



INEFFICIENCY IN THREAT INTELLIGENCE

An exploration of the inefficiency in threat
intelligence sharing platforms

Finbar Downes
B00119229

Contents

| | |
|---|----|
| Table Of Figures | 3 |
| Abstract..... | 4 |
| 1. Introduction | 5 |
| 1.1 Chapter Overview | 8 |
| 2. Literature Review | 11 |
| a. Introduction | 11 |
| b. Scope and Purpose..... | 12 |
| c. Search Strategy | 13 |
| 2.1 What is threat intelligence? | 13 |
| 2.2 What is threat intelligence data?..... | 15 |
| 2.3 Where is threat intelligence data found? | 17 |
| 2.4 How is threat intelligence used?..... | 20 |
| 3. Threat Intelligence Sharing Online Platforms | 25 |
| 3.1 VirusTotal | 25 |
| 3.2 IBM X-Force Exchange..... | 27 |
| 3.3 AbuseIPDB..... | 28 |
| 3.4 Examples of malicious analysis on each platform | 29 |
| 4. Methodology..... | 31 |
| 4.1 Project Development | 31 |
| 4.1.1 VirusTotal API Key | 31 |
| 4.1.2 AbuseIPDB API Key..... | 32 |
| 4.1.3 IBM X-Force Exchange API Key..... | 33 |
| 4.1.4 Python | 34 |
| 4.1.5 Python Flask | 35 |
| 4.1.6 HyperText Markup Language (HTML) | 35 |
| 4.1.7 Cascading Style Sheets (CSS)..... | 36 |
| 4.2 EasyIntel | 36 |
| 4.2.1 EasyIntel Showcase | 37 |
| a. Data Collection | 40 |
| b. Data Analysis | 40 |
| c. Limitations..... | 41 |
| 5. Results And Interpretation..... | 42 |
| 5.1 Testing IP results on VirusTotal, IBM X-Force and AbuseIPDP | 42 |
| 5.2 Spreadsheet of results | 42 |

| | | |
|-------|--|----|
| 5.2.1 | Understanding why these threat intelligence platforms different? | 46 |
| 5.3 | SOC Analyst Threat Intelligence Survey | 47 |
| 5.3.1 | Survey Questions | 47 |
| 5.3.2 | Survey Answers | 48 |
| 5.4 | Results Summary..... | 52 |
| 5.5 | Interpretation..... | 53 |
| 6. | Conclusion..... | 55 |
| 7. | Further Work..... | 57 |
| | Bibliography | 61 |

Table Of Figures

| | |
|---|----|
| Figure 1: Threat Intelligence Data..... | 16 |
| Figure 2 Tactical TI in a SIEM environment..... | 22 |
| Figure 3 Sample SOAR Workflow with Threat Intelligence..... | 24 |
| Figure 4 Virus Total Malicious Scan | 29 |
| Figure 5 IBM X-Force Exchange Malicious Scan..... | 30 |
| Figure 6 AbuseIPDB Malicious Scan | 30 |
| Figure 7 Easy Intel - Login Page..... | 37 |
| Figure 8 Easy Intel - Sign Up Page | 37 |
| Figure 9 Easy Intel - Home Page..... | 37 |
| Figure 10 Easy Intel - IP Reputation Checker Page | 38 |
| Figure 11 Easy Intel - Bad IP Results | 38 |
| Figure 12 Easy Intel - Bad File Hash Results..... | 39 |
| Figure 13 Spreadsheet of IP testing results | 44 |
| Figure 14 All Threat Intelligence Platforms Graph..... | 44 |
| Figure 15 VirusTotal vs IBM X-Force IP Results..... | 45 |
| Figure 16 AbuseIPDB vs IBM X-Force IP Results | 45 |
| Figure 17 AbuseIPDB vs IBM X-Force IP Results | 46 |
| Figure 18 Survey Question - How often do you use threat intelligence? | 48 |
| Figure 19 Survey Question - Select which threat intelligence platforms you use? | 49 |
| Figure 20 Survey Question - Do you find discrepancies in threat intelligence..... | 50 |
| Figure 21 Survey Question - Which platform do you find to be most accurate? | 51 |

Abstract

This thesis explores the landscape of online threat intelligence sharing platforms, focusing on their unique methodologies and the discrepancies that arise among them. It scrutinizes three renowned platforms – VirusTotal, IBM X-Force Exchange, and AbuseIPDB – and sheds light on the intricacies of their operations. Further, it delves into the root causes of the discrepancies observed, ranging from differences in data collection and analysis methodologies to varying levels of user expertise and technological capabilities.

In response to the identified discrepancies, this thesis introduces EasyIntel, a tool designed to bridge the gap among various threat intelligence platforms. EasyIntel assimilates data from multiple platforms, providing users with a comprehensive overview of potential cyber threats.

However, the thesis acknowledges the scope for further work in this domain. It advocates for broadening the research scope to include a wider range of threat intelligence platforms, both established and emerging. Additionally, it recommends a deeper exploration into the causes of discrepancies among different platforms, with the aim of improving the accuracy and relevance of threat intelligence data.

Lastly, the thesis proposes several refinements to the EasyIntel tool. These include enhancing the user interface, incorporating additional functionalities such as file uploading and URL scanning, leveraging machine learning algorithms for better threat detection, and introducing features like a general threat score and bulk IP searches.

The overall objective of this thesis is to provide a holistic understanding of the current threat intelligence landscape, explore the gaps therein, and propose a comprehensive, user-friendly, and effective solution through EasyIntel.

1. Introduction

To safeguard businesses and individuals from potential cyberattacks, effective threat intelligence techniques have had to be developed in the connected world of today due to the rising incidence of cyber threats. It is impossible to emphasize the significance of accurate and trustworthy threat intelligence as hackers become craftier and more persistent in their tactics. The goal of this dissertation is to examine the gaps in threat intelligence online sharing platforms and how they might affect the platforms' overall capacity to strengthen cyber security capabilities.

To achieve this aim, the following objectives have been established:

- **Identify and examine the key features of popular threat intelligence sharing platforms:** This objective involves conducting a thorough investigation of various threat intelligence sharing platforms, such as VirusTotal, IBM X-Force Exchange, and AbuseIPDB. The goal is to understand their strengths, weaknesses, and unique characteristics, which will provide a basis for comparing their approaches to threat intelligence sharing.
- **Analyse the potential discrepancies in threat intelligence data across these platforms:** Building on the first objective, this step entails conducting a detailed analysis of the threat intelligence data provided by each platform. This will involve identifying any inconsistencies, gaps, or contradictions in the information shared, which may hinder the effectiveness of threat intelligence for cyber defence. By pinpointing these discrepancies, the research can shed light on potential areas of improvement for each platform and the threat intelligence sharing ecosystem as a whole.
- **Evaluate the implications of these discrepancies for the users of threat intelligence platforms:** The final objective is to assess the consequences of the identified discrepancies for organizations and individuals who rely on threat intelligence sharing platforms to protect their digital assets. This will involve exploring how these discrepancies might impact decision-making processes, resource allocation, collaboration efforts, and overall cyber defence strategies. Ultimately, this analysis will help determine the extent to which discrepancies in threat intelligence sharing platforms may undermine their effectiveness as a tool for cyber defence.

By addressing these objectives, this dissertation will provide valuable insights into the current state of threat intelligence sharing platforms and their role in combating cyber threats. Furthermore, it will highlight potential areas for improvement and inform the development of best practices for enhancing the effectiveness of these platforms in the ever-evolving landscape of cybersecurity.

The central hypothesis of this research is that discrepancies in threat intelligence online sharing platforms may undermine their effectiveness as a tool for cyber defence. This hypothesis stems from the concern that inconsistent or conflicting information across various platforms could lead to confusion, misinterpretation, or inadequate response to cyber threats. To thoroughly investigate this hypothesis and its implications, the following research questions have been formulated:

- **What is the nature and extent of discrepancies in threat intelligence online sharing platforms?** This question seeks to identify and categorize the various types of discrepancies that may exist across different threat intelligence sharing platforms. Examples of discrepancies could include differences in the classification of threats, inconsistencies in the interpretation of data, or variations in the types and formats of information provided. By understanding the nature and extent of these discrepancies, researchers and practitioners can gain insights into the areas where improvements can be made to enhance the overall effectiveness of these platforms.
- **What factors contribute to the discrepancies observed in these platforms?** This research question aims to uncover the underlying causes of discrepancies in threat intelligence sharing platforms. Potential factors could include differences in the methodologies used to collect and analyse data, variations in the expertise of platform users, or limitations in the technologies and tools employed by each platform. A deeper understanding of these factors can help to inform the development of strategies and best practices for reducing discrepancies and improving the overall quality of threat intelligence sharing.
- **How do these discrepancies impact the efficacy of threat intelligence sharing platforms?** The final research question examines the consequences of the identified discrepancies for the users of threat intelligence sharing platforms.

This includes exploring how discrepancies might lead to inefficient resource allocation, hindered collaboration, or delayed responses to cyber threats. By evaluating the real-world implications of these discrepancies, the research can provide valuable insights and recommendations for enhancing the effectiveness of threat intelligence sharing platforms and, ultimately, improving the cyber defence capabilities of organizations and individuals relying on these platforms.

This dissertation will make several notable contributions to the field of cyber threat intelligence, providing valuable insights and practical solutions for enhancing the effectiveness of threat intelligence sharing platforms. The key contributions are as follows:

- **Comprehensive analysis of discrepancies:** The research will present a thorough examination of the discrepancies in threat intelligence sharing platforms, identifying inconsistencies, gaps, and contradictions in the data provided by these platforms. This comprehensive analysis will serve as a foundation for understanding the current state of threat intelligence sharing and pinpointing areas that require improvement, ultimately leading to more effective cyber defence strategies.
- **Exploration of underlying causes:** By investigating the root causes of these discrepancies, the dissertation will help practitioners better understand the limitations of existing platforms. This exploration will delve into various factors such as differences in data collection and analysis methodologies, variations in user expertise, and technological constraints. A deeper understanding of these factors will enable the development of targeted strategies for mitigating discrepancies and improving the overall quality of threat intelligence.
- **Recommendations for addressing discrepancies:** Based on the findings of the research, the dissertation will provide actionable recommendations for addressing the identified discrepancies. These recommendations will guide practitioners and platform developers in implementing measures to enhance the effectiveness of threat intelligence sharing platforms, ultimately contributing to a more robust cyber defence ecosystem.

- **Development of EasyIntel tool:** As a practical outcome of this research, a tool called EasyIntel has been developed to consolidate results from popular threat intelligence platforms. EasyIntel streamlines the process of gathering threat intelligence data by providing users with a single interface that queries multiple platforms simultaneously. By aggregating information from various sources, EasyIntel facilitates a more comprehensive and efficient analysis of cyber threats, mitigating the impact of discrepancies between platforms. The development and implementation of EasyIntel exemplify a tangible solution for addressing the challenges identified in the dissertation, demonstrating the real-world applicability of the research findings.

1.1 Chapter Overview

Chapter 2:

The literature review delves into a comprehensive analysis of existing research, publications, and industry reports on threat intelligence platforms and their accuracy. This chapter is structured into several sections, each focusing on different aspects of threat intelligence platforms and their evaluation.

This chapter provides an overview of threat intelligence, its importance in the cybersecurity landscape, and the different types of threat intelligence, such as strategic, operational, tactical, and technical intelligence. It also discusses the various sources of threat intelligence, including open-source, commercial, and industry-specific feeds.

Chapter 3:

The Threat Intelligence Sharing Online Platforms chapter offers a comprehensive analysis of three prominent online threat intelligence platforms: VirusTotal, IBM X-Force Exchange, and AbuseIPDB. The chapter delves into the unique aspects of each platform, discussing their main features, benefits, and use cases. By understanding the capabilities of these platforms, cybersecurity professionals can better evaluate and leverage the tools that best suit their needs.

Chapter 4:

This chapter delves into the various tools and techniques employed in the project, starting with the APIs utilized: VirusTotal, AbuseIPDB, and IBM X-Force Exchange. Each API's key features are outlined, from file and URL scanning, access to historical data, IP and domain analysis, to providing comprehensive threat intelligence, promoting collaboration, and offering security tool integration.

The chapter then discusses the choice of Python as the programming language for the project, citing its extensive libraries, simpler code, strong community support, and easy API implementation. Python Flask, a lightweight Python web framework, was chosen for web server hosting due to its simplicity, flexibility, and strong support for user authentication and authorization.

The project's front-end technologies, HTML and CSS, are also presented. HTML, the main language of the internet, is used for organizing and designing the contents of the webpages. Meanwhile, CSS brings style to the HTML structure, contributing to cleaner code and more manageable web projects.

Chapter 5:

This chapter presents the results of tests conducted on three main threat intelligence platforms: VirusTotal, IBM X-Force, and AbuseIPDB. A disparity in flagged IPs between the platforms was observed, indicating differences in sensitivity and accuracy. IBM X-Force flagged fewer IPs, suggesting lower sensitivity, while AbuseIPDB had a more sensitive scoring system. VirusTotal demonstrated a balanced approach, potentially indicating high accuracy.

The chapter also includes a survey among cybersecurity peers and co-workers, shedding light on the popularity and perceived accuracy of each platform. VirusTotal emerged as the most popular choice, while both VirusTotal and AbuseIPDB were perceived as the most accurate. This underscores the importance of using multiple threat intelligence platforms for a comprehensive understanding of potential threats, and not relying solely on popularity or brand recognition.

