Building a Threat Library

Chris Adams

Building a Threat Library is a fundamental capability for cyber threat intelligence teams. Designed and managed correctly, a threat library is the underpinning of decision support systems for the security team and drives operations. At the most basic level, with the restrictions of time and resources, threat intelligence promises to channel focus onto that which matters most.

There are four main categories of activity captured in the use case of operating a threat library:

- Automatic Intel Collection
- Situational Awareness
- Intel Creation & Data Modeling
- Intel Lifecycle Management

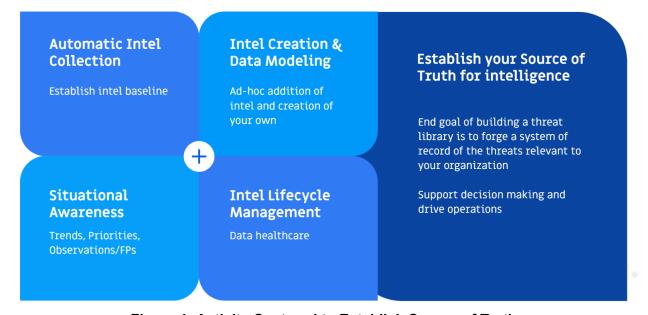


Figure 1: Activity Captured to Establish Source of Truth

We'll walk through each of these segments in a bit more detail, but note that the key objective we aim to achieve for customers is to establish the source of truth, or system of record, for the threats that are most relevant to your organization. The outcome of that objective is to create an environment that supports decision making for analysts and helps drive security operations.

Automatic Intel Collection



Establishing an intelligence baseline means just that - starting out with an initial set of open source or premium feeds that start you in the direction of applying intel to your decision making and driving operations.

ThreatConnect® provides open source intelligence out of the box. Use the <u>Feed Explorer</u> to decide which feeds are available as part of our open source content.

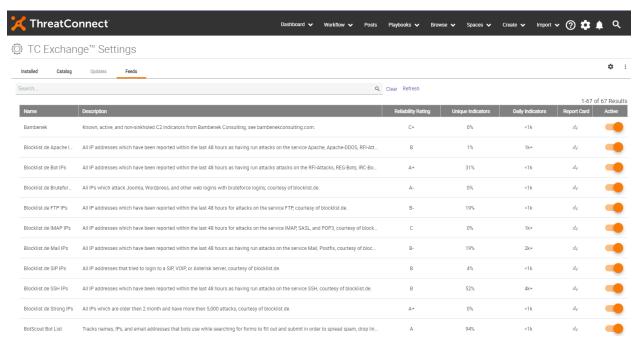


Figure 2: Feed Explorer

<u>Feed metrics and report cards</u> make it simple to review which feeds may be most impactful. This is accomplished using rating qualifiers and common classifiers, as shown in Figure 3.

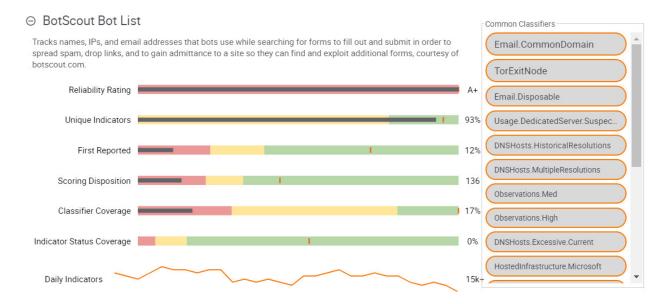


Figure 3: Report Card

Context is king in the world of threat intelligence. ThreatConnect uses data analytics and machine learning algorithms to derive contextual information pertaining to the threat data and presents that context using CAL classifiers. Analysts should be familiar with the CAL Classifiers glossary so that context is understood, the following table shows only a small portion of classifiers.

