Chapter 5 The Disinformation SOC

TL;DR Team Setup

- People:
  - create secure space to discuss incidents in
  - ensure team are aware of safety (mental health, opsec)
  - create / get access to training/ mentoring, if needed
- Process:
  - write team mission
  - detail process that's easy to follow under time pressure
  - set up connections to responders/ other teams
- Tech:
  - create safe shared space for notes
  - create / get access to data storage areas, if needed
  - create / get access to analysis tools, if needed
  - create / get access to information sharing tools, if needed

The Team

If you're working on a distributed disinformation defence, chances are you're either setting up a team or part of a team made up of individuals and/or organisations. We've learned a few things about that.

Team mission

You're setting up a distributed disinformation defence team, but you also need to know what its core mission is, so you can check any new requests against it (because scope creep is real, and you will at some point try to handle all the disinformations at once).

As as example, the CTI Disinfo mission is, "We're here to find, analyze, and coordinate responses to Covid19 disinformation incidents as they happen, where our specialist skills and connections are useful. We find and track new disinformation incidents, work out ways to mitigate or stop disinformation incidents and get information to the people who can do that."

Team needs

A disinformation response team typically needs:

- People
- Information sharing space / channels
- Incident management process
- Connections to responders and/or other teams
- Supporting technology
- Training / knowledge sharing materials
- An easy way to onboard people / help people find all the materials above (we use the team READme for this)

Another way of looking at this is through the process triangle, e.g. we need enough trained people responding fast enough to be able to make a difference to an incident - which includes noticing incidents fast enough, and ways to make those differences happen:

- People
    - Enough people to be able to make a difference to an incident, in the timespan that difference matters (includes noticing incidents in time)
    - Enough connections or levers to make a difference
- Culture
    - Safety processes: mental health and opsec
- Process
    - Understand disinformation, understand threat response
    - Fast, lightweight processes
- Technology
    - Speed - supporting analysis, storage etc
    - Sharing - get data to responders in ways they understand (use whatever works)

We'll cover each of these components, but the most important part of this is people.

Specialisations

Cognitive security needs many different skills - you're likely to find yourself working alongside information security people, but also academics, researchers, technologists, students, and journalists. We've found it useful sometimes to also form specialist teams:

- Leads: makes sure the disinformation team works smoothly and produces value. The lead keeps all the people and pieces in the whole team working well together: guides and supports the team, arranges resources as needed, coordinates team activities, tracks and logs team activities, and keeps an eye on overall team health (we're also sometimes in a difficult environment, and it helps to have someone watching for stresses).
- Incidents: Runs the disinformation incident response. Prioritizes incidents, e.g decides which alerts to respond to, and which incidents to concentrate effort on. Finds and maintains effort (alerting, collection, analysis, mitigation, cleanup) on incidents and decides when and how to hand off or close down incidents.
- Tech: Makes sure disinfo has all the tech it needs to do its job and keeps that tech running. Builds tools as needed. Finds and guides development talent as needed and appropriate for the disinformation team. Ensures tech builds are documented, repeatable, and maintainable. If appropriate, this role might be accompanied by a research or data lead.
- Process and training: Maintains the processes and team skills needed for disinformation team to do its job. Maintains manuals (e.g. the BigBook of Disinformation Response). Manages team training: makes sure there _is_ regular team training, and that the people running it have the resources they need.
- Outreach: Maintains connections between disinformation and other connected teams. Manages alert and data sharing with other teams. Maintains connections to teams feeding data into disinformation, users of disinformation team outputs, and to sister teams whose tech and needs overlap with the disinformation team.
- People: Makes sure disinfo has the people it needs to do its job and ensures there are routes to become a vetted (e.g. triage) disinformation team member. Arranges onboarding (newbie training, buddying, etc.) and vetting for prospective team members, maintains inventory of team skills available and needed, spots potential team trouble (e.g. troll breaches and other incursions), and offboards accounts if needed.

Different teams will have different specialist needs.

Team Channels

One of the features of a team working on disinformation is that it will itself be vulnerable to disinformation and incursion attempts, and will have to design its communications and information sharing channels to accommodate this. In the League, we handled this by running three separate Slack channels for the team:

- Disinformation: an open channel, for people to work on Covid19-related incident tasks, and to learn, engage, and share about Covid-related disinformation and disinformation techniques. Members have full access to open tools, process notes, training materials.
- Triage: a high-trust channel, for people who've been vetted to work on Covid19-related sensitive data and incidents. Members have named access to all data, tools, materials.
- Random: Open channel for anything that doesn’t fit into the other channels, e.g. non-Covid19-related resources and observations, general chat about disinformation.

