zvelo

MAKING THE WEB SAFER
AND MORE SECURE

6 Steps to Driving Quantifiable Value From Cyber Threat Intelligence

zveloCTI WHITE PAPER | JANUARY 2022

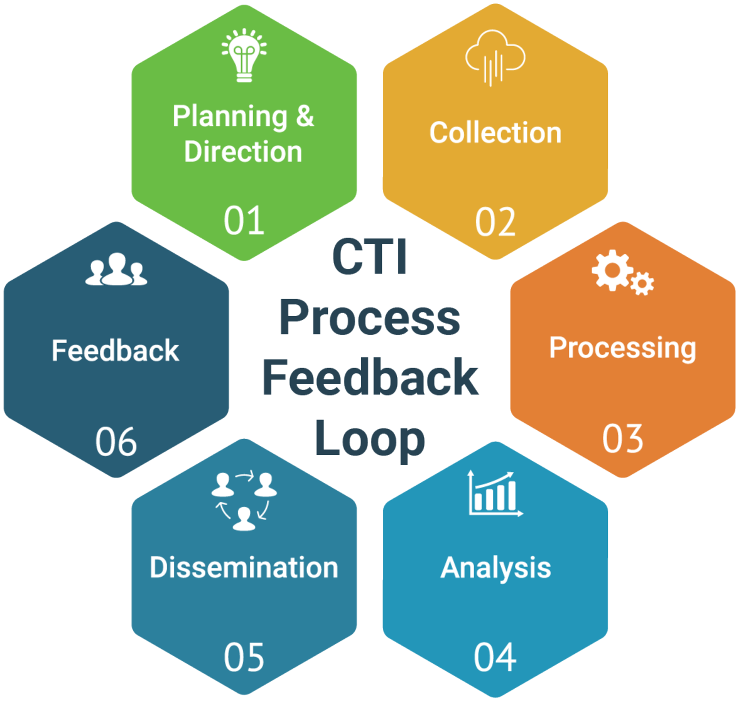
INTRODUCTION

CYBER THREAT INTELLIGENCE

If you have been around the cybersecurity business in the past 5-10 years, Cyber Threat Intelligence (CTI) has probably become part of your vocabulary. CTI is somewhat challenging to define, as its meaning within the Cyber Intelligence community varies by role, as well as by the value it delivers. To some, it is simply a feed of information about current threats. To others, it may be the latest report detailing the actions of an Advanced Persistent Threat (APT). Regardless of its varied meanings, CTI can be a 'firehose' of information that — *if not properly curated* — delivers little value.

So, what is CTI really? To establish a baseline understanding, zvelo defines CTI as a six step process feedback loop, adapted from military doctrine, to deliver quantifiable, and actionable intelligence to decision makers.

Cyber Threat Intelligence is generally defined as the methodical process of gathering and analyzing information about cyber threats. However, CTI is actually far more complex than that because in order to be considered 'intelligence', it must be tailored to each organization — otherwise, it's just information which may or may not be relevant. Furthermore, ***how*** an organization decides to use CTI is just as important as ***why***.



ZVELO SUPPORTS EVERY STAGE OF THE CTI PROCESS FOR MAXIMUM VALUE

1 | PLANNING & DIRECTION
Industry Expertise
zvelo's Internal Threat Intelligence experts collaborate with external sources to guide zvelo's global cybersecurity strategies.

2 | COLLECTION
99.9% ActiveWeb Coverage

- Traffic from 600+ million end users
- Third party data sources
- zvelo proprietary data

3 | PROCESSING
Meticulous Data Curation

- Extensive verification
- High veracity
- High accuracy
- Low false positive rate

4 | ANALYSIS
99%+ Accuracy
AI and Human Supervised Machine Learning assures maximum accuracy for web content categorizations.

5 | DISSEMINATION
Flexible Deployment Options

- API
- Data Feeds

6 | FEEDBACK
Global Partner Network
zvelo's extensive partner network enables a continuous feedback loop for real-time, dynamic updates of the ActiveWeb.