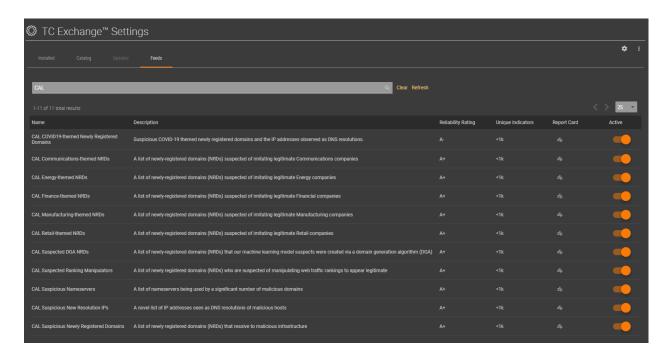
Table 1

Classifiers Glossary

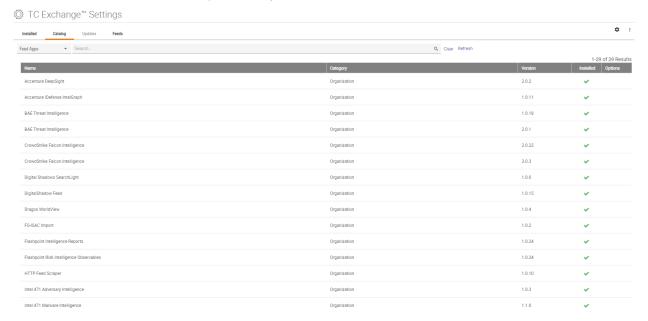
Table 1

Classifier	Valid for Indicator Types	Comments
Active Host	Host	Indicator has a sufficient number of reported observations from ThreatConnect users within the allotted timeframe.
ASN.Invalid	ASN	This ASN does not exist, according to the master list of ASN:CIDR mappings.
CloudHosted	URL	Indicator hosted on a common cloud-hosting domain (e.g., amazonaws.com).
DNSHosts.Excessive.Current	Address	The IP address has an excessive number (3+) of hosts concurrently resolving to it.
DNSHosts.HistoricalResolutions	Address	The IP address has had hosts historically resolve to it, but currently no tracked hosts do.
DNSHosts.Malicious.Current	Address	The IP address has a sufficiently evil host that currently resolves to it.
DNSHosts.Malicious.Historical	Address	The IP address has had a sufficiently evil host resolve to it in the last 30 days, but not currently.

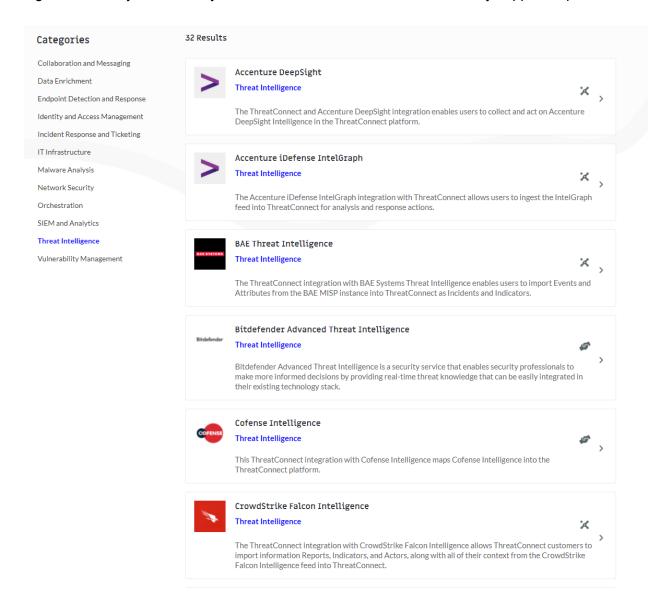
ThreatConnect <u>Collective Analytic Layer Feeds</u> are a combination of the immense CAL dataset and analytics along with the tradecraft of our Research team to identify pockets of intelligence that are fertile hunting grounds. Packaging this intelligence and pushing it directly to your ThreatConnect instance, we're providing a broad set of theme-based and suspicious intel threat teams should be aware of - consider this an added benefit to making ThreatConnect your Threat Library of choice.



Next, consider what you can gain from premium feeds, there are an extensive list of <u>vendors</u> whose feed apps can be <u>deployed</u> simply with a valid license.



Premium feeds can provide great value with sector-specific, or threat-specific context that your organization may need to stay ahead of threats. Here is a list of currently supported providers.



ISACs & government agencies are beginning to adopt STIX/TAXII in greater form so ingesting intelligence from <u>TAXII</u> sources, or by <u>ingesting bulletin emails</u>, is a native capability of the platform. More complex methods for extracting detailed information from emails can also be achieved with <u>email-trigger based playbooks</u>.

