INSTITUT FÜR INFORMATIK
DER LUDWIG–MAXIMILIANS–UNIVERSITÄT MÜNCHEN

Bachelorarbeit

The use of Cyber Threat Intelligence approaches in the context of security monitoring in university networks

Laura Krämer

INSTITUT FÜR INFORMATIK
DER LUDWIG–MAXIMILIANS–UNIVERSITÄT MÜNCHEN

Bachelorarbeit

The use of Cyber Threat Intelligence approaches in the context of security monitoring in university networks

Laura Krämer

Aufgabensteller:	Prof. Dr. Helmut Reiser
Betreuer:	Dipl.-Inform. Stefan Metzger
Abgabetermin:	October 9, 2017

Hiermit versichere ich, dass ich die vorliegende Bachelorarbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

München, den 9. Oktober 2017

. . . . . .
(Unterschrift des Kandidaten)

Abstract

More and more frequently news channels are reporting on successful cyber-attacks. Among those affected are organizations from all areas, including education and research. Universities are vulnerable to cyber-attacks because they are characterized by a very heterogeneous and decentralized administered network and they own valuable information, such as personal data and research findings. One solution that could support organizations in combating the increasingly number of cyber-attacks is Cyber Threat Intelligence (CTI). The purpose of this thesis is to investigate the use of CTI in the context of the security monitoring of university networks. Therefore, the thesis provides an overview over the more and more popular topic of CTI and defines requirements for the use of CTI in the context of university networks. Moreover, the thesis presents a generic concept for the integration of CTI platforms into the security monitoring of higher education networks as well as a prototypical implementation of this concept.

Contents

1 Introduction

Cybercrime is on the rise in terms of both frequency and sophistication. In its current cyber security strategy paper the German Federal Ministry of Interior warns against the growing risk of cyber-attacks and calls for new approaches to ensure IT security. As the cyber threat situation in Germany is more and more characterized by the complexity and interdependence of technology and a constantly changing threat environment, the impact of successful attacks on a social, economic, political and personal level can be immense [Bun16]. As the result of a study published by the IT association Bitkom in 2015, the damage caused by attacks by cyber criminals is estimated at around 50 billion euros annually. Moreover, the report finds that nearly half of all companies in Germany have been victims of digital economic espionages, sabotage or data theft [Bit15]. Looking at daily news and headlines, a decline of this development is not in sight.

The aims of cyber criminals are as diverse as their techniques. While some attackers are driven by purely financial incentives, others pursue political or ethical objectives. However, in most cases data theft plays a significant role. The importance of data in the information era is bigger than ever before. In other words, data pays off. Whether you as an attacker follow political, economic or financial aims, getting hold of secret data such as credit card information, confidential state files, customers' data, or health records will for sure help to accomplish your goals. As an example, the biggest economic damage is attributable to sales losses due to plagiarism and infringement of patent rights [Bit15]. However, the targets of the attackers are by far not limited to companies alone.

With their philosophy of free exchange of information, widely open networks, huge amounts of data being exchanged every day between thousands of network users - who mostly bring their own device - higher education networks are the perfect target for cyber criminals. Moreover, the list of the data to be obtained is long and promising. First, a lot of personal information is at stake. This includes sensitive data such as matriculation numbers, names, birthdays, place of residence or exam results of students, faculty members, and administrative staff. Moreover, due to their research, universities own a huge amount of intellectual property. One successful cyber-attack can demolish years of work by exposing the research on a new product or methodology.

In the United Kingdom (UK), universities have been a popular target of cyber-attacks for a long time now. According to a report by the software company VMware, 79 percent of 75 surveyed IT decision makers in universities across the UK claimed that their institution has experienced damage to their reputation in the wake of a cyber-attack. 74 percent of them even had to halt a research project and a noteworthy number of 77 percent of interviewees assumed that attacks on higher education institutions have the potential to impact national security, due to the sensitive nature of the information which could been disclosed [VMw16].

It is highly unlikely that this development will stop before the gates of German universities. According to the VDE[1] Tec Report 2017, which covers a survey among the 1,300 VDE member companies and universities across Europe, more than half of the interviewees (53 percent) were already affected by cyber-attacks. Large companies (71 percent) and universities (68 percent) were clearly above-average targets of cyber-attacks [Der17]. Already today, German universities are complaining about targeted attacks on their computer networks. A recent example is the politically motivated hacker attack on printers at different universities throughout Germany [Süd16]. As if by magic, network printers and copy machines of the affected universities started printing racist and anti-Semitic pamphlets. Another example is the recently published report on the allegations that the NSA started hacker attacks on several German universities between 2000 and 2010 [FF16].

May they be induced by political or economic objectives, by state interests or by personal motivations as to take revenge on a former teacher, the reasons for an attack on German universities are manifold. As cyber criminals become more and more professional, the challenges faced by higher education institutions as well will constantly grow. It is mainly up to the IT security staff members of higher education institutions to find appropriate answers to these developments.

1.1 Motivation

No German university can survive today without a comprehensive IT security concept. The IT service providers of universities – in Germany these are usually data centers which are subject to a university or a group of universities – have a lot to offer to students and faculty members: Besides various e-learning and e-teaching platforms, digital grade and assignment management tools and encompassing software, higher education data centers provide emails for students and staff, WLAN in all university buildings and facilities, and much more. However, the longer the list of service offerings the more systems to secure. Many of the above listed tools and services can be used via Internet. While this is reasonable in terms of usability, it also opens the door for Internet-based attacks [WH13].

The list of security monitoring measures and tools that are in use to counter these attacks is long: firewalls, intrusion prevention systems (IPS), intrusion detection systems (IDS), security incident and event management (SIEM), and evaluation of NetFlow data are just a few examples. Over the past few years, a lot has been achieved especially by the intake of SIEM software products and services. As integrated management platforms, SIEM solutions allow real-time monitoring and analysis as well as reaction and notification mechanisms such as automated reporting. Furthermore, SIEMs are the perfect tool for the collection of log data and correlation of security events, which, for example, can be found through their intrusion detection devices or firewalls. However, despite the many security measures and monitoring mechanisms, university systems get compromised. To ensure the integrity of German higher education institutions, it is inevitable to constantly improve the existing cyber security in-

[1] The VDE (German: Verband der Elektrotechnik, Elektronik und Informationstechnik) is one of the largest technical and scientific associations in Europe. Having 36,000 members (including 1,300 companies), VDE is involved in research, standardization work, technical knowledge transfers and product testing in the field of electrical engineering and information technology [VDE16].

frastructure and to search for new approaches.

A new approach that is already part of heated discussions among industry experts, is Cyber Threat Intelligence (CTI). According to most cyber security specialists, CTI is the answer to many security problems which the IT industry is facing today. In short, CTI is a proactive measure to combat cyber criminals and to protect one's systems from being compromised. The basic idea is to collect relevant data on security vulnerabilities, attack patterns, malware and so on, and to analyze and correlate that data (thereby transforming them into intelligence), resulting in valuable and actionable information that can be used to protect your own systems. While existing security monitoring systems are primarily suited to detect compromised systems within one's network, to provide notification mechanisms and to support the security staff in finding appropriate reactions, the promise of CTI is that it will help to detect *future* attacks before they result in harm.

According to its proponents, CTI attempts to minimize the risks of potential damage such as interruptions in university operations due to technical disturbances, loss of personal data and research results or damage to the reputation of the university. Therefore, CTI falls under the category of proactive measures, as opposed to detection or reactive measures that take place during and after the attack on a system. Both proactive and reactive measures play a crucial role in ensuring information security which is why it is worth researching new methods and possibilities in both areas. Therefore, the aim of this thesis is to investigate whether CTI is a suitable complement to the steps already taken in the field of proactive and reactive measures within the area of higher education networks and to examine ways for its integration into the security monitoring of higher education networks.

1.2 Problem Formulation

The present thesis examines ways to improve the cyber security monitoring in higher education institutions by incorporating CTI approaches into the existing tools and mechanisms. The principal aim of this examination is to assess whether CTI is suitable to improve the effectiveness of the mechanisms used in the context of security monitoring in higher education networks. Therefore, the cyber security monitoring tools that are currently used by higher education data centers to secure their networks are being presented and examined to determine whether they are suitable for the incorporation of CTI.

The Leibniz Supercomputing Centre (LRZ) of the Bavarian Academy of Sciences and Humanities serves as a reference framework for the practical part of this thesis. The LRZ supports research, teaching, and administrative processes at universities in Bavaria, among which are the University of Munich (LMU) and the Technical University of Munich (TUM) as well as research institutions such as the Max-Planck-Institute. As an IT-service provider, the LRZ offers a great variety of services such as communication infrastructures (e.g. email and web services), e-learning and e-teaching platforms, or file storage. Furthermore, it operates the Munich Scientific Network (MWN) at more than 540 locations.

The LRZ network now has over 200,000 terminals. However, most of them are not managed directly by the LRZ. By placing different sensors at the central X-WiN network transition,

security monitoring is used to detect attacks on systems in the MWN, as well as conspicuous communication behavior of associated malware-infected systems. After a successful detection, attacks are forwarded to a SIEM system where they are correlated and enriched with additional information to keep the number of false positives as low as possible. Currently, this is the most important part of the LRZ security monitoring mechanisms to report to the LRZ Abuse-Response-Team, the network administrator, or the user about serious events.

First, a general overview of the concept of CTI and commonly used approaches are to be presented in this thesis. Besides its overall purpose, current standardization efforts in the field of CTI as well as ways to incorporate it through platforms and feeds are to be examined. After analyzing the specifics of higher education networks as well as monitoring mechanisms that are in use at the LRZ, but which are also typical of many universities throughout Germany, requirements for the extension of these with CTI approaches are to be identified and summarized in a weighted criteria catalog. Moreover, a selection of CTI platforms that could be integrated into the security monitoring systems are to be presented and analyzed according to the defined requirements. Subsequently, a multistage concept for integrating CTI platforms into higher education networks is to be presented. In addition, for demonstration purposes, parts of the developed concept should be prototypically implemented.

1.3 Outline of the thesis

The present thesis is divided into six chapters. Chapter 1 entails the introduction and describes the motivation behind this thesis. Chapter 2 introduces the main topic of this thesis and gives a general overview over CTI approaches. Chapter 3 presents the requirements analysis of this thesis. The first part (Chapter 3.1) of this chapter discusses the specifics of higher education networks. The second part (Chapter 3.2) then gives an overview over the security monitoring that is deployed at the LRZ. The third part (Chapter 3.3) defines requirements for the use of CTI platforms in the context of the security monitoring in university networks. The fourth and fifth parts of Chapter 3 (Chapter 3.4 and Chapter 3.5) present the analysis of CTI platforms as well as the results of this analysis. Chapter 4 describes a multistage concept for the integration of CTI into the security monitoring of university networks. Chapter 5 implements parts of the concept that has been developed in Chapter 4. Chapter 6 concludes the thesis by summarizing the main findings of this thesis and presents an outlook for future work.

2 Overview - What is Cyber Threat Intelligence?

The first thing to realize about CTI is that there is no common understanding of the subject. When reading related articles, papers and studies the most commonly expressed opinion on CTI is that it is "a loosely used term"[Gar14], that "it means different things to different people"[Dal15] and that there exists "a lot of confusion around what threat intelligence is and how it's delivered and consumed" [Sha15].

One of the rarely presented definitions of CTI has been published by the US research and advisory firm Gartner:

> "Threat intelligence is evidence-based knowledge, including context, mechanisms, indicators, implications and actionable advice, about an existing or emerging menace or hazard to assets that can be used to inform decisions regarding the subject's response to that menace or hazard"[Gar14].

Another US research and marketing consulting firm defines CTI as follows:

> "Cyber threat intelligence is knowledge about adversaries and their motivations, intentions, and methods that is collected, analyzed, and disseminated in ways that help security and business staff at all levels protect the critical assets of the enterprise"[JF15].

Being one of the buzzwords of current cyber security debates, CTI has been examined by various organizations, institutes and enterprises, who offer insights both from an academic and business perspective. There are some scholarly articles such as Qamar et al. (2017) [SQ17] and Mattern et al. (2014) [TM14] who underline the proactive nature of using CTI as a source to predict cyber threats and gain knowledge about present as well as future techniques and attack patterns. In their article *Data-driven analytics for cyber-threat intelligence and information sharing* Qamar et al. aim at providing mechanisms that allow automated evaluation of large volumes of CTI data by providing a threat analytics framework based on Web Ontology Language (OWL). The proposed framework is designed to classify the threat relevance and likelihood and thereby evaluate the risk it poses to a network. Burger et al. also address CTI in a scientific way. In *Taxonomy Model for Cyber Threat Intelligence Information Exchange Technologies* they provide a taxonomy model to classify, identify and reveal the differences of exisiting CTI technologies [EWB14].

However, there exist far more researches and studies done by non-academic research, marketing and IT security firms such as Gartner, SANS Institute or Securosis. A comparison of the academic and non-academic research reveals a different focus: Studies that are deriving from a business perspective aim at explaining the benefits of the integration of CTI into an

organization's security monitoring system and how to put it into practice. Scholars, on the other hand, look at ways for automated sharing and gathering of CTI as well as techniques to improve the overall quality of the shared data. Standardization efforts, on the other side, is a topic both academic and non-academic researchers show a great interest in.

The reasons of researchers with various backgrounds and objectives for participating in the debate are linked to their confidence that CTI is a key solution for improving todays cyber security mechanisms. Proponents argue that CTI can help organizations to

- be better prepared for cyber attacks
- be aware of present and future threats
- better detect existing threats
- develop a better understanding of attacks
- develop more accurate defense strategies against attacks

To reach its full potential CTI must meet certain requirements. One of the most commonly mentioned requirements for CTI is that it needs to be **actionable** [Dal15, TM14, SQ17]. By this, researchers underline the fact that there is no use for CTI data if it does not provide a reason upon which some sort of action or decision can be based on. Another important requirement is that CTI needs to be **relevant** [Dal15, TM14]. If an organization gathers huge amounts of CTI of which none is important to them from an institutional or technical perspective, then the content of the collected CTI to that specific organization is useless. A third requirement, which is often mentioned by researchers, concerns the **timeliness** of CTI data [TM14, EWB14]. As adversaries constantly develop new techniques and tools to compromise systems and launch new attacks against networks, the relevance of the corresponding data changes rapidly. Therefore, consumers of CTI need to keep pace and constantly update their records.

2.1 CTI Contents

Data that counts as CTI can be characterized as either strategic or tactical [Far13]. **Strategic CTI** provides information about the motivation behind cyber-attacks and therefore can help to understand why a certain system is or will be under threat (e.g. geopolitical or sector-specific indicators). In addition, strategic CTI can be very useful for preventing future attacks by delivering information on so-called Tactics, Techniques, and Procedures (TTPs) and thereby revealing how an attack is planned and executed. Analyzing strategic CTI can therefore help to improve the overall security infrastructure by revealing vulnerabulities in the systems and processes of a company or organization. **Tactical CTI**, on the other hand, entails indicators on which an organization must focus if it wants to prevent and detect an attack. Typically, tactical CTI entails so-called Indicators of Compromise (IOCs). While strategic CTI is usually produced and consumed by humans, tactical CTI such as IOCs can as well be produced and consumed by machines. Furthermore, tactical CTI is more of a short-lived product since its relevance can change from very high to very low (or irrelevant) within a very short time.

To show the difference between strategic and tactical CTI, Table 2.1 gives an overview of typical TTPs and IOCs.

TTPs	IOCs
• spear phishing with malicious file attachments	• IP addresses
• malware disguised as a Microsoft Office plugin	• Unified Resource Locators (URLs)
• point-of-sale (PoS) intrusions	• email addresses
• dumping cached authentication credentials	• file hashes
• using botnets to perform distributed denial of service (DDoS) attacks	• domain names
• using pop-up windows for phishing attacks	• registry key values
• Cross-site scripting to bypass access controls	• HTTP requests
• ...	• ...

Table 2.1: Overview of TTPs and IOCs

To better understand the difference between strategic and tactical CTI, it can be as well useful to take a look on the so-called Pyramid of Pain which is displayed in Figure 2.1. The Pyramid of Pain was originally developed by David Bianco [Bia13] to illustrate the value of different types of cyber threat information. He therefore classified them in different categories depending on how much pain it will cause an adversary if its target learns about one of the displayed indicators. The four lower levels of the pyramid, which are made up of hash values, IP addresses, domain names and network/ host artifacts, are all IOCs and thus count as tactical CTI. The upper two levels are made up of TTPs and Tools which count as strategic CTI. As one can see from the level description, at the beginning of the pyramid the adversary's pain is the lowest if detected and denied. However, the more up you go in the levels, the more painful it gets for the adversary. While obtaining strategic CTI and thereby detecting and denying adversaries on one of the upper two levels is obviously more painful to the adversary, it is also more difficult to collect and process such data than it is with IOCs. A major reason for this is that tactical CTI can be produced and consumed by machines and therefore is more suitable for automated exchange and consumption. Hence, it is much easier to obtain such data and use it to defend a network against malicious actors.