Chapter 6:

This concluding chapter of the thesis provides a summarization and reflection on the in-depth investigation conducted on the discrepancies within online threat intelligence sharing platforms, primarily VirusTotal, IBM X-Force Exchange, and AbuseIPDB. The importance of accurate and timely threat intelligence in the contemporary digital era is emphasized, with a focus on the role it plays in building solid cyber defense strategies.

The chapter further addresses the discovered inconsistencies within the threat intelligence data provided by the platforms. This inconsistency is presented as a significant issue given its potential to disrupt effective cyber defense strategies. It may lead to misinterpretations and erroneous responses to cyber threats.

Chapter 7:

This chapter highlights the potential for further exploration and improvement in the field of threat intelligence platforms and the EasyIntel tool.

The first section underlines the need to broaden the scope of platform analysis, suggesting that future research includes a wider array of threat intelligence platforms. This expansion could reveal additional insights and discrepancies in the threat intelligence landscape.

The next section addresses the causes of discrepancies among different threat intelligence platforms. It emphasizes the importance of understanding these root causes to enhance the overall quality and relevance of threat intelligence data.

The final section focuses on refining the EasyIntel tool. It proposes several improvements such as enhancing the user interface, incorporating additional functionalities, utilizing machine learning algorithms, and introducing features like a general threat score and bulk IP searches. The ultimate aim is to ensure that EasyIntel continues to evolve and meet the diverse needs of the cybersecurity community.

2. Literature Review

a. Introduction

Threat intelligence has been increasingly important in the realm of cybersecurity in recent years. Organizations must modify their defensive tactics as cyber threats continue to develop and grow more sophisticated in order to safeguard their important assets. Utilizing threat intelligence is one strategy that aims to give organizations the knowledge they need to proactively defend against cyberattacks. Threat intelligence, however, is a vague topic with many different definitions and interpretations in the literature. By addressing a number of important issues, this review aims to give readers a thorough knowledge of threat intelligence.

First, we will examine the numerous definitions of threat intelligence found in the literature as we investigate the idea. By doing so, it will be easier to grasp what threat intelligence is and is not. The sorts of data that make up threat intelligence will next be discussed, along with the main sources from which it is derived. This will make it possible to comprehend the inputs that power threat intelligence systems and procedures better.

Then, to present a more comprehensive picture, we will investigate how threat intelligence is applied in real-world scenarios, highlighting some of its most important uses in the field of cybersecurity. This will involve a review of the most widely used platforms for exchanging threat intelligence as well as a performance and accuracy evaluation of each. We'll also go through the elements that affect these platforms' effectiveness, which can aid businesses in choosing the best solution for their unique requirements.

Finally, a Security Operations Centre (SOC) environment's use of threat intelligence will be discussed in this review. We'll look at how threat intelligence is incorporated into a SOC's routine activities and talk about the advantages and difficulties of doing so. This review seeks to provide a thorough overview of threat intelligence by looking at these various facets of it, allowing organizations to make educated judgments regarding its adoption and application in their cybersecurity plans.

b. Scope and Purpose

With an emphasis on its sources, applications, accuracy, and well-known sharing platforms, this literature review seeks to present a thorough overview of the current state of research on threat intelligence. We want to outline our goals for this review and investigate the breadth of the literature on this subject. The concept of threat intelligence, the sources of threat intelligence data, the practical applications of threat intelligence, the reliability of platforms for sharing threat intelligence, and threat intelligence in a SOC context will be the main topics of our review. We will study the material that has been published over the previous ten years and analyse a variety of sources, including academic journals, trade periodicals, and government reports.

The following aims will guide our review:

- What is threat intelligence, and how is it defined in the literature?
- What constitutes as threat intelligence data?
- Where is threat intelligence data found, and what are the primary sources of this information?
- How is threat intelligence used in practice, and what are its key applications in cybersecurity?
- What are the most used threat intelligence sharing platforms?
- How accurate are threat intelligence sharing platforms, and what factors affect their performance?
- Threat intelligence in a SOC environment

Our review seeks to provide answers to these concerns in order to advance knowledge of the complicated challenges surrounding threat intelligence and to contribute to continuing discussions regarding its application, dependability, and importance in the field of cybersecurity. Understanding the significance of threat intelligence in defending enterprises and people is essential as the digital world develops and cyber-attacks become more complex. The idea and uses of threat intelligence will be clarified through this review, laying the groundwork for further study and advancement in the area.

c. Search Strategy

We started by outlining our research questions and identifying the main ideas and search phrases associated with threat intelligence, cybersecurity, and data sources in order to construct the search strategy. To find information, we used a variety of databases, including Google Scholar, CORE, BASE, and Google search.

The emphasis was on finding literature that covered several facets of threat intelligence, including its definition, sources, uses, and sharing platforms. We looked for papers and publications that discussed the accuracy and dependability of platforms for exchanging threat intelligence as well as assessments and comparisons of well-known platforms in the industry. We also looked for literature examining threat intelligence's larger effects on the cybersecurity landscape.

We were able to identify scholarly literature on the subject using Google Scholar and other academic search engines, and we were able to locate relevant trade journals and government reports using the Google search engine. In order to ensure that our literature evaluation includes a diverse variety of viewpoints and insights into the topic of threat intelligence, a thorough search approach was developed.

2.1 What is threat intelligence?

The first question I felt that I needed to answer was what is threat intelligence? I mainly wanted to answer this question to get a more definitive and stronger understanding of what threat intelligence is down at source level.

Threat intelligence can come from a vary of different sources and can be confusing to really understand what is correct and what is just noise. “*Cyber Threat Intelligence – Issue and Challenges*” by Md Sahrom Abu at the Universiti Teknikal Malaysia Melaka, describes it as “*intelligence from professional perspective as data that has been refined, analysed, and processed and the output must be relevant, actionable, and valuable. Those three requirements can be achieved through logical and analytical process conduct by human that can provide contextual data and produce useable output*” (Abu, 2018). My understanding of this statement is that Md Sahrom Abu characterizes threat intelligence as analyzed, and processed data, yielding relevant, actionable, and valuable output that can be used to determine bad actors. In my opinion, this definition

emphasizes the importance of human involvement in the logical and analytical processes that ensure the contextual relevance of the data, ultimately producing usable output.

However, as I delved deeper into the paper, it became apparent that the lack of a universally accepted definition for Cyber Threat Intelligence (CTI) allows companies and individuals to interpret and define it in their own ways. This lack of clarity may lead to ambiguity and confusion when trying to understand the true meaning of threat intelligence. *“Cyber Threat Intelligence (CTI) has become a hot topic in Information Security (IS) but the lack of literature review on clarifying the concept and companies tend to use their own definition to distinguish their product may lead to some ambiguity”* (Abu, 2018).

The main summary used by Md Sahrom Abu in this paper was *“As ambiguous as it can be, Cyber Threat Intelligence (CTI) can be define comprehensively as evidence-based knowledge, including context, mechanisms, indicators, implications and actionable advice, about an existing or emerging threat that can be used to inform decisions regarding the subject's response to that menace or hazard”* (Abu, 2018), which gives a more clear and concise understanding of what threat intelligence actually is, evidence-based advice, proven by indicators that show their intent.

Looking at another paper *“A Comparative Analysis of Cyber-Threat Intelligence Sources, Formats and Languages.”* by Andrew Ramsdale, Stavros Shiaeles, and Nicholas Kolokotronis with the Plymouth University and The University of Peloponnese, they talk about how CTI is known as *“the information and knowledge gained about an adversary through observation and analysis.”* (Andrew Ramsdale, 2020) Again, to me this reinforces the point that threat intelligence is information gained through different types of feeds.

NIST released a *“Cyber Threat Intelligence – Issue and Challenges”* with code “800-150”, this was a publication on “Guide to Cyber Threat Information Sharing” and offers some good clarity and understand of threat information under section 2 of the paper.

“Threat information is any information related to a threat that might help an organization protect itself against a threat or detect the activities of an actor” (Chris Johnson, 2016) I feel like this is the simplest and most universally acceptable definition of threat intelligence.

We can see that these definitions are widely backed and supported by many different large security companies like Crowdstrike, the first thing we see in this article is *“Threat intelligence is data that is collected, processed, and analysed to understand a threat actor’s motives, targets, and attack behaviours”* (Baker, 2023) this is exactly comparative to what we have discovered earlier through papers, and its coming from one of the biggest threat intelligence companies out there.

In summary, the idea of "Cyber Threat Intelligence" (CTI) is a crucial component of contemporary information security, acting as a vital tool for enterprises to make educated decisions about their response to cyber-attacks.

It is obvious that CTI incorporates evidence-based knowledge obtained from a variety of sources, which is then processed and analysed to give context, mechanisms, indicators, consequences, and actionable advice on existing or new threats. This is true even though there is no single definition that is universally accepted.

The findings from Md Sahrom Abu's study, those from the research paper by Ramsdale, Shiaeles, Kolokotroni and NIST, as well as the definition offered by top security firm Crowdstrike, all highlight the significance of CTI in today's increasingly interconnected digital ecosystem.

2.2 What is threat intelligence data?

I felt like “What is threat intelligence data” is another important question to help with 360 degrees understanding of threat intelligence, if you don’t know what it is, how can you understand it fully?

So, we already know threat intelligence is data collected from a variety of sources that helps understand different threats and threat actors. But what exactly is the data that help do that?

A paper published by Gartner called *“How to Use Threat Intelligence for Security Monitoring and Incident Response”* by Micheal Clark and Augusto Barros, does a really good job at explaining the answer to this question.

They give a great overview example of some of the types of threat intelligence seen:

“Examples of threat intelligence include:

- *IP address of a command and control (C&C) server*
- *MD5 hash of a malicious executable with context*
- *Report on a threat actor who is known to target the financial sector*
- *Network intrusion prevention system (NIPS) rule that will detect malware communications delivered as part of a report” (Michael Clark, 2020)*

We can see that threat intelligence comes in many ways, from simple IP addresses of known IOCs to high level reports on threats.

They continue to explain this threat intelligence in two different categories “Strategic TI” and “Tactical TI”, “*Strategic TI includes reports and other human-readable products on threat actors and their intentions*” (Michael Clark, 2020), meaning that strategic TI is more human designed and human digested information, such as reports, they go on to explain it is usually associated with long term security. On the other hand, Tactical TI, was described as “*Tactical TI often consists of IOCs, such as IP addresses, domains, URL or hash lists, and other system-level or network-level artifacts. These artifacts can be matched to what is observed on information systems*” (Michael Clark, 2020), and this type of TI is usually used in shorter term, to identify immediate threats. They included this nice graph to help better understand this.

| | Strategic | Tactical |
|-----------------------------|---|---|
| Created By | Humans using technical and nontechnical sources | Machines or humans |
| Consumed By | Humans | Machines and humans |
| Delivery Time Frame | Days to years | Seconds to hours |
| Useful Life Span | Long | Usually short |
| Resistance to Change | Durable | Fragile |
| Focus | Planning and high-level decisions | Detection, triage and response |
| Examples | Information targeted, organization affiliation of the threat actor, intentions, preferred tools and threat actor profiles | IP, domain, URL, MD5, hostname and filename |

Source: Gartner (February 2020)

Figure 1: Threat Intelligence Data

Furthermore, to support this idea of “Strategic” and “Tactical”, Microsoft give a similar definition in more detail with 4 headings:

- Strategic
- Tactical
- Operational

- Technical

They explain how “Strategic” threat intelligence is at a higher level, like what was mentioned in the Gartner paper, which is meant for higher level people such as “*C-Suite and non-technical stakeholders*”. This type of information is to help greater understand certain threats in terms of certain business decisions.

They then explain how “Tactical” threat intelligence is more for cybersecurity experts and are usually “*tactics, techniques, and procedures (TTPs) and IOCs*”, like what was mentioned in the Gartner paper also. They explain that this information is usually used to “*make decisions about security controls and create proactive defense strategies*”.

What I liked about the Microsoft definition was they had two more categories, which helped break it down slightly more and give a better understanding.

They spoke about “Operational” threat intelligence “*knowledge about specific threats and campaigns. It provides specialized information for incident response teams about an attacker’s identity, motivations, and methods*”, and “Technical” threat intelligence as “*signs that an attack is happening—such as IOCs*” (Michael Clark, 2020).

I feel these provide a comprehensive understanding of the diverse nature of threat intelligence data. Both the Gartner paper by Clark and Barros and Microsoft's definitions help to categorize this data into distinct types, such as strategic, tactical, operational, and technical intelligence. In my opinion, the Microsoft approach, which includes more categories, offers a more nuanced understanding of threat intelligence data. This further division helps to understand what the data is and how it should be interpreted.

2.3 Where is threat intelligence data found?

I felt like “Where is threat intelligence data found” is another question that I could benefit from answering, it helps to strengthen and back up the earlier definitions of threat intelligence and give a well-rounded understanding of threat intelligence as a whole.