Now that you have some content loaded, take a peek at the data by using <u>Browse</u>. You'll see data that is ingested in normalized, de-duplicated and maintained in each individual source so as to preserve context and scoring.

Intel creation & data modeling

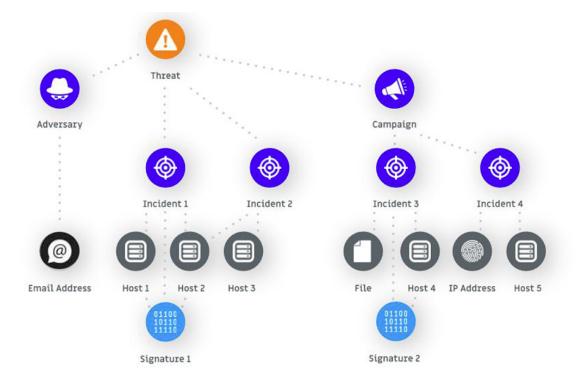
Manual ingestion via utilities, intel creation and threat data modeling

Okay you have a baseline, now what...

Many times the best intelligence is in the analyst's own network, server logs or items that team members share in ad hoc forms. ThreatConnect provides a number of tools to help teams build threat models and collect important information that are a part of your threat landscape.

It's important to review how <u>Threatconnect organizes threat information</u>. ThreatConnect leverages the <u>Diamond Model for intrusion analysis</u> as a principle in that there are both strategic and tactical views of an intrusion that help you understand the root of attacks and give you the best chance toward defending yourself from future attacks.

Simply put, ThreatConnect helps you organize information so that key elements or objects involved in an attack are captured, along with the intent of malicious activity, along with the relationships between those objects. Below is an example of a model ThreatConnect helps create and store.



A lot of what we base our data model comes from industry standards like STIX, where relationships between data elements help forge a sense of strategy - strategic relationship between actors and victims, knowing what tools are used to attack and which infrastructure is used to accomplish the attack - that is what the diamond model is all about.

Behavior is captured using MITRE ATT&CK and attribution can be applied using methods such as Kill Chain or Phase of Intrusion.

One of the most common needs from analysts is an easy way to create intelligence from unstructured sources. For example, Analysts can ingest content by <u>importing</u> it and <u>create</u> a model from it. As threat data is collected or products, analysts are given tools to qualify and enrich the data ingested to give you a clue to its relevancy. ThreatConnect's Collective Analytics Layer (CAL) is a major part of that.

This process is best described in a scenario:

[Note there are also internal training documents that provide materials to demo]

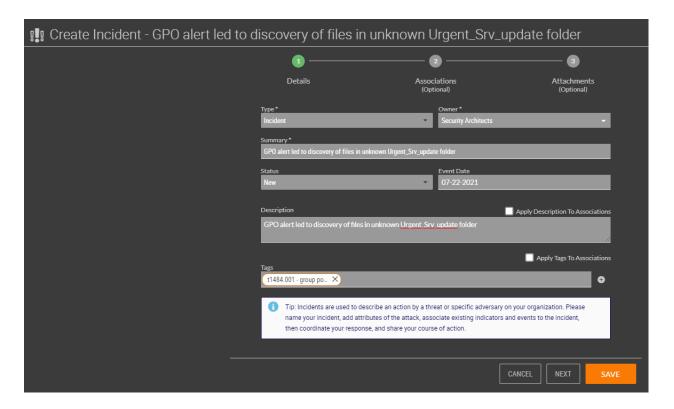
- 1. Alert kicks off because of a policy violation on a windows server, noting an escalation of privs. After inspection suspicious files are found in a folder called *Urgent-Srv-update*
- 2. Analyst sends me a list of tactical threat indicators, concerned they are malicious. Take the list and pull them into the platform, <u>tag</u> them and associate them to an incident. Until we know more, not associating them to any particular actor.

