BLACK CELL
Protecting critical infrastructures

The Role of OSINT in Cyber Threat Intelligence

Table of Content

1. Introduction

1.1 The importance of CTI

Nowadays, when cyber threats are constantly evolving in complexity and scale, the need for strong cybersecurity measures has never been more important in the operation of an organization. Cyber Threat Intelligence (CTI) has emerged as a crucial component in the defense gear of organizations striving to protect their digital assets and maintain operational integrity. This whitepaper delves into the essence of CTI, and describes its operating mechanisms and applications utilized by Black Cell.

Cyber Threat Intelligence is the collection, analysis, and dissemination of information about potential and current threats posing risks to an organization's cyber infrastructure. It encompasses a wide range of data points, including indicators of compromise (IoCs), tactics, techniques, and procedures (TTPs) used by threat actors, and contextual information about the motivations and capabilities of them. The goal of CTI is to provide actionable insights into these areas, which helps organizations anticipate, identify, and respond to cyber threats more effectively.

As adversaries in cyber environment become more sophisticated, relying solely on reactive security measures is no longer enough. CTI empowers organizations to shift from a reactive to a proactive stance, enhancing their overall security posture. By leveraging CTI, organizations gain:

- **Enhanced Situational Awareness**: Real-time insights into the threat landscape enable organizations to stay ahead of adversaries and preempt potential attacks.
- **Improved Resilience**: With timely and actionable intelligence, organizations can better withstand and recover from cyber incidents, minimizing downtime and operational disruptions.
- **Informed Decision-Making**: CTI equips security teams and executives with the knowledge needed to make strategic decisions that align with their risk tolerance and business objectives.

1.2 CTI services provided by Black Cell

Currently, we provide the following CTI services:

Surface Scan
The purpose of the surface scan is to identify threats and vulnerabilities in cyberspace that pose risks to a given organization. The goal is to reduce the attack surface by minimizing the

weaknesses and vulnerabilities found during the reconnaissance phase of a cyberattack against the organization.

Managed Brand Intelligence
Brand monitoring aims to actively monitor and manage an organization's presence in cyberspace. This includes contextual monitoring of online conversations, posts, media, news articles and activities that affect the organization - and its reputation - and identifying threats and vulnerabilities that pose a risk to the organization in cyberspace.

Managed Supply Chain Intelligence
The purpose of supply chain monitoring is to identify threats and vulnerabilities in cyberspace that pose a risk to the partners and suppliers of the organization, and thus to identify the factors that potentially mean threat to the organization itself through an actor in the supply chain. Typical supply chain threats are: typosquatted domain, social engineering, e-mail spoofing, spear phishing, whaling, breaching of sensitive databases, compromise of a user account belonging to a supplier/partner.
Compromise of a partner or supplier may expose the organization to severe financial, intellectual, reputational damage or legal consequences.

Sector Specific MITRE ATT&CK Heatmap
As a part of our MITRE GAP analysis service – which is a type of audit that uses MITRE ATT&CK Framework based data sources, tactics, techniques and mitigations to map the security readiness of an organization's environment, providing actionable intelligence and mitigation steps – we prepare the sector specific heatmap documentation of the organization based on what market sector it belongs to. During this process, we use various open-source resources to discover, identify and assess the cyber threats that a specific market sector faces.

Dedicated Cyber Threat Intelligence Platforms
Dedicated CTI softwares automatically gather, analyze, disseminate, correlate and visualize information about potential cyber threats, helping organizations identify and mitigate risks of their digital assets and IT infrastructure. These services are provided via third-party platforms, which include the subscription-based Recorded Future, Secutec, and the free, open-source MISP Threat Sharing.

2. Methodology

2.1 Open Source Intelligence

On one hand, the foundation of our CTI services is built on the methodology of Open Source Intelligence (OSINT). The method involves the collection and analysis of data from publicly available sources on the internet – including the clear-, deep- and darkweb -, and it has proven to be an invaluable component of CTI, since the severity of the vulnerabilities thus identified

at an organization is illustrated by the fact that these tools, by their very nature, can only extract data that are freely available to the public, allowing a knowledgeable attacker to gain critical information about vulnerabilities in a system with minimal resource investment and without being detected. By leveraging OSINT, organizations can gain a comprehensive understanding of their threat landscape, allowing them to improve their cybersecurity defenses and eliminate vulnerabilities in their IT infrastructure.

The various sources of the data collection includes forums, social media, news articles, academic journals, published reports, technical databases, open source search engines, and more.