Threat intelligence can be found or gathered in many different places, understanding how and where it is from can help further understand it. Upon research, the first method I found to discover threat intelligence data is to extract it from hacker forums on the

web and / or dark web, In “*Extracting Cyber Threat Intelligence From Hacker Forums*” by Isuf Deliu at the Norwegian University of Science and Technology, they talk about the importance of this and using “elite” level hackers to gain good threat intelligence data, “*The most important group, the so-called "elite" constitute a third category who not only can understand and use the existing tools and techniques, but are also able to create new methods that can be exchanged with others*” (Isuf Deliu, 2017). This data can be very valuable and useful but, in my opinion, this can possibly lead to false positives and misinformation depending on the level of hacker, even a high level, it’s possible there is misinformation been spread.

Upon further research, in the paper “*Darknet and Deepnet Mining for Proactive Cybersecurity Threat Intelligence*” by Eric Nunes of Arizona State University, they undertook some experiments on marketplaces on the darknet, “*Darknet marketplaces provide a new avenue to gather information about the cyber threat landscape. The marketplaces sell goods and services relating to malicious hacking, drugs, pornography, weapons, and software services.*” (Eric Nunes, 2016), These marketplaces can help with gaining some threat intelligence data and are rather like forums in my opinion, with both sharing some of the same features and vendors.

In the paper “*A Comparative Analysis of Cyber-Threat Intelligence Sources, Formats and Languages.*” by Andrew Ramsdale, Stavros Shiaeles, and Nicholas Kolokotronis with the Plymouth University and The University of Peloponnese, they talk about internal sources of threat intelligence gathering, “*The CTI obtained from internal sources is comprised of observable events that have happened on an organisation’s internal network and hosts*” (Andrew Ramsdale, 2020), this means that threat intelligence data can be collected from internal events that took place on the network, using logs it is possible to collect important and accurate data of Tactics, Techniques, and Procedures (TTPs). They give examples of some things that are helpful to gain data, such as Anti-Virus Systems, Network Intrusion and Prevent Systems, Web Application Firewalls, SIEMs etc.

Reading further through this paper, we can also learn about external sources of threat intelligence gathering, which can provide priceless information and data to add to an organisations threat intelligence, in the paper, they identify and compare different free to use threat intelligence feeds online, “*These community, open-source IoCs and*

observables typically consist of the observed malicious sources or data, e.g., IP address, domain, URL, file names and hashes” (Andrew Ramsdale, 2020), most of these threat intel feeds, provide really good data and can be used in analysis to further understand a certain actor. Although, from my experience, sometimes the data provided can be inaccurate or unreliable, and needs to be crosschecked.

Again, referring to the same paper again “*A Comparative Analysis of Cyber-Threat Intelligence Sources, Formats and Languages.*”, they talk about open-source intelligence (*OSINT*), this type of threat intelligence, mainly done by human research or machine learning, can include:

“News articles covering ongoing threats

Alerts and advisories

Using search technologies to find vulnerable systems: Google dorks, Shodan, etc.

Information, alerts, news feeds on malware activity and threats

Monitoring communication channels for intelligence: Slack, IRC, Twitter, etc.

Intelligence available directly from the criminal underworld” (Andrew Ramsdale, 2020)

OSINT covers some of the methods we discussed earlier such as, dark net forum and marketplace browsing. They explain and discuss the advisory standards, such as Common Vulnerabilities and Exposures (CVE) and Common Weaknesses Enumeration (CWE), Common Vulnerability Reporting Framework (CVRF) & Common Vulnerability Scoring System (CVSS), All of these are very important to understand what type of threat intelligence you are receiving and the severity of certain vulnerabilities at hand.

In conclusion, the process of gathering threat intelligence is multifaceted, with various methods and sources available for organizations to utilize. Hacker forums and darknet marketplaces offer unique insights into the cyber threat landscape and the tactics, techniques, and procedures (TTPs) employed by malicious actors. However, the reliability and accuracy of information obtained from these sources can be questionable, requiring verification to ensure its utility.

Internal sources, such as logs from network systems and security tools, provide valuable data on events and incidents within an organization's infrastructure. These internal sources can be more reliable, as they are based on observable events within the organization.

External sources, including open-source threat intelligence feeds and open-source intelligence (OSINT), offer diverse insights into the broader cyber threat landscape. OSINT covers a wide range of data sources, from news articles and alerts to information obtained directly from criminal communication channels. Advisory standards, such as CVE, CWE, CVRF, and CVSS, further help organizations in understanding the nature and severity of the threats they face.

While each source and method has its merits and drawbacks, a comprehensive and proactive approach to threat intelligence gathering should involve a combination of these sources. By utilizing multiple methods, organizations can gain a broader understanding of the threats they face, while also being able to cross-verify the information for accuracy and reliability. In my opinion, the source I have found to be the best, comes in the form of Internal Sources, as these are active, legitimate threat actors at the current time.

2.4 How is threat intelligence used?

This is kind of an open-ended question, that I feel could have 1000 answers, but my main objective is to understand in what places is threat intelligence used and how it is used in certain places such as a Security Operations Center.

The first paper I looked at was a SANS Whitepaper called *“Threat Intelligence: What It Is, and How to Use It Effectively”* by Matt Bromiley`. Where he talks about different ways to incorporate threat intelligence into an organisation. The first thing he talks about is incorporating threat intelligence into an organisations posture, explaining different ways it can be used *“For example, a hospitality chain may receive TI about an attack group that is targeting vulnerable payment card systems. The chain can then build continuity and contingency plans around high-value, essential targets”* (Bromiley, 2016), this would come as invaluable data to the given organisation in my opinion and helps demonstrate the value of threat intelligence. Matt also goes on to talk about how

an organization can use threat intelligence to identify critical assets that it had not before considered critical or high value *“Let’s say that the hospitality chain receives TI showing that an attack group is attempting to compromise reservation systems to gain information on potential victims’ whereabouts. Previously, reservations may have been seen as arbitrary data and not a critical asset.”* (Bromiley, 2016) Although I am not 100% sure on this example, as I believe most organisations would consider this highly confidential data.

He then goes on to talk about using threat intelligence to help drive investigations and response, and this is the main topic I wanted to know about, he talks about how threat intel can be used to help distinguish certain hacks or hacking groups, which can help incidence response teams know what to do, or what to look out for in the network “TI can guide the IR team on which hosts to examine, what type of malware to look for, and what methods an attacker might be using to maintain persistence after infected machines have been powered down”. I personally can agree that this threat intelligence is very helpful in my role as a SOC analyst, where I would use threat intelligence to further understand threats.

The last point he raises is to do with using threat intelligence to look into the future and protect against future attacks. He discusses how threat intelligence can be used to identify and understand the trend in the attacks moving forward, referencing how PowerShell has become more adopted by hackers and how organisations can use this information to further protect themselves *“Organizations can use the change in trends via TI analysis to identify what may be the next attack vector”* I feel like threat intelligence can be a seriously good help in this type of use, as in security, its most important to get out in front of as many attacks as possible.

In the paper *“How to Use Threat Intelligence for Security Monitoring and Incident Response”* by Micheal Clark and Augusto Barros, they talk about the way to use “Strategic and Tactical” threat intelligence, which we discussed earlier. They explained how Strategic threat intelligence is rolled out in a more traditional report method and is usually assessed and used to strengthen the security of organisations with the data, *“strategic TI can influence an organization’s security posture, leading to a more qualitative impact on the organization’s overall security position”* (Michael Clark, 2020). They also talk about how, usually this type of intelligence is used at a higher

level and would generally have less quantitative metrics and become more of a story. *“Quantitative metrics become less relevant the higher up the reporting chain you go. Instead, a better strategy is to form a story around the metrics using the context provided from the sources of threat intelligence”* (Michael Clark, 2020).

The use of tactical TI is more complex than traditional strategic TI and can be very useful if it's used in a timely manner. The first thing they talk about in this paper about tactical TI is detection. Using tactical threat intelligence for detection is probably the most advantageous ways of using threat intelligence. They talk about how using TI to feed security, incident event management software information is a common use of this TI, threat intel is compared against logs coming in from different sources, which then triggers an alert, and can even block these IPs in some cases. *“The SIEM is the most common recipient of TI for detection purposes, as it was purpose-built to perform rule matching on log data from many sources”* (Michael Clark, 2020).

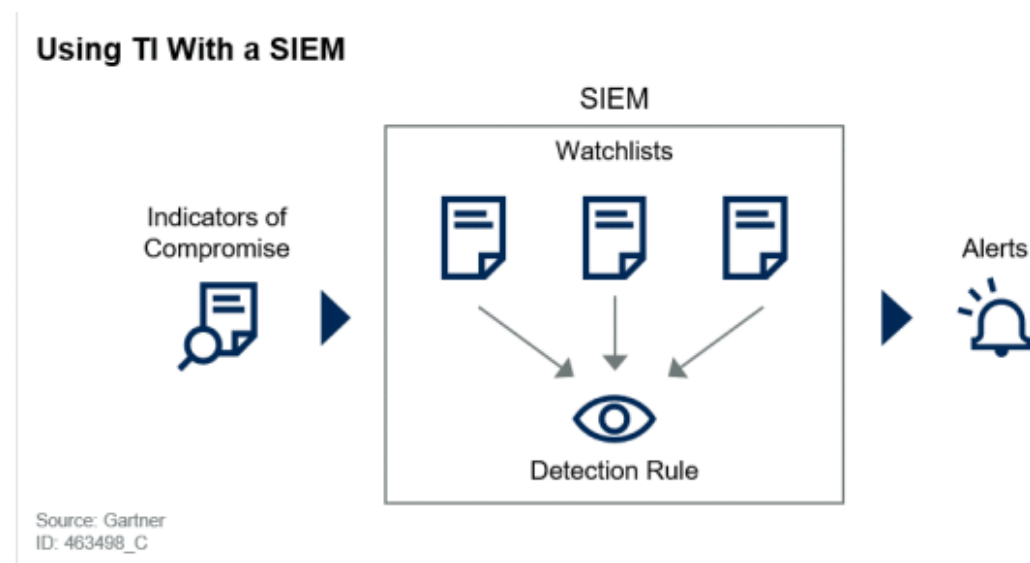


Figure 2 Tactical TI in a SIEM environment

In the figure above, provided from the paper, we can see how this would work in real time, with the SIEM ingesting IOCs onto watchlists and using that information to monitor and pickup suspicious traffic.

They also go on to explain how using tactical TI for prevention can be a risky situation, as sometimes legitimate websites and companies can be on the end of incorrect threat intelligence. *“Using tactical TI for prevention can be a risky proposition. It is not uncommon for a legitimate FQDN, IP address or hash to be included in a threat*

intelligence feed.” (Michael Clark, 2020), although, I would much rather not be able to access some certain websites and be on the safe end of the stick myself. They talked about how some clients who use a common prevention scenario of blocking these threat intelligence IPs on their perimeter firewalls have found some false blocks and network issues in some cases. I feel like this highlights the point of ensuring that the threat intelligence is legitimate and checked to make sure stuff like that doesn’t happen.

Tactical threat intelligence also plays an important role in threat hunting, Some IOCs aren’t necessarily handleable by a tool and that data is still very important, it can be used by threat hunters to understand patterns of activity and methods associated with specific threat actors. *“Then threat hunters will still need to use that information. What’s more important to threat hunters is a type of TI called “tools, techniques and procedures” (TTPs). TTPs can be defined as “patterns of activities or methods associated with a specific threat actor or group of threat actors”* (Michael Clark, 2020), This highlights the idea of having more than just IOCs such as IPs and other technical level data.

Another way that tactical threat intelligence is used is in Alert Prioritization, which links back to SIEM implementation, alert prioritization can help with separating the noisy false positive alerts with the important possible true positive alerts, *“For example, an alert shows an outgoing connection to an uncommon country. While this connection could be legitimate, a TI source has labeled the FQDN as a command-and-control server. With that added information, this alert should be given priority”* (Michael Clark, 2020), This can be extremely useful in enhancing security and response times from the SIEM. They talk about using a Security orchestration, automation and response system (SOAR) to help with this process, SOAR can identify threats, use threat intelligence to enrich the analysis and then send it for human review all by itself, improving efficiency and creating a uniformed approach to certain alerts, *“When the SOAR platform receives the alert from the SIEM, it can check TI to determine whether the destination of the connection is known to be used for malware command and control. If the destination is malicious, the SOAR platform can raise the severity of the alert or issue a ticket”* (Michael Clark, 2020).

Sample SOAR Workflow With Threat Intelligence

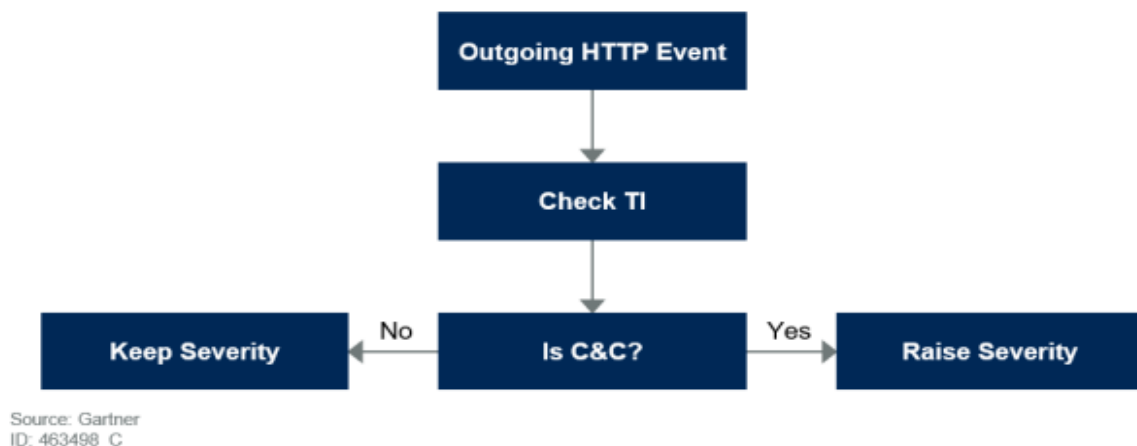


Figure 3 Sample SOAR Workflow with Threat Intelligence

The last point they mentioned is incidence response, discussing how tactical threat intelligence can help understand the incident better, analysts or incident responders can identify certain types of attacks and then proceed with remediating the incident with previous knowledge, *“Once the IR team understands the threat it is dealing with and has verified that the TI is accurate, containment of the incident can happen quickly”* (Michael Clark, 2020), although as they discussed, this can be an unreliable way to use threat intelligence as attackers can slightly modify their attacks, which can lead to confusion and delay of remediation. It’s also important to note that incident response can be used to help build on threat intelligence and further enhance understanding.