PLANNING AND DIRECTION

CTI requires that the data which is collected, processed, and analyzed, be very specific to an organization's unique environment and priorities to derive relevant context for understanding Malicious Cyber Actor (MCA) targets, attacks and motives. Additionally, organizations need to take into account the infrastructure required to utilize CTI. For example, ingesting a CTI feed is a good starting point but then what? Questions that need to be considered include:

- Where would the feed be housed?
- How would the feed be integrated (in SIEM, IDS/IPS, or other system)?
- Could analysts easily use the provided data?
- What would be the feed retention policy?

Regardless of consumption type (feed, API call, executive summary, detailed trend report and so on), the work required to create each follows the same process. If organizations do not have a plan when they procure CTI, there is not much value regardless of how it is designed to be consumed.

Intelligence that is not actionable is nothing more than noise. Related, intelligence that does not support the needs of a decision maker has little to no value. Ultimately, procuring or producing actionable intelligence comes down to well defined requirements. Ultimately, procuring or producing actionable intelligence comes down to well defined requirements developed during this first stage of the CTI Process.

As CTI is rooted in military doctrine, below are some terms which may, or may not, be familiar. Seasoned Military Intelligence personnel will tell you there are three types of requirements which must be defined to understand an Operating Environment (OE):

1. **Priority Intelligence Requirements (PIR).** What you need to know to complete the mission. Specifically, what do you need to know about the enemy?
2. **Friendly Force Information Requirements (FFIR).** What you need to know about your own forces.
3. **Commander's Critical Information Requirements (CCIR).** Key information needed to support decision making.

Intelligence that is not actionable is nothing more than noise. Related, intelligence that does not support the needs of a decision maker has little to no value. Ultimately, procuring or producing actionable intelligence comes down to well defined requirements.

zvelo can offer deep industry expertise during the planning and direction phase. Internal Threat Intelligence experts collaborate with external sources to guide zvelo's global cybersecurity strategies which can then align with partner requirements to maximize value.

PIRs, FFIRs, and CCIRs are tracked from initiation to completion, supporting decision points. Answers to PIRs are termed Essential Elements of Information (EEI). In military operations, these lists of intelligence requirements are considered living documents, as ever-changing as the OE. As requirements are answered, the Military Commander adds the next intelligence question to the list. If the answer to a requirement is not satisfactory, the Military Commander provides feedback to the Intelligence Team, and the cycle continues.

In general, CTI is used by organizations to enhance their defensive posture by understanding threats in relation to their Cyber Operating Environment (COE). With that in mind, the military-based definitions have been adjusted to better describe an organization's requirements for CTI.

1. **Priority Cyber Intelligence Requirements (PCIR).** What are the potential Cyber Threats to your organization? Be sure to include MCAs intent to cause harm.
2. **Friendly Cyber Information Requirements (FCIR).** What do you need to know about your organization's COE? From a Defense-in-Depth (DID) perspective — Perimeter, Networks, Endpoints, and Data — including related vulnerabilities.
3. **C-Suite Critical Cyber Information Requirements (C3IR).** What is the key information executive leadership must have to make decisions? For example, if you have a system with a critical vulnerability, is there a threat with intent to exploit it? And how will you know if, or when, a breach actually occurs?

Organizations looking to utilize CTI successfully, should make a risk-based determination of their Cyber Intelligence needs first, and then look for the feeds or sources which provide the most value. Unfortunately, many organizations buy one, two, or even more CTI feeds or sources without a clear understanding of what makes their COE an appealing target for MCAs. This results in blindspots in coverage, data overload, integration issues and worse — MCAs hanging out, undetected in their environment for 200+ days (per IBM, the global average for breach detection in 2019 was 206 days).

When it comes to CTI, each organization has a unique set of needs and priorities. While many organizations have the same or similar technology stack, the way each organization leverages that technology stack is significantly different. Every organization's COE is truly organic, and integrating CTI should be too. Also, keep in mind that CTI needs are vastly different across industry verticals. For example, the Finance sector

It is crucial to establish your organization's PCIR, FCIR, and C3IR before you even consider either procuring CTI sources/ feeds, or building your own.

Know your organization's COE, and establish the executive leadership's priorities to avoid the old adage, *'if everything is a priority, then nothing is a priority'*.

has distinct requirements which don't necessarily apply to the Energy sector — at least from a Cyber Threat perspective. After establishing a solid understanding of their requirements, every organization should, at minimum, ingest at least one general ***and*** one industry-specific CTI feed/ source.

