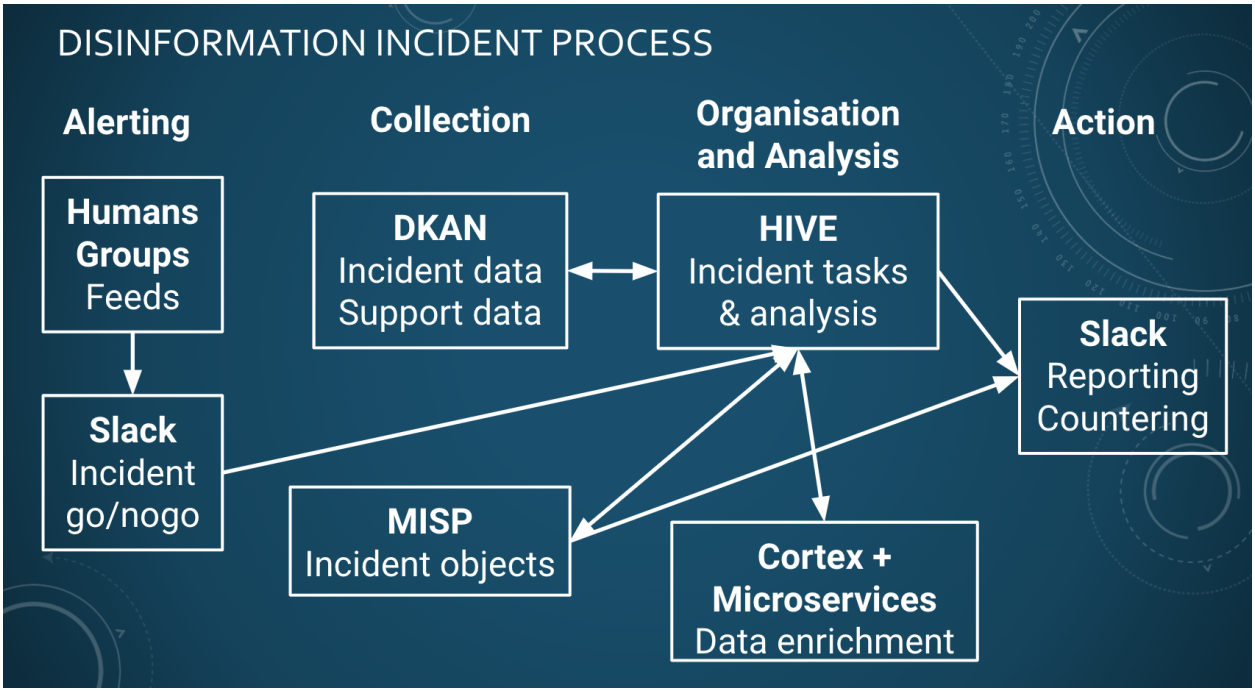
1 Chapter 4: CTI League Incident Workflow

Disinformation workflows include:

- Tracking an incident
- Tracking narrative flows across incidents
- Adding and maintaining supporting disinformation data

1.1 Incident workflow



Incident dataflow

The main workflow in the disinformation team is tracking an incident. We've broken this into 4 stages: alerting, collection, analysis, and action.

1.2 Workflow instructions

Starting a disinformation incident:

A new Covid19-related rumour has started online. You've seen it yourself, someone has sent you an example of it, you've seen another group tracking it - there are a bunch of ways to spot something new happening. Now what? **NB each of these steps can be by different people**