Threat intelligence can be used in a variety of ways to strengthen security, and the distinction between strategic and tactical threat intelligence provides important insights. While tactical threat intelligence is crucial for in-the-moment threat detection, prevention, threat hunting, and incident response, strategic threat intelligence enables businesses to make knowledgeable decisions about their entire security posture.

However, it's critical to recognize the difficulties in efficiently utilizing threat intelligence. It is crucial to ensure the accuracy and relevance of the intelligence data since false positives or out-of-date information may result in the improper blocking of legal traffic or the inaccurate identification of threats. Organizations must have effective processes for confirming the accuracy of the data and integrating it into their current security systems, such as SIEMs and SOAR platforms, in order to reduce these risks.

3. Threat Intelligence Sharing Online Platforms

3.1 VirusTotal

VirusTotal is a free online program that checks files and URLs for the existence of malware, viruses, trojan horses, worms, and other unwanted software. It was founded in 2004 and later acquired by Google in 2012; it operates as a branch of Chronicle, a cybersecurity company owned by Alphabet Inc. (VirusTotal)

The service combines the results of over 70 antivirus engines and other URL scanning services to deliver an in-depth analysis of uploaded content. Users can upload files (up to 650 MB) or URLs for analysis, and programmers can integrate the functionality of the service into their creations by using the public VirusTotal API.

Key features of VirusTotal include:

- **Multi-engine scanning:** To give a thorough analysis of the uploaded content, VirusTotal combines the findings from over 70 antivirus engines and other URL scanning services. Because different engines may have varied detection capabilities and signature databases, this enhances the likelihood that malicious information will be detected.
- **File and URL analysis:** Users have the option of uploading files or submitting URLs for analysis, allowing them to check for the presence of malware in a variety of formats, such as documents, executables, and online content.
- **Public API:** VirusTotal gives developers access to a public API that enables them to include its scanning and analysis tools into their own software and services.
- **Intelligence generated by the community:** Users can offer feedback and comments on the analysis, adding context and details about potential dangers. The accuracy and detecting capacities of the service can be enhanced with the use of this user-generated data.

- **Sandbox analysis:** In addition to static analysis, VirusTotal undertakes dynamic analysis of certain files in a supervised setting, enabling it to watch the actions of potential malware and acquire more data.
- **Historical information:** To help researchers and security experts better understand trends and patterns in virus spread, VirusTotal keeps a library of historical analysis results.
- **Private services:** VirusTotal offers subscription plans with advantages including larger API request limits, quicker file analysis, and access to premium tools for users who need more capabilities and privacy.

3.2 IBM X-Force Exchange

IBM X-Force Exchange is another free online program that is a cloud-based, fully collaborative threat intelligence sharing platform. The platform aggregates and analyses vast amounts of threat data from various sources, including the dark web, malware samples, and other threat intelligence feeds. X-Force exchange offers the ability to access many different threat indicators such as IPs, domains and file hashes, it is frequently updated and provides great understanding for security professionals on certain threats. (IBM X-Force, n.d.)

Key Features of IBM X-Force Exchange:

- **Threat Intelligence collection:** A thorough, regularly updated collection of threat indicators, including IPs, URLs, domains, and file hashes, sourced from numerous international data feeds and IBM's own investigation.
- **Advanced Analytics:** The platform makes use of advanced analytics to find patterns, trends, and connections in the threat data so that users can comprehend the context and applicability of certain risks.
- **Cooperation and Sharing:** The platform promotes community-driven cooperation, enabling security researchers and experts to share their knowledge and ideas and add to a common body of knowledge.
- **Integration with Existing Security Tools:** X-Force Exchange interfaces quickly with a variety of security platforms and tools, improving the effectiveness and efficiency of threat analysis and response.
- **Customized Intelligence Feeds:** Based on their own requirements, users can design customized feeds that let them concentrate on the dangers that are most pertinent to their firm.
- **Incident and Threat Management:** The platform provides tools for managing incidents and threats, allowing users to monitor, look into, and effectively address new dangers.

- **API Access:** Thanks to the platform's availability of APIs, users may easily automate the addition of threat intelligence data to their already-in-use security workflows and solutions.

3.3 AbuseIPDB

IBM X-Force Exchange is another free online program that is a cloud-based, fully collaborative threat intelligence sharing platform. The platform aggregates and analyses vast amounts of threat data from various sources, including the dark web, malware samples, and other threat intelligence feeds. X-Force exchange offers the ability to access many different threat indicators such as IPs, domains and file hashes, it is frequently updated and provides great understanding for security professionals on certain threats. (abuseipdb, n.d.)

Key Features of IBM X-Force Exchange:

- **IP Reporting:** Users have the option of reporting IP addresses linked to illegal activity including spamming, hacking, brute-force assaults, and more. These reports aid in the creation of a thorough database of malicious IPs.
- **Community Ran:** AbuseIPDB is a constantly expanding and changing resource for identifying and preventing online abuse thanks to the participation of its user community.
- **IP Blacklist:** To safeguard their networks and systems from recognized risks, users can download and use AbuseIPDB's updated list of harmful IP addresses. Firewalls, intrusion detection systems, and other security measures can relate to the list.
- **IP Lookup:** Users can use an IP lookup to see if a certain IP address has ever been reported for abusive behaviour.

- **Confidence Score:** Based on the volume of reports and the seriousness of the abuse, each reported IP address is given a confidence score. Users can prioritize which IP addresses offer the greatest risk using the confidence score.
- **Country based lookup:** AbuseIPDB offers a breakdown of abusive IP addresses by nation, enabling users to pinpoint areas with a lot of harmful activity.
- **History based Analysis:** The platform enables users to examine historical patterns of IP misuse, which may be used to spot new risks and trends.
- **API Access:** AbuseIPDB offers an API that enables programmers to incorporate the features and data of the platform into their programs, websites, or network security solutions.

3.4 Examples of malicious analysis on each platform

Below is an example of a malicious IP on VirusTotal, as you can see, it shows a main score of 8/87 vendors, location and host details, below that, there is a more detailed list of the sources and what they have flagged it for.

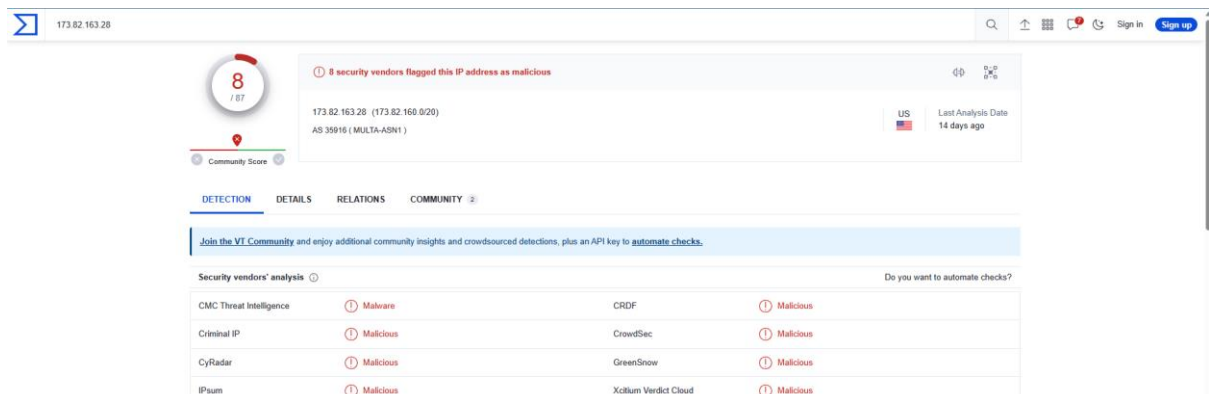


Figure 4 Virus Total Malicious Scan

Below is the IBM X-Force Exchange analysis, as you can see also, it has an overall risk score calculated by an algorithm, it also shows the reason for it being flagged “Scanning IPs” and it shows the location and host details like above.

IBM X-Force Exchange ALL 173.82.163.28

Risk 5.7

X-Force IP Report

173.82.163.28

This report does not contain tags. Add tags via the comment box.

Details

Categorization • Scanning IPs(57%)

Application No known application

Location United States

ASN • AS 35916 : MULTA-ASN1, US

WHOIS Record

Created Nov 14, 2008

Updated Jan 20, 2015

Registrant Organization MULTACOM CORPORATION

Registrant Country or Region United States

Registrar Name ARIN

Email abuse@multacom.com

Figure 5 IBM X-Force Exchange Malicious Scan

Below is the last platform, AbuseIPDB, which gives its confidence rate along with host details and location details and some community reports down below, which can be useful for understanding what the reason behind the blacklist is.

173.82.163.28 was found in our database!

This IP was reported 3,350 times. Confidence of Abuse is 100%: ?

100%

ISP Multacom Corporation

Usage Type Data Center/Web Hosting/Transit

Hostname(s) mlb.outbound.ed10.com

Domain Name multacom.com

Country United States of America

City Los Angeles, California

IP info including ISP, Usage Type, and Location provided by IP2Location. Updated monthly.

REPORT 173.82.163.28 WHOIS 173.82.163.28

IP Abuse Reports for 173.82.163.28:

This IP address has been reported a total of 3,350 times from 735 distinct sources. 173.82.163.28 was first reported on October 10th 2022, and the most recent report was 6 minutes ago.

Recent Reports: We have received reports of abusive activity from this IP address within the last week. It is potentially still actively engaged in abusive activities.

| Reporter | Date | Comment | Categories |
|-------------|---------------|---|--------------------|
| ✓ opcenter | 6 minutes ago | Apr 21 **REMOVED** sshd[11113]: Failed password for root from 173.82.163.28 port 47400 ssh2 Ap ... show more | Brute-Force SSH |
| ✓ ydwis40 | 7 minutes ago | Apr 21 17:19:37 jendela-lb sshd[3186087]: Disconnect | Brute-Force |

Figure 6 AbuseIPDB Malicious Scan

4. Methodology

4.1 Project Development

4.1.1 VirusTotal API Key

VirusTotal API is an indispensable tool that enables developers and security professionals to seamlessly integrate the wide-ranging capabilities of VirusTotal into their applications and systems. As a free online service, by utilizing the API, we can automate scanning processes, access aggregated data from various sources, and gather invaluable information for threat intelligence and security purposes. (VirusTotal)

Key Features:

1. **File and URL Scanning:** Using the API, users can submit files and URLs to be scanned by a variety of antivirus programs and website scanners, which results in a thorough analysis of any potential dangers.
2. **Access to Historical Data:** The API makes it simple for users to track and examine malware trends over time by giving them access to a wealth of historical scan data.
3. **IP and Domain Analysis:** Using the API, users can investigate IP addresses and domain names, providing important information about their connections to criminal actions.
4. **Commenting and voting:** Users can participate in the VirusTotal community by contributing comments and voting on the analysis findings, which enhances the platform's accuracy and overall value.

We will be using the VirusTotal API Key to pull IP information such as ISP and the malicious record of the IP, along with the same for file hashes to implement into the application.

4.1.2 AbuseIPDB API Key

The AbuseIPDB API is a powerful tool, like the VirusTotal API Key, that allows developers and cybersecurity professionals to interact and pull information from this malicious IP database to enhance their security measures. (abuseipdb, n.d.)

Key Features:

1. **IP Address Lookup:** Using the API, you can determine whether a certain IP address has been flagged for any malicious activity. You may receive comprehensive data about the IP, such as the number of reports, the types of abuse, and the reputation score.
2. **IP Address Reporting:** You may help the community by utilizing the API to report suspicious IP addresses and the abuse kinds that go along with them. This aids in keeping a current and trustworthy database of malicious IPs
3. **IP Address Blocklist:** You can download a list of offensive IP addresses via the API, and then use it to block traffic coming from nefarious sources. Based on factors like the reputation score, the sorts of abuse, and the confidence level, you can tailor the blocklist.
4. **Bulk IP address lookup and reporting:** The API allows you to examine or report many IP addresses at once. For larger enterprises with plenty of network traffic, this capability is extremely helpful.
5. **Webhooks:** The AbuseIPDB API supports webhooks, enabling you to get instant alerts whenever an IP address on your watchlist is reported for abuse.

With the AbuseIPDB API key, we will pull out the IP blacklist confidence score to add to our list of vendors in our rating system.