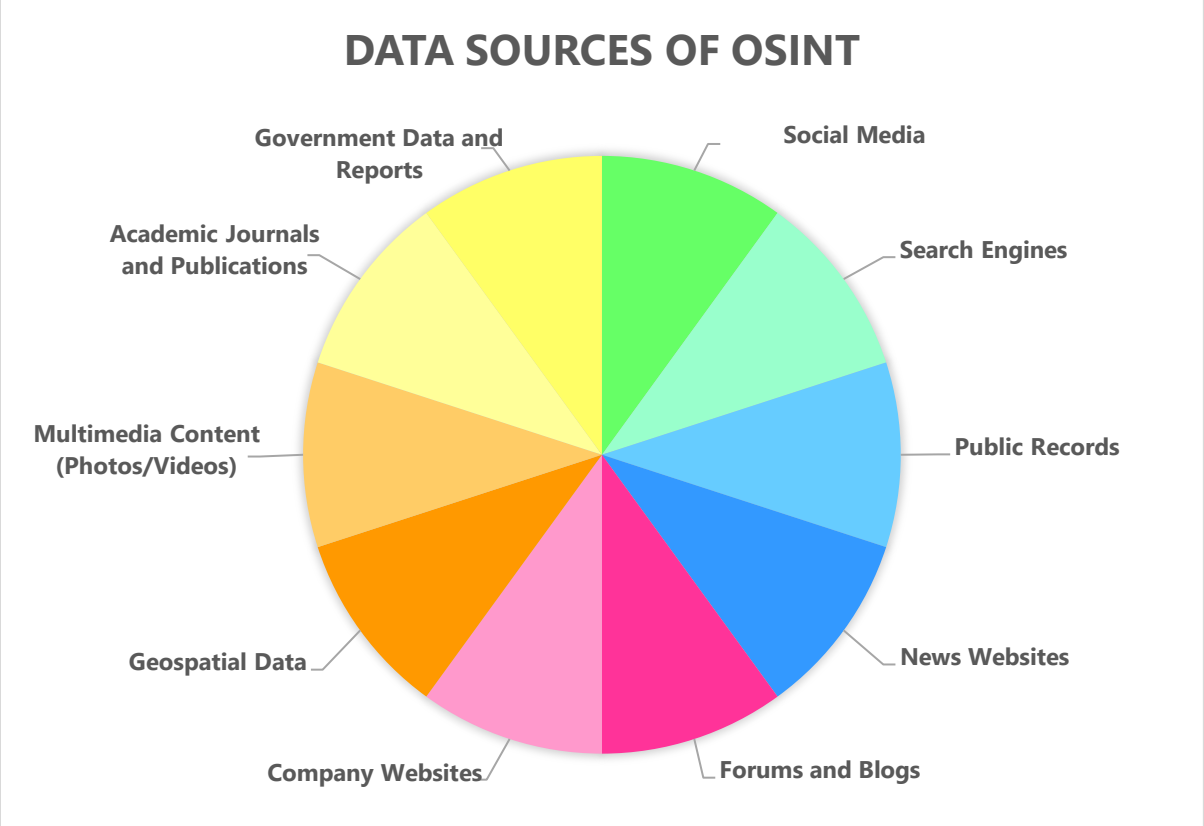

Figure 1: Data Sources of OSINT

The effectiveness of OSINT lies behind the following advantages:

- **Accessibility and Cost Efficiency**: The vast amount of publicly available information means that organizations can gather a diverse range of data without allocating significant financial and human resources to the task.
- **Timeliness**: Information on the internet is continuously updated, providing real-time insights into current and emerging threats. OSINT allows organizations to stay

informed about the latest cybersecurity related trends and adjust their security measures accordingly or detect new vulnerabilities in their digital infrastructure.

- **Early Warning**: OSINT can function as an early warning method by identifying potential threats before they escalate into full-blown attacks. Monitoring discussions in deep- and darkweb forums can reveal planned cyber operations, or following zero-day vulnerability alerts allows organizations to take preemptive measures just in time.
- **Contextual Understanding**: OSINT provides context that is crucial for understanding the broader implications of a threat. For example, technical databases can shed light on the capabilities of specific malware strains or the impact of a particular vulnerability.

