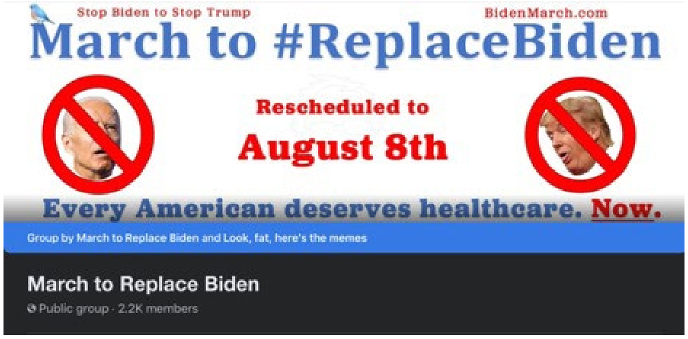


Texans for Biden
Group · 22 members
6 posts a day

Biden2020Texas
Group · 16 members
5 posts a day

March to Replace Biden
Group · 2.2K members
We are in an unprecedented time and we are calling for unprecedented acti...
60 posts a day

Figure 4: Examples of politically oriented Facebook Groups.



Stop Biden to Stop Trump
BidenMarch.com
March to #ReplaceBiden
Rescheduled to
August 8th
Every American deserves healthcare. Now.
Group by March to Replace Biden and Look, fat, here's the memes
March to Replace Biden
Public group · 2.2K members

Figure 5: A Facebook event hosted by a group dedicated to disinformation.

Facebook groups are already being used by disinformation actors in this way. There were plenty of examples exposed during the IRA investigation that demonstrated how Russian actors were organizing rallies and various other political events through these types of groups. In July 2020, [Facebook took down a series of accounts linked to political operative Roger Stone] in a recent example of how these groups can be abused.

It's easy to find groups of people with similar interests, and this is especially true when it comes to political ideology. Additionally, you can restrict these groups to specific cities in specific states. So, if you wanted to target supporters of a specific candidate/party/issue in a specific place, you can do so easily through these groups. Additionally, as was shown in the Roger Stone takedown, you can then leverage these memberships to post content from your "fake news" websites. Not only does this provide a starting point to test disinformation but it also provides a way to get the content in front of people that will then take it to other platforms, like posting about it on their personal Facebook pages or other social media sites. Detecting this type of behavior is increasingly difficult, especially for groups that are private and require approval to join.

To explore this issue in greater detail, we conducted our own case study. We started by searching for Texas-based Facebook groups for Democratic Presidential candidate Joe Biden, where we stumbled upon a group called "March to Replace Biden" (Figure 4).

We found the group's members had a clear preference for Sen. Bernie Sanders and a strong dislike for Biden and U.S. President Donald Trump, as well as content aimed at fostering distrust in the democratic process overall. The group was also organizing a public demonstration (Figure 5) and encouraging users to share disparaging memes and news articles, many of which were being tagged by Facebook as potential misinformation. We identified this group within minutes of starting our search, highlighting the ease at which a disinformation actor could do the same. Moreover, our search did not include private groups, which would significantly increase the number of results. For adversaries, Facebook groups are a quick, easy way to find people receptive to specific types of disinformation, especially in today's world of hyper partisanship and social media echo chambers.

TOOLING

Even before establishing a presence on various platforms and creating and disseminating disinformation, and certainly afterward, the actors need additional infrastructure in the form of hardware, software and scripts. Each area produces its own challenges and specialized tooling to help facilitate success. The sophistication or complexity of the tooling varies widely based on the campaign and the groups responsible. Let's start with the challenges presented with social media usage in disinformation campaigns in 2020.

SOCIAL MEDIA

One of the biggest challenges associated with using social media today is a requirement to associate a phone number with an account. Twitter, for instance, requires you to have a phone number and limits up to 10 accounts to be associated with a single phone number. Commercial software as a service telephony services may at first appear to be a viable means of easily procuring additional phone numbers, but in our experience, social media platforms detect and prevent accounts from using numbers associated with common services. As a result, actors must leverage actual SIM cards. The procurement and management of SIM cards, especially at scale, is a challenge for many of these campaigns.