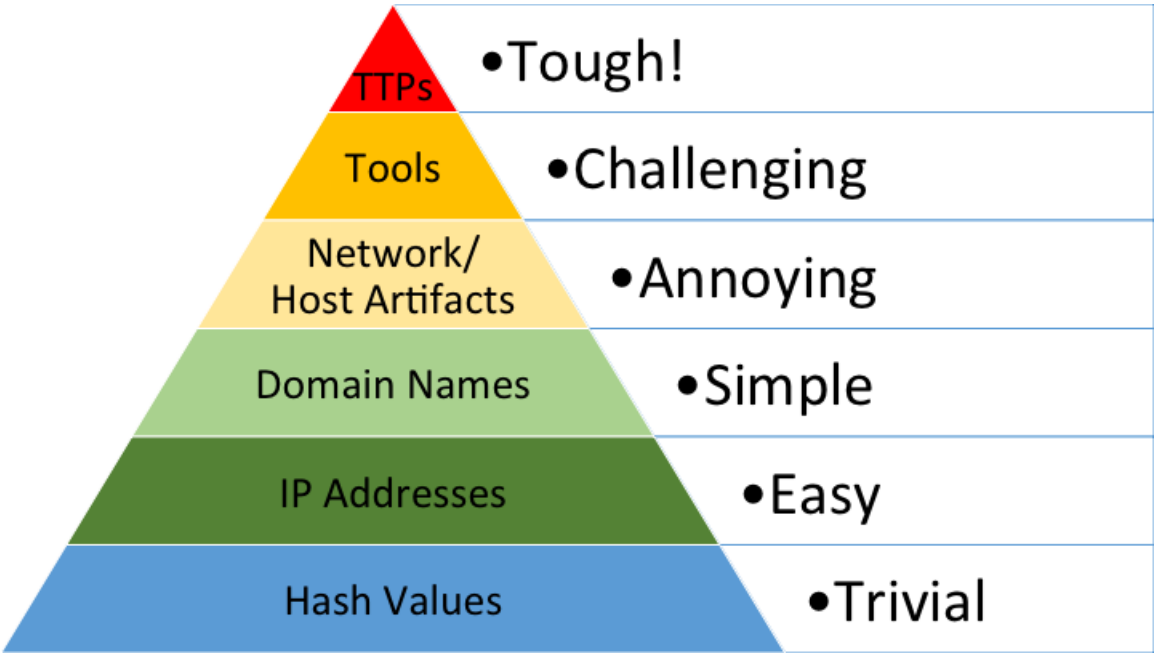

Figure 2.1: Pyramid of Pain
[Bia13]

2.2 CTI Frameworks, Tools and Sources

Usually CTI is submitted via data feeds. Organizations that plan to integrate CTI into their security monitoring system can subscribe to such feeds from multiple sources. These sources can be divided into internal and external sources, the latter being again divided into commercial and open or public sources. Another type are community-driven sources, which are somewhere in between these categories since they rely on the mutual exchange of CTI and can be commercial as well as open-source. Depending on an organization's capabilities and resources it can gather its own, internal CTI from an IDS or its SIEM solution, by analyzing and processing the thereby obtained data in such ways that it fulfills the above discussed requirements. "As enterprises experience exploit kits, malware infections and other daily issues that can seem random and unconnected, they have an opportunity to build a profile of their environment by organizing such information into meaningful content" [Bro16]. Another source for processing your own CTI are search engines for Internet-connected devices, such as Shodan or ZoomEye, that allow its user to detect devices with insufficient security settings as well as other system vulnerabilities. However, this data is only as useful as one's ability to turn it into actionable intelligence. Therefore, the possibility of gathering internal CTI depends on the organization's analyzing capabilities and processing skills. Apart from that, every organization needs to decide for itself whether it wants to share its internal CTI in a community or not.

In addition, there exists a great variety in external CTI sources. As already stated, external sources can be divided into commercial and open-source. Another distinction can be drawn between governmental and non-governmental sources. Organizations can decide

between a great variety of providers and combine CTI from different sources according to their needs and objectives. One way of purchasing and exchanging CTI that is currently gaining in popularity is the use of CTI platforms. More and more providers are offering platform solutions that provide a range of functionalities to improve the exchange of CTI data. Yet, just as CTI is a broad concept, there exists no common understanding or definition of what constitutes a CTI sharing platform [CS17]. Sauerwein et al. (2017) [CS17] have identified five different types of platforms which primarily focus on sharing IOCs. However, they argue that most of the provided data cannot count as intelligence in its strictest sense, since "the majority of tools primarily focuses on data collection and more or less neglects the other activities of the intelligence lifecycle" [CS17]. According to them, CTI sharing platforms "provide limited analysis and visualization capabilities and lag behind comparable knowledge sharing platforms and data mining solutions from other domains" [CS17]. In the end, what constitutes CTI lies in the eye of the beholder. A brief selection of different open-source and commercial CTI (in its broadest sense) frameworks, tools and sources can be seen in Table 2.2.

Open-source or free-to-use	
AlienVault Open Threat Exchange (OTX)	Platform
Collective Intelligence Framework (CIF)	Framework
Collaborative Research Into Threats (CRITs)	Platform
Cymon	Open tracker of CTI information
Hail a TAXII	Repository
Malware Information Sharing Platform (MISP)	Platform
OpenIOC	Framework
PhishTank	Open community for sharing phishing data
STAXX (Anomali)	Platform
VirusTotal	Open community that analyzes suspicious files and URLs
Commercial (partially free-to-use)	
Anomali ThreatStream	Platform
CTX/SOLTRA EDGE	Software
EclecticIQ	Platform
FireEye	various CTI services
IBM X-Force	Platform
Redsocks	Virtual solutions and hardware appliances
ThreatConnect	Platform
ThreatQ	Platform
Governmental	
CERT-Bund	Warning and information service
Computer Emergency Response Team for the EU institutions, bodies and agencies (CERT-EU)	Warning and information service
Industrial Control Systems Cyber Emergency Response Team (ICS-CERT) (USA)	Warning and information service
National Cybersecurity and Communications Integration Center (NCCIC)	CTI assessments and information service

Table 2.2: Overview of CTI frameworks, tools and sources

2.3 CTI Standards

Before subscribing to a CTI data feed or participating in a CTI community, any organization should ensure that the quality of the CTI it plans to consume meets the above discussed requirements and corresponds to the organization's needs. Moreover, it is essential to think about how, when and where the data is sourced. Therefore, before consuming CTI it is important to ensure that it is timely (How old is the provided data?), relevant (e.g. according to the geography or sector of the organization), actionable and effective (e.g. when it comes to reducing false positives or its correlation with other data) [Gar14]. Any security practitioner should have these questions in mind before choosing a CTI source for its company or organization. Considering that a high volume of complex CTI is generated manually, validation of this data becomes crucial. Hence, another problem concerning CTI is that it is often redundant, incomplete or – even worse – incorrect [SQ17]. According to Qamar et al., "This non-uniformity, as well as redundancy of data, makes it more challenging to analyze a sample CTI report for identifying its relevance for a particular network" [SQ17].

Given the fast-changing landscape of cyber threats and advanced persistent threats (APT), mutual exchange of relevant cyber threat information becomes indispensable. In most cases an attack is not designed to just hit one single target, but to exploit weaknesses that can be found in multiple systems, products or networks. Hence, the benefits of looking at how, when and why organizations have been compromised can be enormous. This applies in particular if an organization operates within the same sector or has similar network configurations. Therefore, the mutual exchange of such information counts as a fundamental pillar of CTI.

However, as has been discussed in the previous section, CTI should fulfill certain requirements. At least, it should be actionable, relevant and timely. Besides this, there are some quality issues, such as redundancy, inaccuracy and even incorrectness, which threaten to compromise the effectiveness of CTI. Under the given circumstances, it is inevitable to incorporate standardized methods for producing and sharing CTI data. There are multiple attempts to develop and establish such standards and frameworks on the way. Mostly, they focus on the development of a standardized language and structure for the documentation of cyber threats as well as tools and techniques to enable automated and standardized distribution. The most important endeavors to establish a standardized format for sharing CTI are presented in Table 2.3.

At the moment, Structured Threat Information Expression (STIX) and Trusted Automated eXchange of Indicator Information (TAXII) are the most promising CTI standards on the market. STIX and TAXII are open community efforts which are constantly being further developed. Initially, the MITRE Corporation served as the creator and moderator of the STIX and TAXII community on behalf and under the sponsorship of the U.S. Department of Homeland Security. However, in 2015, the U.S. Department of Homeland Security transitioned STIX and TAXII to the OASIS Cyber Threat Intelligence Technical Committee (OASIS CTI TC) for further development. Today, STIX and TAXII are being widley discussed among the CTI community. Moreover, they have been already incorporated into many CTI platforms and tools, a circumstance that makes them a de-facto standard for CTI sharing and its consumption. Therefore, the following sections provide a brief overview over STIX and TAXII as well as an analysis of the specifics of their adoption by the community.

Acronym	Title	Organization	Further Information
STIX	Structured Threat Information eXpression	Mitre	STIX is a community-driven effort to develop a structured language and serialization format which can be used to represent and exchange cyber threat information in a machine-readable manner.
CybOX	Cyber Observable eXpression	Mitre	CybOX is a standardized language that provides a common structure to represent information about cyber observables. CybOX has been integrated into Version 2.0 of STIX and is no longer a standalone standard.
TAXII	Trusted Automated eXchange of Indicator Information	Mitre	TAXII provides a standard to exchange CTI across organizations. Among other things TAXII defines concepts and protocols to enable automated exchange of cyber threat information. Being initiated by the same organization, STIX, CybOX and TAXII are designed to match perfectly together.
VERIS	Vocabulary for Event Recording and Incident Sharing	Verizon	VERIS provides a structured format and language for describing cyber security incidents in a structured and repeatable manner. Therefore, it defines a set of categories and metrics to describe these incidents. VERIS aims at providing a structured format that can be used to identify trends within the security sector and to overall improve the quality of cyber threat information.
IODEF	Incident Object Description Exchange Format	Managed Incident Lightweight Exchange (MILE) working group	IODEF is a standardized format that provides a framework for the exchange of CTI. Primarily, it was designed for information exchange between Computer Emergency Response Team (CERTs) and Computer Security Incident Response Teams (CSIRTs).
IODEF-SCI	IODEF for Structured Cybersecurity Information	Internet Engineering Task Force (IETF)	IODEF-SCI is an extension of IODEF. IODEF-SCI uses other formats and defines a pattern to consistently embed structured information into IODEF.

Table 2.3: Standardized formats for sharing CTI

2.3.1 STIX

STIX is a community-driven effort to provide a structured language for the representation of cyber threat information. One of its major goals is to enable the automated exchange of high-quality CTI across organizations. According to the STIX community, cyber threat information should be expressive, flexible, extensible, automatable, and readable [Bar14]. STIX goes beyond simple indicator sharing by providing rulesets for sharing whole sets of indicators – the STIX Domain Objects (SDOs). Therefore, STIX provides an architecture (displayed in Figure 2.2) to combine various SDOs trough relationships. Since its first introduction STIX has been redesigned. The specification of its current version, STIX 2.0, defines the following SDOs [OAS7a]:

- **Attack Pattern:** A type of TTP that helps to categorize attacks and provides information on how an attack is performed
- **Campaign:** Adversarial behaviors that include a set of malicious activities and resources, such as infrastructure, Malware, or Tools. Campaign attacks occur over a period of time and can be characterized by a specific set of people or resources they target
- **Course of Action:** A textual description of taken actions to either prevent or respond to an attack that is in progress
- **Identity:** Identification of individuals, organizations, groups, or whole sectors that qualify as targets of attacks or information sources
- **Indicator:** Textual description of patterns that includes context, such as time. Indicators are used to detect suspicious cyber activity
- **Intrusion Set:** A grouped set of adversarial behavior and resources which capture Campaigns or other activities with common properties thereby indicating a shared Threat Actor
- **Malware:** A type of TTP that is designed to secretly compromise the confidentiality, integrity, or availability of data obtained by the targeted person
- **Observed Data:** Information that was observed on systems, such as an IP address, a network connection, a file, or a registry key, without providing further context
- **Report:** Collections of threat intelligence set to be published to provide context based details on threat actors, malware, attack patterns, etc.
- **Threat Actor:** Individuals, groups, or organizations who can be characterized by various indicators, such as motives, capabilities, goals, or resources, and who are believed to have malicious intentions
- **Tool:** Software packages that can be used to perform cyber attacks, such as remote access tools or network scanning tools
- **Vulnerability:** A weak point in a system that can be exploited by a threat actor to get access to a system or network

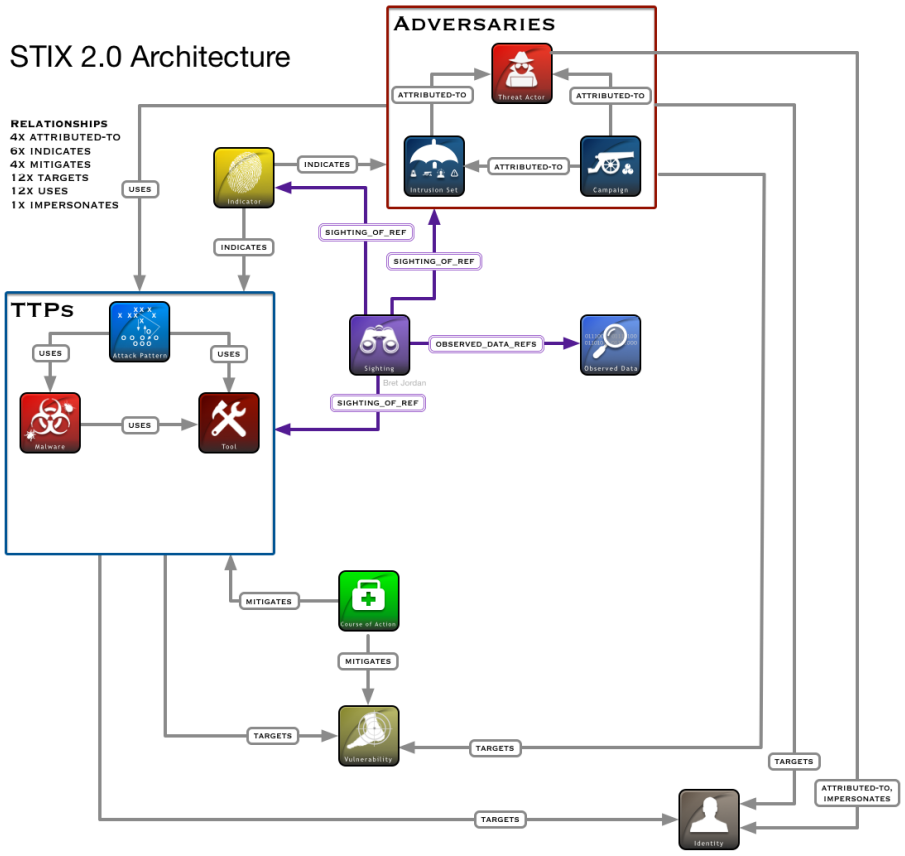

Figure 2.2: STIX 2.0 Architecture
[OAS7b]

First conceived as a language to describe CTI, STIX has undergone a huge transition since its jumpstart in 2012. A huge milestone was reached in 2015, when it was transitioned to OASIS CTI TC, an international standards consortium. This was a major step for becoming accepted as an international standard. In March 2017, the most current version, STIX 2.0, was approved by the OASIS CTI TC. STIX is and was designed to provide a format to capture knowledge about adversaries and their behaviors – such as using spear phishing emails to deliver a Trojan – in a machine-readable manner and to enable automated sharing of this knowledge among organizations. However, while STIX 2.0 builds on the foundation of its previous versions, a few distinctions are worth mentioning.

Unlike earlier versions, STIX 2.0 uses JSON as an exchange format instead of XML. Furthermore, STIX now is a graph-based model, where STIX Domain Objects are tied up together by using STIX Relationship Objects as can be seen in the STIX 2.0 Architecture in Figure 2.2. While previous versions also contained relationships between various components that could be compared to the SDOs of the current version, STIX 2.0 promises to make these connections more explicit and to allow analysts to a) correlate data, even if it first seems that there exists no connection between this data, b) draw connections from IOCs to respective adversaries who initiated the compromise, and c) develop connections over time [Wun7c]. Currently, there are six different STIX relationships: attributed-to, in-

dicates, mitigates, targets, uses, and impersonates. They are represented by gray arrows in the STIX architecture to show the connection between two SDOs.