2.2 The process of OSINT based investigation

When conducting an OSINT investigation, we draw on the same tools, methods and technologies that attackers use in the initial stages of preparing their offensive during the reconnaissance stage of the Cyber Kill Chain, such as mapping the target's infrastructure - both IT and human - from open-source data, looking for potential weaknesses and vulnerabilities. The information gathered in this way allows attackers to understand the defensive measures to be bypassed and to determine the operational techniques required to penetrate the target's system, as well as the possible methods of delivering arbitrary payloads and malware. This method of weak point reconnaissance provides them anonymity through the freely available open-source technical databases and search engines, furthermore makes the process extremely resource-efficient regarding the time and costs that need to be allocated to it.

The most important step in the preparation of an attack is the discovery of the target's infrastructure, which includes the identification of vulnerable domains, web interfaces, applications, open ports, and company related vulnerabilities such as leaked sensitive documents, data or websites that are intended for internal use but are still accessible freely from the Internet. In addition to the technical weak points comes the human risk. It's often said that the weakest link in an organization's cybersecurity is the human being - the employee - who, due to lack of adequate security awareness or negligence, can easily fall victim to social engineering attacks, use weak passwords for their user accounts, use their company e-mail address for private purposes or handle internal sensitive data and documents with insufficient care.

The subject of the investigation encompasses the following areas:

Domain discovery

When conducting an OSINT investigation on organizations, the first and most crucial step of the process is domain discovery, which begins with identifying the root domains related to the specific organization. Using both active and passive DNS discovery engines, we map the structure of the root domains by discovering subdomains and IP addresses associated with them, in addition to the relevant WHOIS information. While passive DNS research tools collect and analyze historical DNS data by observing DNS traffic rather than actively querying DNS servers, active ones engage in sending DNS queries to the domains to gather information about them and their associated records in real-time. This intrusive nature of the latter allows us to discover subdomains that might be hidden from the scope of passive DNS search engines – and the public -, due to being presumably important assets of the organization.

Knowing the root domains, we also run a typosquat-check on them to see if there are fake, often phishing sites that misleadingly resemble the original websites and domain names of the investigated domains in the wild. If the attacker's intention was to steal user data, they could clone a login page found on the original domain and use a website with a similar (typosquatted) domain name to create a targeted phishing campaign to trick employees of the organization into entering their credentials in a fake login interface that looks exactly like the original website. Another example that often occurs in targeted social engineering attacks, is where attackers create an email address with a domain name that is deceptively similar to the original domain but differs in one or two characters (e.g. Google.com - Goog1e.com) and contact their target with the intention to steal sensitive data or even money from them by pretending to be the original entity.

Vulnerability scan

We run an automated vulnerability scan on the IP addresses and subdomains identified during the domain discovery process. The scan checks for open ports, running services, service banners, SSL/TLS certificates, HTTP headers and content, device information, geolocation, operating system, hostnames and domains. Based on the banner information and software versions thus received, it runs them against CVE databases and provides us the possible vulnerabilities that might be found on the investigated system with CVE numbers and description.

Web surface discovery

Based on the result of the domain discovery process, manual investigation of the subdomains is conducted, where we are looking for possible weak points like having a developer, test or internal system, login page or by any means misconfigured and vulnerable site facing the internet. We also compare the result of the passive and active DNS search engines to find out

which subdomains can be discovered only by the enumeration of the root domain, and weigh their exposure to risks by manually investigating if they are accessible freely from the internet, and whether they contain any of the above-mentioned weak points.

We also utilize different types of clearweb crawlers and advanced search strings - which use advanced search operators to find information that is not directly available on a certain website but is still accessible to the public if the path to the information happened to be found out this way - to search for vulnerable web interfaces that could be exploited (e.g. admin or any login pages related to internally used systems, developmental or misconfigured webpages, accidental sharing of sensitive data in source code, or SQL injection web vulnerability).

Nowadays, with the rapidly increasing number of social engineering attacks, sensitive login pages facing the internet pose significant security threats to the organization as they offer several possibilities for an attacker to compromise their target. For example, with knowledge of leaked user credentials they could easily gain unauthorized access to internal systems or even admin permissions undetected, or in the absence of such access, by performing brute-force attacks or exploiting identified web vulnerabilities, thereby, inter alia, gaining access to sensitive information or performing malicious activities on the affected systems.

Sensitive documents and information discovery

Using the clearweb crawlers and advanced search strings mentioned in the previous section, we look for uploaded files on the domains, especially the ones with PDF, Word document, Excel table, text or compressed file extensions. By investigating their content, we are able to find internal, confidential, or other documents containing sensitive information. If the file itself seemingly does not contain sensitive data, it can still provide us interesting information about the organization in its metadata, if it had not undergone metadata cleanup prior to publication. Useful metadata can tell for example the author's full name or username, e-mail address, geolocation, as well as the name and software version of the producer application, which, when interpreted in context with information from other sources – like CVE databases in case of softwares -, could form the basis for sophisticated social engineering attacks.

