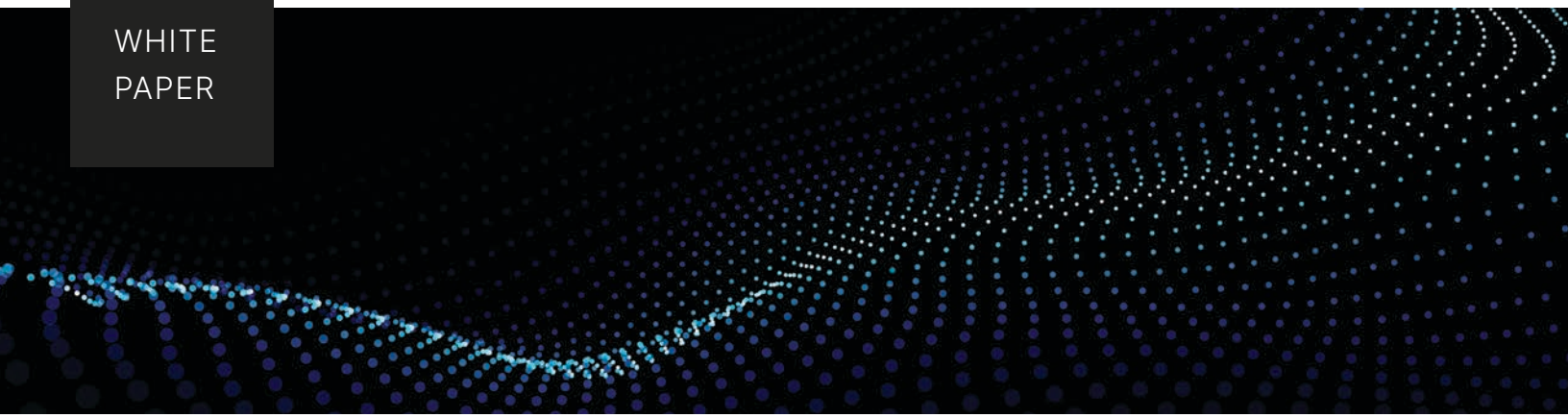
Recorded Future®

WHITE PAPER



The Diamond Model for Influence Operations Analysis

This paper presents a diamond model framework for the analysis of malign influence operations. It is intended for intelligence analysts and incident responders who seek to operationalize the analysis of influence operations (IO) indicators and tactics, techniques, and procedures (TTPs), as well as for researchers who strive for a comprehensive, organized, visual method to understanding malign influence. This framework draws from academic papers, open sources, cyber threat intelligence (CTI) practices, and collaboration within Recorded Future's Insikt Group. Special acknowledgments go to Team T5 and DoubleThink Lab, who have used the diamond model to analyze influence and psychological operations in previous reports and presentations. Additionally, Recorded Future acknowledges the many teams and individuals who have laid the foundation for influence operations analysis with their own frameworks, many of which are recognized in this paper. Special thanks to Sergio Caltagirone for his insightful review and contributions to the original Diamond Model of Intrusion Analysis, from which this framework is derived.

Executive Summary

The diamond model for influence operations analysis is a framework that leads analysts and researchers toward a comprehensive understanding of a malign influence campaign by addressing the socio-political, technical, and psychological aspects of the campaign. The diamond model for influence operations analysis consists of 5 components: 4 corners and a core element. The 4 corners are divided into 2 axes: influencer and audience on the socio-political axis, capabilities and infrastructure on the technical axis. Narrative makes up the core of the diamond.

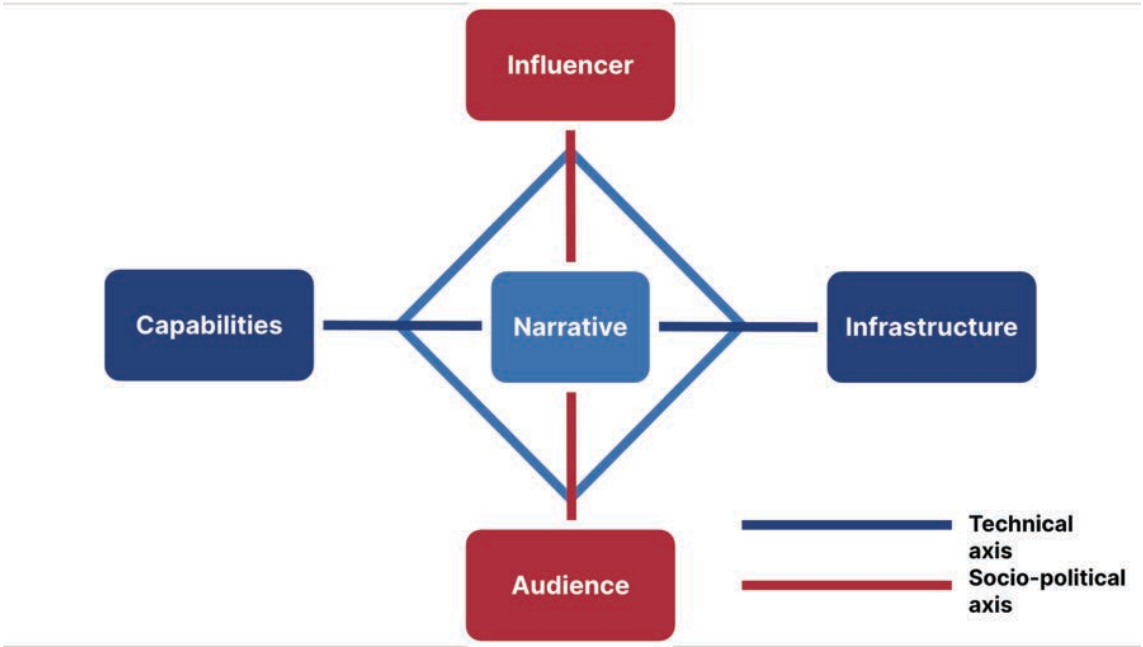


Figure 1: The Diamond Model for Influence Operations Analysis (Source: Recorded Future)

In this paper, we explore the individual components of the diamond model for influence operations analysis, addressing both the socio-political and technical axes of the diamond, the relevance of each axis, and the ways in which the components and axes interact with each other to reveal in-depth understanding of the influencer's objectives in a malign influence campaign.

The diamond model helps analysts advance their understanding of the various components of influence operations and the ways in which influencers change their tactics to appeal to different audiences with core elements of identity and storytelling. Recognizing that much work has already been done to define key concepts of influence operations, this report will not establish new definitions or a standard set of terminology to use in the analysis of influence campaigns; it will instead use established terminology and definitions from across the academic and research communities. Although it includes 2 case studies for purposes of practical application, this paper is not intended to be a comprehensive report on the IO landscape, nor is it intended to address the history or evolution of the IO environment.

Table of Contents

Executive Summary 1

Background 3

- Challenges in Analyzing Influence Operations 4

The Diamond Model for Influence Operations 5

- Socio-political Axis 6
- Technical Axis 8
- Core: Narrative 10

Operationalizing the Diamond Model for Influence Operations 10

- Practical Application 1: China and COVID-19 10
- Practical Application 2: Russian Influencers Posing as Patriots 14

Appendix A: Incorporating Existing Tradecraft 17

- Kill Chain Models 17
- Disinformation ABC (+D) 18
- AMITT (Adversarial Misinformation and Influence Tactics and Techniques) 20
- The 4 Ds of Propaganda 22

Background

Influence operations pose a difficult but vital analytical challenge in efforts to understand today's expanding and dynamic information environment. Several frameworks exist that address separate, isolated parts of influence campaigns, but none of them are holistic frameworks that offer a comprehensive view of an influence campaign. Although various sectors use different definitions for malign influence, alongside terms such as information operations, disinformation, misinformation, and coordinated inauthentic activity, this document defines influence operations as the coordinated, malign, or manipulative use of information and narrative storytelling to achieve a competitive advantage over an adversary or competitor and further one's own objectives.