Figure 2.3 displayes a simple STIX 2.0 JSON example drawn from the STIX 2.0 documentation [OAS7c]. The use case is a malware that gets delivered via an URL. This a very common phishing method where the victim gets directed to a malicious URL via an e-mail. Once the victim clicks on the URL, it gets compromised. Two SDOs are relevant for this uses case: the **Malware** and **Indicator** SDOs. The used relationship is **indicates**. The malicious URL used in the example is http://x4z9arb.cn/4712/. To provide more context about the URL, the Indicator SDO must contain a labels property which is taken from an Indicator Label open vocabulary. In this case, the label (*labels*) property is "malicious-activity". Moreover, Indicator SDOs require another field called *valid_from* which in this case will be the time from when the object was created. Since the URL in this use case delivers a malware, the second SDO to be used is Malware. Again, labels properties are used to classify the SDO. The indicates relationship is then used to link the two SDOs together. To specify the source (*source_ref*) and the target (*target_ref*), the indicator id and the malware id respectively are used.


```
{
  "type": "bundle",
  "id": "bundle--44af6c39-c09b-49c5-9de2-394224b04982",
  "spec_version": "2.0",
  "objects": [
    {
      "type": "indicator",
      "id": "indicator--d81f86b9-975b-bc0b-775e-810c5ad45a4f",
      "created": "2014-06-29T13:49:37.079000Z",
      "modified": "2014-06-29T13:49:37.079000Z",
      "labels": [
        "malicious-activity"
      ],
      "name": "Malicious site hosting downloader",
      "pattern": "[url:value = 'http://x4z9arb.cn/4712/']",
      "valid_from": "2014-06-29T13:49:37.079000Z"
    },
    {
      "type": "malware",
      "id": "malware--162d917e-766f-4611-b5d6-652791454fca",
      "created": "2014-06-30T09:15:17.182Z",
      "modified": "2014-06-30T09:15:17.182Z",
      "name": "x4z9arb backdoor",
      "labels": [
        "backdoor",
        "remote-access-trojan"
      ],
      "description": "This malware attempts to download remote files after establishing a foothold as a backdoor.",
      "kill_chain_phases": [
        {
          "kill_chain_name": "mandiant-attack-lifecycle-model",
          "phase_name": "establish-foothold"
        }
      ]
    },
    {
      "type": "relationship",
      "id": "relationship--6ce78886-1027-4800-9301-40c274fd472f",
      "created": "2014-06-30T09:15:17.182Z",
      "modified": "2014-06-30T09:15:17.182Z",
      "relationship_type": "indicates",
      "source_ref": "indicator--d81f86b9-975b-bc0b-775e-810c5ad45a4f",
      "target_ref": "malware--162d917e-766f-4611-b5d6-652791454fca"
    }
  ]
}
```


Figure 2.3: STIX 2.0 JSON example
[OAS7c]

2.3.2 TAXII

TAXII is a transport mechanism used to exchange CTI in a secure and automated manner across organizations. It is an application layer protocol used to exchange CTI over HTTPS. Although most commonly the data to be shared over the TAXII protocol is represented in STIX, TAXII can be used to share other formats as well. Nonetheless, STIX and TAXII were designed to work together and their development is driven by the same community, which is why they usually are used in combination with each other. TAXII enables organizations to define an application programming interface (API) to share and receive CTI. Therefore, TAXII defines two services, Collections and Channels.

Collections allow TAXII Clients to exchange CTI with a TAXII Server in a request-response manner. The TAXII Client makes an information request to the TAXII Server which then responses to that request by sending the demanded information.

Channels, on the other hand, allow TAXII Clients to exchange CTI with other TAXII Clients in a publish-subscribe manner. Therefore, TAXII Clients can publish messages to Channels, which are maintained by a TAXII Server, with the effect that other TAXII Clients that have subscribed to that Channel will receive the respective messages [OAS7d].

Figure 2.4 illustrates how Collection and Channel based communications are used.

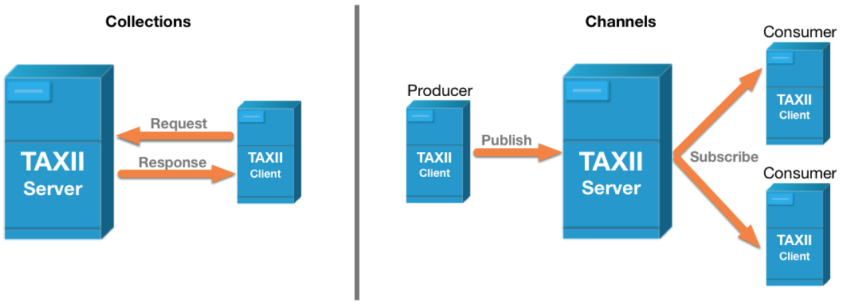

Figure 2.4: TAXII Collections and Channels
[OAS7d]

A TAXII API Root consists of Channels, Collections, and their related functionality. Each API Root is available at an URL and contains a set of Endpoints that can be contacted by a TAXII Client to request information, such as a description of the Collection that is linked to the Endpoint, the Collection's content, or to get objects from this Collection. To allow access control for Channels and Collections used by different groups, a TAXII Server can support several API Roots – one for each user group. TAXII Endpoints can be identified by its URL and HTTP method which a TAXII Client can use to make its request [OAS7d].

3 Requirements analysis

The aim of the requirements analysis is to define prerequisites for the integration of CTI platforms into security monitoring mechanisms of higher education institutions. CTI platforms are very useful if an organization wants to automatically integrate CTI into their security monitoring mechanisms. Since they offer many functionalities to collect, correlate, analyze, and export threat data from multiple sources and formats, the analysis concentrates only on such platforms for the integration of CTI instead of websites, such as PhishTank, which only provide indicators related to a specific attack pattern, such as phishing attempts. The first sections of this chapter examine the characteristics of higher education networks. Despite many similarities, it is still hardly possible to describe a single prototype of a university. Therefore, the application scenario of the MWN will serve as the main reference framework for the requirements analysis. After defining various requirements (**R01 - R26**), a selection of CTI platforms will be analyzed based on those requirements.

3.1 Specifics of higher education networks

The following sections deal with the specifics of higher education networks in contrast to company networks. They are mostly derived from the analysis of higher education networks which has been carried out by a former master student, Michael Steinke, at the LMU in his final thesis [Ste15]. The overall aim of the following sections is to present an overview over the most important pecularities of higher education networks to get a better understanding of the application scencario used in this thesis. The present section consists out of four subsections which deal with different charcteristics of higher educaiton networks and make a comparison to company networks.

To understand the specifics of higher education networks, it is important to look at the organizational structure of universities. Therefore, Steinke did a brief analysis of five universities throughout Germany. The number of students at those universities varied between 14.000 and 42.000 and the number of staff varied between 2.000 and 12.000. Even though there is some variance of user numbers in the respective university networks, they all can be characterized by a great number of users. Among the users of higher education networks are scientists and employees of the individual institutions as well as students and researchers. Additional users are affiliated institutions that also get integrated into the university network, such as scientific institutions, libraries, student residences, or university hospitals. Unlike universities throughout the US, there are only few campus universities in Germany. Consequently, buildings that belong to the university are often spread over the whole city area. The organizational structure of the universities themselves is decentralized. Usually, their structure is split into faculties which again are subdivided into different institutes and departments. The following sections examine the peculiarities of German higher education networks in terms of network management and administrator rights, IT services and network access.

3.1.1 Network management and administrator rights

Mostly data centers are responsible for the planning, provision, and operation of the university network. However, to balance the load on the network, university networks are usually subdivided into subnets. Normally, those subnets are then being administered in a decentralized manner by employees who do not belong to the data center but to a faculty or institute. Therefore, the responsibility of the data center often ends with the data socket of the faculty or institution. One of the most striking characteristics of university networks is that the "bring your own device" (BYOD) philosophy has become a matter of course. This is a great difference to company networks where employees are being provided with devices by their company. Even though BYOD is gaining popularity at companies as well, the extent of its dissemination at companies is yet not comparable to its dissemination at universities. An important consequence of BYOD at universities is that administrator rights are distributed decentralized. Hence, not only employees of the data center can have administrator rights, but employees of the different faculties and institutes or the students themselves. Due to distributed administrator rights, some important security measures that have proven valuable in company networks cannot be enforced. Such measures, for example, include whitelists or blacklists for certain software as well as compulsory installation of software throughout the network [Ste15].

3.1.2 IT services

The primary aim of universities is to support teachers and students and to provide an appropriate framework to drive forward their research and teaching on the highest level. Therefore, a university network must serve as the provider of multiple IT services. These services include both classical (e.g. email services, file systems, or provision of certain software) as well as higher education specific services, such as e-learning and e-teaching platforms or digital grade and assignment management tools. Among the most important services that guarantee the efficient use of a network are DHCP (Dynamic Host Configuration Protocol) and DNS (Domain Name System). By using a DHCP server, the device of a user is automatically assigned an IP address when connected to the network. Usually, the data centers operate the DHCP and DNS servers. However, sometimes a faculty or institution can provide this service as well. In most cases, Eduroam, an international roaming service, is also integrated into the service offering. By using Eduroam, users in research and higher education can gain network access at any institution they visit by using the same credentials as when they access their local network. Another characteristic of university networks concerning the provided IT services is the use of private IP addresses. According to RFC 1918, private IP addresses in certain ranges are identified as private and therefore are not routable to the Internet. Unlike public IP addresses, private IP addresses are not allocated to a specific organization. Since they are private and have not been approved by an Internet registry, the packages sent from a private IP address cannot simply be transmitted through the Internet. Such requests must use a certain technology, such as a network address translation (NAT) gateway to translate the private into a public IP address [Ste15].

3.1.3 Network access

There are some striking characteristics of university networks concerning the access to them. To guarantee independent research and free exchange of views, the reliable availability of

services is one of the most important aims of university networks. Therefore, data centers aim at making the accessibility as easy as possible. There are several ways of getting access to the network. Besides the use of the wireless LAN access, which can be used to register the own devices, most universities provide their students with already configured workstations. Moreover, students take it for granted to be able to access the university network through another network via a VPN tunnel. Notwithstanding the above described possibilities of gaining access to the university network, the user needs to authenticate. Most commonly, this is done by using a combination of a network-wide user recognition and a password. As described in the previous section, if their home institution provides Eduroam, students and teachers can use those credentials to gain network access at any other institution providing Eduroam as well. In addition, the university network can serve as a point of access to larger networks such as the German National Research and Education Network (DFN) or the pan-European research network GÉANT [Ste15].

3.1.4 Summary

Due to the multiple possibilities of gaining access to the network, a huge but also changing user basis, and the widely-spread practice of BYOD, university networks can be characterized as **high-dynamic**. Moreover, university networks are characterized by a high extent of **heterogeneity**. First of all, this is an effect of the BYOD philosophy. Students, staff and teachers bring various devices, such as laptops, tablets, smartphones or smart watches and connect them to the university network. Those devices again operate on various operating systems, such as Microsoft Windows, Android, macOS, iOS, or Linux. Furthermore, the software applications that can be found in a university network are diverse. They range from products from large manufacturers to all types of open-source products, to self-developed applications for scientific purposes as well as for private use [Ste15]. Last but not least, university networks can be characterized as **decentral** due to the distributed administrator rights and responsibilities concerning the subnets. In contrast to company networks, the users of university networks enjoy a greater freedom. Conversely, the ability of the network operator to control and manage the network is lower. While network administrators of company networks often have full access to the devices and even more rights than the user himself, this is only the case for a relatively small part of the university network [Ste15].

3.1.5 Application scenario: the Munich Scientific Network

The application scenario for the use of CTI approaches in the context of security monitoring in university networks is the MWN which is operated by LRZ. The connection of the MWN with the worldwide Internet is established via the German scientific network (X-WiN), which is provided by DFN. The MWN connects various Universities, such as the LMU, the TUM and the University of Applied Sciences Munich (HM). Moreover, it is also used by other scientific institutions (e.g. Max Planck Society, Fraunhofer Society, Museums, and the like). The MWN forms the basis for the communication and cooperation within and between the affiliated institutions as well as with cooperation partners in Germany, Europe and the whole world [Lei16].

Even though most of the affiliated institutions are located around Munich, there are also some facilities at other locations in Bavaria. At present, more than 500 buildings or groups

of buildings are connected to the MWN and up to 180,000 devices are supplied via the MWN. The size of the supplied areas ranges from a single building to an entire "campus area" with more than 30 buildings. At present, 52 student residences with a total of more than 12,600 dormitories are connected to the MWN. The LRZ is responsible for the entire backbone network as well as large parts of the connected institute networks. Only the internal networks of the medical faculties of the Munich universities as well as the computer science institute of the TUM are an exception. They are supported by the respective data centers of the faculties themselves [Lei16].

3.2 Security monitoring at the Leibniz Supercomputing Centre

The LRZ operates the MWN which serves as the application scenario of this thesis. One peculiarity of the MWN is that the LRZ's responsibility ends at the data socket of the institution supplied. The faculties and departments decide themselves which systems they connect to the network and therefore carry the administrative responsibility for the connected terminal devices. However, despite its distributed network structure and the very large number of connected terminal devices, which are usually administered decentrally, the LRZ is trying to protect its own systems as well as the systems that are managed for other facilities from attacks. In addition, it tries to prevent attacks or attack attempts which are launched from LRZ-supervised systems or those systems that are connected to the MWN [SM11].

Since mid-2001, the activity of hackers worldwide has increased dramatically and likewise the abuse cases in the MWN have multiplied. According to LRZ representatives, this is due to the following reasons: First, due to the increasing criminalization of the Internet, the tools of hackers and other cyber criminals are being developed by highly specialized experts and therefore have noticeably improved. Second, the number of malware, such as viruses, worms or Trojans increased drastically and became much harder to detect. Finally, the security awareness and behavior of too many users is still insufficient [HR16]. To meet these challenges, it is crucial that the data that is offered via a CTI platform is of high quality (**R04 - R08**). A good method to assess the quality of CTI is to have some sort of a ranking or scoring system which indicates the threat potential or the trustworthiness of the stated vulnerability or malware (**R04**). Moreover, a platform that not only offers raw data, such as single indicators, but threat information enriched with more detailed information and descriptions, will be much more valuable to organizations that aim at integrating CTI into their security monitoring (**R05**). Finally, as already discussed in Chapter 2, the timeliness of the offered CTI is crucial. The relevance of certain CTI on, e.g., a malicious IP address can change within only a short period of time. Therefore, it is important that the platform that is to be integrated offers near to real-time CTI (**R06**).

There are many security monitoring solutions that are suitable for the integration of CTI within university networks. To derive further requirements, the following sections present a detailed description of the security monitoring measures that are in use at the LRZ to protect the MWN against misuse and criminal behavior.

3.2.1 Security Information & Event Management

Among major security monitoring solutions that became indispensable for organizations concerned with network security are SIEM systems. At the LRZ a SIEM solution has been in use for several years now. At the moment, this is the SIEM QRadar solution from IBM. SIEM solutions are easy manageable tools that allow real-time monitoring and analysis as well as reaction to security events. They aim at reducing false positives and provide notification mechanisms such as automated reporting in case of real threat. Furthermore, SIEMs easily integrate with other solutions. For example, they can be used for the collection of log data and correlation of security events which can be found through other security monitoring tools. To integrate CTI into a SIEM via a CTI platform it is important that certain requirements concerning the CTI collection (**R12** - **R18**) are fulfilled. First of all, it is very important that the platform has export functionalities which enable the user to download, or even better, to automatically export the data into their security monitoring devices (Requirements **R15** and **R16**). Moreover, depending on the deployed SIEM solution, the platform that is to be integrated should offer export functionalities for certain formats (**R15a** - **R15c**).

The SIEM solution used by the LRZ is the central point for the evaluation of security relevant events and the automated response to them. Security alerts detected by the SIEM are mostly reported by email notifications to the network and system administrator. Such automatic notification mechanisms have proven to be quite helpful for the further investigation of and reaction to threats. Automatic notification mechanisms are therefore a useful requirement for the integration of CTI platforms (**R25**) as well. In addition, the system also provides mechanisms for the automatic locking of systems in case of serious emergency [Lei16]. Besides the security notifications sent by the SIEM, the LRZ receives security alerts from the DFN-CERT and CERTBund that can be correlated to other security notifications and known network vulnerabilities. Often LRZ employees want to further research the information on malware and threats they received via these alerts. Therefore, it would be useful if the prospective CTI platform offers (automatic) search functions (**R12** and **R13**). The easiest way to integrate a CTI platform with a SIEM would be if the platform provider and the SIEM provider already work together as "integration partners" or if the platform in other ways offers direct integration options with the deployed SIEM (**R21**).

3.2.2 Intrusion Prevention System

As has been described in Chapter 3.1, the use of private IP addresses is a characteristic of university networks. The LRZ also assigns such private IP addresses. Since packages from the Internet cannot simply reach a private IP address, they are a good shield against hacker attacks. On the contrary, to make use of internet services and communicate outside the MWN, private IP addresses usually need to be translated into a public IP address. This is done by means of a so-called NAT gateway, which at the LRZ is known under the project name Secomat. However, the scope of the Secomat's function goes beyond the simple translation of private into public IP addresses. For example, it can limit the data transfer rate or identify attack scenarios, such as port scans, denial of service, or spam. Therefore, the Secomat observes the number of packets from and to certain destinations. Based on predefined rules in combination with a penalty point system, user activities get rated. Once a certain limit is exceeded within a specific timeframe, the system assumes that the respective

computer might have been compromised. In consequence, this computer gets automatically blocked until the number of penalty points falls below a certain limit again [Lei17].