1. Tell people
    a. Put a message in slack #4-disinformation, with the artefact you found and a short description.
        - Start with "NEW RUMOR" so we will be able to track them
        - Any supporting information or links (under that rumor) should be posted in a thread off that initial NEW RUMOR post
        - This will make documenting and adding objects and observables to the incident and analysis log easier to track, and also keep everything a little more tidy
2. Decide whether to start an incident
    a. Do a quick check that it's a rumour. One sighting doesn't make an incident. 15 copies of the same message on Twitter, or 3 friends sending you the same strange DM, and you're probably onto something.
3. If it's significant, start an incident
    a. Give it a name. Names help.
    b. Add a row to the incidents spreadsheet
    c. Create a folder in the googledrive INCIDENTS folder for notes and anything that won't fit into the DKAN
    d. Start adding data to the DKAN (learn more about DKAN here)
4. Investigate the rumor
    a. Look for related artefacts, accounts, urls, narratives etc
5. Investigate ways to close down the rumor / repeater sites etc.
6. Report on the rumor
    a. Add an incident to the MISP instance for this rumor (learn more about MISP here)
        - The incident must include some relevant observables such as a Tweet, social media username or URL.
    b. Write and send notes/reports to the people who can respond
    c. Close down the rumor and move onto the next one (there's always a next one)

Help with a disinformation incident

1. The master document for what we're doing on incidents is the incidents spreadsheet. Look at the status column - the priority is live incidents, then monitor long-term, then

"keep an eye on it" (the potential 'zombie' incidents that are probably dead but might restart)
2. Check back in the slack channel, and in the incident README in the googledrive INCIDENTS folder to see what's been done with this incident recently. As we get things together, we'll probably have incident-specific tasks in the github issues list, but we're still working on that.
3. Find articles and artifacts, investigate the ones we have, put results into the slack channel for harvesting by the bots, and/or discussion with the team.
4. If you spot something significant (new objects tied to the incident etc, new things of interest), update the incident README.

1.3 Alerting

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

- Alerts from disinformation team members
- The covid19activation slack group (the Tedx team feed)
- The covid19disinformation slack group (the Atlantic Council team feed)
- CTI League Phishing inputs - maybe not so much; lots false positives
- Phone honeypots
- disinfo@ctileague.org - reporting hotline
- Feeds potential from other groups - e.g. peacetechlabs have offered a feed
- Mitre covid19 feed - might be in wrong direction; needs to be symmetric
- Sniff EuVsDisfo - is slow (narrative based) - SJ's dataset/ data stream list
- Sniff hamilton68 dashboard for themes
- Sniff botnet feeds for themes
- Set up reporting from Facebook, twitter etc
- Ask Facebook for feeds from them
- New data coming into the DKAN

At the moment, all the team's feeds are manual; team members check other slack channels etc, or CTI League members post alerts in the 4-disinformation slack channel. 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 #4-disinformation 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. An incident needs to be within the team's scope, and large enough to be worth expending team effort on.

- CTI League is Covid19. Do we just cover Covid19? No - can include politics. Don't care about aliens though.

- Anybody in the disinformation team can start an incident, but the group decides what it reports on.

When we see an alert, we have some questions:

- Is this an incident, e.g. is it a large coordinated disinformation incident, or an isolated piece / few pieces of disinformation?
- Is this disinformation suitable for processing by the disinformation team (e.g. 419 scams might be better handled by the Phishing team, but might also contain information about incidents that we should check out too)?
- Is this disinformation already being handled by platform teams or other specialist teams (we might want to check in with them just in case, for instance referring to #3-medical-sector-supporting if it is healthcare-related, or issuing a #4-takedown-request because of a finding)?
- Is this incident something that we should track?

"Is this incident something that we should track?", e.g. how do we choose which incidents to track?

- We don't track incidents for fun or interest. We track the ones that we have a reasonable chance of doing something useful about - whether that's raising the alarm to groups or organisations that can respond to the incident, asking them to take specific actions (like taking down a disinformation account or site), or taking actions ourselves (like amplifying counternarratives).
- We also track and counter incidents that we believe give us the best chance of a positive effect, and in the Covid19 deployment, ideally one that impacts health.
- Yes health. We prioritise that over other incidents, although we will include disinformation around current events where they impact populations.