Examples of posts and where they should go include:

- A fun post about disinformation in the gas industry = Random
- Non-medical political disinformation = Random
- Extremist disinformation = Random, unless they’re part of a disinfo incident
- An alert about a (non-sensitive) potential new incident = Disinfo
- Articles on health-related disinfo = Disinfo
- Cool tools and ideas = Disinfo (unless sensitive)
- Announcements = Disinfo
- News from other groups = Disinfo
- “How’s everyone doing” check-ins = Disinfo and Triage
- An alert about a (sensitive) potential new incident = Triage
- A post about a sensitive incident that we need to keep in triage = Triage

Other channels used in the League include:

- Data channels: useful for streaming supplementary input data that we don't want flooding the main human channels
- User channels: useful for finding us the people and places we need to get assistance, to report to (e.g. to find a specific Twitter group representative), to request takedowns, etc.
- External team channels: other teams (e.g. darknet) who work alongside us, sometimes on the same artifacts
- Output channels: stream clean outputs from teams

The channel norms we used in the League include:

- Follow the League code of conduct https://cti-league.com/cti-league/code-of-conduct
- Don’t put disinformation into the disinformation channel without a warning that it's disinformation
- Incidents should be threaded, so please add posts related to incidents in the relevant thread

- Machine safety:
    - Defang your urls (i.e., www\[.\]google\[.\]com)
- Human safety:
    - Keep anything potentially triggering in threads
    - Content Warnings: Live by the rule of “First, do no harm”
    - You can use “CW” to indicate that there’s a content warning
    - Types of content that require content warnings: Violence, Hatred, self harm, etc.

Onboarding: coming in to help

If you're joining a disinformation team, it's on the team leads to make make the processes simple to use and tools, notes, training etc easy to find. Some of the helpful things you might need to know include:

- The main work of the disinformation team is incident tracking and response. Live incidents are listed in the incident tracking system, and new ones are flagged in team channels as they’re added.
- Everything starts at the team README, which lists where to find tasks and information, the team channels, tech stack, and a cut-down description of the team's incident process.
- Information sources include the BigBook of disinformation response, team-specific explainers, and regular training / knowledge sharing sessions run by the team.
- Triage is a high-trust team, so we vet everyone who joins it. To join triage, fill out the disinformation team survey, so we have enough information to check you're not a bot, sockpuppet or similar.
- When in doubt, ask a team lead: we've created a Slack group, @disinfo-leads, so you can always reach a lead. Otherwise, checking social media to see if a new incident is brewing is a never-ending job.

How Disinformation fits into an Infosec Threat Response team

Reading through the CTI League handbook, the league stresses “Our members prioritize efforts on helping hospitals and healthcare facilities protect their infrastructures during the pandemic and creating an efficient channel to supply these services”. The disinformation team should do this too.

It lists services as:

1. Neutralize malicious activities in the cyber domain with takedown, triaging, and escalating relevant information for sectors under threats.

2. Prevent attacks by supplying reliable, actionable information (IoCs, vulnerabilities, compromised sensitive information and vulnerabilities alerting).
3. Support the medical sector and other relevant sectors with services such as incident response and technical support.
4. Act as clearinghouse for data, a connection network and a platform for facilitating those connections
5. Neutralize malicious activities in the cyber domain with takedown, triaging, and escalation relevant information for sectors under threats.
6. Prevent attacks by supplying reliable, actionable information (IoCs, vulnerabilities, compromised sensitive information and vulnerabilities alerting).
7. Support the medical sector and other relevant sectors with services such as incident response and technical support.
8. Act as clearinghouse for data, a connection network and a platform for facilitating those connections

There are disinformation equivalents to these:

1. Neutralize. Disinformation incident response: disinformation triage, takedown, triage and escalation.
2. Clearinghouse. Collate and share incident data, including with organizations focusing on response and counter-campaigns (the “elves” who fight the “trolls”).
3. Prevent. Collate disinformation vulnerabilities and indicators of compromise (IoCs), and supply these to the organizations that we work with.
4. Support. Assess the possibility of direct attacks, and ways to be ready for that. For example, prepare resources that could be used in countering campaigns that target specific facilities, groups and high-profile individuals.

For the neutralization part, the league lists as examples:

- Infrastructures used by a threat actor that is exploiting the pandemic – malicious command and control server / DDoS servers / domains / IPs / etc.
- Exploiting legitimate services (such as open port in a legitimate website or compromised website used by hackers) and relevant to our stakeholders can be used to deploy attacks

The disinformation equivalents here would include:

- Hashtags, groups, networks, botnets, information routes, etc. used by disinformation actor groups to create and run incidents. We can map several of these ahead of time, monitor them for new events forming (e.g. qanon checkins), file abuse complaints to registrars, notify companies hosting botnets and command and control accounts, etc.
- Medical events (e.g. vaccination rollouts) that we know will trigger disinformation incidents

For prevention and support, the league lists examples:

- Alerting about vulnerabilities / compromised information and infrastructure to our stakeholders
- Creating a database of malicious indicators of compromise for blocking (via both MISP and GitHub repository)
- Alerting about trends and uneventful events regarding the pandemic in the cyber domain
- Creating a database of hunting queries for alerting systems.
- Create a safe and secure infrastructure for CTI League activities
- Create reports dedicated for the stakeholders and update them about ongoing trends of attack vectors regarding their organizations, such as significant information from underground-based platforms (darknet).

This is more detailed work, but as we track more incidents and become more familiar with the methods and tools used by incident creators, some measure of prevention activities become possible.

Red Team: Learning how the other half thinks

If you're defending against a team's actions, it's useful to understand how that team thinks. Whilst it's possible to reach out and speak to disinformation creators, that's not always advisable, so CogSecCollab runs weekly Disinformation Red Team sessions. We've learned a lot from these sessions, often changing the way we work to match insights like "yeah, of course that's why it happens that way".

Some topics you might like to try include:

- running "disinformation as a service"/alternative marketing companies,
- running hostile social media platforms,
- running hybrid disinformation / traditional infosec incidents,
- extending an existing campaign, to predict where it might go next