



Vertex

# Synapse Bootcamp

Module 18

Reports, Articles, and the Spotlight Tool

---

v0.4 - May 2024



# Objectives

- Describe "reports" as they relate to Synapse
- Know the model elements used to represent reports
- Understand options to ingest reports into Synapse
- Know the key features and use cases for the Spotlight Tool



# Reports in Synapse



# Reports and Threat Intelligence

- We **consume** numerous reports for threat intel
  - Threat reports (blogs, whitepapers...)
  - Vulnerability reports
  - News reports / articles (breaches, compromises, current events...)
  - Internal reports (incident, audit...)
- How can we:
  - Get various types of reports into Synapse
  - Make the report content more useful / accessible for our analysis?



# Reports vs. Feeds

- With reports, the content is typically **prose**
  - Largely **unstructured**
  - Link data of interest to the **report** or **article**

```
media:news -(refs)> *
```

- With a feed, content is typically **data**
  - **Structured** data, often from API
  - Typically IOCs, sometimes with labels for context
  - Link data of interest to the **data source**

```
meta:source -(seen)> *
```



# How Does Synapse Model the Data?

Real World Data to Represent	Examples	Synapse Model Element
The source of the reporting	An online blog (HTML) or whitepaper (PDF) A news article A list of indicators (CSV, TXT) An internal report or document	media:news
	An online post (social media, forum, etc.)	inet:service:message
	A conference presentation	ou:preso
A <b>source</b> contains or talks about ("references") something	A blog lists IOCs A whitepaper contains MITRE ATT&CK elements A news article describes a person, attack, conference, event...	-(refs)>
The thing the source talks about	A set of IOCs A specific attack or campaign A targeted company or individual A conference	Any relevant form "referenced" by the source



# What Can I Model?

Kind of Data	Examples	Example Forms
Indicators	Hashes, domains, IP addresses, URLs... File names, file paths Registry keys, mutexes, commands executed...	hash:md5, inet:fqdn, inet:url file:base, file:path it:dev:regkey, it:dev:mux, it:cmd
Related reports	Additional links (references, citations)	inet:url
TTPs	MITRE ATT&CK elements CVE numbers / vulnerabilities	it:mitre:attack:* it:sec:cve, risk:vuIn
Threats	Threat group names Malware family names Campaign / operation names	ou:name it:prod:softname ou:campname
Targets, attacks	Companies, people, web sites Attacks / campaigns / compromises	ou:name, ps:name, inet:url risk:attack, ou:campaign, risk:compromise
Data of note	People, organizations, conferences, generic events	ps:person, ou:org, ou:conference, meta:event

The amount of data / information you want to extract and link to reporting is up to you!



# Automating Report Ingest

- So far we have mostly pasted in / created indicators
  - Have not created / modeled reports
- To **automate** ingest, we may want to:
  - **Model** the report (create a `media:news` node)
  - **Extract** and **link** key information
  - **Tag** nodes for context
  - **Index** the report content
- Synapse **Power-Ups** and **Tools** can help





# Synapse-RSS

- Download and process feeds (RSS/Atom)
- Specify feed URLs
  - o Tag URLs for simplicity (e.g., #vtx.auto.rss)
- Synapse-RSS will:
  - o Create a `media:news` node
  - o Save content as a `file:bytes` node
  - o Use `synapse-fileparser` to extract IOCs
  - o Link IOCs to **both** the `media:news` and `file:bytes`
- Options to:
  - o Save as PDF (vs HTML), disable parsing

NODE	ALL TAGS	ALL PROPS
media:news		d25195c363f8d40aa6c633573d166f39
:file		sha256:134e90a658c2990da9a4d88adf5f5dc...
:published		2023/11/13 11:00:23
:rss:feed		http://feeds.feedburner.com/Unit42
:summary		In July 2023, pro-Russian APT Storm-09...
:title		in-depth analysis of july 2023 exploit...
:url		https://unit42.paloaltonetworks.com/ne...
:url:fqdn		unit42.paloaltonetworks.com
:created		2023/11/13 12:20:03.569
:seen		(2023/11/13 12:20:03.581, 2023/11/17 0...
+ Add Tags		



# Other Power-Ups

- Some vendors include APIs that provide reports
  - Synapse-AlienVault
    - AlienVault pulses
  - Synapse-MISP
    - MISP events
  - Various commercial vendors
- "What" Synapse can model depends on the data provided by the vendor



# Power-Ups Demo



# Spotlight Tool



# Automation

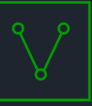
- Automation is better than hand-modeling reports!
- Drawbacks:
  - Can only extract / link:
    - Objects that can be extracted by the FileParser
    - Objects returned by an API (usually atomic indicators only)
  - Can only apply tags returned by an API
- This limits:
  - The types of data (nodes) automation can represent
  - What tags (if any) are available to provide context to data

How do we automate report ingest  
while capturing additional data and rich context?



# Spotlight Tool

- Automate ingesting and processing reports
- Ingest any PDF or URL
  - Identify (or create) associated media :news node
  - Save content as PDF
  - Extract and link common indicators for review
- Analyst can:
  - Review data
  - Add to Synapse
  - Tag indicators
  - Leverage Research Tool features directly from Spotlight
    - Query, run Node Action...



# Spotlight Demo



# Summary

- Leveraging **reports** is a key CTI process
- Synapse ingests reporting using:
  - **Power-Ups**
    - synapse-rss, synapse-alienvault, synapse-misp...
  - **Spotlight Tool**
    - Ingest, extract, review, and tag relevant data
- Automation is best for **throughput**
  - Load more data faster
- Analyst processing is best for **detail**
  - A human captures more / richer information than a process
- **Spotlight** combines both methods