CTI COLLECTION

Collection can be defined as gathering raw data from across the COE in various disciplines — Signals Intelligence, Open-Source Intelligence, Geospatial Intelligence, etc. The data collection strategy can vary depending on the data's intended purpose, as well as the Intelligence discipline used.

DATA COLLECTION METHODS AND SOURCES

In general, there are three ways to conduct CTI Collection: Third party data, self-sourced data, or a combination of the first two options.

Third-Party Data

Organizations can procure third party data feeds and information. These feeds can be open source or free but may come with the risk of higher False Positives (FPs). Alternatively, when accuracy is critical, third party data may be purchased from a trusted vendor for a premium.

Self-Sourced Data

Organizations often decide to gather data on their own from various internal sources, sensors, honeypots/nets, crowdsourcing, etc. The downside of this method is that the data collected must be stored, which can be costly.

Third-Party and Self-Sourced Data Combination

Many organizations employ a mix of feeds plus their own proprietary data. In theory, this may be the most effective approach to maximize CTI coverage. Research by the zvelo Cybersecurity Team shows both gaps and differences across the unique third-party feeds available. And, while combining data sources is unlikely to result in 100% CTI coverage, organizations should maximize sources as much as possible.

zvelo collects billions of data points across the ProActiveWeb, ActiveWeb, and InActiveWeb, in addition to its own proprietary threat detections, zvelo integrates multiple proprietary sources, open source and third party threat feeds so your organization benefits from having a single source collecting as much threat intelligence data as possible.

THE KEY TO COLLECTION IS VOLUME, VISIBILITY, AND LOCATION

Volume

It requires billions of raw data points to produce just a few hundred pieces of actionable intelligence. According to Cisco, Zettabytes (~1 Billion Terabytes) of data transits the internet every day. Since collecting the data from the whole of the internet is untenable, CTI providers target and gather the raw data crucial to specific needs — previously defined and documented in Priority Cyber Intelligence Requirements (PCIR). The figure below shows the relationship between data, information, and intelligence.

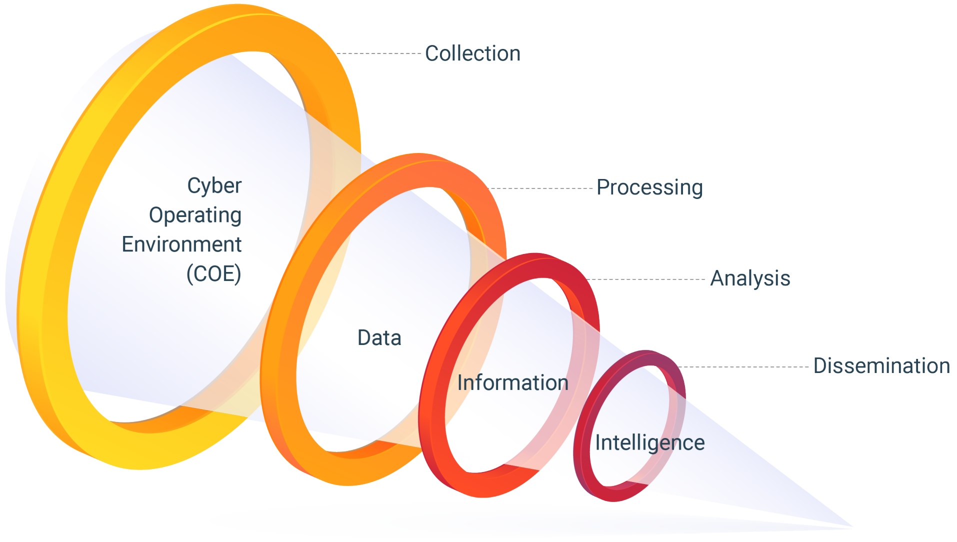
Visibility

Visibility is crucial because you must be able to “see” what it is that you want to collect. Unfortunately, Malicious Cyber Actor (MCA) obfuscation techniques, such as browser header tracking to redirect standard bots, often hinder visibility for CTI.

Location

Geographical location matters. For example, it is entirely possible that a site may show offline in North America but appear as perfectly functional in Europe — information which may result in actionable intelligence later in the process.