In the IRA example from 2016, published reports say that twitter identified 2,752 bot accounts linked to IRA and another nearly 36K Russian bots in total being used in relation to the 2016 campaign. Doing some simple math, you would need more than 275 unique SIM cards to register and use those accounts, which is a staggering number. The cost and logistics required to acquire those SIM cards would require specialized hardware and a dedicated team. In a recent takedown of a suspected Russian bot farm in Ukraine, photos show a plethora of SIM-related hardware. Using OSINT, we identified multiple GSM SIM gateways, and estimate approximately $10,000 in hardware costs alone, just from the devices identified in the photos.

Another big hurdle associated with using this amount of SIM cards is actually obtaining the SIM card itself. In the U.S. and many countries in Europe, it is increasingly difficult to get SIM cards without providing legal ID and various other pieces of information. The idea of obtaining hundreds of them seems unlikely. This is another clear example of how state-sponsored and other well-funded groups will have an advantage when it comes to disinformation campaigns.

If an adversary could pass the logistical challenges around SIM cards and phone numbers, there's the challenge of creating, sending and amplifying their message. In this case, adversaries fall back on the open-source community. A simple internet search for things like Twitter bots will reveal a nearly never-ending list of open-source projects that are designed to allow users to build their own bot. Beyond that, as we discussed previously, there are huge amounts of libraries in scripting languages like Python that would allow someone with basic scripting skills to create their own specialized bot without much additional effort.

A simple internet search for things like Twitter bots will reveal a nearly never-ending list of open-source projects that are designed to allow users to build their own bot.

One of the key goals of disinformation is the tracking the effectiveness of the message they are pushing. This allows adversaries to push further down avenues that increase engagement and potentially abandon those that don't. Regardless, it requires tracking.

There were multiple ways these disinformation campaigns have operated over the years and some of it included purchasing ads on social media platforms. These ad buys provide the purchaser a wealth of information about what the engagement rates are with the various promotions and feedback on the groups they are reaching. As you can imagine, this type of data would be invaluable to actors crafting a disinformation campaign. However, that only covers the content that is paid. What about the unpaid organic data? We began looking at solutions and uncovered some interesting data related to Cambridge Analytica and the work they did in the 2016 campaigns and beyond. A leaked Github repository, used by a related company, AggregateIQ, walked through some of the tools they were leveraging. Based on the tools described in the report, we began searching for various open-source alternatives to see if someone could build this campaign without the heavy lifting associated with software development and testing.

For example, one of the tools listed in the publication was called “Peon” and was designed to ingest and utilize various data sources for tracking purposes. This allowed the customer to understand how effective a campaign or effort was through data. We began looking for, and quickly found, multiple open-source projects that cover similar capabilities. As an example, here is a project focused on similar tasks specifically related to Twitter data. The project includes detailed instructions on how to gather, ingest and analyze the data associated with various keywords and brands. These types of projects were available for virtually every tool uncovered in the publication, laying out the groundwork for what needs to be done to run an information campaign for a political candidate.

CAMPAIGN EFFICACY

This is far from the only tooling we found around these campaigns. Additional capabilities enabled actors to identify social media connections, account activity and engagement activity for disinformation content and advertisement telemetry. The tools also allowed comparisons with external data sources, such as voter records obtained from the previously mentioned NationBuilder service. There are some common capabilities shared between these tools. First, actors need a method of ingesting and organizing data. We commonly found that to be related to taking messages out of data queues, the most common implementations we saw revolved around Amazon Web Service (AWS) Simple Queue Service (SWS). For nearly every tool type we identified, there are either open-source equivalents or scripting libraries that could be extended to handle these types of tasks without undue burden.

AMPLIFIERS

Most of the tooling we observed was built around social media platforms, as that was the primary method of disseminating and amplifying content. One additional set of tools we commonly found were page loaders. Page loaders or page viewers are designed to increase traffic to web pages. We saw these types of tools used commonly in the IRA campaign in 2016. These tools are primarily pointed at the content that is generated to increase its popularity. This all feeds into the goal of getting the disinformation to be picked up by larger publications and amplified organically. These page loaders are relatively simple to create using basic scripting or open-source projects to generate traffic. Typically, it’s not enough to just use an off-the-shelf tool to generate clicks, as some platforms have improved detection of automated clicks. We found some groups bypassing these restrictions with VPNs and proxies.

CONNECTIVITY

Virtual Private Networks, or VPNs, are a common way for people to obfuscate the true origin of their traffic and additionally prevent nosy onlookers from sniffing the network traffic the user is generating. As you can imagine, the groups involved in these campaigns can use VPNs for a variety of reasons. There are endless services that offer VPN connections, but they can also easily be set up by leasing a server from a data center and leveraging something like the OpenVPN project to create VPN endpoints.

