Chapter 9 Detective work: Asking the Right Questions

TL;DR The Right Questions

Start by listing questions for the team to answer. Most of the time we're doing detective work, looking for evidence that we can send over to another team. Typical questions:

- What artifacts are we starting with? What are they connected to? Have we seen any of these before?
- Who is involved (groups, accounts, etc)?
- What are we trying to do: track artifacts and incident back to origins? Track its spread outwards? Work out where the influence points, places it might be stopped, ways it might be countered or mitigated?
- What are we looking for: just artifacts, or narratives and techniques to add to the incident report?
- Are there other groups already tracking this? Are there related datasets we can pull in and use?
- Who should we be alerting, when, and what do they need?

OSINT, data science and intelligence analysis all have methods that can be useful here.

That's the top-level questions: each artifact will have questions connected to it too, for instance: do we want to map out any networks connected to URLs we find? Are hashtags being amplified etc. These are covered in the next chapter.

Framing Questions

Always start with the questions you want to answer.

- What are we starting with?
    - What's our initial artifact, theme, narrative, lead
- What's our "research question"?
    - What do and don't we care about here?
    - What's more and less important to us (if we have limited resources)?
- What are we trying to produce and for whom?
    - Enough evidence that we can identify who to pass it to, and give them enough to either act or start their own investigation
    - Enough evidence and information to take action ourselves

The Five Ws

In a lot of information gathering and sharing work, we want to know the five Ws: who, what, when, where, why, and how. (https://en.wikipedia.org/wiki/Five_Ws)

Artifact questions

Once we start an incident, our first job is to gather enough information to determine whether we should act, hand this information over to another party, stand down, or not act, but keep a watch on this area. This is usually a mixture of artifact-based activity analysis, network analysis and fact-checking.

- Activity analysis
    - Track artifacts (messages, images, urls, accounts, groups etc), e.g.
        - find artifact origins
        - track how an artifact moves across channels, groups etc
        - find related artifacts
    - Detect AMITT Techniques, e.g.
        - Detect computational amplification
        - Detect, track and analyze narratives
- Network detection
    - find inauthentic website networks (pink slime)
    - find inauthentic account and group networks (including botnets)
- Credibility/ Verification
    - Fact-checking: verify article, image, video etc doesn't contain disinformation.

    - Source-checking: verify source (publisher, domain etc) doesn't distribute disinformation.

Fact checking is hard, and usually needs a team of fact-checkers, up-to-date knowledge etc. Source checking isn't the same as article classification, although it does include article classifications. Source-checking is why we label and track URLs. Several groups already publish labelled lists of domains. Fake news creators often run multiple, seemingly unconnected, sites, so finding already-labelled sites in a network can help a lot.

Other related activities include things like looking for signals of inauthentic accounts, inauthentic amplification, and other inauthentic online behaviours, e.g. by looking at patterns of account creation dates for popular messages.

Artefact Tasks

Most alerts start with an artifact: an image, a URL, a piece of text etc. We usually have a lot of starter questions about these too:

- Have we seen this artifact before? Is it related to an artifact we've seen before (e.g. a variant of an earlier artifact).
- Is the artifact's context similar to earlier artifacts (e.g. are the networks and accounts pushing it the same as the ones in earlier incidents?).
- We see a lot of repeat offenders - is this a known scam? Are the actors behind it people already associated with known scams?

Attribution

Attribution - working out who's responsible for a disinformation incident - is hard. You don't have full access to data, and there are incentives for people to obfuscate and hide who they are. At best, attribution is probabilistic, but even a hint can help us assess potential moves, and countermoves.

Further Reading

- Heuer, Structured analytic techniques for intelligence analysis
- Beebe, Pherson, Cases in Intelligence Analysis
- Joint Forces Smartbook, JFODS5