3. List of indicators taken from a windows server file_hashes.txt, that were stored in a folder called Urgent-Srv-update

File_hashes.txt

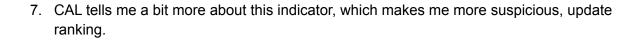
ec4cdc752c2ecd0d9f97491cc646a269 edb648f6c3c2431b5b6788037c1cd8ef ee3e297abd0a5b943dce46f33f3d56fb ee4862bc4916fc22f219e1120bea734a ef14448bf97f49a2322d4c79e64bb60b ef2738889e9d041826d5c938a256bc45 ef6fcdd1b55adf8ad6bcdf3d93fd109e

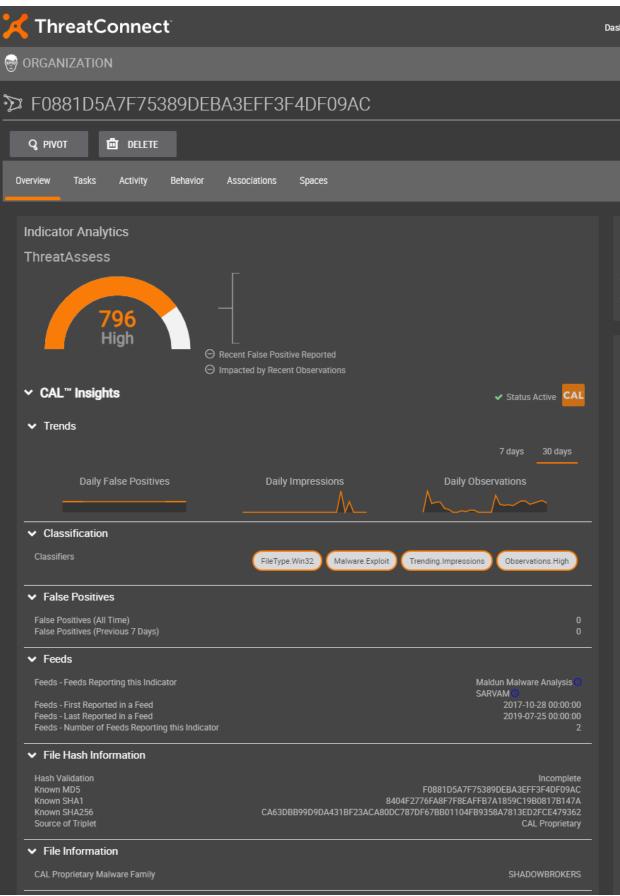
- 4. Incident response team that were discovered on a server which some security software triggered an alert when escalation of privs [T1484], domain policy modification
- 5. Step 1.
 - a. Create an incident
 - b. Summary: GPO alert led to discovery of files in unknown Urgent_Srv_update folder
 - c. Reference <u>MITRE ATT&CK</u> to determine the appropriate TTP. In this case, we are looking for group policy modification T1484.001