Relationship of Data, Information, and Intelligence



zvelo’s supported coverage includes active web traffic from 600+ million users across the globe, supplying the volume, visibility, and location data required to observe MCAs and potential threats as early as possible. These potential early detections provide key insights as the data moves into the processing phase.

CTI PROCESSING

Once an organization has amassed the necessary raw data about the COE, the next step is to process it. Joint Pub 2-0: Intelligence defines Processing as "a system of operations designed to convert raw data into useful information." The key words in this statement are 'useful information'. What makes information useful goes all the way back to the Planning & Direction stage, where the PCIR were defined. Similar to Collection, where it is not possible to gather every single piece of information in the COE, processing all of the raw data with the same degree of granularity would ultimately take so much time that information produced would no longer be actionable.

CTI Processing starts with data aggregation and normalization. An organization must aggregate all the data into a single information set which can then be placed into a common schema. This is a critical step as the CTI Analysis stage is dependent upon aggregation and normalization for the pattern analysis, assessment, and scoring which will ultimately produce actionable intelligence from the raw data.

PROCESSING METHODOLOGIES

There are multiple ways to process data in CTI. The three most common are Rules-based, Artificial Intelligence/Machine Learning (AI/ML), and Manual Analysis.

Rules-Based

Rules-based processing can be viewed as simple "business-logic". If a URL/domain meets specific criteria, certain actions are performed.

Artificial Intelligence/Machine Learning (AI/ML)

AI/ML is increasingly popular in the CTI arena — largely because humans just cannot keep up with processing billions of data points. However, unsupervised AI/ML engines are potentially subject to biases and False Positives (FPs), so it is best to deploy an AI/ML engine which incorporates human supervised Machine Learning.

Manual Analysis

As the most tedious and resource intensive method, manual analysis is used for edge cases when neither the rules-based nor AI/ML engines can reliably process an input. Manual analysis can be complicated and time consuming which does not meet the speed of today's MCAs.

CTI processing is a critical step as the CTI Analysis stage is dependent upon aggregation and normalization for the pattern analysis, assessment, and scoring which will ultimately produce actionable cyber threat intelligence from the raw data.

zvelo's methodology for CTI processing uses a balanced mix of rules-based, human supervised AI/ML, and manual analysis. The blended approach to CTI processing and validation ensures that zvelo's meticulously curated cyber threat intelligence feeds deliver more unique detections, faster and more accurately than other threat feeds on the market.

THE KEY TO PROCESSING IS DATA VALIDATION

Data validation is a crucial step necessary to derive actionable intelligence from the raw data. Unfortunately, it's also frequently skipped. Just because (pick an input, internal source, third-party feed, crowdsourced, etc.) one source claims something is malicious/phishing/evil, it may or may not be accurate. The old saying "trust but verify!" is particularly true when it comes to CTI processing.

THE ROLES OF THE COLLECTION AND PROCESSING STAGES IN THE CTI CYCLE

The CTI Collection and Processing stages are crucial to the Analysis stage, where the focus is to produce actionable intelligence. In addition to providing the necessary foundation for the following stages of the process, the CTI Collection and Processing stages reveal vast Intelligence opportunities hidden in 'Big Data'.

CTI ANALYSIS

Analysis is the stage of the CTI Process which directly follows Processing. As we have mentioned in previous posts, the roots of CTI are found in military intelligence, where we find the definition of analysis in Army Techniques Publication (ATP) 2-33.4: "Examine relevant information using reasoning and analytic techniques to reach a conclusion/determination."

For Analysis in CTI the key word in that definition is relevant. Assuming the useful information derived from the raw data was collected based on PCIR, FCIR, andC3IR — *which goes all the way back to Planning & Direction* — relevant information should be in hand.

The goal of Analysis is to take the useful information from Processing and extract Intelligence, ready for dissemination. So what happens to all the additional information which doesn't support PCIR, FCIR, or C3IR and, therefore, deemed not relevant? Is it simply thrown away? In short, it can be tossed, however, a better strategy is to store that information for later analysis as time allows. When trying to discover a needle in the proverbial haystack, an analyst needs stacks of hay to search.