3.2.3 Network Intrusion Detection System

A network intrusion detection system (NIDS) monitors network traffic for malicious activity or unauthorized behavior and alerts the network administrator or sends reports to the SIEM. Ideally, it is installed at strategic points within the network to scan inbound and outbound traffic. The NIDS operated by the LRZ, a signature-based intrusion detection system based on Suricata, is placed at the central X-WiN network transition. Mainly, it is used to detect misuse such as port scans, denial-of-service attacks or spamming. Those are reliable indicators to identify malware within the network, such as Trojans or ransomware. In the past, the system has proven to be useful to detect malware, such as XcodeGhost which infects apps on mobile devices with malware [Lei16]. Regarding the integration of CTI platforms, direct integration options with the deployed NIDS (**R22**) and a NIDS rules export functionality (**R15d**) would be useful.

3.2.4 Notifying and blocking mechanisms: NeSSI

The LRZ uses notification and blocking mechanisms in conjunction with their SIEM solution QRadar. Network administrators can use a web portal called NeSSI (Network Self Service Interface) to unblock IP addresses after a detected problem is solved. By providing various notification and unblocking functions to the network administrators within their managed address areas, the NeSSI portal helps to reduce the workload of the LRZ Abuse-Response-Team [Lei16].

3.2.5 Security and network management: Nyx

Nyx is a security and network management tool that allows its users to locate computers within the MWN. After entering a specific MAC or IP address, Nyx provides the switch, port and data socket where the computer is connected using this address. In addition, network administrators can retrieve Nyx-Data for their managed address areas via the NeSSI portal [HR16].

3.2.6 Summary: Integrated Security Incident Management

After detecting security relevant events via the above described monitoring mechanisms, incidents get forwarded to the centralized SIEM where they get analyzed and correlated to one another. Based on predefined policies, an automated response mechanism starts. First, the responsible administrator gets notified and the concerned machine's internet access gets suspended. Thereafter, the computer security incident response team (CSIRT) gets notified and a trouble ticket is generated. In addition to the detection of security incidents via the deployed security monitoring mechanisms, incidents can also be reported by university staff or security experts manually, e.g. by phone or email, as well as through third party services, such as the DFN-CERT alerting service. As can be seen in Figure 3.1, the LRZ tries to combine the different reporting channels and embed them in in an integrated security incident management process that allows quick reactions and responses [SM11].

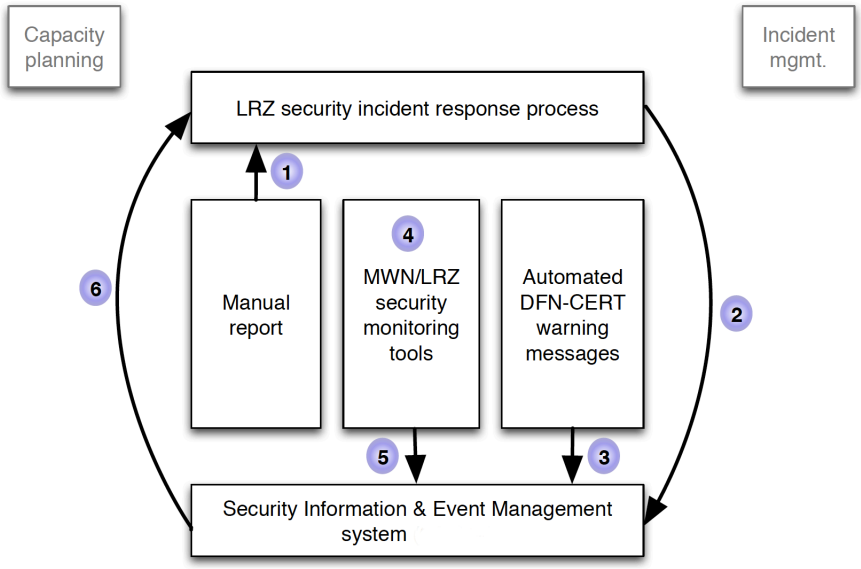

Figure 3.1: Integrated management of security incidents (revised figure) [SM11]

The LRZ uses a three-level approach: After detecting a conspicuity of a system, the respective administrator or user gets a first warning sent via email. Once detected again, the same person receives a second warning which is sent combined with a penalty of the suspension of the system. After the third conspicuity, the respective IP address or user account gets blocked. However, internal SSH attackers get blocked immediately. On basis of the NAT address, LRZ employees search the concerned private IP address and lock it with the Secomat. Since it is not always possible to locate an offending system only by its IP address, an automated query to Nyx is used to determine the switch port to which the system is connected [SM11].

3.3 Requirements for the use of CTI platforms in the context of security monitoring in university networks

This section summarizes and complements the above described requirements for the integration of CTI platforms into the security monitoring of higher education institutions. In order to define the requirements, the overview chapter as well as the previous sections on the characteristics of higher education networks and the application scenario of the MWN have been taken into account. Table 3.1 shows an overview of all the requirements together with a prioritization. To find the most suitable platform(s), a weighting with three levels was introduced:

high: If a requirement is essential for the successful integration of the platform into the security monitoring of higher education networks

medium: If a requirement is quite useful but not necessary for the successful integration of the platform into the security monitoring of higher education networks

low: If a requirement is just a nice add-on for the successful integration of the platform into the security monitoring of higher education networks

General requirements

R01: Open-source **low**

Open-source projects are developed by an open (free to join) community and free-to-use. The requirement is met, if the platform is released under an open-source license, if its development is community-driven and there are no costs for using the platform.

Priority: As it is not critical for the successful integration of CTI platforms, whether or not it is an open-source product, the priority of R01 is classified as low.

R02: Free-to-use **medium**

This requirement is met, if there are no costs for using the platform. It does not matter if the platform is developed and provided by an enterprise or an open-source community.

Priority: Since many organizations must make cost savings, it is always good to have a technology that is free-to-use. At the same time, however, it is also important (especially in the IT security sector) that the systems used offer a high quality and functionality, which is why the priority of R02 is classified as medium.

R03: Language of instruction **high**

German and English are the most commonly spoken languages at German Universities. Therefore, this requirement is met if the platform instructions are in either one of those languages.

Priority: To use the platform properly, instructions must be in English or German. Therefore, the priority of R03 is classified as high.

Quality of CTI

R04: Ranking **high**

University employees who are responsible for the IT security monitoring rely on correct and valuable data. Therefore, a ranking or scoring system would help to evaluate the usefulness of the provided data. The requirement is met if the platform uses a scoring or ranking system to evaluate the threat potential and/or reliability of the concerned vulnerability, malware, or indicator.

Priority: For the successful integration of CTI into the security monitoring of university networks, it is very important that the data has a high quality and that it meets the requirements described in Chapter 2. A ranking or scoring system can help to better assess the quality of the data. The priority of R04 is therefore classified as high.

R05: Further information **high**

When university employees, who are responsible for the IT security monitoring, receive information on potential threats, e.g. via DFN-CERT alerts, it would be helpful if the CTI platform offers further information and detailed descriptions on the latest threats and malware. Requirement R05 is met if each threat information entailed in

the platform contains a detailed description of the concerned indicators or the kind of malware it refers to.

Priority: For the successful integration of CTI into the security monitoring of university networks, it is very important that the data has a high quality and that it meets the requirements described in Chapter 2. Further information and detailed descriptions can help to make better use of the data. The priority of R05 is therefore classified as high.

R06: Real-time CTI **high**

Only real-time CTI that is constantly being updated can help to defeat cyber criminals. Therefore, it is important that the platform provider and the involved community members constantly update old CTI and post new threat information on a near to real-time basis.

Priority: For the successful integration of CTI into the security monitoring of university networks, it is very important that the data has a high quality and that it meets the requirements described in Chapter 2. One of those requirements, which influences the relevance of the respective data immensely, is timeliness. Therefore, the priority of R06 is classified as high.

R07: Various indicators **high**

As discussed in Chapter 2.1, a lot of different information count as CTI. Requirement R07 is met, if the platform offers threat data concerning at least 5 different kinds of indicators, such as IP addresses, hash values, domain names,...

Priority: Different types of data counts as CTI. For the successful integration of CTI into the security monitoring of university networks it can be useful to have as many different indicators as possible. Therefore, the priority of R07 is classified as high.

R08: TTPs **medium**

As likewise discussed in Chapter 2.1, there is a difference in the usefulness of CTI data. On top of the pyramid of pain, and therefore the most painful for the adversary, is the detection of their TTPs. This requirement is met, if the platform offers information on TTPs.

Priority: TTPs are very valuable data. However, they are very difficult to obtain and their processing is more complex than that of IOCs. TTPs are useful but not necessary for the successful integration of the platform into the security monitoring of higher education networks. Therefore, the priority of R08 is classified as medium.

Support of standards

R09: STIX **high**

As discussed in Chapter 2.3.1, STIX has become a de-facto standard for the structured description of CTI data within the CTI community. To be able to engage oneself within this community and to exchange CTI in a standardized format, it is important that the platform that is to be integrated supports the STIX format.

Priority: The support of widely accepted standards, is very important for the successful integration of CTI into the security monitoring of university networks. Therefore, the priority of R09 is classified as high.

R10: TAXII **high**

As discussed in Chapter 2.3.2, TAXII has become a de-facto standard for the mutual exchange and distribution of CTI data within the CTI community. To be able to engage oneself within this community and to exchange CTI in an easy and automated manner, it is important that the platform supports the TAXII standard.

Priority: The support of widely accepted standards, is very important for the successful integration of CTI into the security monitoring of university networks. Therefore, the priority of R10 is classified as high.

R11: Other standards **low**

Besides STIX and TAXII there are other standards that are useful for the structured description and distribution of CTI. Requirement R11 is fulfilled, if the platform also supports other standards than STIX and TAXII.

Priority: STIX and TAXII have become de-facto standards for the mutual exchange and distribution of CTI data. The support of other standards than STIX and TAXII would be a nice add-on but would not necessarily create added value for the successful integration of a CTI platform into the security monitoring of higher education networks. Therefore, the priority of R11 is classified as low.

CTI collection

R12: Search function **high**

When university employees who are responsible for the IT security monitoring want to further research into specific threats or malware, it is useful if the platform offers a good search functionality. Requirement R12 is fulfilled, if the platform offers such a search functionality to find specific CTI by using key words or by looking through predefined categories.

Priority: The more an employee of the IT security staff knows about potential threats the better. A search function can be very usefull to further research into known indicators, malware or incidents. Therefore, the priority of R12 is classified as high.

R13: Automatic search function **high**

Since resources in terms of time and personnel are limited at most universities, it would be a plus if the platform that is to be integrated would offer ways to use the search functionality automatically.

Priority: An automatic search function would allow further researching into known indicators, malware or incidents but would be less time consuming than a manual search. As is the case with R12, the priority of R13 as well is classified as high.

R14: Filter function **medium**

Due to the amount and diversity of the offered CTI, it would be useful if users of the platform can reduce the amount of displayed data by using a filter function.

Priority: A filter function to reduce the amount of data displayed to the user would be useful but not necessary for the successful integration of the CTI platform into the security monitoring of higher education networks. Therefore, R14 is classified as medium.

R15: Export function **high**

The main idea about integrating a CTI platform into the security monitoring of higher education networks is to gain further information on potential threats and therefore be better prepared to detect and react to threats. Therefore, it is inevitable that the platform offers an export functionality. Requirement R15 is met, if the platform offers manually export functions so the user can download CTI in various formats.

Priority: Since the export of CTI into security devices is one of the main ideas behind the integration of CTI platforms into the security monitoring of university networks, the priority of R15 is classified as high.

R15a: XML export **medium**

Depending on the deployed security monitoring solution, it is important that the platform supports the export of CTI in various formats. Since XML is a widely-used format, it would be good if the platform that is to be integrated offers an export function so the user can download CTI in the XML format.

Priority: XML is a widely-used format, but it is not essential in the field of CTI. An existing XML export would be useful but not necessary for the successful integration of CTI into the security monitoring. Therefore, the priority of R15a is classified as medium.

R15b: STIX export **high**

Since STIX is not only a widely-used format within the CTI community but also gains in popularity within the whole IT security community, it would be good if the platform that is to be integrated offers an export function so the user can download CTI in the STIX format.

Priority: STIX is a widely-used format, that becomes more and more essential in the field of CTI. An existing STIX export would be important for the successful integration of CTI into the security monitoring. Therefore, the priority of R15b is classified as high.

R15c: CSV export **high**

CSV as well is a widely-used format within IT security solutions. Therefore, it would be good if the platform that is to be integrated offers an export function so the user can download CTI in the CSV format.

Priority: CSV is a widely-used format, especially when it comes to IT security devices. An existing CSV export would be important for the successful integration of CTI into the security monitoring. Therefore, the priority of R15c is classified as high.

R15d: NIDS rules export **medium**

Depending on the deployed NIDS solution, it would be helpful if the platform that is to be integrated offers an export function so the user can download NIDS rules.

Priority: A NIDS rules export would be quite useful for the integration of CTI into a NIDS but not necessary for the general integration of a CTI platform into the security monitoring of university networks. Therefore, the priority of R15d is classified as medium.

R16: Automatic export function **high**

Since resources in terms of time and personnel are limited at most universities, it would be important that platform offers automatic export functions so the user can download CTI automatically in various formats.

Priority: An automatic export function would allow the integration of CTI into security devices but would be less time consuming than a manual export. As is the case with R15, the priority of R16 as well is classified as high.

R17: Flexible subscribe/follow functionalities for feeds/sources **high**

As discussed in Chapter 2.2, CTI is submitted via so-called feeds or similar. With regard to integrating CTI platform(s) into the security monitoring of higher education networks, it would be useful if the user can flexibly subscribe/follow and unsubscribe/unfollow different feeds and sources.

Priority: Not all CTI is equally important to an organization. To reduce the amount of data and avoid system overload, R17 is essential for the successful integration of CTI into the security monitoring of university networks. Therefore, the priority of R17 is classified as high.

R18: Flexible integration of external CTI feeds/sources **medium**

The selection of CTI sources is immense. Therefore, it would be a plus if the prospective platform offers ways to flexibly integrate external CTI feeds or sources.

Priority: The range of CTI offerings by various vendors is immense. If the user of a CTI platform could flexibly integrate external CTI feeds or sources it would be useful but not necessary for the successful integration of the platform into the security monitoring of higher education networks. Therefore, the priority of R18 is classified as medium.

CTI correlation and analyzing functionalities

R19: Links to related data **high**

To get the most out of CTI, users must further analyze and correlate the queried data. Therefore, it would be useful, if the platform, that is to be integrated, directly shows links between the viewed and related threat data.

Priority: The more context around the CTI data is available, the better users can integrate the data into their security monitoring. Since links to related data can help to get a better picture over the overall security situation, the priority of R19 is classified as high.

R20: Correlation and analyzing functionalities **high**

In addition to showing links between the viewed and related data, it would be a plus if the platform offers correlation and analyzing functionalities, such as investigation links or a functionality that evaluates uploaded files for indicators that are already known to the community.

Priority: The more context around the CTI data is available, the better users can integrate the data into their security monitoring. Since correlation and analyzing functionalities allow further investigation and therefore support a better understanding of threats, the priority of R19 is classified as high.

CTI integration and distribution functionalities

R21: Direct integration with SIEM **medium**

The easiest way to integrate the prospective platform with a SIEM solution would be if the platform offers direct integration ways with the respective SIEM vendor. Requirement R21 is met, if the platform explicitly advertises its direct integration functionality with specific SIEM vendors.

Priority: The provider-based support for the integration of a CTI platform into different security devices, such as SIEM or NIDS, is essential for the success of this integration. Direct integration options are useful (since it is the easiest way of integration) but it is not necessary for a successful integration. Therefore, the priority of R21 is classified as medium.

R22: Direct integration with NIDS **medium**

Liekwise, the easiest way to integrate the prospective platform with a NIDS solution would be if the platform offers direct integration with the respective NIDS. Requirement R22 is met, if the platform explicitly advertises its direct integration functionality with specific NIDS solutions.

Priority: The provider-based support for the Integration of a CTI platform into different security devices, such as SIEM or NIDS, is essential for the success of this integration. Direct integration options are useful (since it is the easiest way of integration) but it is not necessary for a successful integration. As is the case with R21, the priority of R22 as well is classified as medium.

R23: Integration via API **high**

Another way to integrate the prospective platform would be via an API. Requirement R23 is fulfilled, if the platform offers an API so the provided CTI can be integrated into existing security solutions automatically via the API.

Priority:An API offers great functionalities that are essential for automatic integration of CTI platforms into the security monitoring of higher education networks. Therefore, the priority of R23 is classified as high.

Group and sharing functionalities

R24: Sharing of own CTI **medium**

One characteristic of CTI platforms is that they depend on a vivid community that mutually exchanges CTI. If an organization experiences a cyber attack, it can share related indicators or information on the type and specifics of the respective incident. To profit from the experiences that other organizations have had with cyber crime, the platform should offer ways to share one's own CTI with other users of the platform.

Priority: Since organizations experience similar threats, the sharing of CTI can be very useful in the fight against cyber-attacks. However, since the sharing of CTI is not a necessary condition for the successful integration of a CTI platform into the security monitoring of university networks, the priority of R24 is classified as medium.

R25: Possibility to form and join groups **medium**

In a large community some threat information will be more relevant to an organization

than other. Therefore, the platform should offer ways to share one's own CTI with members that have shared interests. The best way to do so, would be if users can form and join groups that share common interests. For example, universities that use the same platform could form groups where they exchange their own CTI specifically concerning higher education networks.