1.4 Hive lists for starting an incident

When you create a disinformation incident in HIVE:

- Create a new case. Use case template "Influence Operation Incident".
- Name the incident (use this name in all the tools)
- Create an event in MISP for the incident:
    - https://misp.cogsec-collab.org
- List the risks and potential real-world consequences from this incident
- List any time bounds on the incident, e.g. are there events that it's gearing towards etc
- List any geographical or demographic targets in this incident
- Create a DKAN directory for the incident

MISP list for starting an incident

- List actors and other objects that are important in this incident - we're using a combination of STIX and DFRlab's Disinformation Dichotomies standard for this. Add these to the Clean MISP

- List the tactics and techniques that are being used in the incident - we're using AMITT for this (the version that comes as standard in MISP). Add these to the MISP event.

1.5 Organisation

Documenting analysis:

- We have DKAN and MISP, but also useful to have a google folder for each incident for other things that don't fit into those, like research notes
- Classifications: if it's openly available online, then it's okay to put through e.g. Tableau; if it's come though internal routes (e.g. SMS), then keep it off public internet (don't share).
- looking for related artifacts, urls, narratives etc

Who we communicate to:

- Report when something significant happens - e.g. see this main effort for this new line
- Report on time period… if big, a daily report; if smaller a weekly report
- No report goes out without at least 2 people beyond the editor going over it
- End users are also watching the MISP

Who makes decisions:

- Depends on decisions
- Need a board - vote via slack; person calling for vote does @channel to board, or emails them
- Who can add an incident? Anyone can start an incident.
- Who can release a report -
- Who can talk to customer/ victim? Needs to be agreed on

1.6 Collection

DKAN holds data we don't want to lose, and data that's raw and large: it's the in-tray
Data inputs for DKAN

- Potential starts of incidents
    - Feeds from messenger dms (about 30) - on personal facebook/messenger
    - Data in covid19disinfo team slack repository channel
    - Data in covid19activation disinfo-watch channel
    - <sms honeypots>
    - <emails to disinfo email address>
    - <feeds from other groups>
- Analysis datasets
    - Covid5g twitter data (5 directories) - on pc
- Supporting datasets
    - Narrative lists (CMU etc)
    - Narrative descriptions (EuVsDisinfo etc)

MISP hold objects of interest and the relationships between them, so we can quickly look up things we've seen before etc

Data we build up in MISP

- Incidents
- Narratives
- Actors
- Urls

1.7 Action

What we want to do with an incident is disrupt it as much as possible. If we can stop it completely, that's a big win, but generally, we're after disruption. CogSecCollab has a long-list (here: https://github.com/cogsec-collaborative/amitt_counters/blob/master/tactic_counts.md) of the things we can do to disrupt incidents at different stages of the disinformation killchain (https://github.com/cogsec-collaborative/amitt_framework - that, and DFRlab's object labels https://github.com/DFRLab/Dichotomies-of-Disinformation are what we're using in the MISP reporting), but frankly it's still messy so at this stage it's better to put our hacker hats on and think "which artefacts (observable objects) do we have in this incident, and what can we do to make them less effective?"

Examples: are there URLs pushing out covid5g disinfo? Are there social media accounts and groups pushing out covid5g disinfo? If we gather evidence on these, we can get that to the social media companies. Are there botnets involved (yes, yes, I said the b word, but they're part of this too)? Can report those too. Etc etc (and I suspect many of you have etcs CogSecCollab didn't think of when they created that counters repo).

This is the practical part of incident handling. We track an incident until the underlying incident stops or slows significantly (or the event it's building up to has passed), or until we've done as much as we believe we can to counter it, or know that there are other teams dealing with it.

Disinformation counters are much more than "remove the botnets" and "educate people". For most incidents, there are a variety of things that can be done about the incident, its creators, the objects used in it, and the tactics and techniques used. We've collected a few (well, a couple of hundred) suggestions for technique-level counters at https://github.com/cogsec-collaborative/amitt_counters - we're expecting to uncover a bunch more as more infosec people do disinformation.