Once data has been aggregated and normalized, zvelo processes it into information bins for web content categorization, malicious detections, and phishing detections, where the data will then be validated.

Throughout the course of a single day, zvelo analyzes several hundred million URLs using a hybrid analysis model. As part of the analysis, zvelo includes rich metadata attributes which are valuable for contextual relevance, forensic lookback, and incident response — just to name a few. zvelo maintains a multi-disciplinary team to conduct manual analysis when automated analysis is unable to provide an acceptable outcome.

ANALYSIS METHODOLOGIES

There are three basic types of analysis methodologies: Manual, fully-automated, and hybrid.

Manual Analysis

A human analyst takes the information provided and uses software tools, combined with intuition and experience, to discover the nuggets which will be used to create intelligence products. This methodology is not only time consuming, but fraught with the biases that any human being brings to the table. It is rare to find this type of analysis as the sole methodology used by organizations today.

Fully-Automated Analysis

Queue the machines…sort of. Fully-automated analysis takes advantage of a machine's inherent ability to sift through vast amounts of information very quickly. Fully-Automated Analysis can be rules-based, meaning the machine applies pre-determined guidelines to produce usable intelligence. Alternatively, Fully-Automated Analysis can utilize Machine Learning (ML) and/or Artificial Intelligence (AI) algorithms to deliver actionable intelligence. It is important to note that both concepts take time to get right and may still include the biases of those who created the algorithms.

Hybrid Analysis

Is a mix of automation and manual analysis techniques. In Hybrid Analysis, machines do the bulk of the work, while humans perform spot checks. When the machines cannot make a determination, it asks for "help" from a human analyst to do what people do best — focus on a single issue rather than an entire haystack. Most organizations today employ hybrid analysis.

CTI DISSEMINATION

After analysis is complete and the CTI products have been created, the next step is to disseminate those CTI products to consumers. Dissemination is the delivery of CTI products according to the format and timelines specified by the user. The explanation here is simple — if the customer updates their system with new CTI products every 24 hours, then delivery must support that requirement. If the customer is providing near 'real-time' protection to their users, then delivery must support that construct.

Once the data moves through zvelo's analysis process, the end result is high veracity threat intelligence feeds which can be immediately ingested into customer products or network defense systems to break the cyber kill chain by disrupting a variety of MITRE ATT&CK techniques in their Cyber Operating Environment (COE).

Related to delivery timeliness is product format. For example, CTI feeds are typically delivered in one of three formats:

Structured Threat Information Expression (STIX™)

STIX™ is divided into Domain Objects which organizes the data into manageable parts (https://oasis-open.github.io/cti-documentation/stix/intro.html).

Malware Information Sharing Platform (MISP)

MISP is an Open Standard for Threat Information Sharing which provides an attribute and categories/types construct for organization data (https://www.misp-project.org/).

Custom Schemas

Custom schemas defined by the organization providing CTI products for consumption. This means that the organization receiving products using a custom schema may have to transform them prior to use in their environment.

DISSEMINATION METHODOLOGIES

In the ever changing world of cybersecurity vulnerabilities, vectors, and threats, CTI products must be delivered in a timely manner. A variety of dissemination options make expedient delivery possible:

Flat File Downloads

CSV, JSON, spreadsheet, text files are made available to the customer.