4.1.3 IBM X-Force Exchange API Key

IBM X-Force Exchange API provides a powerful platform that streamlines access to actionable threat intelligence, enabling organizations to enhance their security posture. IBM X-Force Exchange API, again like the two previous examples, provides a powerful platform that streamlines access to actionable threat intelligence. (IBM X-Force, n.d.)

Key Features:

1. **Comprehensive Intelligence:** The platform provides a wide spectrum of threat intelligence, including information on malware, phishing schemes, vulnerabilities, and threat actors. Organizations may better comprehend the threat landscape and keep ahead of new dangers by integrating this intelligence into their security technologies.
2. **Collaborative Setting:** IBM X-Force Exchange promotes teamwork among businesses and security researchers. Users can collaborate with one another to develop a collective defence against cyber-attacks by sharing their findings, threat indicators, and best practices.
3. **Security Tool Integration:** The API may be easily linked with an organization's current security infrastructure, including firewalls, SIEM systems, and intrusion detection systems. Through this integration, threat intelligence may be automated, and responding to security issues is made simpler.
4. **Customizable:** By allowing customers to tailor their feeds, IBM X-Force Exchange API enables them to concentrate on the precise threat intelligence that is most pertinent to their organization. This personalization makes sure that security teams obtain information that is specifically relevant to them, cutting down on noise and boosting the efficacy of their security activities.
5. **Scalable and secure:** Organizations can depend on the availability and integrity of the threat intelligence they get because the platform is built to securely handle massive amounts of data.

4.1.4 Python

Python is an extremely versatile and powerful high-level programming language, designed with simplicity and readability in mind. Created by Guido van Rossum in 1991, it has quickly become one of the most popular languages among developers and is used for a wide range of applications. (geeksforgeeks, 2022)

The reason I chose to use python for my project came down to several reasons:

- **Extensive Libraries:** The standard library and third-party packages cover a wide range of functionalities, from web server hosting to login managers, which were necessary for my application.
- **Simpler Code:** Python offers a much more simplistic and high-level coding platform than other languages such as Java which requires more in-depth use cases.
- **Strong Community Support:** Python boasts a large and active community that continually contributes to its development and offers support through forums, mailing lists, and conferences. This was a big factor, as the support would help me with developing the application.
- **API Implementation:** The API's being used in this application have good documentation and implementation with python which was the main factor in choice.

4.1.5 Python Flask

A lightweight Python web framework called Flask offers a quick and easy way to construct websites. It adheres to Python's "batteries-included" philosophy while facilitating quick prototyping and development of web apps. In the Python ecosystem, Flask stands out for its simplicity, flexibility, and adaptability. I chose to use Flask for my webserver hosting as it made the most sense, to keep everything in one place and didn't require as much setup or maintenance as other options such as a third-party host, also I am planning on running this on a local server for now.

Through its built-in modules and third-party extensions, Flask offers strong support for implementing user authentication and authorization.

Popular plugin Flask-Login makes user authentication and session management simple. For managing user sessions, including login, logout, and session persistence, it offers a simple user interface. I choose to also use this for my user authentication system due to the fact it kept everything consistent with what was being used already. (Python, n.d.)

4.1.6 HyperText Markup Language (HTML)

The industry-standard markup language for creating and designing webpages is HTML, or HyperText Markup Language. All of the contents on a webpage are organized and designed using HTML, the main language of the internet. HTML was created in 1989 by Tim Berners-Lee, a physicist at CERN (the European Organization for Nuclear Research). HTML5 is the most current version of HTML and is the most used language to develop webpages throughout the world. I decided to use HTML due to the fact I have vast experience in it since a child and I feel like there is also a lot of options with it to develop a good application. (Wikipedia, n.d.)

4.1.7 Cascading Style Sheets (CSS)

CSS, or Cascading Style Sheets, is a stylesheet language used to define the look and formatting of web documents, primarily HTML. It allows web developers to separate the presentation layer (style) from the structure layer (HTML), resulting in cleaner code and more manageable web projects. CSS was first proposed by Håkon Wium Lie in 1994 to bring styling capabilities to HTML. The World Wide Web Consortium (W3C) recognized the need for a standardized styling language and released CSS1 in 1996. CSS3, now the most modern version of CSS is the one I decided to use to help style my application, I feel like CSS offers a simple but great solution to a bland HTML page and allows bring some life to the application. (wikipedia, n.d.)

4.2 EasyIntel

Easy Intel is a powerful and user-friendly application designed to provide comprehensive threat intelligence by leveraging multiple APIs for accurate and real-time information. Built with Python and Flask, this application is an invaluable tool for anyone interested in keeping track of potential threats and gathering intelligence on malicious IPs or file hashes.

Easy Intel seamlessly integrates with various APIs, such as VirusTotal, IBM X-Force Exchange, IP-API, and AbuseIPDB, to provide users with detailed information on IP addresses and file hashes. By combining data from these sources, the application offers insights into the reputation, geolocation, and security risks associated with a given IP address or file hash.

Users can easily sign up and log in to access the application's features. With the user's API keys for the different services, Easy Intel fetches and displays key information like malicious activity score, country, ISP, and abuse confidence score for IP addresses. Similarly, for file hashes, the application gathers and presents the malicious activity score from both VirusTotal and IBM X-Force Exchange.

All the necessary code to run EasyIntel can be found with a guide to run it here:

[FinbarDownes23/EasyIntel \(github.com\)](https://github.com/FinbarDownes23/EasyIntel)

4.2.1 EasyIntel Showcase

In the first screenshot we can see the first page you are greeted with once you open the web application, the log in page, here the users enter their details to be granted access to the application or can sign up with the signup button.

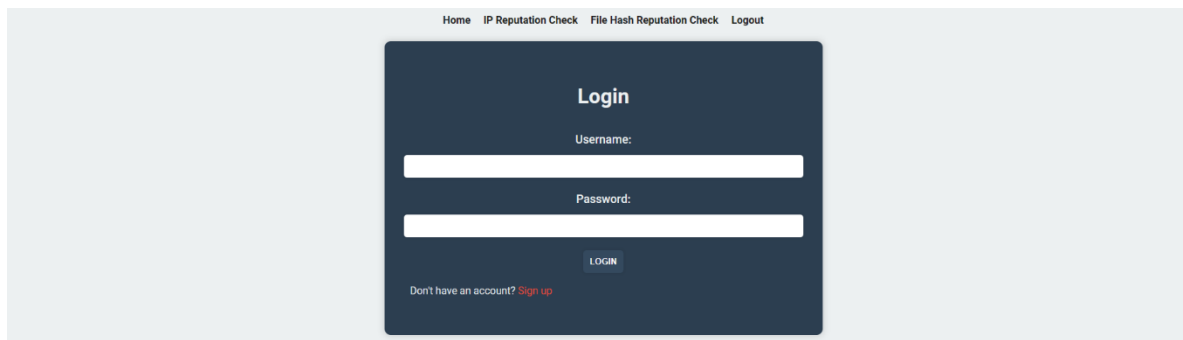


Figure 7 Easy Intel - Login Page

The signup page is simple but effective, asking for all the information necessary to setup a user account to function with the API calls.

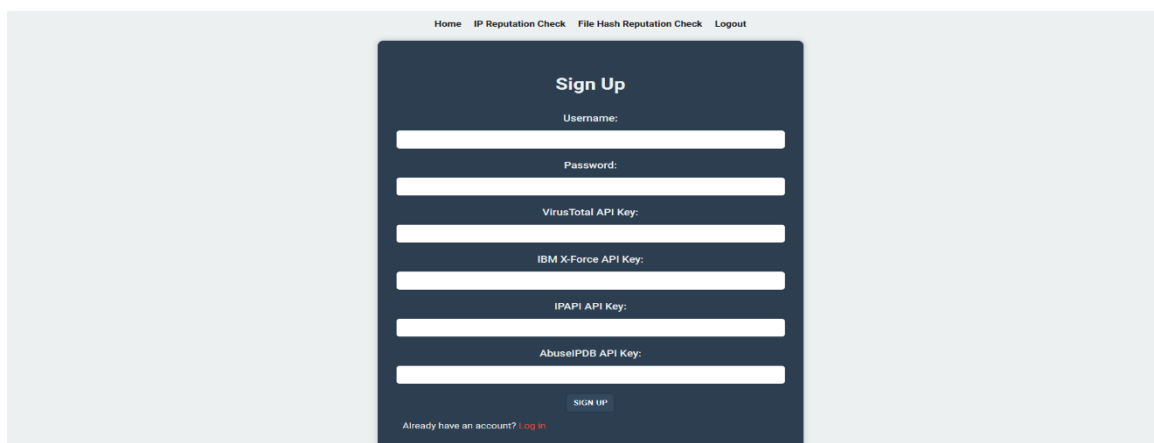


Figure 8 Easy Intel - Sign Up Page

Once signed up and logged in, the user is directed to the home page, which is actually the first page opened with the application, but of course without being logged in, it will always redirect to the login page. The homepage is simple and has links to both the IP Reputation Checker and the File Hash Reputation Checker.

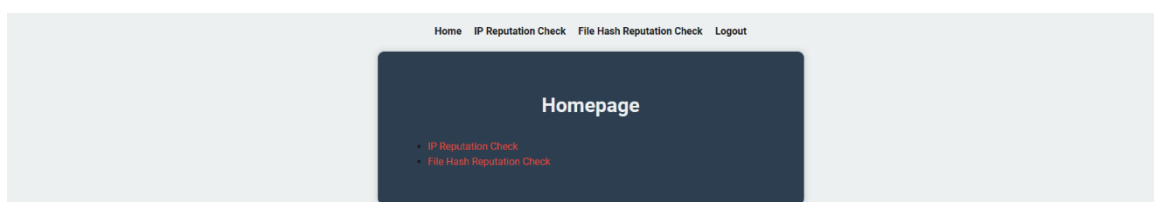


Figure 9 Easy Intel - Home Page

Looking at the IP Reputation Checker First, we are met with a simple form page asking for the IP we want to scan.



The screenshot shows a web application interface with a light gray background. At the top, there is a navigation bar with links: "Home", "IP Reputation Check", "File Hash Reputation Check", and "Logout". The main content area features a dark blue rectangular box. Inside this box, the title "IP Reputation Check" is displayed in white. Below the title, the text "Enter an IP address:" is shown. Underneath, there is a white input field with the placeholder text "Enter an IP address...". Below the input field is a dark blue button with the word "CHECK" in white.

Figure 10 Easy Intel - IP Reputation Checker Page

After entering an IP and letting the application run, it returns the results in this format:

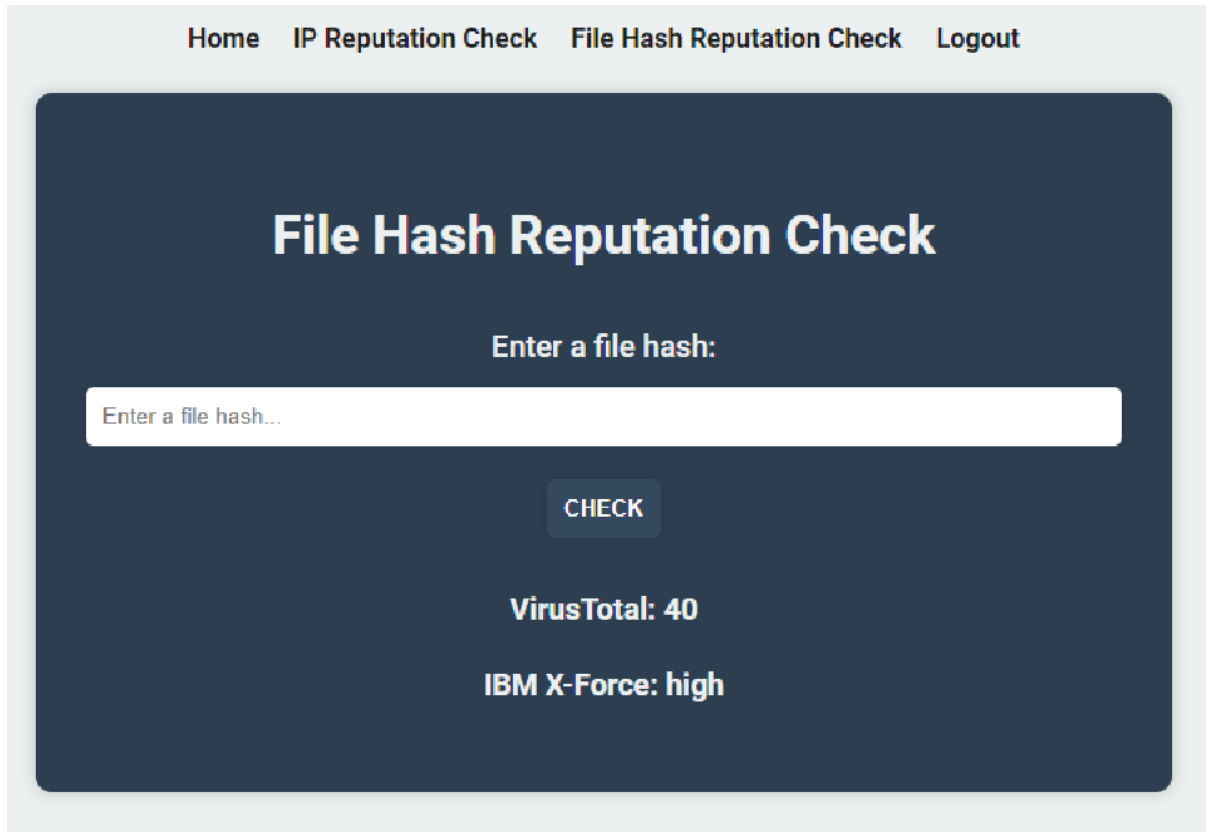


The screenshot shows the same web application interface as Figure 10, but with the results of an IP scan displayed. The navigation bar remains the same. The dark blue box now displays the following information in white text: "IP Reputation Check" at the top, followed by "Enter an IP address:". Below this is a white input field with the placeholder text "Enter an IP address...". Underneath the input field is a dark blue button with the word "CHECK" in white. Below the button, the results are displayed in red text: "ISP: UCLOUD INFORMATION TECHNOLOGY HK LIMITED", "Location: Hong Kong SAR China", "VirusTotal: 5", "IBM X-Force: 2.9", and "AbuseIPDB: 100".