Other interesting source of useful information about an organization's IT infrastructure is their job advertisement. If it contains a job description that tells what databases, softwares or other IT systems and solutions are present in the organization, it can be a starting point for an attacker to look for vulnerabilities in that certain system or to create a phishing email using the properties of the application (e.g. if a vulnerability is identified in a piece of software, the attacker may try to exploit it or, in the case of phishing attack, could use for example, a fake

software update or password change request to trick employees into providing them sensitive information).

Data Leaks

Data leaks are generally shared on the deep- and darkweb in various text collections - for free or in exchange for payment. Using deep- and darkweb search engines with the proper search keywords, we identify the e-mail addresses belonging to the investigated domain that were affected by data breach.

By investigating the content of these data leak collections, we are often able to identify the source of the breach, which in most cases originate from the company itself – e.g. as a result of database compromise -, third-party websites which the user had registered on with company e-mail address, or the user's own device infected with credential stealer malware.

News, forums and social media

By using the previously mentioned deep- and darkweb search engines along with the clearweb crawlers and advanced search strings, we are able to monitor the mentions of the organization on different platforms like news articles, forums and social media pages. For example, we can observe the reputational impacts of opinions and events or the planning phase of possible attacks on the organization, or even employees displaying compromising behavior publicly online, such as exfiltrating sensitive data while admitting to having worked for a specific organization.

2.3 Sector Specific MITRE ATT&CK Heatmap

The main goal of the documentation is to identify and evaluate those cyber threat tactics, techniques, and procedures (TTPs) that pose potential risk to the organization by targeting or having history of major attacks in the market sector it belongs to. Knowing these specific TTPs, the organization can prepare preventive and reactive measures in advance to improve its security posture.

As the very first step, information gathering about the potential threats takes place in the clear-, deep- and darkweb, using various CTI platforms, crawlers and specialized query languages. This involves identifying the most relevant sector specific cyber incidents and threat actors, tools, malwares and attack procedures used during the attack, and determining the inadequacies of the victim.

The next step is quantifying and evaluating the identified threats based on the collected information. Each cyber threat is given an Impact, Evasion, Complexity and Successfulness score on a scale of 1-5, which results are then summed and multiplied with an Accuracy multiplier on a scale of 0.5-1.5, based on how certain and confident we are in our findings. This will be the final score of the specific threat and the TTPs belonging to it.

The last step is mapping all these final scores to MITRE ATT&CK techniques and normalizing the end results to a scale of 1-7 by severity, then displaying the end result on a graphical heatmap.

Execution	Persistence	Privilege Escalation	Defense Evasion	Credential Access	Discovery
Cloud Administration Command	Account Manipulation	Abuse Elevation Control Mechanism	Abuse Elevation Control Mechanism	Adversary-in-the-Middle	Account Discovery
Command and Scripting Interpreter	BITS Jobs	Access Token Manipulation	Access Token Manipulation	Brute Force	Application Window Discovery
Container Administration Command	Boot or Logon Autostart Execution	Account Manipulation	BITS Jobs	Credentials from Password Stores	Browser Information Discovery
Deploy Container	Boot or Logon Initialization Scripts	Boot or Logon Autostart Execution	Build Image on Host	Exploitation for Credential Access	Cloud Infrastructure Discovery
Exploitation for Client Execution	Browser Extensions	Boot or Logon Initialization Scripts	Debugger Evasion	Forced Authentication	Cloud Service Dashboard
Inter-Process Communication	Compromise Host Software Binary	Create or Modify System Process	Deobfuscate/Decode Files or Information	Forge Web Credentials	Cloud Service Discovery
Native API	Create Account	Domain or Tenant Policy Modification	Deploy Container	Input Capture	Cloud Storage Object Discovery
Scheduled Task/Job	Create or Modify System Process	Escape to Host	Direct Volume Access	Modify Authentication Process	Container and Resource Discovery
Serverless Execution	Event Triggered Execution	Event Triggered Execution	Domain or Tenant Policy Modification	Multi-Factor Authentication Interception	Debugger Evasion
Shared Modules	External Remote Services	Exploitation for Privilege Escalation	Execution Guardrails	Multi-Factor Authentication Request Generation	Device Driver Discovery
Software Deployment Tools	Hijack Execution Flow	Hijack Execution Flow	Exploitation for Defense Evasion	Network Sniffing	Domain Trust Discovery
System Services	Implant Internal Image	Process Injection	File and Directory Permissions Modification	OS Credential Dumping	File and Directory Discovery
User Execution	Modify Authentication Process	Scheduled Task/Job	Hide Artifacts	Steal Application Access Token	Group Policy Discovery
Windows Management Instrumentation	Office Application Startup	Valid Accounts	Hijack Execution Flow	Steal or Forge Authentication Certificates	Log Enumeration