Priority: The forming of groups for the mutual exchange of CTI between groups that have the same interests can be useful but is not necessary for the successful integration of a CTI platform into the security monitoring of university networks. As is the case with R24, the priority of R25 as well is classified as medium.

Notification mechanisms

R26: Information/alerting via email **medium**

As discussed in the chapter on the application scenario, LRZ employees receive security alerts from the DFN-CERT and CERTBund. Likewise, it would be good if the prospective platform offers notification mechanisms for CTI that the user is interested in (e.g. feeds that the user has subscribed to). Requirement R25 is met, if the platform offers a notification service so that users can receive these information via email.

Priority: Receiving alerts and notifications on security alerts is quite useful for employees of the IT security staff. It would be useful but not necessary that the CTI platform that is to be integrated offers such a functionality. Therefore, the priority of R26 is classified as medium.

The following table lists the above described requirements together with their weighting and ID in a weighted requirements catalogue.

Requirements ID	Name of requirement	Priority
General requirements		
R01	Open source	low
R02	Free-to-use	medium
R03	Language of instruction	high
Quality of CTI		
R04	Ranking	high
R05	Further information	high
R06	Real-time CTI	high
R07	Various indicators	high
R08	TTPs	medium
Support of standards		
R09	STIX	high
R10	TAXII	high
R11	Other standards	low
CTI collection		
R12	Search function	high
R13	Automatic search function	high
R14	Filter function	medium
R15	Export function	high
R15a	XML export	medium
R15b	STIX export	high
R15c	CSV export	high
R15d	NIDS rules export	medium
R16	Automatic export function	high
R17	Flexible subscribe/follow functionalities for feeds/sources	high
R18	Flexible integration of external CTI feeds/sources	medium
CTI correlation and analyzing functionalities		
R19	Links to related data	high
R20	Correlation and analyzing functionalities	high
CTI integration and distribution functionalities		
R21	Direct integration with SIEM	medium
R22	Direct integration with NIDS	medium
R23	Integration via API	high
Group and sharing functionalities		
R24	Sharing of own CTI	medium
R25	Possibility to form and join groups	medium
Notification mechanisms		
R26	Information/alerting via email	medium

Table 3.1: Weighted catalogue of requirements for the integration of CTI platforms

3.4 Analysis of CTI platforms

In *Threat Intelligence Sharing Platforms: An Exploratory Study of Software Vendors and Research Perspectives* Sauerwein et al. have identified 22 CTI platforms: Accenture Cyber Intelligence Platform, Anomali ThreatStream, Anubis Networks Cyberfeed, BrightPoint Security Sentinel, Collaborative Research into Threats (CRITs), Comilion, Facebook Threat Exchange, Falcon Intelligence Crowdstrike, MANTIS Cyber Threat Intelligence Management Framework, Malware Information Sharing Platform (MISP), McAfee Threat Intelligence Exchange, Microsoft Interflow, Open Threat Exchange (OTX), Soltra Edge, HP ThreatCentral, ThreatCloud IntelliStore, ThreatConnect, ThreatQ, ThreatTrack ThreatIQ, Eclectic IQ, IBM X-Force Exchange, Collective Intelligence Framework (CIF) [CS17].

They are all different regarding the type of data they provide, the standards they support or their licensing models. To not go beyond the scope of this thesis, the author reduced the number to 4 platforms (IBM X-Force Exchange [IBM17], Malware Information Sharing Platform (MISP) [MIS7b], Open Threat Exchange (OTX) [Ali7b], and ThreatConnect [Thr17]) that are being examined in detail in the following sections. The selection was made after researching all 22 platforms via the Internet. The main reasons that influenced this selection were a) missing technical requirements, b) limited functionalities and c) difficulties to get a test version of the other platforms.

3.4.1 IBM X-Force Exchange

IBM X-Force Exchange is a cloud-based CTI repository that allows its users to research, aggregate and share human- as well as machine-generated CTI. The only prerequisites are an IBM ID, which can be obtained freely via the IBM website, and a browser to login to the platform. IBM declares that the the X-Force Exchange platform is dynamically updated every minute and that they create up to 2.5 quintillion bytes of data every day [Ste17]. Therefore, the IBM X-Force Exchange platform gets the full score for **R06**. According to the company's own information, the CTI available in IBM X-Force Exchange includes [IBM6a]:

- various IOCs, such as IPv4, IPv6, URL, FileHash-MD5,...
- reputation scores for over 860k IP addresses
- URL reputation for over 32 billion web pages and images
- intelligence on over 100,000 vulnerabilities
- malware indicators and associated hash values
- intelligence, e.g. on TTPs, generated by security experts (such as IBM X-Force researchers and IBM Security professionals) by adding context to machine-generated data

Due to the broad variety of different indicators and the provision of some TTPs, the IBM X-Force Exchange platform gets the full score for **R07** and one point for **R08**. In addition, IBM X-Force Exchange users can create Collections (for an example see Figure 3.3) which

consist of both unstructured and structured content. Besides a description on the Collection's content, the Collection is characterized by a type (e.g. malware or vulnerability), as well as its associated threat indicators that are relevant to that particular Collection[IBM6a]. For the detailed description of the Collection and its indicators, the IBM X-Force Exchange platform reaches the full score for **R05**. Moreover, users can supplement the CTI available in the X-Force Exchange platform by integrating open sources, including e.g. PhishTank, RiskIQ, or VirusTotal (see Figure 3.5). Therefore, the IBM X-Force Exchange platform as well reaches the full score for **R18**. IBM X-Force Exchange offers various correlation and analyzing functionalities. For example, users of IBM X-Force Exchange can upload files to automatically analyze whether the contained indicators are as well entailed in the IBM X-Force Exchange platform. In addition, IBM X-Force Exchange displays links between related collections. Users that wish to have more analysis functionalities can buy the IBM X-Force Exchange Malware Analysis as a Cloud service. However, for the already entailed correlation and analyzing functionalities the IBM X-Force Exchange platform gets the full score for **R19** and **R20**.

The Dashboard (see Figure 3.2) shows, among other things, recommendations for the newest Collections that have been composed by the IBM X-Force researchers, groups that a user has created or joined (full score for **R25**), shared research results of the community or the latest global security risks. Moreover, there is a keyword search where users can enter a text string and search for indicators, vulnerabilities or incidents which is why the platform reaches the full score for **R12**. In addition, there is an extensive filter function where search results can be grouped or filtered by type, risk score, or up-to-dateness. Therefore, the platform as well reaches the full score for **R14**.

Users can programmatically access the provided CTI through a RESTful API (full score for **R23**) which supports STIX and TAXII standards and returns information in JSON. This allows users to easily integrate the provided intelligence into their own security solutions, including software and hardware. Moreover, users can directly poll Collection data via STIX/TAXII by simply using the Collection URL. Due to its wide support of the STIX and TAXII standards, the IBM X-Force Exchange platform reaches the full score for **R09**, **R10** and **R15b**. However, since no other standards are supported besides STIX and TAXII, the platform gets no points for **R11**. The costs for using the API depends on the amount of data consumed. There are no costs for an API access up to 5,000 records per month. Users that exceed the monthly limit can buy a commercial API subscription. Therefore, **R02** only counts as partially fulfilled. The IBM X-Force Exchange platform can be directly integrated with the IBM SIEM solution QRadar. Therefore, **R21** counts as well as partially fulfilled. Since there are no direct integration options with NIDS, the platform gets no points for **R22**.

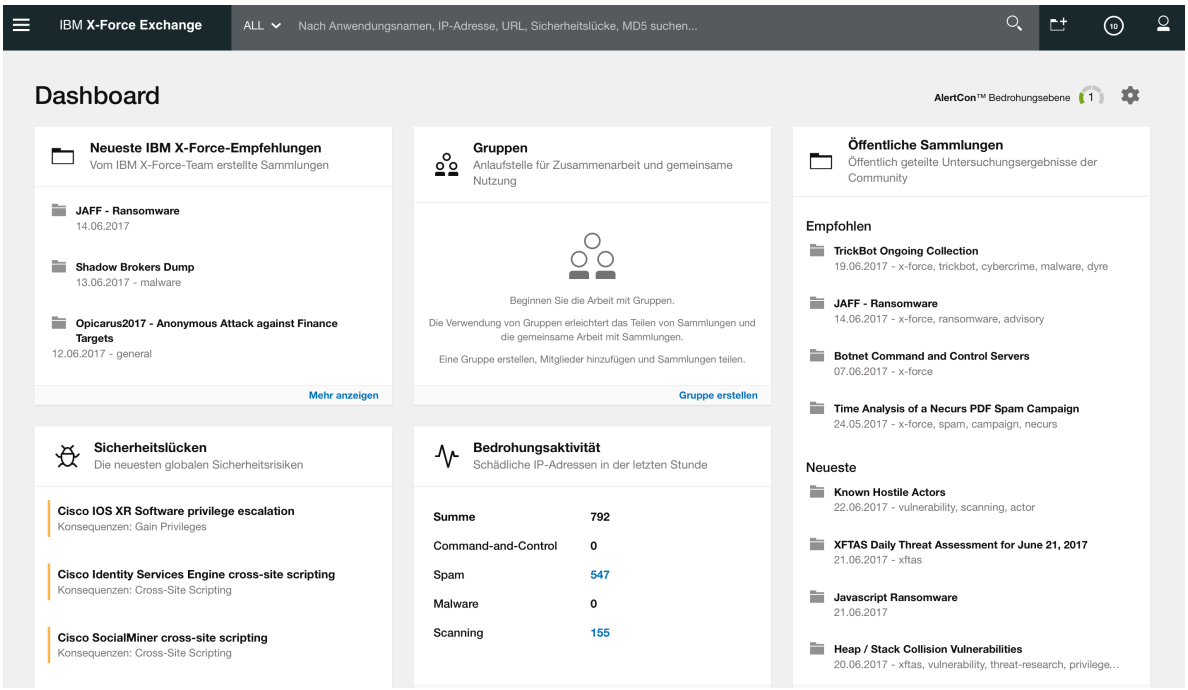

Figure 3.2: IBM X-Force Exchange Dashboard

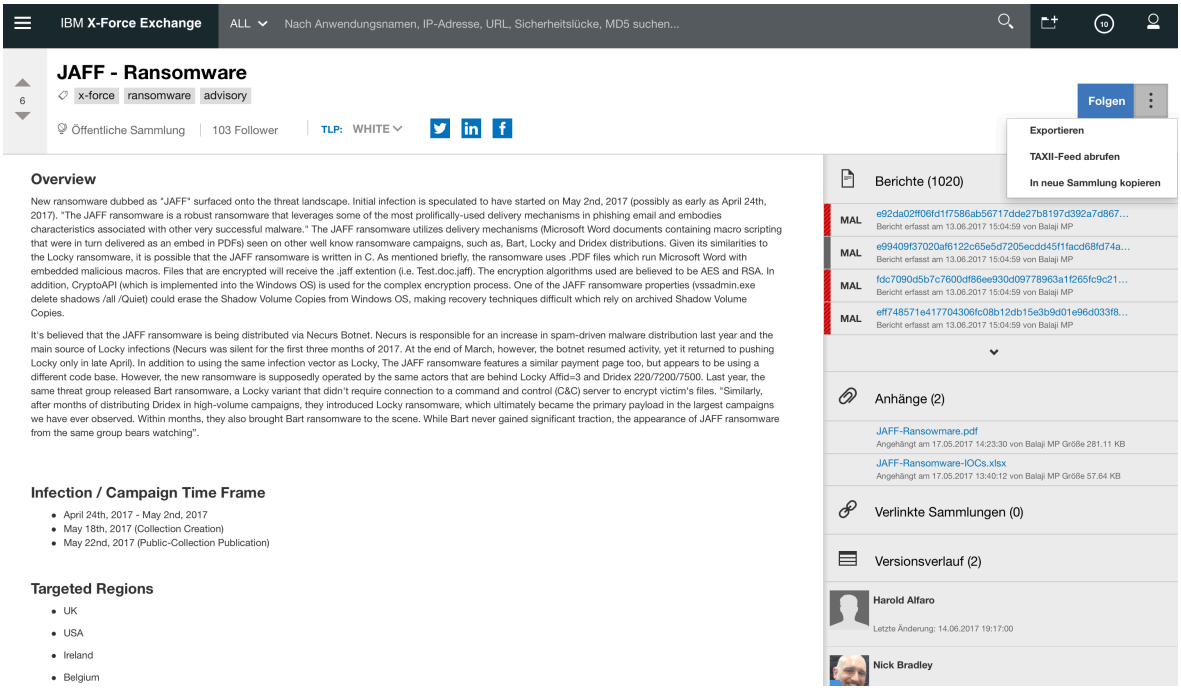

Figure 3.3: IBM X-Force Exchange Collections view: JAFF - Ransomware

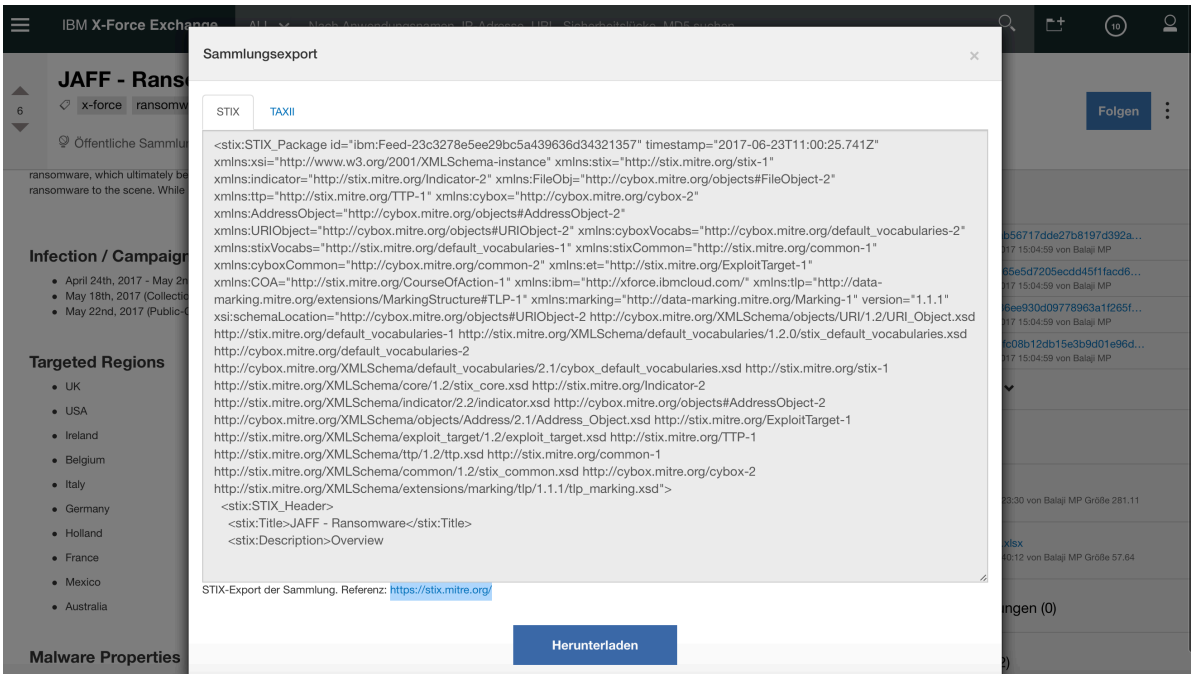

Figure 3.4: IBM X-Force Exchange Export view

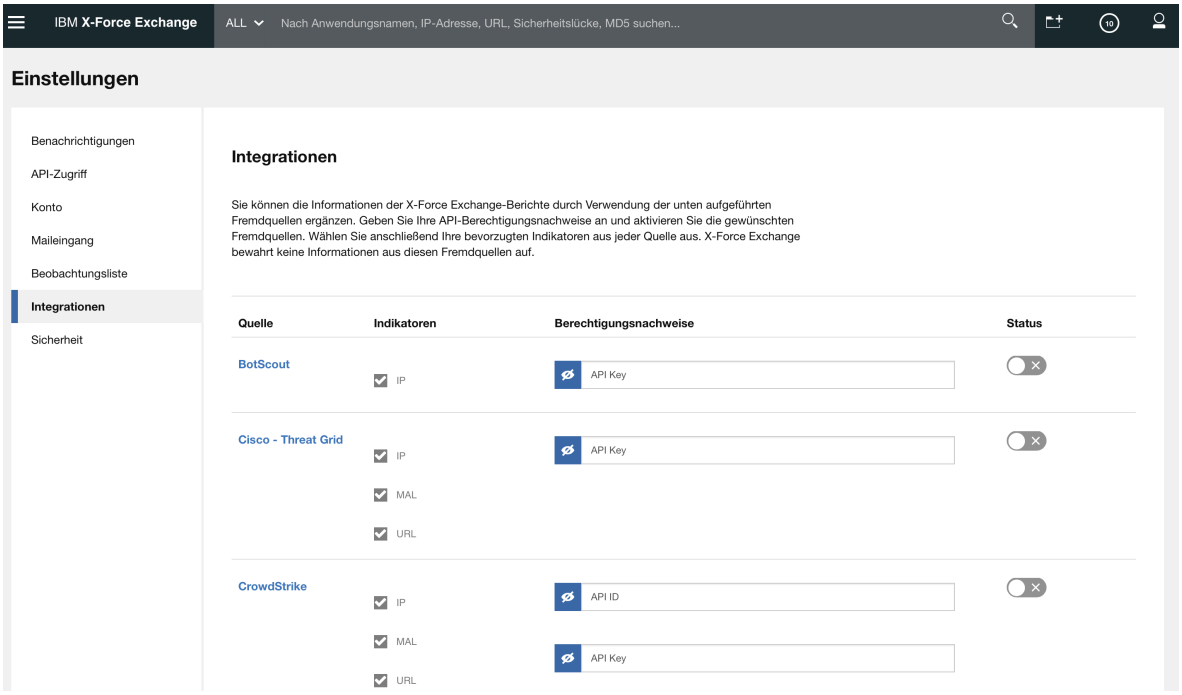

Figure 3.5: IBM X-Force Exchange Integrations view for open source integration

3.4.2 MISP - Malware Information Sharing Platform

The Malware Information Sharing Platform (MISP) is an open source software solution that can be installed on any standard GNU/Linux distribution. The platform serves as a distributed IOC database and allows to share, store, and correlate this data. CTI can be stored locally in the database and, if required, can be shared by synchronization with other instances. According to the MISP documentation [MIS7a], exchanging IOCs within the community leads to faster detection of targeted attacks, improved detection ratios, and reduced numbers of false positives.