6. CAL clues me into something regarding hash

F0881D5A7F75389DEBA3EFF3F4DF09AC

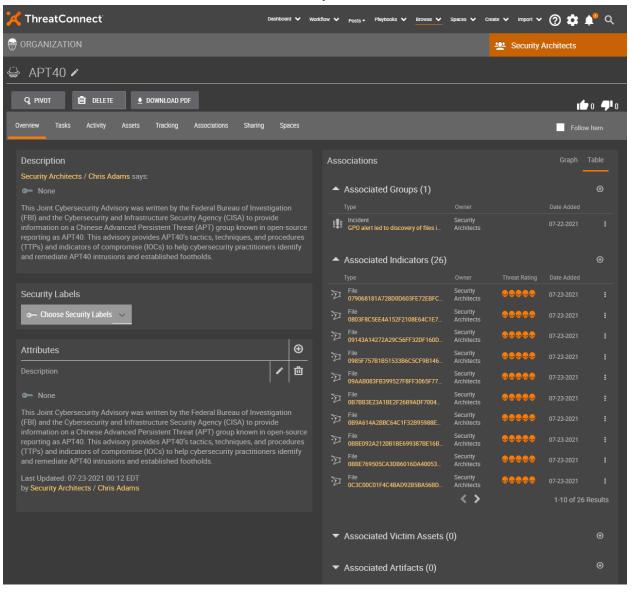




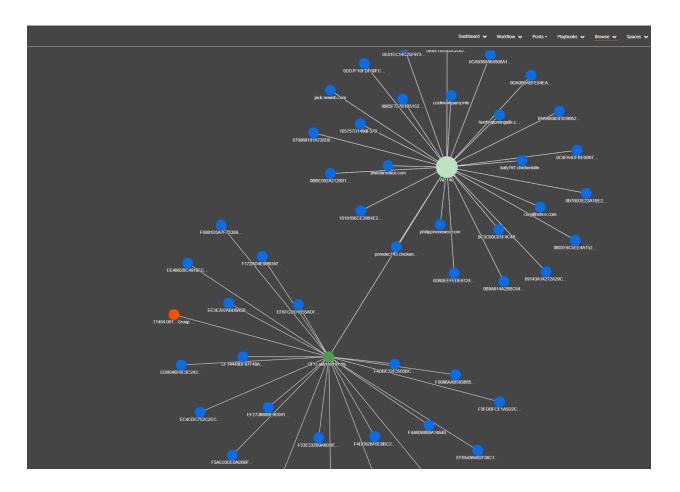
Google search brings me to https://us-cert.cisa.gov/ncas/alerts/aa21-200a

So I pop open my favorite browser, loaded with the <u>ThreatConnect Browser Extension</u> and begin to scan the page. There are a large number of indicators, hashes and URLs, that I should probably know about since they could be related to the indicator. I'm going to collect those indicators and associate them to the threat actor described in the alert.

8. Create a new Adversary APT40 and associate the collected indicators, hashes and URLs, then associate the Incident to the adversary.



You can now present your findings so far and begin the analysis of determining what actions to take based on your discovery.



The goal of this exercise was to get a sense for the utilities that ThreatConnect offers in order to collect intelligence from Ad Hoc sources and organize the data in a way that helps analysts understand the relationship between the items. As the last step using the visualizer depicts, many relationships can be formed at deeper levels that have associations that allow analysts to draw conclusions.

Situational Awareness

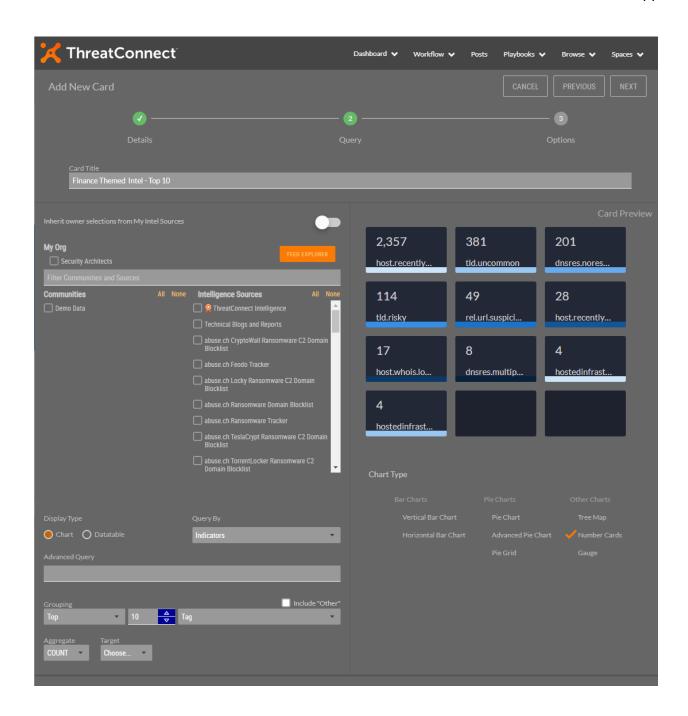
Trends, Priorities, Observations/FPs

Identify trends and stay abreast of what's important to your organization.

You want to track intel that can have an impact on your organization and keep abreast of trends in the industry, across all of your sources of data, including the intelligence created by your own teams. In the previous example, we collected files from a folder that seemed suspicious and were able to determine potential relationships with well-known threat actors in a few number of steps. The next obvious question becomes how do I present or track the information I'm processing for the remainder of the security team?

<u>Dashboards</u> give you the ability to do all of that and share findings, trends, and priorities with team mates and leadership. Dashboard cards can be built with wizards to easily produce the content or use <u>TQL</u> to create custom queries that get to specific perspectives.