Following definitions from the RAND Corporation,[1] influence operations is an umbrella term encompassing military information operations and both state-sponsored and non-state influence operations. Information operations is a subset of influence operations that originated in military environments. The US Department of Defense's (DoD) definition of information operations[2] encompasses 5 pillars of IO: computer network operations (CNO), psychological operations (PSYOP), electronic warfare (EW), operations security (OPSEC), and military deception (MILDEC).[3] Influence operations further include a variety of actors, tactics, infrastructure, and objectives, which will be outlined further in this paper.

Manipulating information, creating fear or doubt in the mind of an opponent, controlling public opinion and sentiment, and playing "mind games" are tactics as ancient as warfare itself.[4] However, the advent of the World Wide Web and the development of online media platforms have created an environment ripe for manipulation by state and non-state actors alike.[5] The Russian government's influence operations fall under the umbrella of active measures — the use of offensive political warfare tactics such as disinformation, propaganda, cyber-enabled attacks on digital infrastructure, deception, sabotage, destabilization, subversion, and espionage.[6] China's leadership also practices an unrestricted approach to asymmetric warfare involving alternatives to direct military confrontation, including international policymaking, influence operations, economic warfare, and cyber-enabled attacks on digital infrastructure and networks. While Russia and China carry out these activities in different ways, and for different objectives, they both target the vulnerabilities in the Western model of open, free, and unrestricted information.[7]

The governments of Russia and China have isolated their respective internet infrastructures and monopolized local media in ways that allow for authoritarian control over information, censorship, and surveillance within their own countries. Russia's "RUNet"[8] has been in development for years and is now being tested in response to sanctions for the war in Ukraine. China's "Great Firewall"[9] started in the 1990s and is now being used as a tool to tighten control over domestic narratives and to censor all content that the party-state finds objectionable.

1 Eric V. Larson, Richard E. Darilek, Daniel Gibran, Brian Nichiporuk, Amy Richardson, Lowell H. Schwartz, and Cathryn Quantic Thurston, Foundations of Effective Influence Operations: A Framework for Enhancing Army Capabilities. Santa Monica, CA: RAND Corporation, 2009. https://www.rand.org/pubs/monographs/MG654.html.
2 https://sgp.fas.org/crs/natsec/IF10771.pdf
3 https://jfsc.ndu.edu/Portals/72/Documents/JC2IOS/Additional_Reading/1C3-JP_3-13-4_MILDEC.pdf
4 https://www.thoughtco.com/psychological-warfare-definition-4151867
5 https://carnegieendowment.org/2020/06/10/challenges-of-countering-influence-operations-pub-82031
6 https://www.intelligence.senate.gov/sites/default/files/documents/os-kalexander-033017.pdf
7 https://www.europarl.europa.eu/RegData/etudes/STUD/2021/653635/EXPO_STU(2021)653635_EN.pdf
8 https://www.atlanticcouncil.org/in-depth-research-reports/issue-brief/reassessing-runet-russian-internet-isolation-and-implications-for-russian-cyber-behavior/
9 https://www.hrw.org/news/2020/09/01/china-great-firewall-changing-generation

Conversely, the democratic free world operates in an online environment that is, for the most part, open to the free exchange of information, and is thus also more vulnerable to manipulation.[10] And while the leadership in China and Russia have developed sophisticated capabilities in asymmetric warfare, they are not the only countries practicing malign influence operations. Malign influence is a global problem. In May 2021, Facebook (now Meta) reported that it had detected over 150 coordinated inauthentic behavior (CIB) networks conducting malign influence operations emanating from 50 countries between 2017 and 2021.[11] Those networks included both state and non-state actors seeking to manipulate global public discourse, disrupt political processes, manipulate markets and consumers, and incite conflict, both at home and abroad.

The core element of every influence operation is a story, or narrative. Narrative is vital to identifying the influencer's intentions and objectives. Humans are innately predisposed to understand information with assigned meaning, and storytelling is the most effective way to assign meaning to information.[12] The most successful influence campaigns use narratives with components of shared identity meant to appeal to a target audience. With identity and meaning, a story can change how people think and behave, which is likely the end goal or objective of an influence operation. Narrative warfare is the weaponization of a narrative and the battle over the meaning and identity of information.[13] This plays a central part in the diamond model for influence operations analysis, discussed further in the "Core: Narrative" section.

Challenges in Analyzing Influence Operations

As influence operations become more complex and the digital infrastructure used for these campaigns expands, it becomes more difficult to detect and analyze IO campaigns for purposes of intelligence gathering. Cyber threat analysts have a unique set of relevant skills and tools to help advance the analysis of influence operations because malign IO and cyber intrusion campaigns have several threads in common, including some overlapping actors, tactics, infrastructure, and audiences.

Threat actors' use of artificial intelligence, machine learning, deep fake videos and imagery, profile photos created by generative adversarial networks (GAN), sophisticated document forgeries, and inauthentic account registrations on a massive scale make detection and analysis of influence operations challenging. Defenders are challenged by several limitations: a lack of clarity on what poses a threat, limited detection and alerting technology to find key narratives and identify new campaigns, and a lack of funding for advanced tools or professionals who understand psychological operations. Influence operations are also deeply social and psychological battles for the minds of the audience, and researchers are not immune to these effects. The cybersecurity and intelligence sectors will benefit from integrating sociological and psychological research and frameworks into fusion reports and products aimed at countering malign influence.

Segregation and Hyperfocus

Most of the existing frameworks for analyzing influence operations, which will be addressed further in the "Existing Tradecraft" section, address a single element of an IO campaign:

- Kill chain models outline the process that the influencer uses to carry out coordinated campaigns.
- The AMITT and 4D models track granular TTPs and make it easier to identify and track tactics.
- The ABC+D method demonstrates the need to fuse intelligence on actors, behavior, content, and technical distribution for an integrative approach to remediation.

10 https://www.journalofdemocracy.org/articles/the-2016-u-s-election-can-democracy-survive-the-internet/
11 https://about.fb.com/wp-content/uploads/2021/05/IO-Threat-Report-May-20-2021.pdf
12 Ajit Maan. (2018). *Narrative Warfare*. Narrative Strategies Ink.
13 Ibid.

Each of these models, and many others that exist today, offer insightful guidance for analysts and researchers to more closely study influence operations, and each can be layered on top of the diamond model for a holistic view of the campaign.

However, the element of narrative warfare — influence through storytelling and shared identity — which is core to every successful influence campaign, is largely missing from these frameworks. The diamond model addresses that gap and places narrative at the center of the framework, serving as a vital component that deserves its own focused analysis.

The Diamond Model for Influence Operations

The diamond model for influence operations is a framework inspired by the Diamond Model of Intrusion Analysis, created by Caltagirone, Pendergast, and Betz for the US Department of Defense in 2013.[14] The DoD framework addresses 4 core elements of a cyber intrusion campaign and how they intersect: the adversary and the victim along the socio-political axis; and the infrastructure and the capabilities along the technical axis. This framework has been widely adopted across the cyber threat intelligence community and is familiar to many analysts who have served in tactical, operational, and strategic roles within the cybersecurity industry. The original Diamond Model of Intrusion Analysis was created with adaptation and evolution in mind, which inspired this new diamond model for influence operations: "The model is purposefully generic and thereby expandable and flexible. It accurately captures the essential concepts of intrusion analysis and adversary operations. These attributes enhance the model's utility, allowing it to grow and encompass new ideas and concepts."[15]

The diamond model presented in this report is adapted for the particulars of analyzing coordinated influence operations. Versions of this framework were used in a 2020 report by DoubleThink Lab for examining Chinese information operations,[16] and in a January 2022 presentation by TeamT5 on the analysis of video-based information operations.[17] DoubleThink Lab's uses an "extended diamond model" to highlight 2 primary components of the operation: the actors conducting the activity and their modus operandi. The center of the diamond is labeled "identity/saving face", which likely indicates the centrality of identity in that particular case study, but is not elaborated upon with respect to the diamond model.[18] TeamT5's presentation adapts the traditional Diamond Model for Intrusion Analysis to examine threat actors, victims, capabilities, and infrastructure, but does not address narrative.[19]

The aim of this paper's adapted diamond model for influence operations is to fuse together the essential components of an influence campaign, with a focus on the core of the diamond: the narrative element that aims to change the behaviors of the target audience. These elements tie together all other elements of a campaign. Some labels were changed to more accurately reflect the nuances of influence operations. On the socio-political axis, the term "adversary" is replaced with "influencer," and "victim" with "audience". The technical axis remains unchanged, with "capabilities" addressing TTPs and "infrastructure" addressing the various technical and physical infrastructure used for production and dissemination of influence materials. The key change in this diamond model for influence operations is the addition of a core (middle of the diamond) element of "narrative", which is a vital but often overlooked socio-psychological aspect of successful influence operations.