Figure 11 Easy Intel - Bad IP Results

Which gives all the information to further understand if this IP is malicious or not.

Moving onto the File Hash Reputation Checker, we can see a similar simple form, to enter a filehash, once entered, this is the results we receive, from IBM X-Force and VirusTotal.



The screenshot displays a web application interface for a File Hash Reputation Checker. At the top, there is a navigation bar with links: Home, IP Reputation Check, File Hash Reputation Check, and Logout. The main content area has a dark blue background with the title "File Hash Reputation Check" in large white text. Below the title, it says "Enter a file hash:" followed by a white input field containing the placeholder text "Enter a file hash...". A dark blue button labeled "CHECK" is positioned below the input field. The results are displayed in white text: "VirusTotal: 40" and "IBM X-Force: high".

Figure 12 Easy Intel - Bad File Hash Results

a. Data Collection

To conduct this study, we employed a two-pronged approach to gather the necessary data. The first method involved obtaining a list of public IPs available online and running them through the three main threat intelligence platforms—VirusTotal, IBM X-Force, and AbuseIPDB—to compare their results. The second method involved distributing a survey to peers and co-workers in the cybersecurity field to gain insights into their experiences with and opinions on the accuracy of these platforms.

b. Data Analysis

To analyze the data collected, we performed the following steps:

Compare the results from the three threat intelligence platforms and tabulate them in a spreadsheet. We created conditions for the cells based on the different grading metrics of each platform's scoring parameters to provide a clearer understanding of the differences in their findings.

Analyze the collected data to identify patterns and trends. This includes observing the number of flagged IPs, the sensitivity of the scoring systems, and discrepancies between the results provided by each platform.

Analyze the survey responses to gauge the popularity and perceived accuracy of each platform. We tabulated the number of votes for each platform and analyzed the patterns and trends in the responses.

Synthesize the findings from both the comparison test and the survey to draw conclusions on the accuracy, popularity, and discrepancies between the threat intelligence platforms.

c. Limitations

The following limitations of the study's methodology should be considered when interpreting the findings:

Small sample size: The number of IPs used in the comparative test and the number of survey participants might not be sufficient to represent the overall effectiveness and precision of the threat intelligence systems in full.

Selection bias: The threat intelligence sources from which the comparison test's IPs were chosen may not fully represent the range of IPs that might be encountered in real-world situations.

Subjectivity: Individual prejudices and preferences may have an impact on how accurately the platforms are perceived to be, according to poll respondents.

Variability in scoring: The various scoring methods used by threat intelligence platforms can make direct comparisons difficult and may have an impact on how the results are interpreted.

Future research might expand the IP sample size, recruit a more diverse and representative sample of survey respondents, and create a consistent grading system to enable a more realistic comparison of the threat intelligence platforms to overcome these constraints.

5. Results And Interpretation

5.1 Testing IP results on VirusTotal, IBM X-Force and AbuseIPDB

I decided I wanted to do to test the accuracy of these main threat intelligence sharing platforms we are focusing on.

The first test ran was to collect public IPs online and run them through all three of these threat intelligence platforms to see what the results would yield; I feel like this test will give the clearest results in terms of accuracy.

5.2 Spreadsheet of results

Conditions for cells:

| | VirusTotal (Capped at 10): | IBM X-Force: | AbuseIPDB (All values divided by 10): |
|---------------|---------------------------------------|---------------------|--|
| Green | 0 – 1 | 0 – 3 | 0 – 4 (40) |
| Orange | 2 – 6 | 4 – 6 | 4.1 (41) – 7 (70) |
| Red | 7 + | 7 – 10 | 7.1 (71) – 10 (100) |

These conditions were set based on the different grading metrics on the scoring parameters of each threat intelligence sharing platform.

VirusTotal was capped to 10 detections max, as 10 is a sufficient number for severe blacklist.

AbuseIPDB was divided by 10 as it goes in percentage and is up to 100, dividing by 10 makes it even across the board so the results are more presentable.

The first thing noticeable looking at this spreadsheet of results, is the large amount of green in the IBM X-Force column, which would indicate a low amount of flagged IPs, comparing this to the other 2 columns, we see a noticeable jump in red and orange, which indicates a higher amount of flagged IPs, essentially showing that IBM X-Force doesn't flag as many IPs as the other two platforms we looked at.

Also noticeable in the spreadsheet, is a rather high number of reds in the AbuseIPDB column, which looks to over double the amount of the next highest in VirusTotal, it shows that AbuseIPDB have a more sensitive scoring system than VirusTotal or IBM X-

Force, and depending on what you're looking for, could be too sensitive to consider accurate.

VirusTotal looks to have a good balance between highs and lows, which could mean that the accuracy of the results is high.

| Malicious IP | VT Score | IBM Score | AbuseIPDB Score |
|-----------------|----------|-----------|-----------------|
| 121.239.32.3 | 2 | 1 | 9.6 |
| 81.88.53.6 | 6 | 1 | 10 |
| 67.205.177.222 | 4 | 1 | 10 |
| 43.156.238.11 | 7 | 1 | 10 |
| 40.92.65.81 | 0 | 1 | 2 |
| 198.235.24.112 | 2 | 1 | 10 |
| 173.82.163.28 | 8 | 4.3 | 10 |
| 162.243.141.15 | 7 | 4.3 | 10 |
| 121.238.155.154 | 1 | 1 | 9.1 |
| 40.92.18.63 | 0 | 1 | 0.5 |
| 182.227.18.158 | 5 | 1 | 10 |
| 128.199.208.187 | 8 | 4.3 | 10 |
| 43.134.170.254 | 9 | 5.7 | 10 |
| 40.92.22.106 | 0 | 1 | 0.5 |
| 223.82.86.2 | 5 | 1 | 10 |
| 218.241.249.38 | 0 | 1 | 3 |
| 181.13.173.21 | 1 | 1 | 5 |
| 121.226.230.213 | 6 | 1 | 10 |
| 116.55.184.228 | 0 | 1 | 3.5 |
| 103.171.181.11 | 2 | 1 | 10 |
| 64.62.197.11 | 6 | 7.1 | 10 |
| 40.92.40.27 | 0 | 1 | 0.4 |
| 183.237.164.206 | 7 | 5.7 | 10 |
| 121.224.3.197 | 10 | 1 | 10 |
| 112.197.89.209 | 0 | 1 | 1.4 |
| 40.92.22.86 | 1 | 1 | 4 |
| 184.105.247.194 | 9 | 8.6 | 10 |

| | | | |
|-----------------------|----|-----|----|
| 78.128.113.250 | 3 | 1 | 10 |
| 61.177.172.114 | 10 | 1 | 10 |
| 60.199.224.55 | 9 | 4.3 | 10 |

Figure 13 Spreadsheet of IP testing results

The first thing noticeable looking at this spreadsheet of results, is the large amount of green in the IBM X-Force column, which would indicate a low amount of flagged IPs, comparing this to the other 2 columns, we see a noticeable jump in red and orange, which indicates a higher amount of flagged IPs, essentially showing that IBM X-Force doesn't flag as many IPs as the other two platforms we looked at.

Also noticeable in the spreadsheet, is a rather high number of reds in the AbuseIPDB column, which looks to over double the amount of the next highest in VirusTotal, it shows that AbuseIPDB have a more sensitive scoring system than VirusTotal or IBM X-Force, and depending on what you're looking for, could be too sensitive to consider accurate.

VirusTotal looks to have a good balance between highs and lows, which could mean that the accuracy of the results is high.

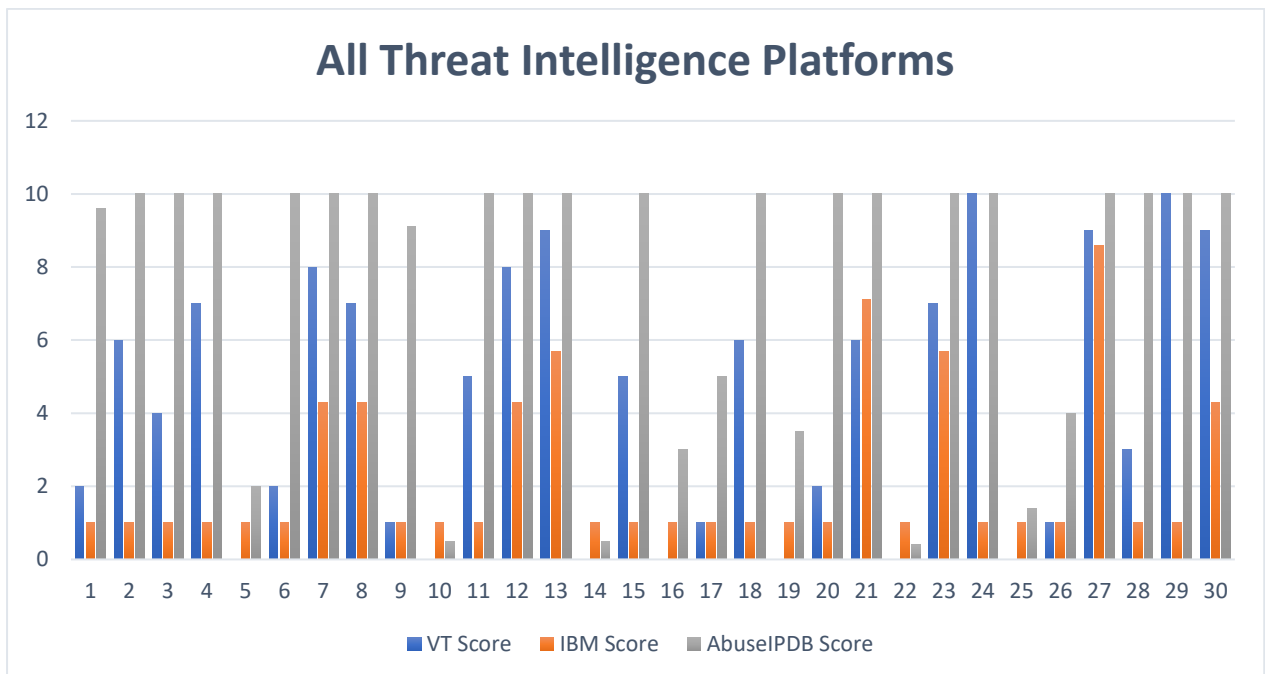


Figure 14 All Threat Intelligence Platforms Graph

Looking through this graph, we can see in a more visual approach that much of the time, VirusTotal and AbuseIPDB are significantly higher than IBM X-Force, although there is still a big gap between VirusTotal and AbuseIPDB.

Looking at the graphs below gives a clearer view of the gap between X-Force and the other 2 platforms. The majority of the time X-Force is coming in with 1 risk score when VirusTotal can be higher at up to 6 risk score, and AbuseIPDB up to 10 at times.

