Chapter 7 Monitoring

TL;DR Monitoring

Monitor groups, narratives and artifacts (e.g. articles on URLs) in the team's area of interest, so when an incident starts, there's a body of knowledge supporting it.

Monitoring

When an alert comes in, the incident workflow starts (see the next next chapter for details of this). But whilst the sexy incident responses and "took down a huge operation" responses exist, incident response isn't all that a disinformation response team does.

We build out our knowledge bases and communities, write code to speed up our responses, test tools and processes, and work on a lot of the background things that help an incident response run smoother.

Workflows support the top-level mission and goals of the group. Activities to support those goals include incident response, but they also include:

- monitoring for public safety issues that we can report before they become harmful;
- mapping harmful narratives as they emerge;
- monitoring known disinformation feeder channels.
- chasing down disinformation tactics and counters;
- Adding and maintaining supporting disinformation data

Monitoring work includes spreader analysis - looking for infrastructure and accounts that are set up in advance of incidents, including sock puppet accounts “laundered” and left to mature.

Monitoring Alert Feeds

A team has many places it can potentially get disinformation alerts from. These include:

- Alerts from disinformation team members
- Feeds from other groups
- Phone honeypots
- Reporting hotline (dedicated email address)
- Sniff disinformation report lists, dashboards and botnet feeds for themes
- Set up reporting from social media (Facebook, twitter etc)
- Ask social media companies for feeds from them
- New data coming into the DKAN

We learn about potential incidents from several places:

- Teams connected to this one, e.g. Covid19activation and covid19disinformation, who are watching for disinformation online
- Team members spotting online disinformation and raising the alert in the team slack channel
- Team members spotting alerts from other disinformation tracking teams
- Other CTI channels telling us about disinformation in their feeds

Important: An alert isn’t the same as an incident. It's an indicator that something might be worth investigating and starting an incident response for.

Monitoring Narratives

Narratives are part of incidents - each incident might have multiple narratives involved, or just one, but there’s usually an identifiable narrative somewhere in there, that you can use to see if there are related incidents already tracked or dealt with etc.

The other thing about narratives is that they, like incidents, have lifetimes. Some narratives appear as a result of a world or local event (or upcoming or anticipated event), and are only useful whilst that event is in peoples’ minds. Example: using the Stafford Act to make everyone stay indoors was a narrative we tracked a month ago, before the stay-at-home orders started and it was a lot clearer about what states could, couldn’t, would and wouldn’t do.

Other narratives appear for a while, go dormant, then reemerge in different forms. Example: 5G, which was originally part of the radiation-of-all-forms-will-do-bad-things-to-you narratives, and has now come back in a mixup with covid19.

Organising Narratives

What we need is a way to log all the narratives that we know (or care) about, whilst keeping a smaller list handy of “currently alive” narratives that we can check incoming disinformation against.

There are a lot of narratives: we've seen hundreds of them in Covid19 alone. We've also seen that these can be grouped into themes; we've used mindmaps to group and organise narratives into hierarchies, making them easier to read and manage. We've also used spreadsheets to share narrative lists.

Identifying new narratives

Part of our work is to identify new threats before they become widespread. One way to do this is to identify emerging narratives from our existing asset collection.

First, we need to establish a baseline understanding of the current threat landscape in our area of interest (e.g. anti-mask, covid5g etc). The places we look to start this work include:

- Master narratives lists
- Existing lists of persistent threats known to carry disinformation: known bots, sources (e.g. disinformation websites), and canaries (accounts or hashtags with a high probability of carrying disinformation in this area)
- Regular threat streams: known disinformation feeds, subscriptions and platforms.

Once we have a baseline, we can establish persistent and repeatable monitoring:

- Identify data sources to monitor, e.g. googlenews, twitter, facebook, news aggregation sites etc
- Create saved or formatted searches for each platform, e.g. twitter = '#disinformation covid qanon boogaloo'; google = google hack formatted with a time parameter, e.g. 'disinformation and covid when=1d'
- Where api access is difficult, use other platform collection resources where possible, e.g. tweetdeck, crowdtangle

Other ways to find outlier or new narratives include watching for one or more of:

- merging and/or reemerging narratives being pushed by usually opposing groups, or old narratives that are reactivating
- local or world events, e.g. protests, changes in an area's status around specific dates (holidays etc)
- anomalous or significantly-sized online activity, e.g. in trending hashtags

Once narratives are found, you'll need to analyse them:

- evaluate source biases (is this state-owned media, an opinion article, social media etc)
- find additional sources with the same and/or competing narratives
- compare and contrast your findings: what's the same - is this fact or opinions? What's different - why? What's the intent and/or agenda behind the narrative - is it political, influence, harm, designed to confuse, distract, disrupt?
- How could this be used for bad (you might want to red team this)
- What would the impact be if this narrative is leveraged for bad?