14 https://apps.dtic.mil/sti/pdfs/ADA586960.pdf
15 Caltagirone, Pendergast, and Betz. "The Diamond Model of Intrusion Analysis". Department of Defense. 2013. https://apps.dtic.mil/sti/pdfs/ADA586960.pdf
16 https://medium.com/doublethinklab/deafening-whispers-f9b1d773f6cd
17 https://youtu.be/I2gMDEYo2Bo
18 https://medium.com/doublethinklab/deafening-whispers-f9b1d773f6cd
19 https://teamt5.org/en/posts/sans-institute-cti-summit-2022/

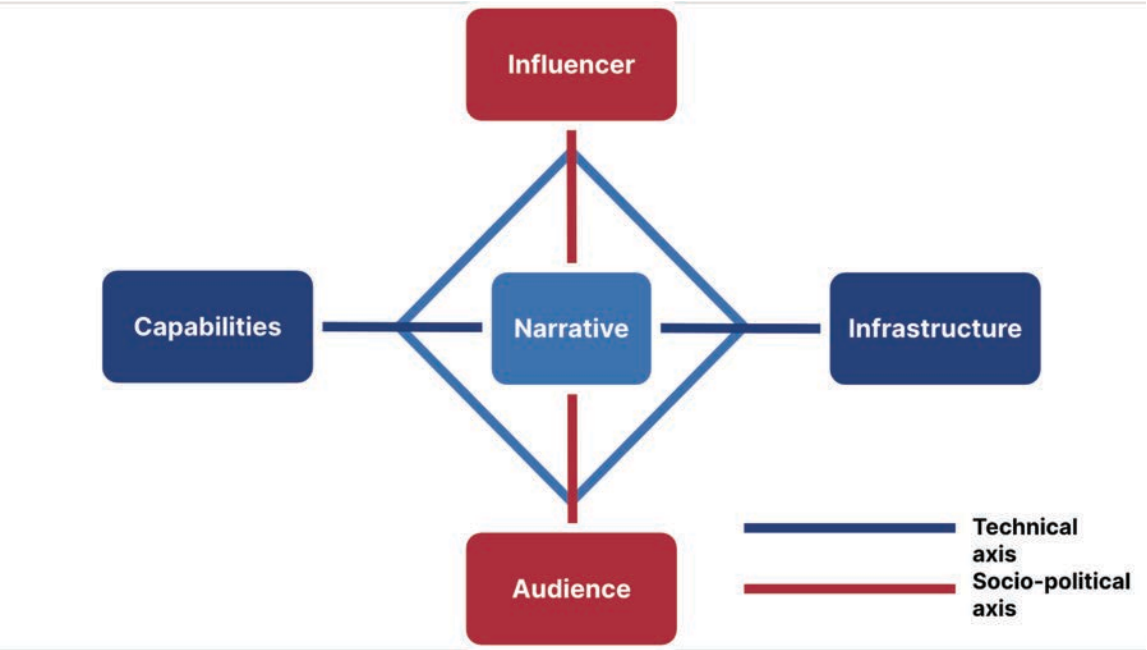


Figure 2: The Diamond Model for Influence Operations Analysis (Source: Recorded Future)

The diamond model can be used to keep track of all cumulative intelligence an analyst gathers on the influencer — however, for the practical applications presented in this paper, it's best to focus on a single campaign or operation. Once an analyst starts logging and tracking the influencer's capabilities and infrastructure, the collection will grow substantially and will assist in the analysis of further campaigns by the same influencers. Familiarity with the influencer will then allow analysts to grow their reporting beyond tactical applications and defenses.

Socio-political Axis

The socio-political axis (y-axis) consists of the malign influence actor, the influencer, at the top of the diamond, and the targeted audience at the bottom of the diamond.

Influencer

The influencer is an individual or organization that is conducting malign influence activity. As researchers have become keenly aware of coordinated overt and covert operations, some threat actors outsource influence activity[20] to marketing and public relations companies in the private sector, or to criminal actors who advertise their services in the underground. Because some of these third-party organizations may be unaware of the malign nature of the content they are hired to promote, it's preferable to call these actors influencers instead of adversaries. It is also important that analysts are open-minded and unbiased in their discovery of influence actors in order to differentiate between those that unknowingly share false information and those that do so on purpose. Most case studies would label the former group as an audience because they were first exposed to the disinformation by a malign actor. However, if it serves the analysis to track the amplification by those same victims, they can be added to the influencer section.

20 https://www.lawfareblog.com/outsourcing-disinformation

Malign influencers are actors that purposefully engage in sharing false information in order to create chaos or panic or to otherwise draw the attention of the audience. They range from overt agents like outspoken politicians to more covert, low-quality inauthentic accounts, called "trolls". Financially motivated criminals also knowingly engage in spreading false information, but for profit. There are many disinformation-as-a-service organizations that advertise their services in the criminal underground or on the dark web. In 2019, Recorded Future's Insikt Group documented this type of activity in a report called The Price of Influence: Disinformation in the Private Sector.

Audience

The audience is the intended target of the influence operation. The audience can range in size from a single individual to a large international audience. However, the most effective and resourced influence campaigns will likely target a specific demographic using a customized narrative. Audiences derive their information from a variety of print and digital sources today, including but not limited to TV, newspapers, social media, messenger applications, streaming video platforms, podcasts, radio, and online discussion boards. Influencers tailor their narratives for particular demographics in particular sources. Some influence operations target multiple audiences across multiple platforms in a single campaign.

Knowledge of geopolitics will assist analysts in understanding the strengths and vulnerabilities of various online and physical communities, and understanding a target audience's linguistics, cultural values, religious norms, social skills, and demographics helps analysts and influencers alike identify vulnerabilities in that audience. For example, the Communist Party of China (CCP) is clear in its intentions to learn the ways of young generations around the world in order to more fluently target them with pro-China narratives.

Influencer-Audience Relationship

The socio-political relationship between the influencer and audience is key to understanding the influencer's objective for the campaign and the audience's vulnerabilities to that particular influencer's narrative. The influencer is attempting to change the mindset, and ultimately the behavior, of the audience. Significant effort should be placed on understanding the audience, their belief system, political system, and social habits. Sophisticated influencers dedicate extensive resources to studying and investigating the target audience, and researchers should as well. Fluency in the audience's language, immersion in the audience's preferred media platforms, and knowledge of the audience's political climate are all helpful for understanding what inspires audience behaviors. Influencers might practice on the target audience for weeks or months to learn what defenses they will face and create new processes to work around them. The test period can also determine how receptive the audience is to the narrative and gauge which topics are most effective. Correspondingly, defenders benefit from monitoring changes in an actor's TTPs when defensive actions are taken.

Influencers as Threats

While much focus is placed on state-sponsored influence actors and their objectives, there are many malign influencers around the world that are aiming to achieve their own objectives. And because the volume of disinformation being spread across the internet can be overwhelming, analysts should understand their intelligence requirements and focus on the malign influencers that pose a threat to their organization, sector, or demographic. A threat exists when the influencer has the capability, opportunity, and intent to target their audience with malign content. While capability and intent are usually determined by the influencer, opportunity often depends on the audience's vulnerability to the influencers' TTPs. For example, societies that value open and free speech are inherently more vulnerable to malign influence than societies where autocrats closely monitor, censor, and restrict participation and content. If an influencer attempts to infiltrate the social media platforms of an autocratic country that has strict control of its Internet, defenders may be able to quickly detect the harmful narratives and identify the influencer, remove the account, and neutralize the malicious information. However, in more open environments where few restrictions exist, the malign influencer may go undetected for long periods of time, building a following and covertly blending in with the target audience.