Proxies similarly provide a way for people to obfuscate their true origins or restrict access, creating a pivotal point for analysis. In the campaigns that we saw, adversaries leveraged proxies to hide where the traffic’s origins and increase page views. Similar to VPNs, there are a huge amount of open proxy services available on the internet or with a leased server and some software, akin to the process for creating VPN endpoints.

SOFT TOOLS

Finally, there were a couple of additional layers of tooling that we found in some of the very large campaigns like the IRA in 2016. Employees at this domestic state-funded company were subject to training on things like grammar and politology, which outlines the proper Russian point of view on current events. These types of tools and capabilities would likely only be reserved for the most highly funded and well-organized campaigns, as it shows a level of maturity that is unlikely in smaller, less organized campaigns.

POLITICIANS’ EXCEPTION

While social media companies have been taking steps to mitigate the threat of disinformation, all people aren’t necessarily treated equally in this space. A major loophole that exists on Facebook and Twitter is a policy exception for politicians that allow elected and government officials to post content that would otherwise meet the criteria for removal under normal circumstances.

If a user wants to publish any ads related to social issues, elections, or politics, a vetting and approval process is required. This process includes things like verifying your identity and enabling two-factor authentication. Once the vetting process is completed and disclaimer banners are placed identifying them as being paid ads, the process is complete. Facebook has come under increased scrutiny as of late due to the decision to not fact-check these posts, allowing politicians to potentially spread disinformation.

Likewise, Twitter has exceptions clearly stated in their terms of service related to politicians. Instead of providing a blanket ability for politicians to continue to advertise and post, they implemented a public-interest exception. The concept being that despite this content being potentially false, it is in the public interest to leave it public. These do come with some caveats, however. For instance, the account needs to be verified, have at least 100,000 followers, and currently hold or be actively pursuing an elected or appointed position. These approaches differ, but still provide an opportunity for politicians to skirt the same restrictions that average users need to abide. This creates potential problems, especially as a growing number of political candidates begin to embrace conspiracy theories and groups such as QAnon while seeking or sitting in office.

TALOS CASE STUDY

Advertising accounts are even more restrictive. Talos recently experienced some of the platform security controls while attempting to publish advertising material promoting our election security research. Because we wanted to promote content including voting, election security, and anti-disinformation, we were required to register as a Cause-Based Advertiser, one step beneath a Political-Based Advertiser. While Cause-Based Advertisers can still target things like users based on demographics or location, once registered, these entities are limited in their ability to track and serve ads based on keywords and cannot target users' followers – both known tactics of previous disinformation campaigns. As a last line of defense, platform users are empowered to escalate tweets for review by a team who will remove any ads that run afoul of policy. One stark difference, however, is that Twitter does not allow politicians to advertise directly on its platform.

OUTLOOK

Talos assesses that state actors will increasingly incorporate disinformation operations as part of their strategy to advance foreign policy and national security objectives. Based on our study of disinformation campaigns that have become public knowledge, actors carrying out such activities frequently achieve considerable gains with lasting effects. These operations are implemented relatively easily and with little or no consequence, increasing the appeal to other actors looking to do the same.

Furthermore, we assess that social media platforms will remain one of the most effective ways for actors to create and spread disinformation for the foreseeable future. Given the low cost and massive reach of social media disinformation campaigns, bad actors will almost certainly continue to see such platforms as a primary method for promoting their narratives. **Additionally, we foresee deepfake technology becoming an increasingly common and challenging problem in the months and years ahead.** Deepfake videos and photos depict fictitious narratives and imagery, portraying people saying or doing things that did not actually happen. They often look incredibly realistic, making it difficult for audiences to recognize that they are being duped. Deepfake content is already growing at a rapid rate, frequently targeting high-profile individuals like former U.S. President Barack Obama and Facebook CEO Mark Zuckerberg. This technology will likely become a dangerous and influential tool for disinformation actors.