1.8 Managing an incident response

An individual can track an incident on their own - open up some notebooks, fire up the coffeemakers and mainline chocolate for a couple of days. That's - not sustainable over time and large numbers of incidents, any more than it is for other infosec incidents.

The short instructions for managing a response are in the team readme. This is some of the thinking around them:

We haven't worked out exactly how to fit cognitive security / disinformation response into a SOC yet, but here's where we are at the moment on starting an incident:

- Incidents need names. Yes, yes, I know that's a slippery slope that ends up in a cute mascot and a dedicated website, but a name makes it easy to quickly identify what you're working on, find the right folder to put things into etc.
    - Action: Make up a name: make it short but descriptive - you're going to be typing it a lot, but you also want to remember what it was about a week later.
- The team needs to know you started an incident - both the team who are around at the time (and can help look for artifacts, add their specialist skills etc), team members who are coming in looking for things to do later, and leads who are trying to balance the load on the team overall. Best way to do this is to add a note to the team chat and an entry in the team log.
    - Action: add a note to the team slack channel, naming the incident and asking for help with it (if needed). If you have a starting artefact, add that too. Adding the word "NEW" will make it easier to find by people looking in on the channel later.
    - Action: add an entry in the team log, saying you're starting an incident response. At the moment, this is the incidents spreadsheet - this is likely to shift to adding a case to an incident tracking tool like TheHive.
- You, and the team, are going to start producing notes and artifacts as you track through the incident. Create a place to put them, that's accessible to the team
    - Action: create a space to put images, artifacts etc in. At the moment, that's creating a folder for the incident under the INCIDENTS googlefolder - this is likely to shift to directly uploading to a tool like TheHive or MISP.
    - Action: create a notes log for the incident. At the moment, that's a README file in the incident googlefolder - this is likely to stay the same for the moment. In the log, write a short description of the incident, and how you started tracking it (e.g. what the first artefact(s) you saw were.

Here's where we are on managing investigating the incident:

- You, and the team, are going to investigate the incident
    - Action: Look for related artefacts, accounts, urls, narratives etc
    - Action: add artefacts to the space you set up for collecting images, artefacts etc. You'll find it helpful if you number the images, because they're difficult to reference otherwise (aka "the yellow poster again" isn't as specific as "image001_yellowposter")
    - Action: keep the flow of investigation moving - keep a list of actions related to the artefacts, and/or direct the team to areas that need further research
- You'll also need to translate that into an incident description that can go out as an alert to other teams, and be used to look for potential counters

    - Action: add incident to alert tools. We're using MISP here, so adding a MISP object for the incident, and attaching the objects important to it is appropriate here.
    - Action: map artefacts seen to tactics and techniques. MISP includes AMITT - you can use the ATT&CK navigator to click on all the tactics and techniques you can see in this incident.
    - Action: Investigate ways to close down the rumor / repeater sites etc. We're working on tools for this too, but for now it's discuss this with the team, and check the lists below.
- Oh, and yes, you get to be scribe for the team too, making sure you keep a record of the investigation:
    - Action: keep the incident log updated with any significant findings, notes, things to do etc.

And here's where we are on managing responding to the incident:

- You need to get information about the incident out to other teams that could do something about it:
    - You've already added an incident to MISP; make sure it's ready to go (question: is there something we need to do to get it out on the feeds?).
    - Write and send notes/reports to the people who can respond
- If you found ways to respond, decide what to do, and check whether you did it
    - If the team found ways it could respond - triage them. Find ways to do the ones you can.
    - Also check on the things you were going to do. Was something done? Chase it up.
- And finally, know when to stop.
    - If you've done as much as you sensibly can, close down the rumor and move onto the next one (there's always a next one).

There are always more incidents, although we're often lucky enough to have a few days without anything major going on. Every morning, one of the leads (often SJ) looks through the list of incidents and decides which ones should continue to be 'live', which we should move to just keeping an eye on, or keep a longer-term watch on in case they flare up again, and which we can close down as unlikely to be active again.