Additionally, unstable societies and political systems are more vulnerable to malign foreign influence. This presents challenges for countries whose populations are deeply divided over issues such as politics, religion, war, or pandemic response. The free exchange of ideas produces representative political systems, but also makes those systems more vulnerable to actors who conduct malign influence operations. On the other hand, societies like China may be highly resistant to foreign influence because they have limited access to foreign media due to increasing restrictions and censorship. China's citizens may be particularly vulnerable to the CCP's propaganda because they live in a media echo chamber that prevents access to and discredits Western news sources.

Technical Axis

The technical axis (x-axis) consists of the **capabilities** (TTPs) on the left side of the diamond, and the **infrastructure** on the right side of the diamond. By identifying the capabilities and infrastructure used by the influencer and potential methods of defense, decision-makers are empowered to take action against malign influence operations. The capabilities and infrastructure corners of the diamond can be the most revealing intelligence gathered on the influencer. Both of these elements are likely to evolve over time, especially if target audiences, platform administrators, or regulators crack down on the activity, resulting in the influencer pivoting to new or unregulated infrastructure and undetected tactics. Well-resourced influencers will be persistent when they encounter obstacles. The smaller, less-resourced influencers will be easier to identify, will make mistakes more often, and will take significantly longer to overcome obstacles and set up new infrastructure.

Capabilities (TTPs)

Capabilities are the influencer's TTPs. Studying the way influencers plan, test, and execute their operations can enable analysts to be more proactive in defending against malign influence and to discern how to neutralize harmful narratives when they are identified. As discussed earlier in this report, there are already many frameworks established for the identification, collection, and analysis of tactics, techniques, and procedures of influence operations.

- **Tactics** are high-level descriptions of behavior.[21] They are the macro actions the influencer will take to try to achieve their objectives. Some examples include disinformation, propaganda, and the use of foreign proxies.[22]
- **Techniques** are specific tactics or actions,[23] such as using forged documents in a disinformation campaign; using GAN images[24] to obfuscate attribution of fake profiles and trick users into believing an account is real; and using fake news media outlets to launder narratives.
- **Procedures** are highly detailed descriptions of a technique.[25] Examples include using a specific social media application to reach an intended audience; spreading a conspiracy theory on a fake news website; or automating bot activity[26] with a computer program.

Infrastructure

The infrastructure used by influencers can include print media, television, digital platforms like websites, mobile phones, mobile applications, and more. While military information operations sometimes involve local approaches like paper flyers and face-to-face encounters ("winning hearts and minds" of the intended targets),[27] in modern state-sponsored operations, influencers are often working overtly and covertly on computers and mobile devices in an office environment across the world from the intended audience. Technical infrastructure includes all technical data that is obtained during the investigation of an influence operation, including domain names and WHOIS registration data, social media URLs and creation dates, metrics around social media account usage, website hosting information, activity on social media platforms, satellite use, cellular signals, and information on physical buildings and addresses. Details about the infrastructure used can be indicators of the scope, scale, and sophistication of the operation and the operational budget. It can also reveal information about who is involved and who sponsors the operation.

An influencer might monitor how open certain platforms are and find the social spaces where people feel they can most honestly express themselves. If an audience becomes fearful of speaking their mind because a platform censors them, they will likely shift to an alternative platform with like-minded individuals and less censorship. A well-resourced influencer (and analyst) goes to those spaces to target (or monitor) the preferred audience, out of the eye of censors and regulators. It is vital to consider both technical and psychological aspects when identifying infrastructure and when forecasting what shifts may take place. This also ties in with the socio-political axis, which together with the technical infrastructure can reveal valuable intelligence about the influencer.

21 https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-150.pdf
22 https://github.com/cogsec-collaborative/AMITT/blob/main/amitt_red_framework.md
23 https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-150.pdf
24 https://cset.georgetown.edu/wp-content/uploads/CSET-AI-and-the-Future-of-Disinformation-Campaigns-Part-2.pdf
25 https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-150.pdf
26 https://arxiv.org/pdf/1801.06863.pdf
27 https://www.warhistoryonline.com/instant-articles/winning-hearts-and-minds.html

Core: Narrative

The core element of every successful influence operation is a story or narrative. Narrative warfare is the weaponization of a narrative — the battle over the meaning and identity of information.[28] The study of narrative warfare and the communication of meaning between influencer and audience is essential for modern-day IO analysis. Analysts should learn to identify the elements of identity and narrative in a campaign to understand more about all of the factors: influencer, audience, capabilities, and infrastructure.

The narrative is often key to identifying who would be affected by the story and who would be motivated to propagate that particular message. For example, if the narrative is a false assertion that an election was manipulated by fraudulent voting,[29] an analyst can determine who benefits most from the false assertion (the candidate that lost the election and their party), what action that narrative may be intended to inspire (crowdsourced opposition to the election results and a recount of votes), and who would lose the most if that narrative does inspire a change in behavior (the candidate that won the election and their party). The influencer may also simultaneously target the political opposition to incite strong emotional reactions to the same disinformation and to the audiences that believe the disinformation. A strong enough emotional response can incite behavioral change in voters and inspire changes to voting laws by legislators.

Operationalizing the Diamond Model for Influence Operations

Using the diamond model to visualize the overlaps and connections between influencer, audience, capabilities, and infrastructure can help analysts solve complex problems. As the analyst discovers new information about a campaign, the information should be sorted according to the appropriate corner of the diamond, while narratives and elements of identity should be listed in the middle of the diamond. Analysts should work on intelligence leads that can reveal more about the 5 elements of the diamond, then start researching the relationship between the various elements. The more researchers and analysts understand the use of narrative and identity, the influencer-audience relationship, and the capabilities-infrastructure relationship, the more enabled they will be to recommend defensive actions and forecast the influencer's next moves. Below are 2 case studies analyzed with the diamond model for IO.

Practical Application 1: China and COVID-19

In March 2020, just months after the COVID-19 pandemic broke out in Wuhan, Chinese officials took to popular global social media platforms to convince the world that the COVID-19 virus originated in the US.[30] The conspiracy theory posited that 300 US military members brought COVID-19 from the US to China in October 2019 during the 7th Military World Games in Wuhan.[31]

After recognizing that this new narrative was trending among China's state-affiliated influencers, Recorded Future investigated further and found hundreds of posts across multiple platforms amplifying this same narrative. On March 13, 2020, Zhao Lijian, a spokesperson for China's Ministry of Foreign Affairs, posted on social media, "This article is very much important to each and every one of us. Please read and [share] it. COVID-19: Further Evidence that the Virus Originated in the US." The post linked to a conspiracy theory report from notorious disinformation outlet globalresearch[.]ca. Note the use of shared identity used in this statement: "... each and every one of us." This language is important because the influencer is using a shared sense of victimhood to gain sympathy for this conspiracy theory during a time of intense scrutiny towards China's government's handling of the COVID-19 outbreak.

28 Maan, *Narrative Warfare*.
29 https://www.reuters.com/world/us/trumps-false-claims-debunked-2020-election-jan-6-riot-2022-01-06/
30 https://www.buzzfeednews.com/article/ryanhatesthis/chinese-diplomats-are-pushing-conspiracy-theories-that-the
31 http://www.xinhuanet.com/english/2019-10/15/c_138473332.htm

Figure 3: Chinese disinformation alleging a US COVID-19 origin (Source: GlobalResearch[.]ca)

By redirecting the blame towards the US, and emphasizing that China's leaders are seeking the truth, Zhao is attempting to include Chinese citizens as victims of this alleged conspiracy by the US. Furthermore, by encouraging his audience to read the disinformation article and share it, he is activating his audience to become amplifiers of the disinformation. The conspiracy report was a fraudulent study by globalresearch[.]ca, a known disinformation outlet used in previous malign influence campaigns.