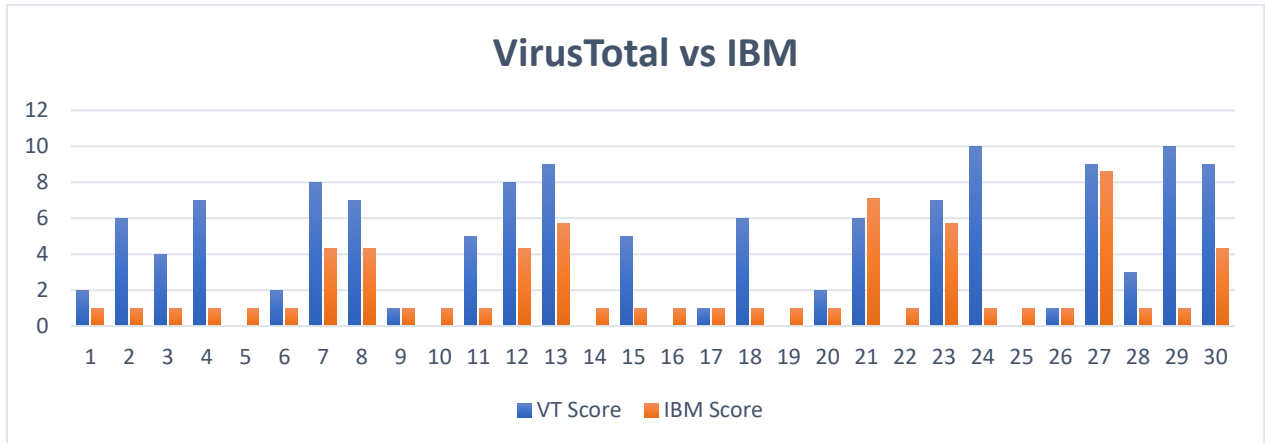


Figure 15 VirusTotal vs IBM X-Force IP Results

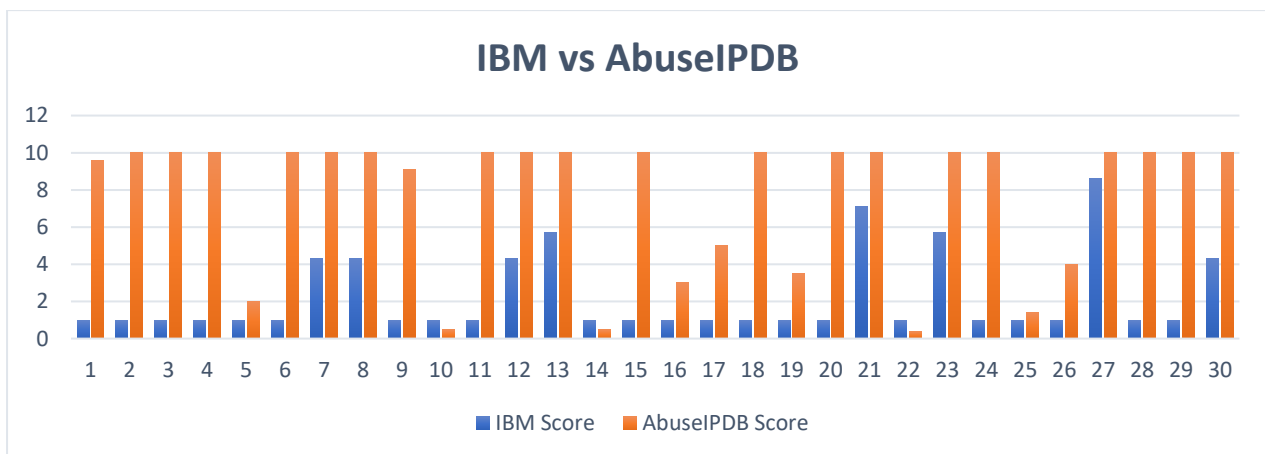


Figure 16 AbuseIPDB vs IBM X-Force IP Results

Looking at the graph of VirusTotal vs AbuseIPDB, its clear to see there are a lot more in common between the two than the other platforms relationships, but still, we can see a good deal of gaps and differences in scoring.

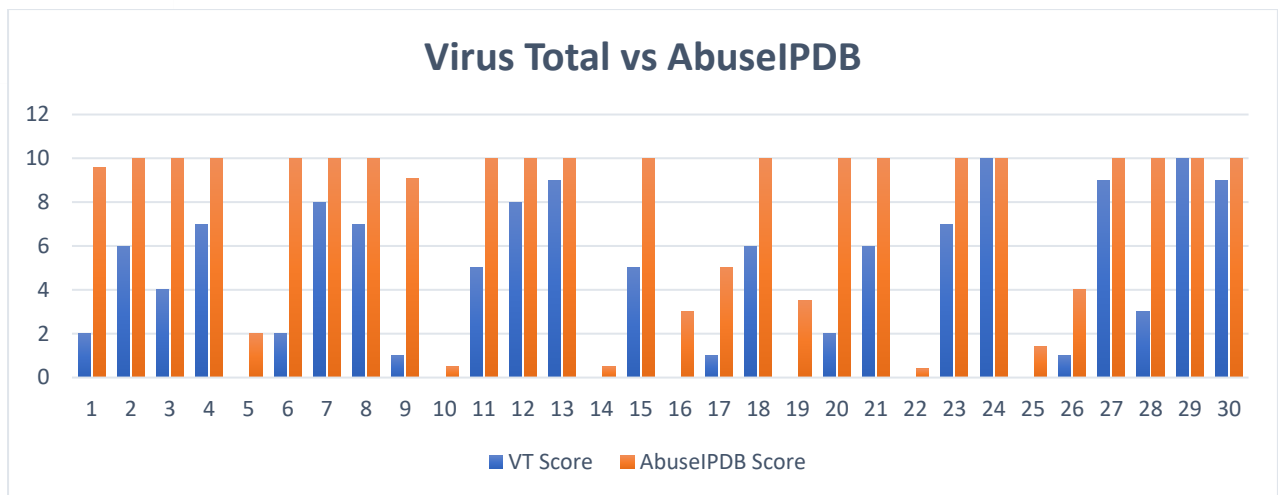


Figure 17 AbuseIPDB vs IBM X-Force IP Results

The outcome of this test demonstrates that, even only between three widely used threat intelligence sharing systems, there are some significant variations. Obviously, all of these platforms are of a high calibre, but this raises the question of threat intelligence accuracy. It's crucial to cross-reference between numerous platforms to guarantee the correct information is being digested because we can't simply accept any one platform as correct and occasionally even multiple platforms can be mistaken.

5.2.1 Understanding why these threat intelligence platforms different?

It's important to understand why these three platforms have different results, and only with experience can you decide which one suits your use case the most. Some of the reasons why these threat intelligence platforms differ are:

- **Data sources:** Each platform may rely on different data sources and partners to gather information about malicious activities. As a result, the breadth and depth of coverage can vary between these platforms, leading to discrepancies in their results.
- **Detection methods:** The three platforms use different detection methods and algorithms to identify and analyse potential threats. These methods may include signature-based detection, behavioural analysis, or heuristic analysis, among others. The difference in detection methods can lead to discrepancies in the results provided by these platforms.
- **Update frequency:** The frequency at which each platform updates its database and detection mechanisms can also contribute to the differences in their results.

Some platforms may update their data more frequently, resulting in more up-to-date and accurate information, while others may update less frequently.

- **Focus areas:** Each platform may have a different focus or specialization, such as certain types of malwares, phishing, or IP addresses. As a result, they may prioritize different types of threats or analyse them in greater detail, leading to differences in their findings.

In summary, differences in data sources, detection methods, update frequency, false positive/negative rates, and focus areas can all contribute to discrepancies in the results provided by VirusTotal, IBM X-Force, and AbuseIPDB. To get a more comprehensive understanding of a potential threat, it's often recommended to use multiple tools and compare their results.

5.3 SOC Analyst Threat Intelligence Survey

The second test that was conducted was a survey among my peers and co-workers on threat intelligence. The survey was created to better enhance the understanding of what people use on a day-to-day basis, which gives a real life understanding and helps further evaluate which threat intelligence platforms are accurate.

5.3.1 Survey Questions

The questions asked:

1. **How often do you use threat intelligence? ***
 - Daily
 - Weekly
 - Monthly
 - Other
2. **Select which threat intelligence sharing platforms you use? ***
 - VirusTotal
 - IBM X-Force
 - AbuseIPDB
 - MxToolbox
 - Other

3. Do you find discrepancies between different platforms? *

- Yes
- No

4. Which platform do you find to be most accurate? *

- VirusTotal
- IBM X-Force
- AbuseIPDB
- Other

5.3.2 Survey Answers

1. How often do you use threat intelligence?

| How often do you use threat intelligence? | |
|---|----------|
| Daily | Monthly |
| Daily | |
| Daily | |
| Daily | |
| Daily | |
| Daily | |
| Daily | |
| Daily | |
| Daily | |
| Daily | |
| 11 | 1 |

Figure 18 Survey Question - How often do you use threat intelligence?

Looking at the results of the first question, It is clear that the majority of the people surveyed, use threat intelligence on a daily basis, meaning that the results of this survey should be accurately sound and be based on real life day-to-day experience.

2. Select which threat intelligence sharing platforms you use?

| Select which threat intelligence sharing platforms you use? | | | | | |
|---|-------------|-----------|-----------|----------|--------------------|
| VirusTotal | IBM X-Force | AbuseIPDB | MxToolbox | Ipvoid | Talos Intelligence |
| VirusTotal | IBM X-Force | AbuseIPDB | MxToolbox | | |
| VirusTotal | IBM X-Force | AbuseIPDB | MxToolbox | | |
| VirusTotal | IBM X-Force | AbuseIPDB | | | |
| VirusTotal | IBM X-Force | AbuseIPDB | | | |
| VirusTotal | IBM X-Force | AbuseIPDB | | | |
| VirusTotal | IBM X-Force | AbuseIPDB | | | |
| VirusTotal | IBM X-Force | AbuseIPDB | | | |
| VirusTotal | IBM X-Force | | | | |
| VirusTotal | | | | | |
| 11 | 10 | 9 | 3 | 1 | 1 |

Figure 19 Survey Question - Select which threat intelligence platforms you use?

The breakdown of votes for each threat intelligence platform, as displayed at the bottom of the table, provides an interesting snapshot of the preferences and popularity of these platforms within the cybersecurity community. VirusTotal emerges as the clear leader, garnering 11 votes, followed by IBM X-Force with 10 votes, and AbuseIPDB with 9 votes. MxToolbox trails behind with only 3 votes. Additionally, the "Other" category

introduced us to a couple of lesser-known platforms, namely IPVoid and Talos Intelligence, each receiving one vote apiece.

These results indicate that VirusTotal enjoys a prominent position as the most widely used and popular threat intelligence platform among peers. IBM X-Force and AbuseIPDB also demonstrate significant usage, though they both lag slightly behind VirusTotal. It is worth noting that in the earlier comparison test, IBM X-Force was found to be somewhat less accurate than both VirusTotal and AbuseIPDB. This finding suggests that popularity and usage do not always correlate directly with accuracy and effectiveness.

3. Do you find discrepancies between different platforms?

| Do you find discrepancies between different platforms? | |
|--|----|
| Yes | No |
| Yes | |
| Yes | |
| Yes | |
| Yes | |
| Yes | |
| Yes | |
| Yes | |
| Yes | |
| Yes | |
| Yes | |
| Yes | |
| 11 | 1 |

Figure 20 Survey Question - Do you find discrepancies in threat intelligence

It is evident from the poll results that 11 people have voted in favour that there are indeed discrepancies between various threat intelligence sources, which serves to reinforce the findings from our initial analysis that highlighted significant differences between these sources.

4. Which platform do you find to be most accurate?

| Which platform do you find to be most accurate? | |
|---|-----------|
| VirusTotal | AbuseIPDB |
| VirusTotal | AbuseIPDB |
| VirusTotal | AbuseIPDB |
| VirusTotal | AbuseIPDB |
| VirusTotal | AbuseIPDB |
| VirusTotal | AbuseIPDB |
| 6 | 6 |

Figure 21 Survey Question - Which platform do you find to be most accurate?

Analysing the responses to the final survey question, it becomes apparent that all votes were cast in favor of either VirusTotal or AbuseIPDB, with no votes allocated to IBM X-Force or the other alternative options that were presented. Both VirusTotal and AbuseIPDB received an equal share of the votes, with 6 apiece, which serves to further reinforce the notion derived from our previous test that these platforms are perceived as the most accurate among the available choices.

Interestingly, we encounter an unusual situation where respondents have indicated their preference for using IBM X-Force, as evidenced by the earlier question, yet they do not seem to regard it as being as accurate as VirusTotal or AbuseIPDB. This apparent discrepancy could be attributed to several factors, such as the user-friendliness of the platform, the breadth of information provided, the frequency of updates, or even the brand recognition associated with IBM as a whole.

5.4 Results Summary

The analysis of the data collected from both the comparison test and the survey of peers and co-workers in the cybersecurity field led to several key findings.

These findings include the identification of discrepancies between the results provided by the three main threat intelligence platforms (VirusTotal, IBM X-Force, and AbuseIPDB), as well as insights into the popularity and perceived accuracy of each platform.

The key takeaways from this analysis are as follows:

- IBM X-Force generally flagged fewer IPs as malicious compared to VirusTotal and AbuseIPDB, suggesting a lower sensitivity in its scoring system.
- AbuseIPDB appeared to have a more sensitive scoring system than both VirusTotal and IBM X-Force, with a higher number of flagged IPs.
- VirusTotal demonstrated a balanced approach in terms of flagged IPs, which could be indicative of a high level of accuracy.
- VirusTotal emerged as the most popular threat intelligence platform among survey respondents, followed by IBM X-Force and AbuseIPDB.
- The majority of survey respondents acknowledged the existence of discrepancies between different threat intelligence platforms.
- Both VirusTotal and AbuseIPDB were equally perceived as the most accurate platforms by survey respondents, with no votes allocated to IBM X-Force in terms of accuracy.

5.5 Interpretation

The results of this study underscore the significance of utilizing multiple threat intelligence platforms in order to gain a comprehensive understanding of potential threats in the rapidly evolving cybersecurity landscape.

VirusTotal, IBM X-Force, and AbuseIPDB, being some of the most reputable platforms in the industry, each have their unique methodologies, data sources, focus areas, and update frequencies that can lead to discrepancies in their findings. This presents a challenge for cybersecurity professionals who rely on these platforms for accurate and timely threat intelligence.

The survey conducted as part of this study offers valuable insights into the preferences and experiences of cybersecurity experts when it comes to using these platforms.

It is evident that VirusTotal enjoys a prominent position as the most widely used and popular threat intelligence platform among the surveyed group. IBM X-Force and AbuseIPDB also demonstrate significant usage, but they both lag slightly behind VirusTotal in terms of popularity.