MISP is developed as a free software by the Computer Incident Response Center Luxembourg (CIRCL) together with developers from the Belgian Defense and NATO. Although MISP is free-to-use, organizations need to request access from CIRCL in order to join the community. Therefore, **R02** only counts as partially fulfilled. Once organizations have access to the CIRCL MISP instance, they can use the application via a web interface and do not necessarily need to install their own MISP instance. According to CIRCL, MISP is designed to [Com16]:

- facilitate the storage of technical and non-technical CTI in a structured format
- create automatically relations between malware and their attributes
- automatically feed detection systems, e.g. generate rules for NIDS
- share malware and threat attributes with trusted partners within the community

Multiple export functionalities allow the integration with other security tools, such as NIDS or SIEMs (see Figure 3.8). Using the export functionality, one can automatically generate signatures for NIDS, such as Snort or Suricata. Therefore, MISP reaches the full score for **R15d**. On the other hand, various functionalities allow to easily import data to the repository, which is why MISP as well reaches the full score for **R24**. MISP offers IOCs containing IP, emails, hashes, URLs, and many more technical and non-technical information which is why the platform reaches the full score for **R07**. However, MISP does not offer further information or any detailed description of the indicators. Therefore, the platform gets no points for **R05**. No information was found on the up-to-dateness of the data or whether the data can count as real-time CTI, which is why the MISP does not get any points for **R06** either. Although MISP offers a filter function, the amount of data cannot easily be reduced and as clearly arranged as within the other platforms. Therefore, **R14** only counts as partially fulfilled.

Unlike the other platforms, the CTI within MISP is not displayed in the form of feeds or similar. Hence, users cannot subscribe to sources which they are specifically interested in or get email notifications on specific CTI (only gerneral email alerts are possible). Therefore, MISP gets no points for **R17** and only one point for **R26**. The core sharing format of MISP is JSON, however, it also supports other formats. Regarding the CTI standards, MISP supports STIX (full score for **R09**) but does not offer TAXII support (no points for **R10**). Instead MISP supports OpenIOC which is why the platform reaches one point for **R11**. MISP displays links to related events of the viewed event (full score for **R19**) but does only offer limited analyzing functionalities (correlation graph) which is why **R20** only

counts as partially fulfilled. To automatically update one's security devices, MISP provides a free-to-use RESTful API as well as a Python library (called PyMISP) to access the API. For its provision of an API access, MISP reaches the full score for **R23**. Users can share their own CTI as well as create and manage sharing groups for the distribution of events (full score for **R24** and **R25**).

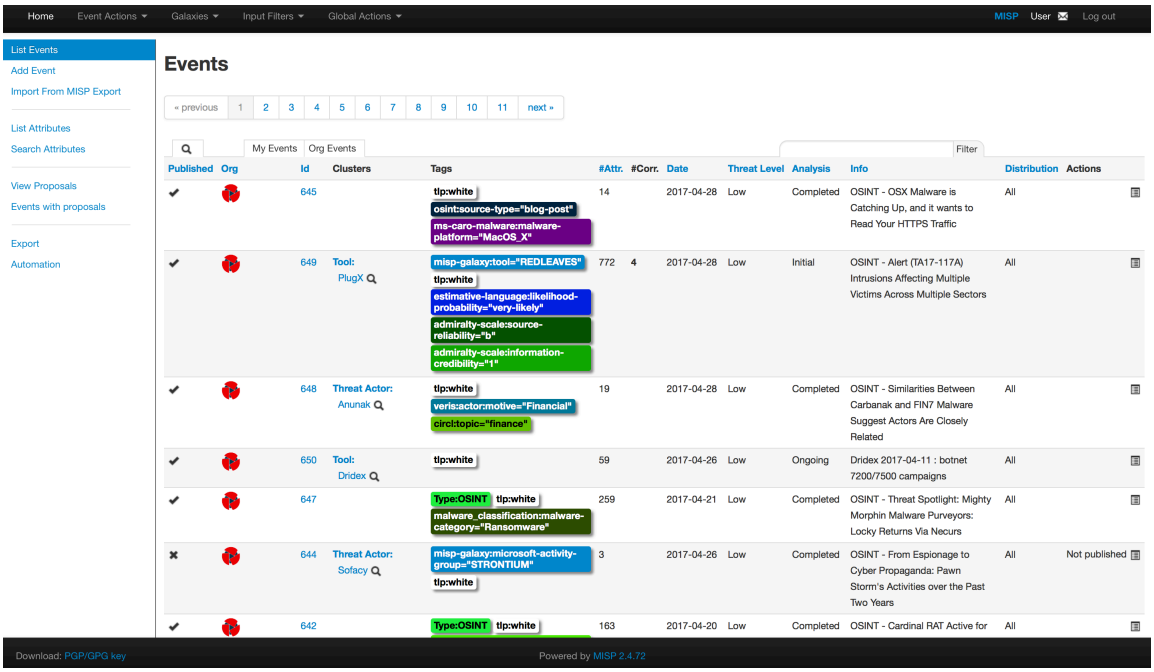

Figure 3.6: MISP Dashboard

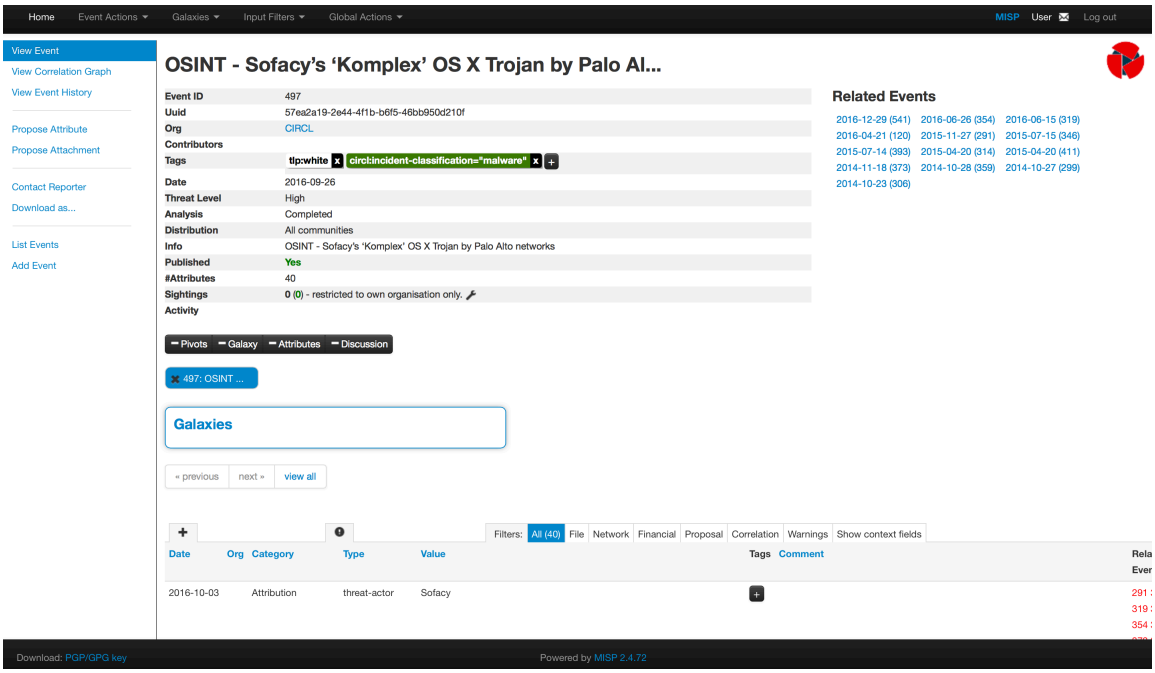

Figure 3.7: MISP Events view

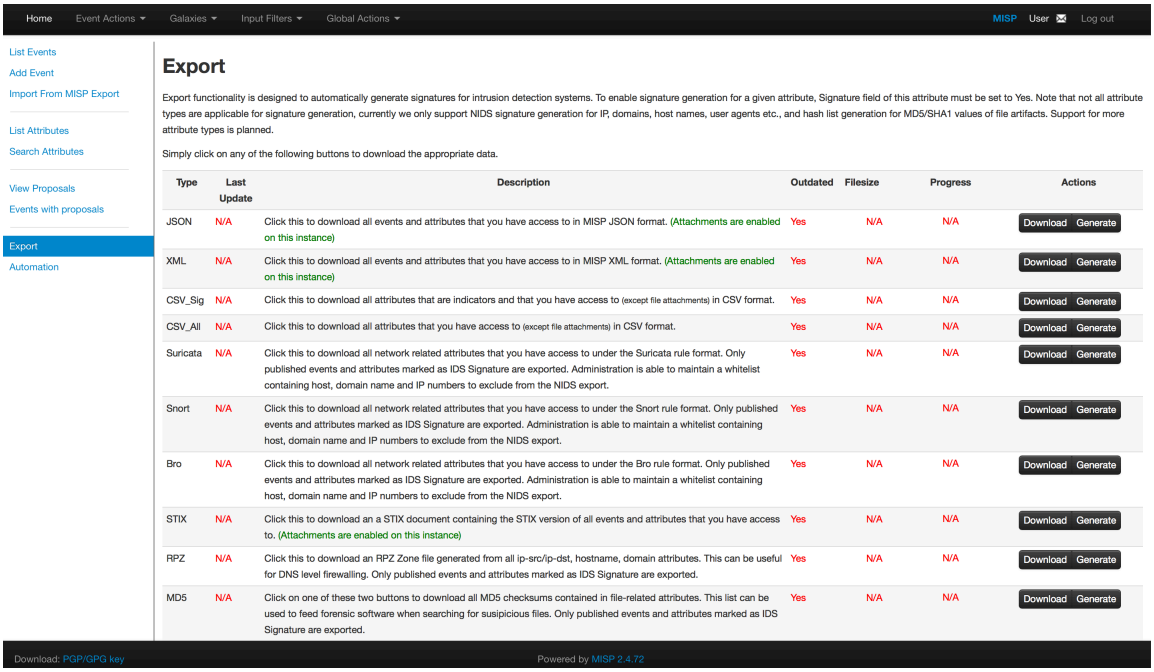

Figure 3.8: MISP Export view

3.4.3 AlienVault - Open Threat Exchange

Open Threat Exchange (OTX) is a cloud-based CTI platform, created and developed by AlienVault, but free-to-use (full score for **R02**). AlienVault promotes a global community of more than 50,000 participants in 140 countries, who contribute millions of threat indicators daily [Ali7a]. According to the platform providers, OTX offers real-time CTI which is why the platform gets the full score for **R06**. CTI is shared in the form of *pulses* (see Figure 3.10) which contain several IOCs with threat potential. Users can search for pulses on the main dashboard (see Figure 3.9) by entering a text string into the search field. Therefore, the platform get the full score for **R12**. Moreover, the dashboard shows some statistics, groups the user has joined and a list of pulses which can be sorted, e.g. according to the number of subscribers or their topicality.

Among the multiple IOC types associated with OTX pulses are:

- IPv4
- IPv6
- domain
- hostname
- email
- URL
- FileHash-MD5, -SHA1, -SHA256
- ...

Due to the broad variety of different indicators, the OTX platform gets the full score for **R07**. Unlike other platforms, OTX offers its users to look for industry-specific pulses. However, as can be seen in Figure 3.11, the number of pulses related to the education sector is not very high. The platform shows links to related pulses and therefore gets the full score for **R19**. **R20** counts as partially fulfilled since the platform sometimes refers to external sources for further information but does not offer any further anaylzing functionalities.

By subscribing to a feed, OTX members can receive these pulses through the OTX Activity feed or receive the information via email. Therefore, the OTX platform reaches the full score for **R26**. The CTI from OTX can be leveraged within AlienVault but also within third-party security monitoring systems. Integration can be easily achieved by using the AlienVault OTX DirectConnect API and DirectConnect SDK (full score for **R23**). Moreover, AlienVault currently provides direct connect agents for Bro-IDS, TAXII, and Suricata. Due to the well support for the integration of the platform with some NIDS, the OTX platform reaches the full score for **R22**. Since OTX serves as a TAXII server, the platform as well reaches the full score for **R10**. Besides STIX and TAXII, OTX supports OpenIOC which is why the platform reaches one point for **R11**. **R21** counts as partially fulfilled, since OTX can only be directly integrated with AlienVault's own SIEM solution OSSIM.