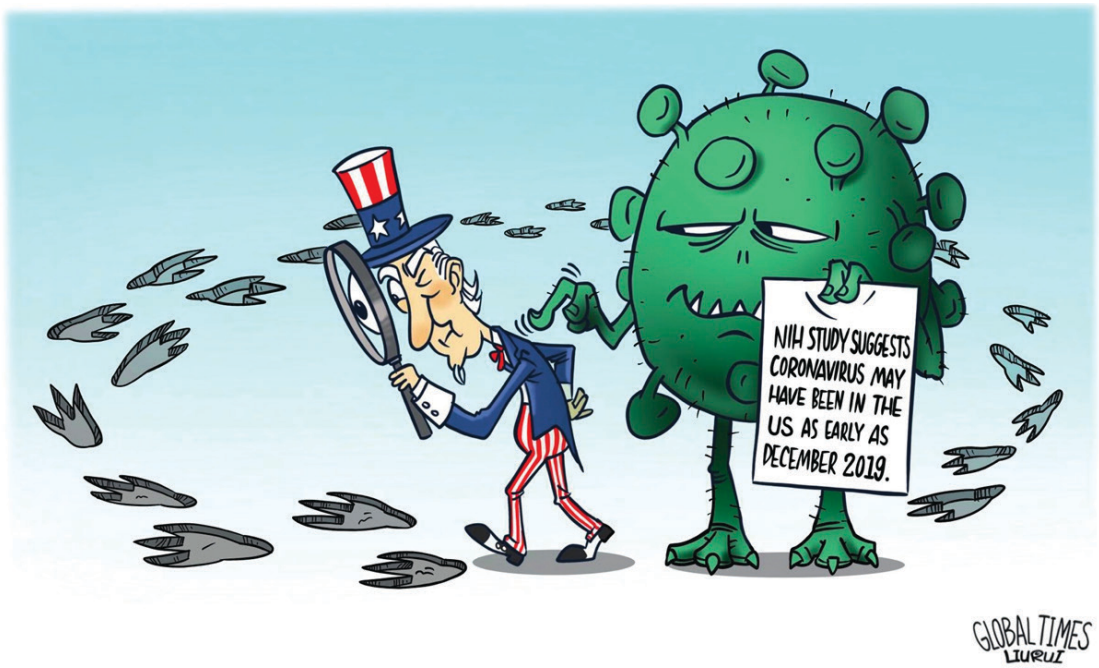


Figure 4: Political cartoons implying that COVID-19 may have started in the US (Source: Global Times)

Zhao's social media post received over 12,700 shares and over 20,000 likes and gained international attention—and scrutiny. Over several weeks, the Chinese government exposed global audiences to countless English-language memes, cartoons, press briefings, and social media posts accusing the US of covering up a Maryland-based COVID-19 origin story.

Figure 5: Cartoons, images, and advertisements used by Chinese state-affiliated influencers (Source: CCTV, Global Times, Recorded Future)

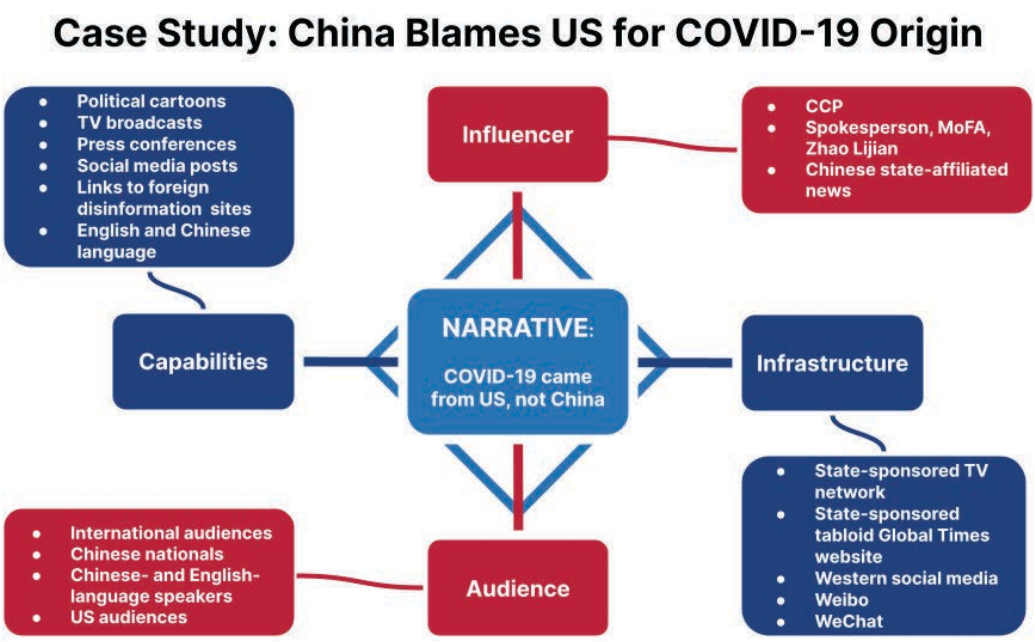


Figure 6: A list of known indicators related to a China state-sponsored influence campaign in early 2020 (Source: Recorded Future)

Narrative

In our investigation of this campaign, we determined that the **core narrative** (center of the diamond) was **"COVID-19 came from the US, not China"**. Variations of that narrative were observed, but they supported this central message. Note that Zhao Lijian's post said, "This article is very much important to each and every one of ***us***." This self-referential language is used to convince the target audience that the influencer shares their struggle and can empathize with them. At the same time, the influencer attempts to disparage the US government and create animosity toward the US. The tactic being used here is distraction, one of the 4 Ds in the framework discussed later in this paper. In this case, because the world believed that COVID-19 originated in Wuhan, China, a spokesperson for the Chinese government instead redirects the blame toward the US, pointing the finger back at the accuser.

The geopolitical context is also important — at the time this narrative emerged, then-US president Donald Trump was calling the virus the "China flu" and the "Wuhan virus", and denying that COVID-19 was a problem in the US. Chinese government officials reacted with harsh rebukes and then responded with this alternative COVID-19 origin narrative. This context can be mapped onto the socio-political axis. If you take notes on these relationships in the margins of those axes, it will assist you in making those strategic links and will also help you determine if the influencer has the intent to share malign content.

Influencers and Audience

Open-source tracking revealed that the primary influencers involved in this campaign were party-state officials (many in diplomatic positions) and spokespersons, as well as state-owned news outlets. The narratives targeted international and US audiences as well as the Chinese diaspora community. Because the same message was being amplified to China's native audiences across popular Chinese social media platforms like Weibo and WeChat, these entities are also listed in the audience section.

Infrastructure

Next, a broad open-source investigation revealed all of the infrastructure being used to disseminate the narrative. Key terms and rhetoric repeated across multiple mainstream social media and video platforms were used to track the narrative and see how the narrative was functioning on those different platforms. The investigation revealed that: the majority of the content had been seeded and amplified overtly through state-affiliated news outlets around the world and at home; mainstream social media was a primary method of dissemination; and much of the engagement that occurred on Western social media was likely automated and inauthentic. These indicators were added to the infrastructure axis of the diamond model.

Capabilities

The influencers shared links to known disinformation outlet websites, social media pages, cartoons, memes, and screenshots of fake news that alleged that COVID-19 emerged from a laboratory in Fort Detrick, Maryland. The influencers were also pushing the narrative on TV programs, digital newspapers, magazines, tabloids, video platforms, and press conferences. Although these should all be listed as infrastructure, they also indicate China's capability of advancing propaganda across all major media outlets. Most of these were English-language outlets, but there was also some coordinated inauthentic amplification in Mandarin Chinese. These indicators fit in the capabilities section of the diamond.

Once the diamond model is filled in, it's time to study **the relationships between each of the elements** and to apply other frameworks to each of the elements of the model. For example, the AMITT framework can help identify the specific tactics the influencer used, and apply those to the capabilities corner of the diamond model. Kill chain frameworks are excellent for breaking down each phase of an influence campaign — applying the diamond to each phase of the kill chain model will lead to a more granular analysis. Be aware that while direction and planning may come from the top levels of a government propaganda program, there is likely to be a variety of actors involved as the narrative is amplified across various platforms. It is important to differentiate which threat actors are involved with specific phases of the campaign or operation.