API

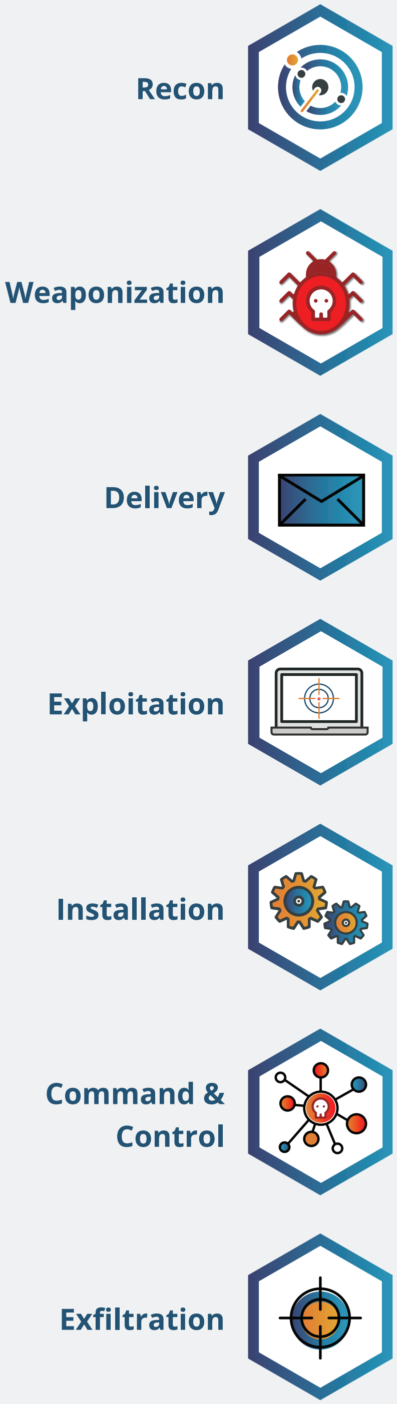
Programmatic access to products allowing customer access to pull the data they want based on types, time ranges, and other parameters.

Feeds

Automatically pushes products to customers in an agreed upon format.

Accessing any of these dissemination methods should be done via the appropriate secure authentication capabilities.

zveloCTI Solutions are designed to target and break the Kill Chain to disrupt attacks *before* they make it to the ActiveWeb.



CTI FEEDBACK

The final piece of the CTI Process is Feedback. In practical terms, feedback is a dialogue between the intelligence producer and the intelligence consumer, or customer. Feedback should be collaborative. Producers must be willing to listen to feedback and the customer must be willing to offer it. Feedback should also be push and pull. In other words, the customer should push feedback to the producer and the producer should proactively reach out to the customer to pull feedback.

Feedback can also be viewed as helping the CTI producer to assess the value of the products delivered. This assessment is typically viewed in two bins:

- **Measure of Performance (MOP):** Quantitative metrics on how much/ how many CTI products were produced and delivered.
- **Measure of Effectiveness (MOE):** Qualitative assessment of how well the CTI products produced and delivered fulfilled organization PCIR, FCIR, and/or C3IR.

Analyzing both MOP and MOE feedback can help the CTI producer tune their pipeline to ensure the most actionable intelligence is delivered.

zvelo is a highly collaborative organization that regularly reviews feedback from our customers. For example, zvelo maintains 24/7 coverage of our web content categorization efforts to quickly review and correct any classification errors. zvelo's threat feeds include methods for both interactive and offline feedback from customers. We measure both MOP and MOE for threat feeds and drive that information back into the CTI process for loop continuation. This ensures that zvelo continuously refines collections to stay ahead of Malicious Cyber Actors (MCAs).

zveloCTI™ SOLUTIONS

zveloCTI provides curated, high veracity, actionable data on phishing and malicious domains that Threat Intelligence teams can ingest for analysis and enrichment.

PHISHBLOCKLIST

zvelo's AI-powered Phishing Intelligence detects phishing threats within the ActiveWeb traffic and other sensor-based data streams to deliver a richly packaged feed of highly curated and validated phishing threats that are enriched with additional metadata attributes like date detected, targeted brand, and other crucial data points.

MALICIOUS DETAILED DETECTION FEED

zvelo's Malicious Intelligence identifies, confirms, and enriches malicious URLs for direct action by defenders and analysts. The data feed includes specific malicious URLs, as well as a range of metadata attributes such as date detected, malware family, and other significant intelligence data attributes.

THREAT INTELLIGENCE REPORTING

zvelo delivers meticulously curated, actionable Threat Intelligence data with industry leading accuracy and low false positive rates for maximum precision, efficacy and protection against evolving and emerging threats.

- Malicious Intelligence Reporting
- Phishing Intelligence Reporting
- Brand Threat Intelligence Reporting
- Exposed/Exploited Assets Reporting
- Additional Intelligence Reporting Available

ABOUT ZVELO

zvelo's passion is to make the internet safer and more secure by providing the industry's premium Cyber Threat Intelligence and web classification data services.

LEARN MORE

To learn more about zveloCTI or to request an evaluation, please contact us at **sales@zvelo.com** or visit **www.zvelo.com**.

zvelo

www.zvelo.com
cybersecurity@zvelo.com