Looking forward, we expect that threat actors will use many of the loopholes and tactics outlined in this report to avoid detection. This includes working with shell companies to obfuscate operations and exploiting Facebook groups to enlist unwitting people to amplify their messages. In particular, we assess that adversaries will increasingly use private social media groups while also leveraging the policy exceptions built into many platforms for elected officials. In practice, the latter would include threat actors masquerading as politicians to circumvent the rules against paid advertisements and content or co-opting a person or group, either wittingly or unwittingly, who currently holds such status. Alternatively, we could see an increase in legitimate politicians embracing disinformation narratives to appeal to certain voters and cast doubt on democratic institutions. This has been the case in several recent U.S. congressional races, in which candidates have promoted QAnon conspiracy theories.

Despite the seemingly bleak outlook, there are some encouraging developments that point to increased public awareness about disinformation. The relative failure of the Secondary Infektion campaign, discussed earlier in this report, is evidence of this. Most of the operation's fake stories never gained traction, with many being either ignored or mocked by forum users, highlighting users' increased awareness and skepticism about fake content. While there is still more to be done, social media companies are gradually implementing controls to identify and remove inauthentic behavior campaigns. These efforts have resulted in many disinformation campaigns being exposed and have allowed researchers the ability to retroactively assess their scope and outcome.

Leading up to and after the U.S. presidential election, major disinformation players like Russia will probably dedicate a substantial amount of resources to circumventing newly established practices and protections put in place by U.S. election officials, social media companies, and other related entities. Given the amount of focus on protecting the election from foreign interference, adversaries are likely going to pay close attention to how we implement and carry out changes in security practices. This means operations like intelligence gathering and reconnaissance will probably be prioritized. Additionally, as major social media platforms begin to implement certain levels of protections, such as verification or notification of a non-reputable source, adversaries may try to go around that by establishing fake news sites that target a smaller but more easily influenced audience. These can take the form of conspiracy theory sites and alternative news sites.

We also expect that threat actors will remain keen on recognizing the latest social media trends, such as new or up-and-coming platforms, so that they can quickly begin establishing a presence for influence. TikTok's quick ascendance as a globally popular social media site is evidence that this threat space continues to change, and adversaries will undoubtedly look to leverage the latest platforms to reach large audiences.

Russia's ability to carry out well-orchestrated, successful disinformation campaigns is encouraging other actors to adopt similar techniques. China, for example, has also been evolving their disinformation operations. Whereas previous Chinese information operations primarily focused on propagating a single narrative extolling the virtues of the CCP, Beijing now employs tactics that are more commonly associated with Russian threat actors. For instance, China has begun disseminating multiple conflicting narratives, which can increase distrust in targeted countries. Examples of this include COVID-19 disinformation in which Beijing put forth narratives that the virus began in a U.S. military lab, was found in Italy months before it appeared in China, or that it simply did not originate in Wuhan, China. We will explore this topic more in a forthcoming Cisco Security blog on the global political impact of COVID-19 disinformation.

Despite the seemingly bleak outlook, there are some encouraging developments that point to increased public awareness about disinformation.

For now, Beijing's information operations do not appear as effective or sophisticated as similar Russian operations. As the Australian Strategic Policy Institute writes, "while these efforts are sufficiently technically sophisticated to persist, they currently lack the linguistic and cultural refinement to drive engagement..." However, as it contends with a number of existential threats, including border disputes, increasing tensions with the U.S., COVID-19, and the accompanying economic fallout, the PRC will likely continue to evolve, emphasize, invest, and promote influence operations abroad and at home.

Our next election security report will explore the psychology of disinformation campaigns, including their domestic impact, the effect of certain behavioral and psychological factors, and information hygiene practices, such as how to identify false content. Some questions we'll attempt to answer include: Why do individuals fall for these campaigns? How do their social media-based trust chains impact the spread of false content? And what can they do to not fall victim to disinformation?

About Talos

The Talos Threat Intelligence Group is Cisco Security's threat intelligence organization, an elite group of security experts devoted to providing superior protection for our customers, products, and services – as well as a vast collection of open source security products and tools. Talos is among the largest threat research teams in the world, encompassing seven key areas: Community & Open Source, Detection Research, Engineering & Development, Incident Response, Intelligence and Interdiction, Outreach, and Vulnerability Research & Discovery.

Talos detects and correlates threats as they happen, pushing coverage to customers globally within minutes to protect against known and emerging cyber security threats. With great visibility comes great responsibility – Talos also supports open-source security and often undertakes interdiction efforts to mitigate threats in the wild that pose significant risk to the internet at-large.

For more information, visit www.talosIntelligence.com.