Practical Application 2: Russian Influencers Posing as Patriots

In June 2021, Graphika researchers reported[32] that suspected state-sponsored Russian influencers covertly infiltrated alternative social media platforms with politically divisive narratives and disinformation to appeal to American far-right communities. This campaign used tactics that

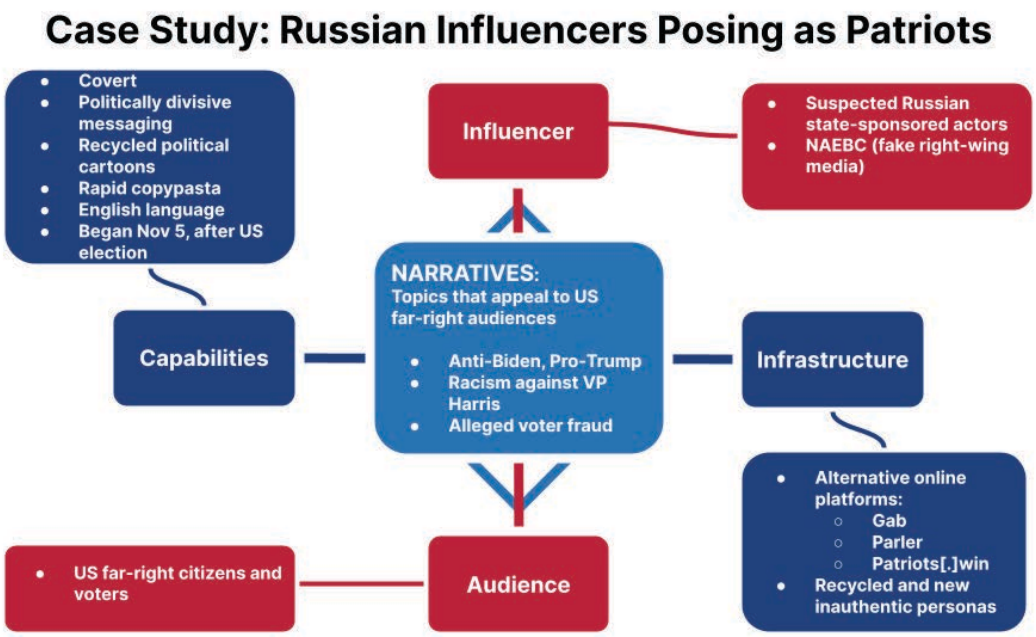


Figure 7: A list of known indicators related to a Russian state-sponsored influence campaign (Source: Recorded Future)

Narrative

The core narrative in this campaign was that the 2020 election was stolen from President Trump through widespread and systemic voter fraud. This core narrative was supported by the introduction and amplification of divisive political messages related to tensions over racial inequality and police violence, criticism of the US government's response to COVID-19, racist attacks on Vice President Kamala Harris, and accusations of senility and pedophilia directed at President Joe Biden.

32 https://graphika.com/reports/posing-as-patriots/
33 https://public-assets.graphika.com/reports/graphika_report_step_into_my_parler.pdf

Influencers and Audience

The influencers in this campaign were assessed to likely be the same Russian state-sponsored actors behind NAEBC, a fake right-wing media outlet previously exposed by Graphika and Reuters[34] in October 2020 as part of a covert Russian state-sponsored influence operation targeting an audience of US voters ahead of the presidential election. By observing the schedules and working patterns of the influencers, Graphika researchers discovered that the covert influence accounts were inactive during the New Year period when many Russians take extensive time off from work, and that their daily schedules closely matched Russian time zones for daytime workers. According to the Graphika report, the actors behind that campaign hired freelance writers to produce content, and ran a fake left-wing media outlet called Peace Data to provoke political debate and emotional responses from both sides of the American political spectrum.[35] Subsequently, in May 2021, Facebook attributed the NAEBC campaign to the Internet Research Agency (IRA), a Russian troll farm orchestrated by financier Yevgeny Prigozhin.[36]

Infrastructure

Following the 2016 US presidential election, social media companies increased enforcement of community standards, including prohibiting the use of inauthentic accounts to spread election disinformation. Many of the previously used inauthentic accounts were detected and shut out of mainstream social media platforms, prompting Russian state-sponsored influence actors to pivot to alternative **infrastructure** immediately after the 2020 election. Throughout the end of 2020 and the beginning of 2021, the target audience (conservative, far-right US voters) migrated to alternative platforms to avoid being censored and deactivated for spreading political disinformation. Many in the audience parroted rumors claiming that widespread voter fraud had occurred. Gab, Parler, and patriots[.]win (which moved to thedonald[.]win) became popular alternative platforms used by the target audience, presenting a ripe opportunity for the malign actors to exploit. These alternative platforms are widely unregulated, claiming to allow free expression and very little intervention from administrators. Many of the platforms were sponsored and promoted by American politicians and First Amendment advocates as tools for circumventing alleged censorship (removal of disinformation) happening on more mainstream social media platforms. Graphika researchers determined that many of the fake accounts used on Gab and Parler were previously used in earlier NAEBC operations.

Capabilities

The influencers' **capabilities** indicated a coordinated effort involving inauthentic social media accounts, rapid sharing of copied and pasted messages ("copypasta"), creation of pro-Trump and anti-Biden political memes and cartoons, and fake news alleging widespread voter fraud. Although the influencers used the English language to target American audiences, the researchers noted grammar mistakes that are common for native Russian speakers. While this campaign did not directly incite the insurrection event of January 6, 2021, in Washington DC, the influencers in this campaign had already seeded the "stolen election" narratives in the platforms where their targeted audiences spent most of their digital social time.

34 https://www.reuters.com/article/usa-election-russia-disinformation/exclusive-russian-operation-masqueraded-as-right-wing-news-site-to-target-u-s-voters-sources-idUSKBN26M5OP
35 https://public-assets.graphika.com/reports/graphika_report_ira_again_unlucky_thirteen.pdf
36 https://about.fb.com/wp-content/uploads/2021/05/IO-Threat-Report-May-20-2021.pdf

Figure 8: A patriots[.]win post by a Russian influencer alleging that Donald Trump won the election and President Biden lost (Source: Graphika)

While it is important to note the influencers' capabilities, it is also valuable to note any effects an influence campaign may have had on its target audience. Graphika researchers noted in their report that this campaign's reach was limited by the platforms it appeared in: "11 posts gather[ed] more than 100 comments and interactions from authentic users. But any efforts to seed content in the wider right-wing community seem to have failed." Furthermore, they noted that they were "only able to identify four instances of content being organically shared on other platforms, none of which had any success."[37] This is an important distinction that can play a role in determining the scope of a campaign, detecting whether there was an attempt to move the campaign to mainstream platforms, measuring public traction of the narratives, and identifying which audiences relate to the elements of identity used in a particular narrative. In this instance, it appears that the campaign's efficacy was limited to the far-right audience within specific platforms, and that engagement was highest with posts that had shock value (such as presenting evidence of election tampering or evidence of a scandal involving a political actor).

37 https://graphika.com/reports/posing-as-patriots/

Appendix A: Incorporating Existing Tradecraft

Current models for detecting, analyzing, reporting, and defending against influence operations have established a strong foundation from which to work. However, many of the existing models only address isolated elements of influence campaigns — the content (the media produced), the voice (who is involved), and dissemination (the way the content spreads). To understand the benefits of the diamond model and how to integrate other frameworks into it, it's critical to first understand existing frameworks, in terms of the various elements of propaganda and influence operations they address, and how they can be integrated into the diamond model framework.

Kill Chain Models

Kill chain frameworks analyze the lifecycle of a campaign and help isolate and identify steps in the adversary's process in order to recommend preliminary actions to neutralize the threat proactively. Many in the cybersecurity industry are familiar with the 7-step Lockheed Martin Cyber Kill Chain® framework,[38] which enhances visibility into a threat actor's TTPs. Various kill chain frameworks have been developed to analyze the spread of false information as well. 1 example is MITRE's disinformation kill chain,[39] which proposes 7 steps that an influencer takes to achieve desired objectives in a disinformation campaign.

Figure 9: *The MITRE Corporation's disinformation kill chain framework (Source: US Department of Homeland Security)*

Step 1: Recon

The disinformation kill chain follows a sequential process from left to right, starting with the influencer, or threat actor, conducting reconnaissance on the target audience, analyzing the information environment, and designing an operational plan. This is where the influencer will attempt to discover exploitable vulnerabilities, including technical infrastructure weaknesses such as the ability to create fake social media accounts on the target audience's platforms, as well as socio-political vulnerabilities such as religious, racial, or political divisions within the audience's society.