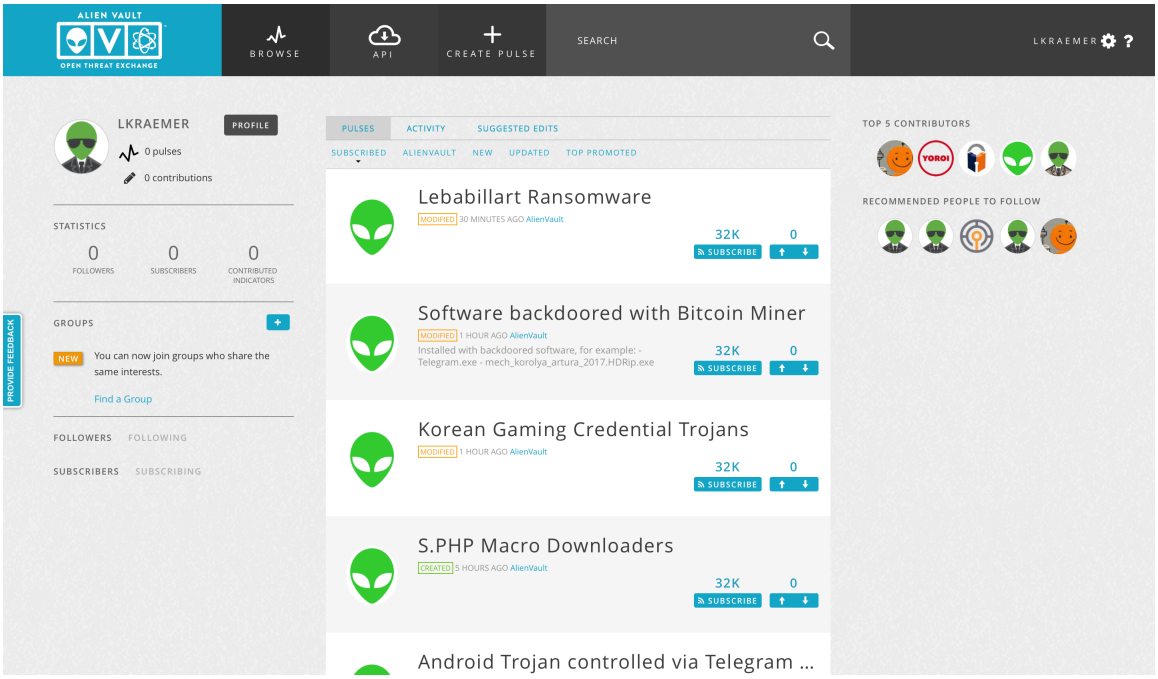

Figure 3.9: OTX Dashboard

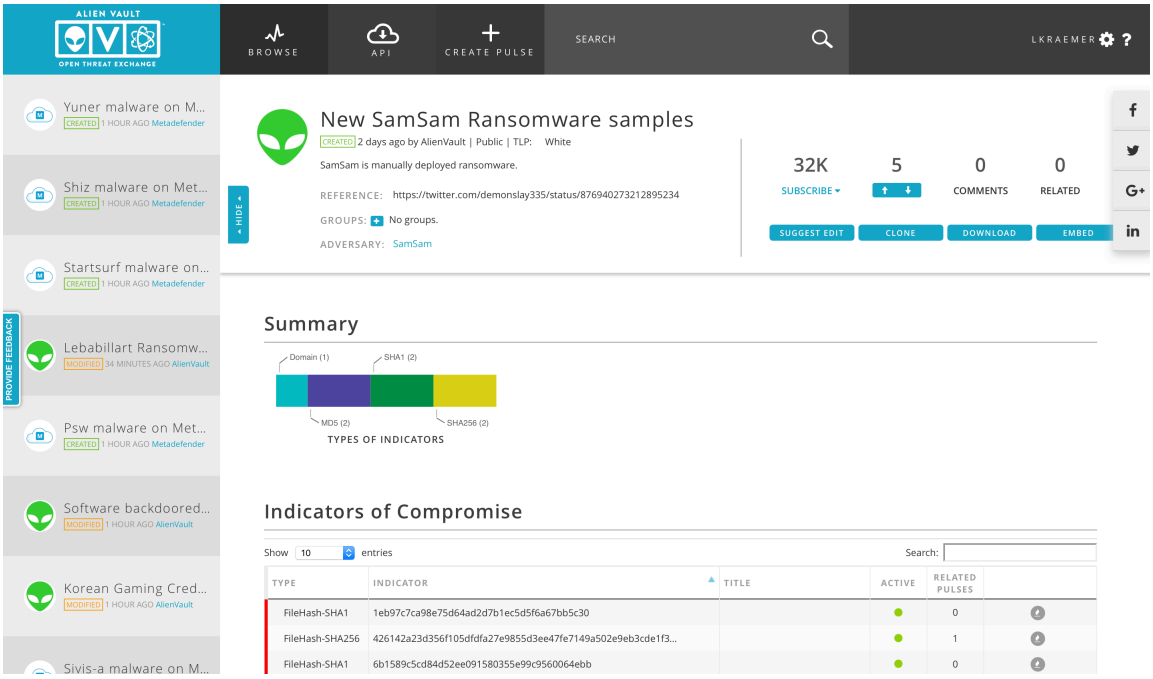

Figure 3.10: OTX Pulse view

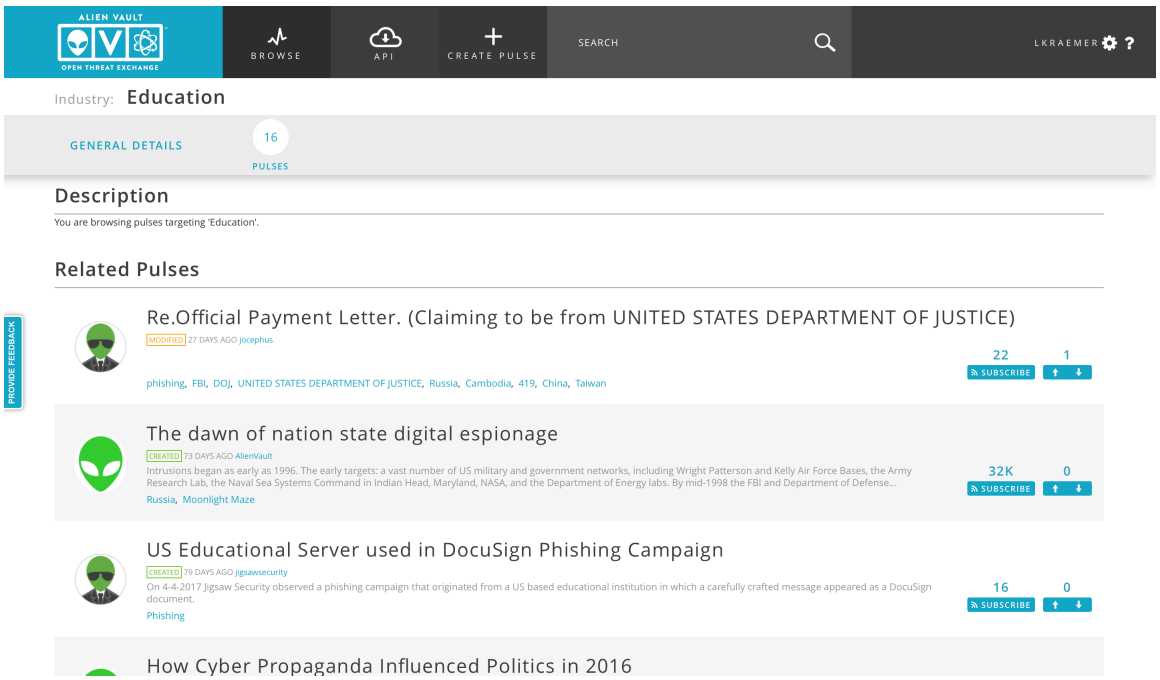

Figure 3.11: OTX Pulse view showing pulses in relation to the education sector

3.4.4 ThreatConnect - TC Open™

TC Open™ is a free version of four different commercial CTI platform solutions offered by ThreatConntect (TC). Since TC Open only offers limited functionalities compared to the commercial versions, **R02** only counts as partially fulfilled. Users of TC Open can see and share open source CTI within the TC community. The available CTI in TC Open includes many different indicators, such as email, file names, host names, URLs, ASN, CIDR, and much more. Therefore, the TC Open platform reaches the full score for **R07**. The free version is a cloud-based platform and includes:

- 1 user license
- Access to 100+ open source intelligence feeds
- Access to threat, incident, and adversary data
- Ability to collaborate or consume indicators, incidents, and threats
- Validation with peers in the TC Community

The dashboard (see Figure 3.12 and Figure 3.13) displays the recent history of searched and viewed CTI, some statistics, comments by the TC community and some links that when clicked direct the user to information pages on the external sources for the CTI included in TC Open. The Analyze menu allows users to upload and automatically parse files of various formats that contain indicators. The Playbooks menu contains templates that allow users to automate cyber defense tasks, such as data enrichment, malware analysis, and blocking

actions, via a drag-and-drop interface. However, this functionality is only available in the TC Manage and TC Complete product versions.

The Browse menu (see Figure 3.14) can be used to search for specific CTI via filters, e.g. indicators, threat rating, or tags, as well as by entering a text string (full score for **R12**). Clicking on a source (see Figure 3.15), users can get a lot of details on the type, threat rating, and associated indicators which can also be exported into the CSV file format with a limit of 5,000 indicators at a given time. NIDS rules can also be searched and downloaded. However, since the availability of NIDS rules is limited, the platform gets one point for **R15d**. The Create and Import menu allow users to share their own CTI with the TC Community (full score for **R24**). TC works together with many integration partners, such as IBM, Cisco, or FireEye, however, the joint solutions are only available for paying users. Therefore, **R21** only count as partially fulfilled. Users of TC Open can subscribe to sources which they are specifically interested in or get email notifications on these selected CTI feed. Therefore, TC Open gets the full score both for **R17** and for **R26**. The premium products offer support for STIX, TAXII, and other standards. However, these standards are not supported within the free account. Therefore **R09**, **R10** and **R11** only count as partially fulfilled. The same applies for automatically exporting data to security tools, such as SIEM, Firewalls, IPS/IDS, etc. This cannot be done with the free accounts which is why **R16** and **R23** as well only count as partially fulfilled. The possibilty to form and join groups is another premium feature. Therefore, TC Open only gets one point for **R25**.

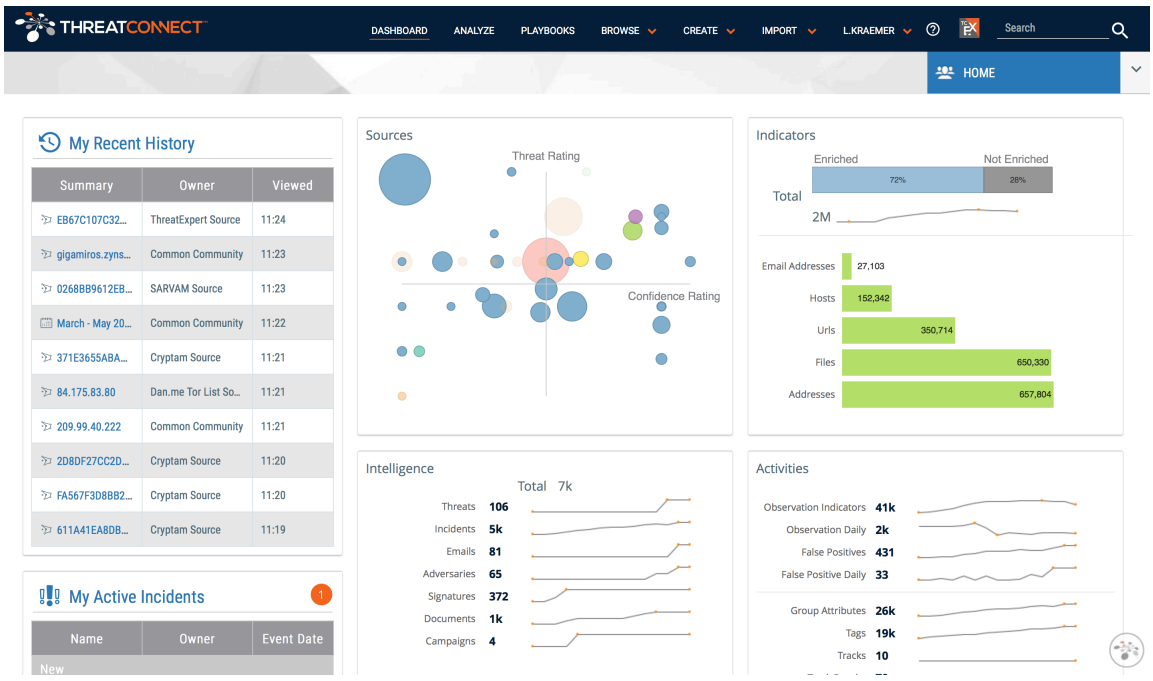

Figure 3.12: TC Open Dashboard 1/2

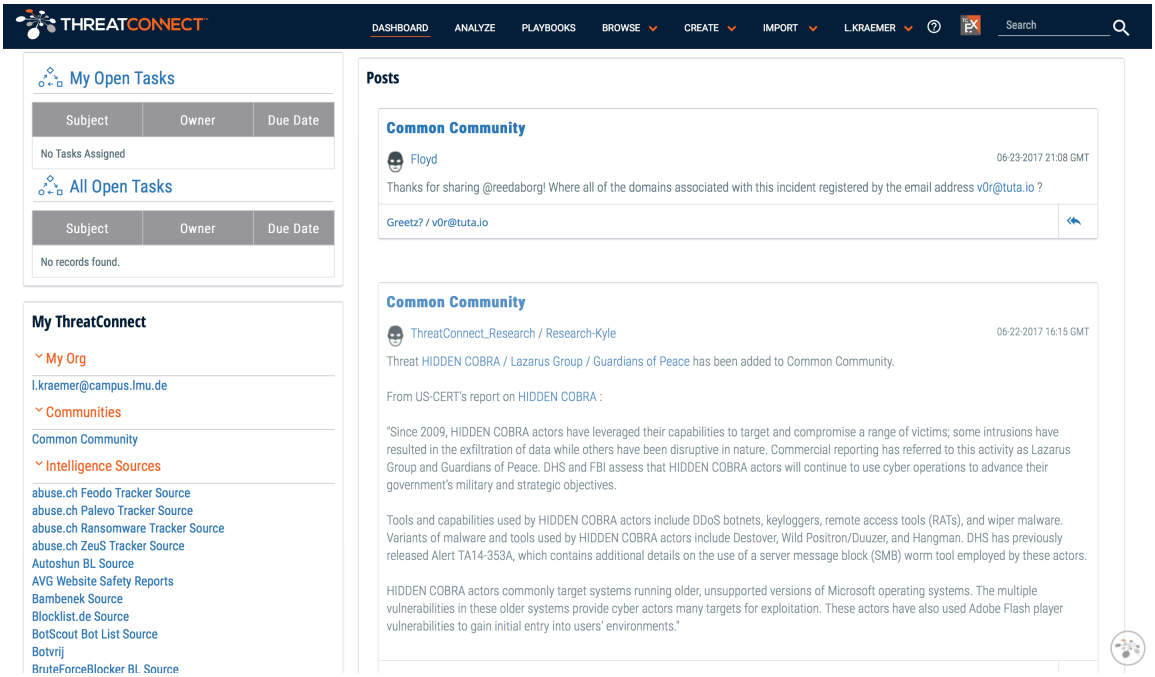

Figure 3.13: TC Open Dashboard 2/2

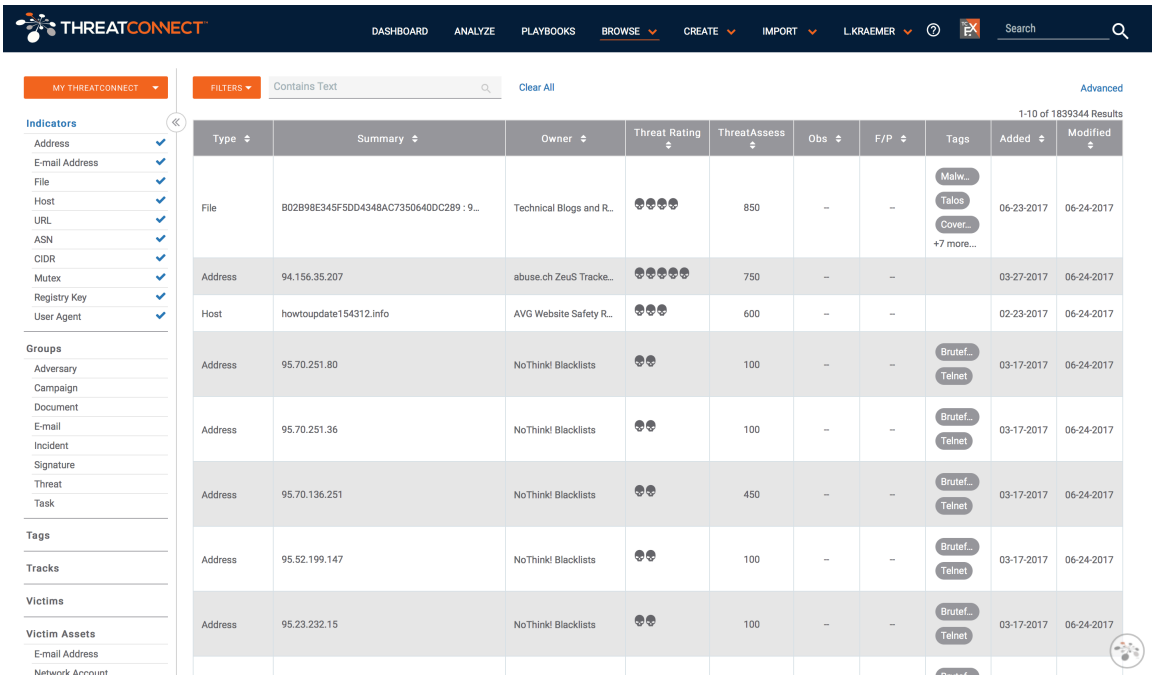

Figure 3.14: TC Open Browse view

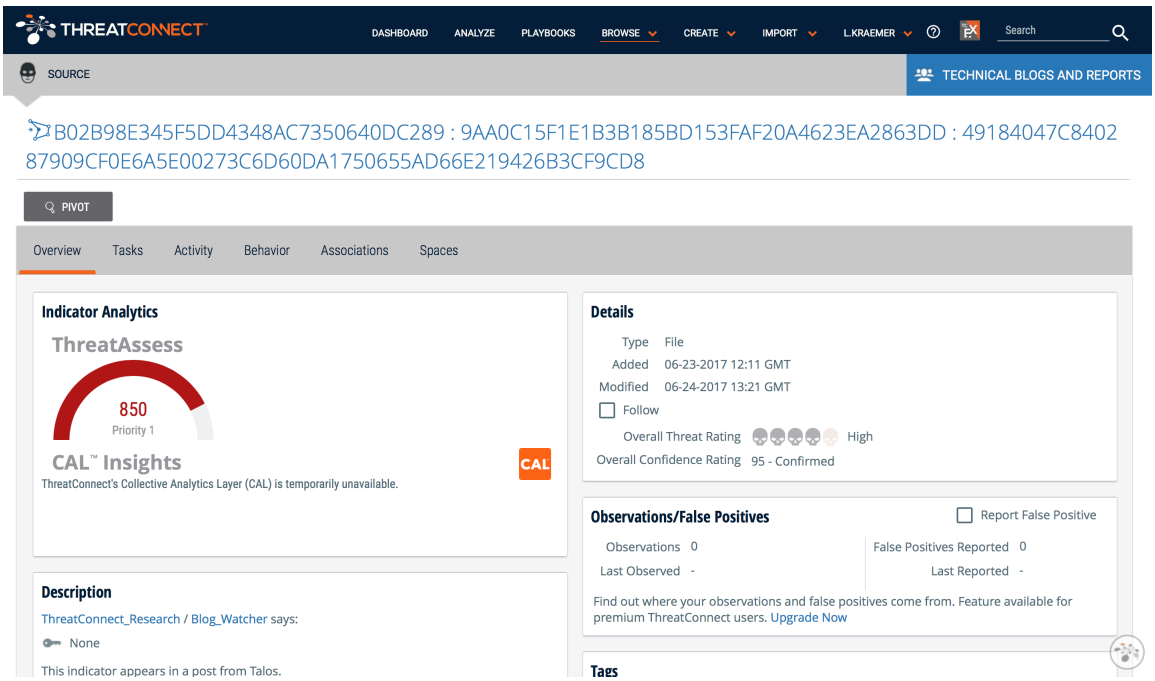

Figure 3.15: TC Open detailed Sources view

3.5 Results of the requirements analysis

Table 3.4 shows the results of the requirements analysis in detail. For each requirement, it was evaluated whether the assessed platform completely fulfills, partially fulfills or not fulfills the respective requirement. For illustration purposes the following characters were used:

++ If a requirement was completely fulfilled

\+ If a requirement was partially fulfilled

\- If a requirement was not fulfilled

To assess the platforms according to their fulfillment of the requirements, the following point-based evaluation scheme was used:

Priority / Fulfillment	low	medium	high
-	0	0	0
+	1	2	3
++	2	4	6

Table 3.2: Evaluation scheme

Table 3.3 shows for each requirements category how many points the respective platform has scored as well as the maximum number of points that were reachable. In total, it was possible to reach 146 points. The IBM X-Force Exchange platform received the most points (116 points), closely followed by the OTX platform (114 points). The TC Open platform is on the third rank (101 points) while MISP occupies the last place in the ranking (91 points).