Example of something I'm looking for (Search - saved) - data set unique for the sector or level of interest

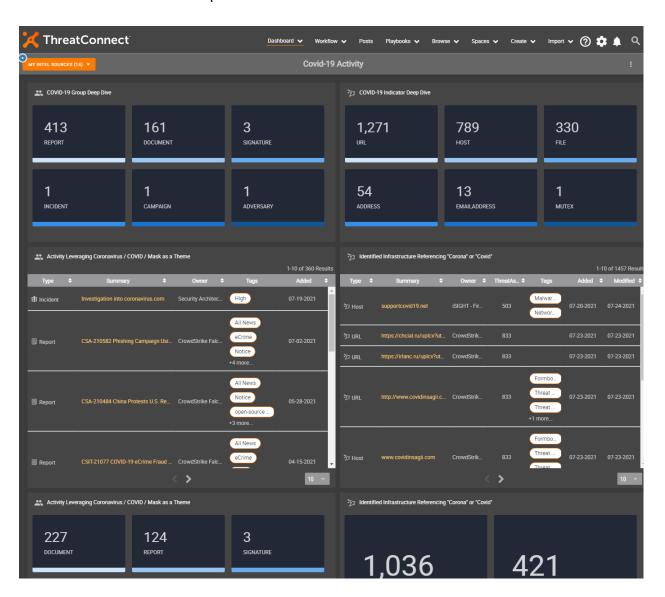


Identification of Cyber Threat Intelligence Requirements

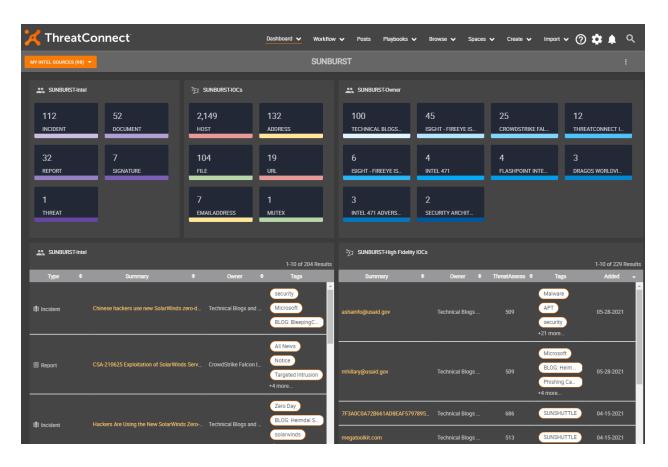
One of the primary functions of any threat intelligence team should be the successful establishment and fulfillment of individual priority intelligence requirements (PIRs). For example, what unique profiles, threat vectors, actors, exploits are you most concerned with that can impact your organization? Analysts can create content filters to collect and present the data sought for unique and focused intel cards.

Example, if your organization's security IAM solution is based on Microsoft AD, then clearly you want knowledge of AD, Windows-type vulnerabilities so you can know what to look for, get COA on how to defend or protect against them.

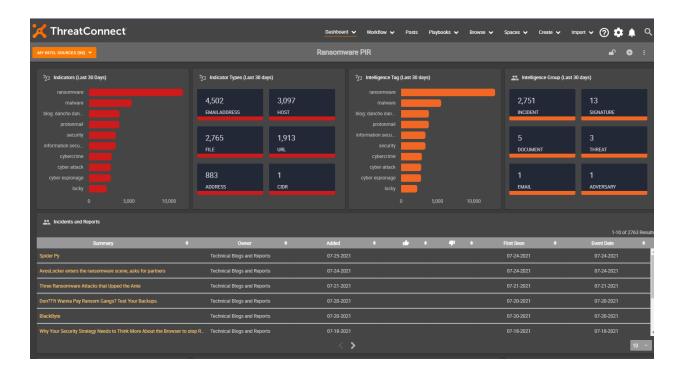
Perhaps your business is involved in the pharma sector and are concerned about threat actors using COVID-19 topics as a means to draw attention and clicks. Create a dashboard that centers on intel related to the pandemic.



Or perhaps your infrastructure may be vulnerable to the Sunburst threat vector.