Step 2: Build

The second step involves the operators setting up the necessary infrastructure to achieve the desired outcome. This step involves a wide swath of digital infrastructure, including but not limited to fake websites, artificial personas, bots, forged documents, deep fake videos and

38 https://www.lockheedmartin.com/en-us/capabilities/cyber/cyber-kill-chain.html
39 https://www.dhs.gov/sites/default/files/publications/ia/ia_combatting-targeted-disinformation-campaigns.pdf

photos, and GAN-generated profile pictures. This step can be conducted by a threat actor with enough resources, or outsourced to third-party marketing, PR, or criminal organizations, which can also obfuscate attribution. In this phase, advanced influencers develop core narratives (stories), assign identity traits to personas, and activate their artificial personas (bots, trolls) for ease of amplification and coordination. State-sponsored operations often coordinate propaganda messaging and make role assignments in this phase.

Step 3: Seed

The third step is the official campaign kickoff. In this step, influencers begin designing content and planting the narrative across social media, such as by sharing images in target social media groups and seeding forged "evidence" such as fake photos and forged documents.

Step 4: Copy

The fourth step involves laundering the false information. Influencers will begin copying and amplifying the false narrative across all platforms, including fringe and fake news sites, state-affiliated media news sites, blogs, vlogs, social media, and chat applications. This step is key to creating distance between the influencer and the audience to obfuscate attribution to the original source.

Step 5: Amplify

The fifth step in the disinformation kill chain is the amplification of content and narrative across the target audience's key infrastructure. By this point, inauthentic accounts have created the illusion of content popularity, increasing likes and shares of the content on social media. During this step, both covert and overt operators are engaging the target audience, which begins to share the false content within its social circles, unaware of the origin or malign intentions of the false information.

Step 6: Control

In the sixth step, the influencers attempt to shape and control the emotional and behavioral reactions of the target audience. With the false information planted and spreading across all intended platforms, the established fake personas, trolls, and bots insert themselves into comment sections with divisive and emotionally-charged messages. If target audiences become suspicious of the information, influencers will adamantly deny being a part of the campaign, deflect blame onto their opposition, and redirect audiences to fake evidence to justify their stance.

Step 7: Effect

In the final phase of the disinformation kill chain, the influencers have ideally achieved their objectives. Objectives could include disrupting an election, creating social upheaval in a target country, or even changing the worldview or belief system of the target audience.

Disinformation ABC (+D)

In September 2019, the Transatlantic Working Group highlighted 3 "vectors of viral deception" to guide industry and regulatory responses to disinformation in a paper called "Actors, Behaviors, Content: A Disinformation ABC".[40] This straightforward framework outlines 3 essential elements of viral deception: Actors, Behaviors, and Content. The framework aims to reconcile approaches to disinformation across various disciplines and stakeholders. While the public sector may be primarily concerned with the identity of the actors behind an influence campaign, the cybersecurity

40 https://science.house.gov/imo/media/doc/Francois%20Addendum%20to%20Testimony%20-%20ABC_Framework_2019_Sept_2019.pdf

industry may be more concerned with the behavior and technical infrastructure, and social media platforms are likely more focused on moderating harmful content. This framework demonstrates the overlaps between the 3 vectors and guides various industries in helping to remediate disinformation campaigns.

Manipulative Actors

Manipulative actors are those that knowingly engage in viral deception campaigns. The authors of the Transatlantic Working Group paper distinguish between the less threatening individual trolls who create anonymous accounts online, and the more sophisticated, well-resourced state-sponsored groups. In this framework, the focus is on what cybersecurity expert Clint Watts calls the "advanced persistent manipulator" (APM),[41] a nickname borrowed from the cybersecurity term "advanced persistent threat" (APT). The APM is persistent in the pursuit of their objectives, is not easily deterred by account takedowns and defensive actions, and maintains enough resources and personnel to pivot to new techniques and technologies when challenges appear.

Deceptive Behavior

As the paper notes, deceptive behavior involves a variety of techniques that aim to "enhance and exaggerate the reach, virality, and impact" of a malign influence campaign. This includes the use of automated tools, artificial intelligence, and manual deception. Bot armies, automated amplification, troll farms, paid engagement, and obfuscation are all examples of deceptive tactics.

Harmful Content

Traditionally, content that promotes violence, hate, extremism, and terrorism is deemed harmful. However, with the advent of digital disinformation, the definition of harmful content is expanding. Health misinformation, political disinformation, and the spread of some conspiracy theories are all known to cause real-world harm to various individuals and organizations.

Information Distribution

In a Brookings Institution blog in April 2020, Alexandre Alaphilippe, the executive director of the EU DisinfoLab, suggested adding a "D" for information distribution to the ABC framework.[42] Alaphilippe pointed out that despite its successes, the ABC framework does not account for the way that "structural factors" inform disinformation campaigns. From a technical perspective, how a platform is structured is "crucial to understanding which actors will use it, how they will behave, and the kinds of content they will generate." This suggested update to the framework addresses the ephemeral nature of digital infrastructures such as network structure, functionality, algorithmic filtering, and "datafication of online platforms". Alaphilippe proposes that the way digital infrastructure is built affects how information is distributed and its ability to reach the intended audience.

41 https://securingdemocracy.gmfus.org/advanced-persistent-manipulators-part-one-the-threat-to-the-social-media-industry/

42 https://www.brookings.edu/techstream/adding-a-d-to-the-abc-disinformation-framework/

AMITT (Adversarial Misinformation and Influence Tactics and Techniques)

The AMITT framework,[43] originally created in early 2019 by a group of researchers called Misinfosec,[44] is an open-source framework designed specifically for describing and understanding disinformation incidents. Based loosely on the MITRE ATT&CK framework, AMITT is a large collection of both offensive TTPs and defensive measures, called "Red Framework"[45] and "Blue Framework"[46], respectively, on its dedicated GitHub page. AMITT objects are available in STIX™ format to facilitate tactical application for security practitioners and defenders that use standards like TAXII™ to integrate data into threat intelligence platforms. The AMITT framework is especially useful in the detection, tracking, and analysis of influence campaigns due to its granular breakdown of specific TTPs. Disinformation researchers are encouraged to contribute emerging tactics to the framework through the official GitHub page.[47]

Red Team Framework

Much like red team operations in cybersecurity, the AMITT red team framework is a chart that breaks out specific adversary TTPs used in disinformation campaigns. It is organized into 12 stages, labeled TA01 (Strategic Planning) through TA12 (Measure Effectiveness).