Figure 2: Close-up detail of a heatmap from green being the lowest to red being critical severity techniques in a specific sector

2.4 Dedicated Cyber Threat Intelligence Platforms

2.4.1 Recorded Future

Recorded Future is an automated threat intelligence cloud platform focused on collecting and analyzing open-source information on cybersecurity threats. It works with various data sources such as public forums, news portals, social media or the darkweb, and then analyzes, aggregates and visualizes the information extracted from these sources.
Its services thus enable users to detect and interpret events and trends in cyberspace, and to forecast potential threats and attacks. The service is designed to help organizations actively defend against threats in cyberspace and respond to incidents in a timely manner.

Three modules of the platform are utilized by Black Cell:

- **Threat Intelligence** enables threat analysts to perform their most important functions, including identifying the actors who most actively threaten the organization, understanding attackers' motive and targets, investigating and documenting their TTPs, and tracking macro trends that affect the organization, including trends relevant to its industry and the regions where it operates. It pinpoints the most relevant threats, reduces the time analysts spend researching them, and generates more intelligence about them — often from sources that would be difficult or impossible for analysts to find and access on their own.[1]
- **Brand Intelligence** helps organizations detect and mitigate external risks to their brand. Real-time alerts on these external brand-related threats are packed with valuable context including screenshots, automatic logo detection, and technical assessments of DNS records, WHOIS data, and certificate data, and more to ensure organizations are able to respond confidently and quickly.[2]
- **SecOps Intelligence** focuses on providing tactical and operational insights, focused on the immediate threat landscape, including data collection from the darkweb and other sources. It helps in identifying potential threats, tactics, techniques, and procedures used by cyber attackers.[3]

2.4.2 Secutec

SecureSIGHT is Secutec's managed threat intelligence service, which monitors and collects cyber threat risk factors for an organization through automated monitoring and manages them with a dedicated security team.

The monitoring components are the following:

- **Attack surface management**: identifying and managing all potential vulnerabilities and weaknesses in the organization's digital environment that could be exploited by cyber attackers to reduce the overall threat risk of the organization.

[1] https://www.recordedfuture.com/products/threat-intelligence
[2] https://www.recordedfuture.com/products/brand-intelligence
[3] https://www.recordedfuture.com/products/secops-intelligence

- **Leaked Credentials Monitoring and Darknet Monitoring**: identifying signs of impending attacks and potential data leaks to proactively prevent account abuse of the organization.
- **Active Threat Hunting**: actively hunting for indicators of compromise to detect advanced threats often overlooked by conventional systems, and continuously adjusting to the evolving cyber threat landscape for rapid response to security incidents.
- **Managed XDR services**: round-the-clock advanced threat detection, real-time visibility, rapid incident response, compliance, and ease of use.[4]

2.4.3 MISP

MISP[5] (Malware Information Sharing Platform & Threat Sharing) is a free, open-source threat intelligence platform for sharing, storing and correlating IoCs of targeted attacks, threat intelligence, vulnerabilities, malware analysis, financial fraud or even counter-terrorism information. Its functionalities can be used to detect and prevent attacks, frauds or threats against ICT infrastructures, organizations or people while being shared within and between the connected organizations and communities.

MISP's key features are the following:

- **IoC Database and Correlation**: it is able to store technical and non-technical information about incidents, attackers, malware samples or other intelligence, and automatically correlate these collected data if relationship is found between them to help identify complex attack campaigns or emerging threats.
- **Extensive Data Types**: it supports a wide range of data types related to threats, e.g. IP addresses, domain names, e-mail addresses, file hashes, URLs, and more detailed contextual information about threat actors and attack campaigns.
- **Flexibility and Customizability**: it provides metadata tagging, custom feeds import, visualization, dashboards and even allows users to integrate with other tools for further analysis due to its data formats and open protocols.
- **Integration and Automation**: it offers APIs and integration capabilities that allow organizations to automate the ingestion of the threat intelligence data into various

[4] https://secutec.eu/securesight/
[5] https://www.misp-project.org/

security tools and systems, such as SIEMs, intrusion detection systems (IDS), and endpoint protection platforms.