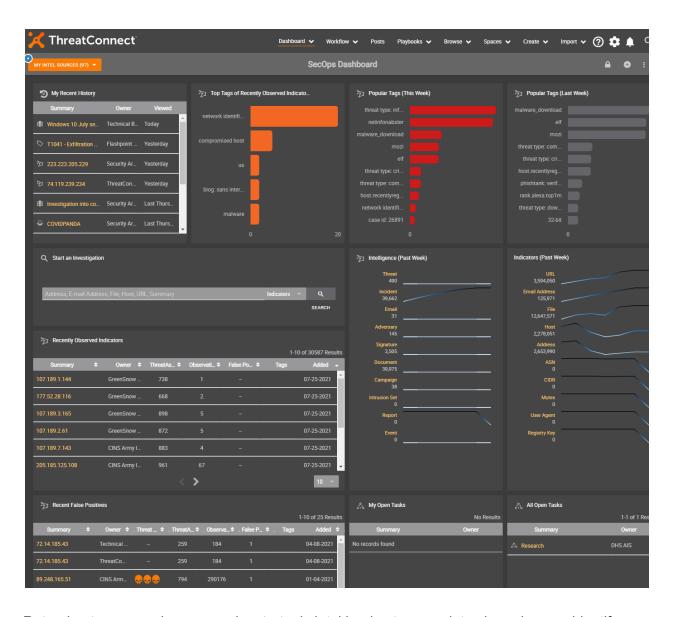
Or perhaps one of the main concerns from leadership is that a business partner just experienced a ransomware attack. Create a PIR dashboard that focuses on that topic of choice and recreate others using some of the same saved queries used to generate previous cards.



Threat intelligence must be able to answer these questions with actionable responses, enabling analysts to divide resources where they are most needed in a timely manner.

However, not that Operationalizing intelligence is where the rubber starts to meet the road. What intelligence is most relevant to your organization? Answer: That which is discovered in your network.

Observed indicators (referred to as *sightings* in STIX) signify that an IOC was found in security infrastructure and reported, in this case, in this Security Operations dashboard card named **Recently Observed Indicators**.



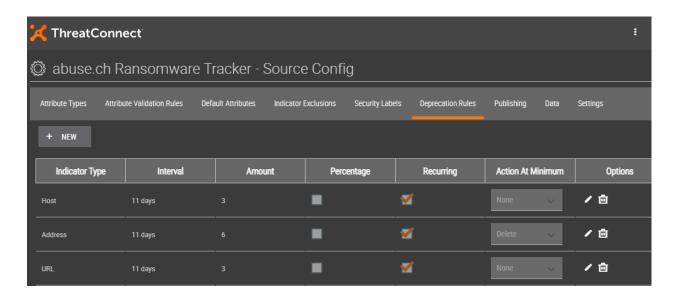
Returning to our previous example, strategic intel begins to come into view when we identify indicators observed in our network that have been associated with threat actors. That type of information becomes relevant to senior leadership, especially if the actor is known to exploit particular assets identified in your asset risk profile.



Managing a threat repository involves maintenance at the system-level, activities that are typically performed on the back end by DevOps for those who leverage dedicated cloud instances, and at the intel data-level. Data maintenance involves deprecation of stale data, altering the state of data between being active or inactive, tuning feeds based upon sightings and false positives, and monitoring feed growth over time.

<u>Indicator confidence deprecation</u> is a great way to allow ThreatConnect® Indicators to drop in Confidence Rating over time or be deleted if the Confidence Rating is not being maintained and updated. Confidence deprecation is used in the case of an Indicator, such as an IP Address, that is no longer being used for any malicious activity for a certain amount of time.

Below is an image of the configuration page where confidence deprecation is set, to the point of deletion if desired.



<u>Indicator Status</u> categorizes Indicators in ThreatConnect as being in one of two states:

• Active: The Indicator is considered to be an Indicator of Compromise (IOC) at the current time and should be treated in accordance with its ThreatAssess score.

• Inactive: The Indicator is not currently considered an IOC, but is being kept in ThreatConnect for historical accuracy rather than being deleted.

Knowledge of whether an Indicator is active or not allows ThreatConnect users to make more informed analytical choices and helps prevent them from wasting time and resources on Indicators that do not have any recent activity or are outdated.

When noting observables and FPs we begin to get a sense of what to focus on from an intelligence perspective, which leads us to characterize the sources of data set up in the instance. Reducing impact of false positives gives time back to your team to focus on things that matter, hence formulating a strategy for tracking where FPs are most commonly occurring should be considered by threat intel teams. Recall that administrators can adjust their consumption of feeds by turning them off within Feed Explorer should it be discovered over time that certain feeds are more of a nuisance than benefit.

Lastly, administrators may want to monitor the growth of feeds over time to get a sense of which feeds are providing in terms of volume. It's not uncommon for externally sourced feeds to go offline when their back-end systems are interrupted or fail, so knowing when they do is important.