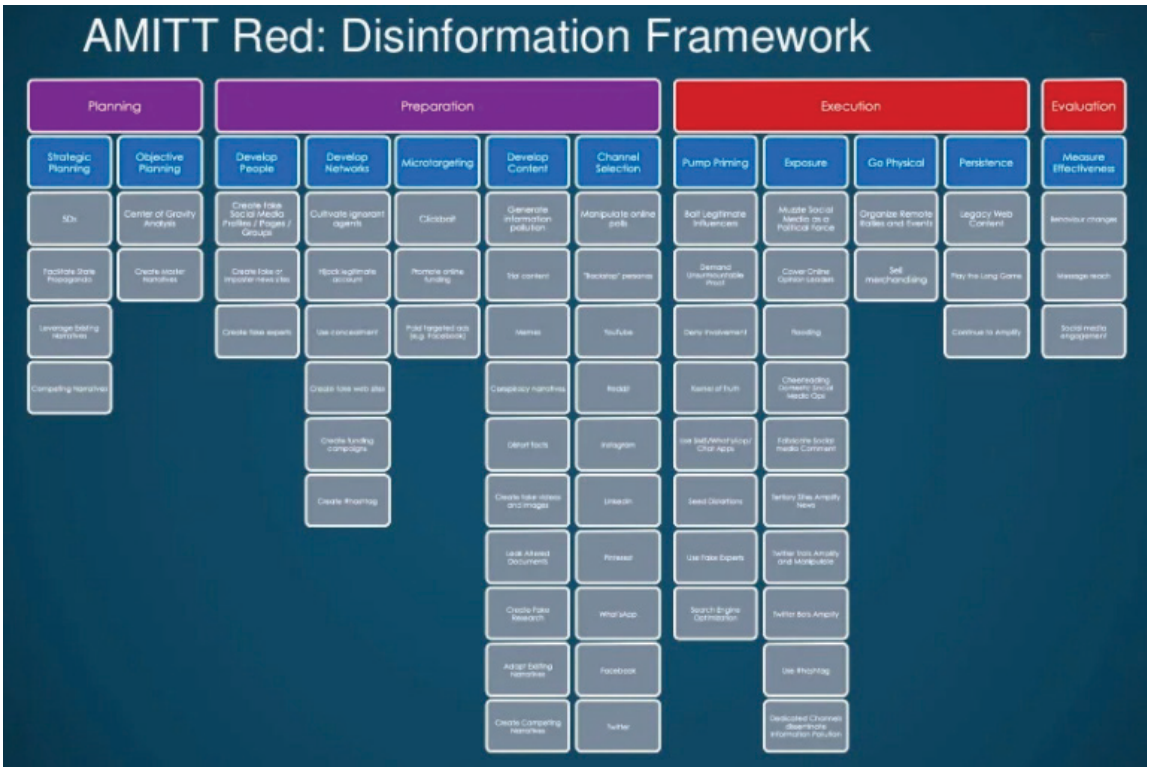


Figure 10: *AMITT Red Team Framework (Source: SJ Terp, Pablo Breuer, CogSecCollab)*

43 https://github.com/cogsec-collaborative/AMITT/blob/main/AMITT_HISTORY/2019-03-06_misinfosec_sent_for_publication.pdf
44 https://medium.com/@credibilitycoalition/misinfosec-framework-99e3bff5935d
45 https://github.com/cogsec-collaborative/AMITT/blob/main/amitt_red_framework.md
46 https://github.com/cogsec-collaborative/AMITT/blob/main/amitt_blue_framework.md
47 https://github.com/cogsec-collaborative/AMITT

Blue Team Framework

The blue team framework is a chart that breaks out specific mitigations that defenders can implement at each phase of the attack lifecycle, and mirrors the red team framework's 12 stages.

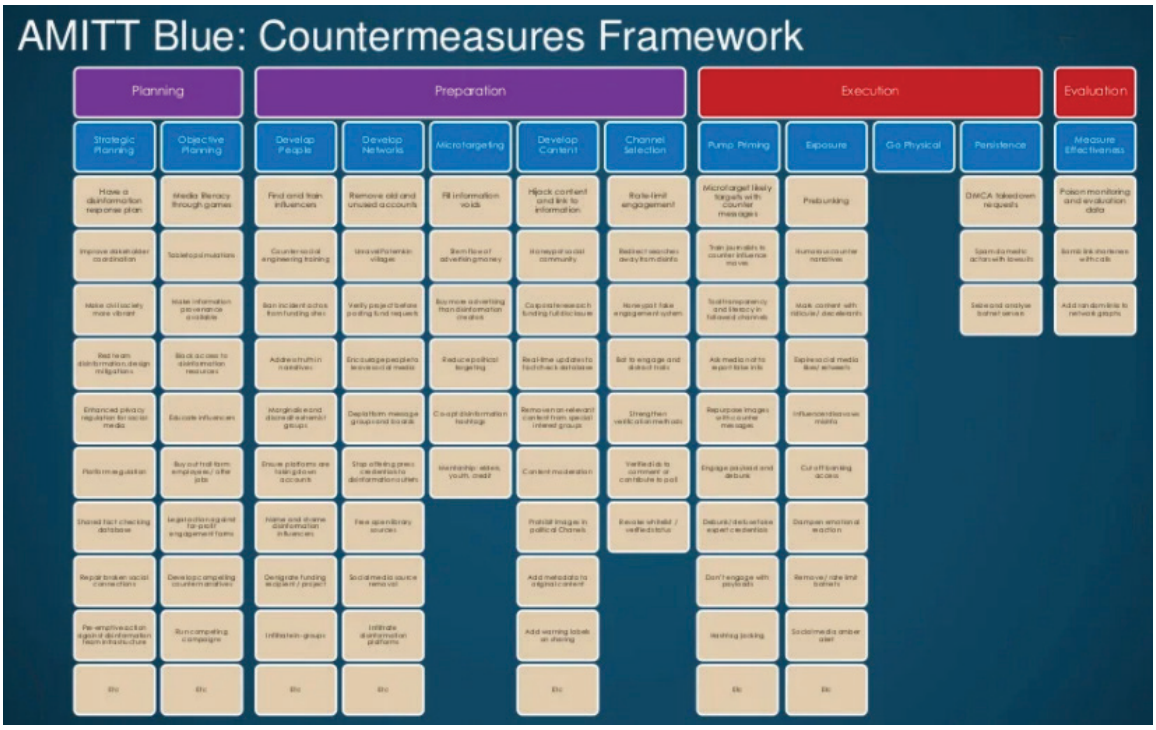


Figure 11: AMITT Blue Team Framework (Source: SJ Terp, Pablo Breuer, CogSecCollab)

The 4 Ds of Propaganda

In July 2017, Ben Nimmo (at that time a Senior Fellow for Information Defense at DFRLab, currently Global IO Threat Intel Lead at Meta), presented his novel framework called the 4 Ds of propaganda, specifically drawing from his experience analyzing Russian disinformation throughout his career. The 4 Ds — dismiss, distort, distract, and dismay — are foundational tactics of Russian propaganda, and are valuable for identifying state-sponsored propaganda from across the world.

Dismiss

This tactic often involves name-calling, attacking with stereotypes, and reinforcing rumors. Example: "_X_ country knows nothing of true democracy and should not insist that other countries implement their model of democracy given what a fraud their recent election was".

Distort

This tactic distorts the facts until the target audience pays attention. Example: "The election was stolen from us. The opposition staged a widespread, coordinated, and fraudulent election. We must expose the many ways in which the opposition pulled off this heist".

Distract

This tactic, which distracts by turning an accusation back against the accuser, is also known as "whataboutism". Example: "Who is _X_ to accuse us of human rights violations when that country cannot even manage racial injustice at home?"

Dismay

This tactic is about discouraging an action or behavior by making it seem frightening. This may entail a disproportionate threat or an exaggeration of the unwanted behavior in an effort to prevent the opposition from taking the intended action. Example: "We will act swiftly and decisively if _X_ pursues national independence. Those that play with fire will get burned. Efforts to assist _X_ to become independent from us will result in war."

Further resources on the collection and analysis of influence operations intelligence can be found on the Recorded Future resources website.

About Insikt Group®

Insikt Group is Recorded Future's threat research division, comprising analysts and security researchers with deep government, law enforcement, military, and intelligence agency experience. Their mission is to produce intelligence on a range of cyber and geopolitical threats that reduces risk for clients, enables tangible outcomes, and prevents business disruption. Coverage areas include research on state-sponsored threat groups; financially-motivated threat actors on the darknet and criminal underground; newly emerging malware and attacker infrastructure; strategic geopolitics; and influence operations.

About the Author

Charity Wright
Expert Threat Intelligence Analyst, Insikt Group®

Charity Wright is a threat intelligence analyst with over 15 years of experience at the US Army and the National Security Agency, where she started her career as a Mandarin Chinese linguist and intelligence analyst. Charity now specializes in Chinese influence operations and strategic intelligence at Recorded Future.

About Recorded Future®

Recorded Future is the world's largest intelligence company. The Recorded Future Intelligence Platform provides the most complete coverage across adversaries, infrastructure, and targets. By combining persistent and pervasive automated data collection and analytics with human analysis, Recorded Future provides real-time visibility into the vast digital landscape and empowers clients to take proactive action to disrupt adversaries and keep their people, systems, and infrastructure safe. Headquartered in Boston with offices and employees around the world, Recorded Future works with more than 1,400 businesses and government organizations across 60 countries.

Learn more at recordedfuture.com and follow us on Twitter at @RecordedFuture.