It's crucial to keep in mind that popularity and accuracy aren't necessarily related. Despite being the second-most well-liked choice in the survey, IBM X-Force received 0 votes for accuracy. On the other hand, VirusTotal and AbuseIPDB both received an equal number of votes and were the most accurate platforms. This apparent mismatch emphasizes the necessity for cybersecurity professionals to thoroughly assess the benefits and drawbacks of each platform and avoid basing their decision on a threat intelligence platform merely on popularity or brand awareness.

These findings reaffirm the need for cybersecurity specialists to consider a variety of information sources when evaluating possible hazards. They can lessen the possibility of running into false positives or negatives and ensure an accurate representation of the threat environment by cross-referencing data from several platforms. This method can assist professionals in prioritizing threats, allocating resources, and putting in place the necessary defensive measures in a more informed manner.

In conclusion, this study stresses the significance of utilizing different platforms and considering each one's unique strengths and shortcomings when approaching threat

intelligence. It is crucial for experts to maintain alertness and modify their techniques as the cybersecurity landscape changes and threats become more complex.

Cybersecurity professionals may better safeguard their firms from the constant risk of cyberattacks by leveraging a variety of threat intelligence sources and regularly evaluating their accuracy and relevance.

6. Conclusion

This thesis embarked on a comprehensive journey to investigate the discrepancies existing within online threat intelligence sharing platforms, particularly focusing on VirusTotal, IBM X-Force Exchange, and AbuseIPDB, which are among the leading platforms in the field. The significance of threat intelligence in today's digital landscape cannot be understated. As cyber threats continue to evolve in their sophistication and frequency, the need for accurate, timely, and actionable threat intelligence has become paramount. These intelligence insights provide the very foundation upon which robust cyber defence strategies are built, enabling organizations to anticipate, prepare, and respond effectively to cyber threats.

However, the exploration of the three threat intelligence sharing platforms brought to light certain discrepancies in the threat intelligence data they provide. These inconsistencies and gaps in threat intelligence data are not trivial, given their potential to significantly impact the effectiveness of cyber defence measures. They may lead to confusion among cybersecurity professionals, result in misinterpretation of threats, and consequently, drive inadequate or even erroneous responses to cyber threats. The fact that these platforms, despite their popularity and widespread use, harbour such discrepancies necessitates a closer look into their operational mechanisms and the broader threat intelligence sharing ecosystem.

The reasons for these discrepancies are multifaceted, stemming from several underlying factors. For instance, the methodologies employed by these platforms for data collection and analysis can differ substantially. Some platforms may prioritize certain types of threats, while others may have different threshold criteria for classifying a cyber event as a threat. Moreover, the expertise of platform users can also contribute to variations in reported threat data. Users with limited cybersecurity knowledge may not accurately report or interpret threat indicators, leading to inconsistencies across platforms. Additionally, the technological capabilities of each platform, including the algorithms they use for threat detection and their ability to process and analyse large volumes of data, can also influence the accuracy and completeness of the threat intelligence they offer.

Despite these discrepancies, it is essential to note that these platforms still provide significant value in the fight against cyber threats. They offer a wealth of information

and insights that can greatly enhance an organization's ability to understand and respond to cyber threats. The discrepancies do not invalidate these platforms but rather highlight the need for users to adopt a more comprehensive approach to threat intelligence. By leveraging multiple platforms in tandem, users can potentially offset the limitations of one platform with the strengths of another, leading to a more rounded and accurate threat assessment.

This research underscores the importance of continuous improvement and innovation in the threat intelligence sharing ecosystem. One significant contribution in this regard is the development of the EasyIntel tool, a tangible outcome of this research. EasyIntel, by consolidating threat intelligence data from multiple platforms, offers a practical solution to the discrepancies identified. By providing users with a unified interface for querying multiple platforms simultaneously, it allows for a more efficient and comprehensive analysis of cyber threats, thereby enhancing the overall effectiveness of threat intelligence in cyber defence.

The findings also emphasize the importance of education and awareness among cybersecurity professionals. As revealed in the survey results, relying solely on the popularity or brand recognition of a threat intelligence platform can lead to a less comprehensive understanding of potential threats. It is crucial for users to understand the unique features, strengths, and limitations of each platform and to use a balanced combination of platforms to formulate effective cyber defence strategies.

In conclusion, the discrepancies within online threat intelligence sharing platforms present both challenges and opportunities. They highlight areas that require attention and improvement in the field of cyber threat intelligence, but they also pave the way for innovation, as demonstrated by the development of EasyIntel. Understanding the unique features, strengths, and limitations of these platforms and using them in combination can lead to a more comprehensive and accurate assessment of cyber threats. Future research should continue to explore ways to reduce these discrepancies and enhance tools like EasyIntel, ultimately contributing to the ongoing evolution and improvement of the cybersecurity landscape.

7. Further Work

Expand the scope of platform analysis: This study successfully dissected the intricacies of three renowned online threat intelligence sharing platforms, namely VirusTotal, IBM X-Force Exchange, and AbuseIPDB. These platforms were chosen because of their prevalent usage within the cybersecurity community and their distinct methodologies in threat intelligence sharing. However, the realm of threat intelligence is vast and ever evolving, with a multitude of other platforms playing pivotal roles in shaping the landscape.

Future research could potentially expand to include a wider range of threat intelligence platforms, both popular and emerging ones. This could encompass a diverse array of open-source platforms, commercial platforms, and those developed by government agencies or academic institutions. By broadening the scope of analysis, we can acquire a more holistic understanding of the threat intelligence landscape. Each platform carries its unique methodologies for data collection, threat analysis, and intelligence sharing, which can further add to the richness of the dataset under scrutiny.

Inclusion of more platforms can also reveal additional discrepancies and unique features, which have not been explored in the current study. For instance, some platforms might use artificial intelligence and machine learning techniques for threat detection and analysis, while others might rely on human expertise and manual investigation. Some might prioritize real-time threat intelligence, while others might focus more on historical data analysis. These differences can lead to variations in the quality, accuracy, and relevance of the threat intelligence data provided, thereby contributing to the discrepancies observed.

Moreover, the exploration of emerging platforms can shed light on innovative approaches and cutting-edge technologies being employed in the field of threat intelligence. These could include advanced data analytics techniques, automated threat hunting tools, predictive intelligence capabilities, and more. Understanding these new developments can not only reveal additional discrepancies but also offer novel insights and ideas for improving existing platforms and tools, such as EasyIntel.

Deep dive into discrepancy causes: The factors contributing to discrepancies among threat intelligence platforms are manifold and intricate, each affecting the output data in different ways. The data collection and analysis methodologies, for instance, are foundational to the output that each platform generates. Some platforms might rely on a vast network of sensors and honeypots scattered across the internet, which provides them with a broad and diverse set of data points. Others might focus on more targeted data collection, such as specific IP ranges, which can yield highly specialized but potentially less comprehensive threat intelligence. As such, variations in these methodologies can lead to substantial differences in the threat landscape presented by each platform.

Moreover, the level of user expertise and technological capabilities of the platform can also greatly impact the quality and reliability of threat intelligence data. Platforms that prioritize user contributions might generate highly detailed and context-rich intelligence, but the data's reliability can be influenced by the varying expertise levels of the contributors. On the other hand, platforms that heavily rely on advanced artificial intelligence and machine learning techniques for threat detection and analysis can produce highly accurate and fast data. However, these may lack the nuanced understanding of a human analyst, thereby potentially missing out on subtle indicators of compromise.

A deeper understanding of these contributing factors can inform more targeted strategies for reducing discrepancies among platforms. For example, by recognizing the potential limitations of certain data collection methodologies, platforms can work towards broadening their data sources or refining their data collection techniques. Similarly, acknowledging the influence of user expertise on data reliability, platforms might implement more rigorous vetting processes for user-contributed intelligence or invest in further training and development for their users. Understanding the strengths and weaknesses of different technological capabilities can also guide decisions about technology investments and development priorities. By taking such an informed approach, it is possible to enhance the overall quality, accuracy, and relevance of threat intelligence data across the board.

Refinement of EasyIntel: While EasyIntel has been developed as a tool to bridge the discrepancies among various threat intelligence platforms, there is immense scope for continuous improvement and refinement. Given the fast-paced and evolving nature of the cyber threat landscape, it is crucial that EasyIntel remains at the forefront of innovation and usability.

One of the key areas of future focus can be enhancing the user interface. An intuitive, user-friendly interface not only improves the user experience but also facilitates more efficient and effective use of the platform. This could include the development of a more streamlined dashboard, improved data visualization tools, and more efficient navigation. By refining the user interface, we can ensure that both technical and non-technical users can leverage EasyIntel's features to their maximum potential.

Incorporating additional functionalities such as file uploading, and URL scanning could significantly broaden EasyIntel's utility. By allowing users to upload files or scan URLs directly through the platform, EasyIntel can provide immediate threat intelligence, making it a more proactive and versatile tool. This feature would essentially allow users to perform checks against potential threats in real-time, enhancing their defense capabilities.

The introduction of machine learning algorithms to build a database of malicious IPs could dramatically improve EasyIntel's threat detection and analysis capabilities.

Machine learning models can learn from historical data, identify patterns, and make predictions, thereby significantly enhancing the platform's ability to detect potential threats. As the database grows and the model continues to learn, the quality of threat intelligence provided would become increasingly reliable and accurate.

Another potential enhancement could be the development of an algorithm that generates a general threat score based on data from all the included platforms. This score could offer users a quick, composite understanding of a potential threat, saving them time and effort in interpreting data from multiple sources. To cater to diverse user preferences, EasyIntel could also provide an option for users to enable or disable specific platforms as per their liking.

Lastly, for users dealing with large data sets, the ability to run bulk IP searches could be a significant value addition. This feature would allow users to analyse multiple IP addresses simultaneously, drastically reducing the time and effort required for individual searches. This could prove particularly beneficial for organizations and security professionals dealing with large networks or numerous potential threats.

By focusing on these improvements, EasyIntel can continue to evolve as a comprehensive, user-friendly, and highly effective threat intelligence platform that caters to the diverse needs of the cybersecurity community.

Bibliography

- Abu, M. S. (2018, March 18). *Cyber Threat Intelligence – Issue and Challenges*. Retrieved from researchgate.net:
https://www.researchgate.net/publication/322939485_Cyber_Threat_Intelligence_-_Issue_and_Challenges
- abuseipdb. (n.d.). *AbuseIPDB APIv2 Documentation*. Retrieved from docs.abuseipdb:
<https://docs.abuseipdb.com/?python#introduction>
- Andrew Ramsdale, S. S. (2020, May 16). *A Comparative Analysis of Cyber-Threat Intelligence*. Retrieved from researchgate:
https://www.researchgate.net/publication/341454393_A_Comparative_Analysis_of_Cyber-Threat_Intelligence_Sources_Formats_and_Languages
- Baker, K. (2023, March 23). *WHAT IS CYBER THREAT INTELLIGENCE?* Retrieved from Crowdstrike: <https://www.crowdstrike.com/cybersecurity-101/threat-intelligence/>
- Bromiley, M. (2016, September). *Threat Intelligence: .* Retrieved from SANS:
https://nsfocusglobal.com/wp-content/uploads/2017/01/SANS_Whitepaper_Threat_Intelligence__What_It_Is__and_How_to_Use_It_Effectively.pdf
- Chris Johnson, L. B. (2016, October). *Guide to Cyber Threat .* Retrieved from NIST:
<https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-150.pdf>
- Eric Nunes, A. D. (2016, September). *Darknet and Deepnet Mining for Proactive*. Retrieved from researchgate:
https://www.researchgate.net/publication/310498262_Darknet_and_deepnet_minin_g_for_proactive_cybersecurity_threat_intelligence
- geeksforgeeks. (2022, November 26). *History of Python*. Retrieved from geeksforgeeks:
<https://www.geeksforgeeks.org/history-of-python/>
- IBM X-Force. (n.d.). <https://api.xforce.ibmcloud.com/doc/>. Retrieved from IBM X-Force Exchange API: <https://api.xforce.ibmcloud.com/doc/>
- Isuf Deliu, C. L. (2017, December). *Extracting Cyber Threat Intelligence From Hacker Forums: Support Vector*. Retrieved from researchgate:

https://www.researchgate.net/publication/322515698_Extracting_cyber_threat_intelligence_from_hacker_forums_Support_vector_machines_versus_convolutional_neural_networks

Michael Clark, A. B. (2020, February 24). *How to Use Threat Intelligence for Security*.

Retrieved from Gartner :

<https://emtemp.gcom.cloud/ngw/eventassets/en/conferences/hub/security/documents/how-to-use-threat-intelligence.pdf>

Microsoft. (n.d.). *What is cyber threat intelligence?* Retrieved from Microsoft:

<https://www.microsoft.com/en-us/security/business/security-101/what-is-cyber-threat-intelligence>

Python. (n.d.). *Welcome to Flask*. Retrieved from flask.palletsprojects:

<https://flask.palletsprojects.com/en/2.3.x/>

VirusTotal. (n.d.). *VirusTotal API v3 Overview*. Retrieved from developers.virustotal:

<https://developers.virustotal.com/reference/overview>

wikipedia. (n.d.). *CSS*. Retrieved from wikipedia: <https://en.wikipedia.org/wiki/CSS>

Wikipedia. (n.d.). *HTML*. Retrieved from wikipedia: <https://en.wikipedia.org/wiki/HTML>