Requirements Category	IBM X-Force Exchange	MISP	Open Threat Exchange	TC Open	Max. points
General requirements	8	10	10	8	**12**
Quality of CTI	26	9	18	24	**28**
Support of standards	12	7	13	7	**10**
CTI collection	38	40	40	35	**58**
CTI correlation and analyzing functionalities	12	9	9	12	**12**
CTI integration and distribution functionalities	8	6	12	5	**14**
Group and sharing functionalities	8	8	8	6	**8**
Notification mechanisms	4	2	4	4	**4**
Total	**116**	**91**	**114**	**101**	**146**

Table 3.3: Overall assessment of the requirements analysis

Starting from the requirement categories, the OTX platform even wins in more categories (6) than the IBM X-Force Exchange platform (4). In the category **General requirements**, MISP and OTX get the most points. Regarding the **Quality of CTI**, IBM (26 points) is the winner, however, the platform is closely followed by TC Open (24 points). Exactly the opposite is true with the category **Support of standards**. Here OTX tightly wins with 13 points compared to 12 points which the IBM platform scores. With 40 points, MISP and OTX share the first place in the category **CTI collection**. This was the category where the most points could be reached (58 in total). In the next category, **CTI correlation and**

analyzing functionalities, IBM X-Force Exchange and TC Open share the first place with 12 out of 12 points. Both platforms offer good analysis functionalities that suppert IT security staff in further researching into known and unknown threats. Regarding the **CTI integration and distribution functionalities**, again OTX claims the first place. The platform offers great support for the integration of its CTI into security monironing systems via TAXII, direct connect agents or its API. In the category **Group and sharing functionalities**, three providers share the first place: IBM X-Force Exchange, MISP and OTX reach 6 out of 8 points. Finally, in the category **Notification mechanisms**, again three platforms (IBM, OTX, and TC Open) share the first place.

Taken as a whole, all four tested platforms offer good functionalities and should be closely considered if an organization wants to integrate CTI into their security monitoring. IBM X-Force Exchange and TC Open offer great analysis functionalities while OTX makes integration of CTI quite easy. The strength of MISP on the other hand lies in its export functionalities and its support for different formats. Considering that TC Open is just a platform with limited functionalities compared to the commercial versions of ThreatConnect, it already got a good result. Therefore, it might be useful to take a closer look at the commercial versions as well.

Which platform is the most suitable, or if a combination of platforms should be considered, depends mostly on the specifics of the organization that wants to integrate those platforms into their security monitoring. For example, a combination of the IBM X-Force Exchange and the OTX platform would bring the effect that within each requirements category one of the "winners" platforms would be deployed. However, depending on the organization that wants to integrate one or more CTI platforms, some requirements will be more or less important. For example, if the organization has great financial resources, R02 will not be very important. Moreover, it depends on the security monitoring devices that are already deployed within the security monitoring of that specific organization. Depending on the SIEM vendor or the NIDS solution, one or the other platform will be more suitable. In addition, the goals of the organization are a key factor for the selection of a CTI platform. In some cases, the results of the requirements analysis for the category **Group and sharing functionalities** will have a major impact on the selection of a platform while for another organization the category **CTI correlation and analyzing functionalities** would be more important. The categories **General requirements** and **Notification mechanisms** on the other hand, might be in general less significant for the selection process, while the categories **Quality of CTI**, **Support of standards** and **CTI integration and distribution functionalities** should be taken into account in any case.

To draw a general conclusion from the requirements analysis is therefore not particularly useful. Rather, the requirements analysis helps to show the strengths and weaknesses of the individual platforms according to the introduced categories. It thus forms the basis for a later, individual conducted selection process.

Requirements ID	IBM X-Force Exchange	MISP	Open Threat Exchange	TC Open	Priority
General requirements					
R01	-	++	-	-	low
R02	+	+	++	+	medium
R03	++	++	++	++	high
Quality of CTI					
R04	++	+	-	++	high
R05	++	-	++	++	high
R06	++	-	++	++	high
R07	++	++	++	++	high
R08	+	-	-	-	medium
Support of standards					
R09	++	++	++	+	high
R10	++	-	++	+	high
R11	-	+	+	+	low
CTI collection					
R12	++	++	++	++	high
R13	-	-	-	-	high
R14	++	+	++	++	medium
R15	++	++	++	++	high
R15a	-	++	-	-	medium
R15b	++	++	++	-	high
R15c	-	++	++	++	high
R15d	-	++	-	+	medium
R16	++	++	++	+	high
R17	++	-	++	++	high
R18	++	-	-	+	medium
CTI correlation and analyzing functionalities					
R19	++	++	++	++	high
R20	++	+	+	++	high
CTI integration and distribution functionalities					
R21	+	-	+	+	medium
R22	-	-	++	-	medium
R23	++	++	++	+	high
Group and sharing functionalities					
R24	++	++	++	++	medium
R25	++	++	++	+	medium
Notification mechanisms					
R26	++	+	++	++	medium

Table 3.4: Overview of the requirements analysis

4 Concept for the integration of CTI into the security monitoring of university networks

The objective of this chapter is to present a general concept for the integration of CTI into the security monitoring of university networks. As examined in the previous chapter, there are some CTI platforms on the market that offer good functionalities and support a lot of the requirements that have been developed in Chapter 3.3 in accordance to the specifics of higher education institutions. In this chapter, a general implementation concept is presented which points out the major steps for the successful integration of CTI platforms. The concept is structured in several stages and considers the necessary steps towards a mature CTI concept to be deployed by universities. The individual steps build on one another and are presented precisely in that order in the following chapters.

Figure 4.1 is a simplified representation of the overall process of the automated integration of CTI into the security monitoring of higher education networks. A detailed description of the individual steps is given in the following sections. As discussed in Chapter 3 it is hardly possible to describe a single prototype of a university. Since the organization of the network can vary from university to university, it is likewise difficult to present a general concept for the integration of CTI into the security monitoring of university networks. For example, in the present application scenario of the MWN, the security monitoring takes place at the central X-WiN network transition. However, this is a characteristic of the LRZ who functions as the IT service provider of LMU. At another university, a central network transition might not exist. In this case, the individual university departments might have different network transitions that need to be monitored individually. In such a case, the concept presented below must as well be applied differently. One solution would be, for example, that each institute deploys the concept to its individual needs. Another possibility would be to task a central office, which makes pre-processing and offers CTI as a service for the individual institutes. Figure 4.2 illustrates how the process of the automated integration of CTI into the security monitoring of higher education networks could look in this case.

If the security monitoring does not take place at a central network transition as in the case of the application scenario of the LRZ, the responsible persons must consider whether each institute carries out such an implementation on its own (Figure 4.1) or if the data center provides CTI as a service (Figure 4.2). In the latter case, the service provider would be responsible for the collection, cleansing and storage of the data and then distribute it to the security devices of the individual institutes. The steps discussed in the following sections of this concept would be divided in this case. Step 1 would have to be carried out by the institutes themselves, except for the selection of one or more CTI platforms, which would be the task of the service provider. With the requirements elaborated in step 1, they could then contact the central IT provider and commission the procurement of CTI. Step 2 and 3 would then be the task of the entity that occurs as the central CTI provider. The CTI

provider would be responsible for the CTI collection, analysis, and sharing as well as the individual working steps included in these three processes, such as format conversion and CTI storage.

Besides, some decisions that lead to one or another implementation of the concept might change over time. This is especially true when organizations are gaining more and more experience with the integration of CTI into their security monitoring and want to rethink their current implementation of this concept. Figure 4.1 and 4.2 cover all three use cases discussed under step 2. However, it is also conceivable that an organization only implements one of the three use cases to gain experience in the field of CTI and then expands the implementation step by step at a later stage. In such a case, the concept can function as an iterative process and the responsible persons can go through the individual steps of the concept again to expand their existing CTI integration.

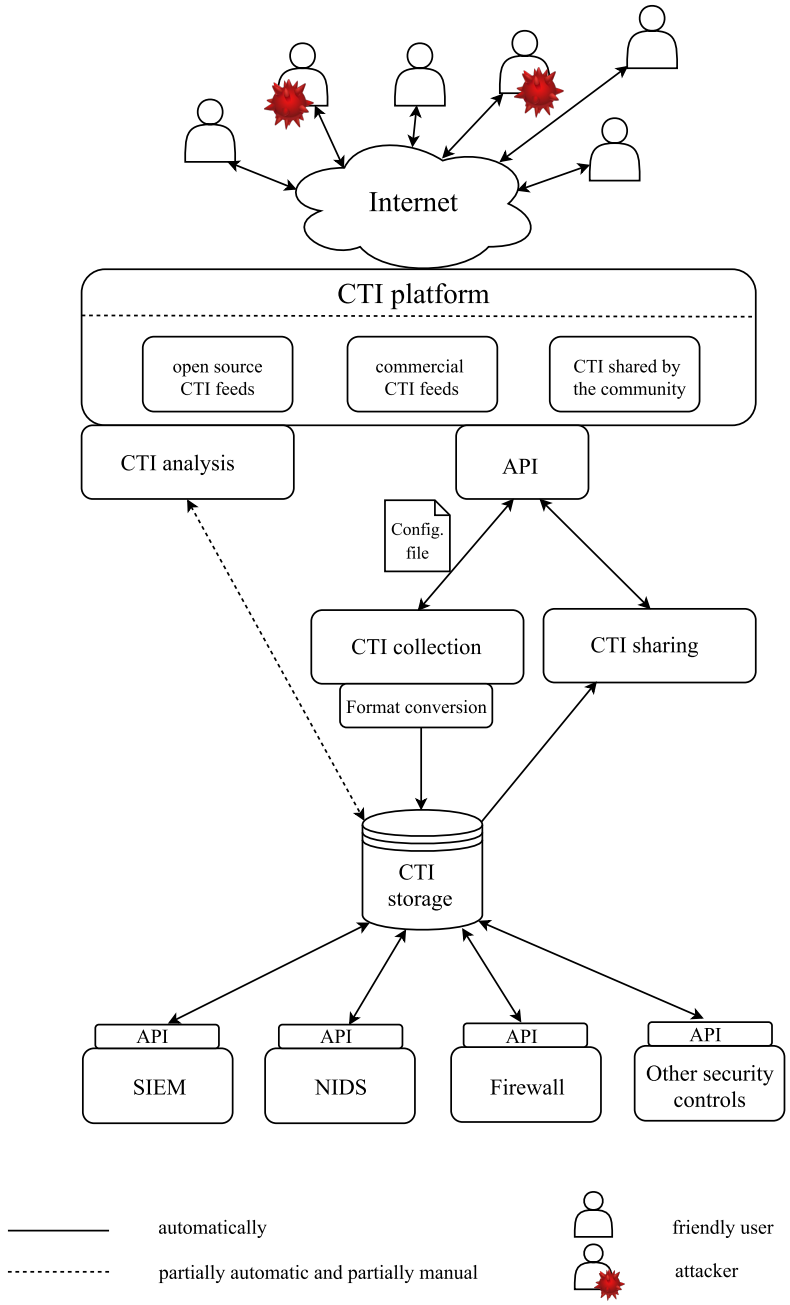

Figure 4.1: Automated integration of CTI into the security monitoring of higher education networks in the case of a centralized security monitoring

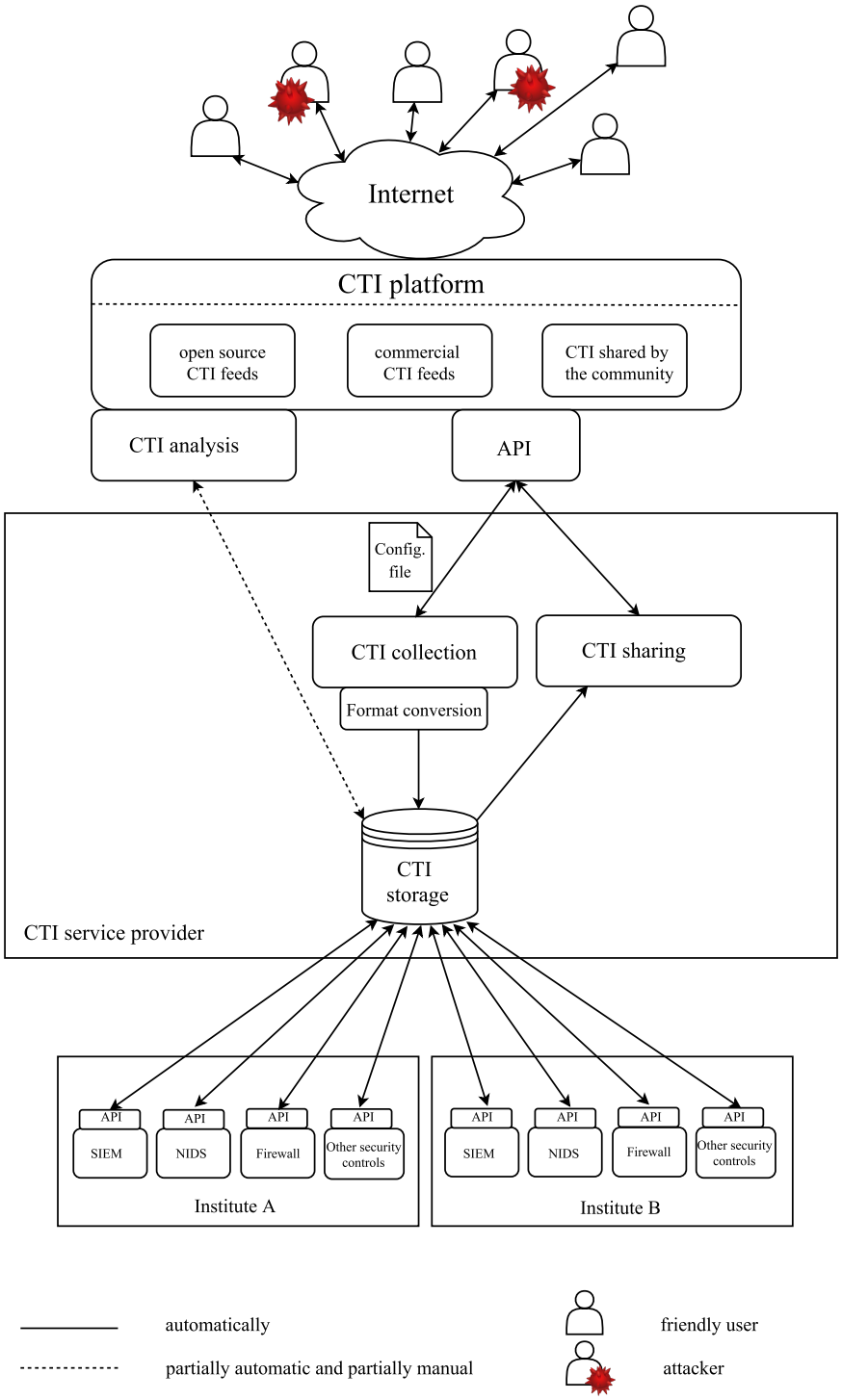

Figure 4.2: Automated integration of CTI into the security monitoring of higher education networks in the case of a decentralized security monitoring

4.1 Step 1: Definition of goals and strategy development

The previous chapter deliberately refrained from selecting a specific platform as *the* best solution. This is mainly because, depending on the individual organizational characteristics (e.g. resources, such as money or staff, organizational goals and internal structures), a CTI platform is more suitable for one or the other organization. For example, already deployed security monitoring solutions can be decisive for selecting one or the other platform. The same is true for the implementation process. For this reason, as a first step, it is very important to be aware of one's own goals and internal structures. Depending on the different expectations and needs, the measures taken for the integration of CTI platforms into the security monitoring of university networks can have totally different focal points. Moreover, downloading CTI from a platform without strategic pre-selection of relevant data risks creating a huge number of false positives and could easily lead to an overload of the systems.

Therefore, people concerned with the integration of a CTI platform should think about as well as define attainable objectives. Questions that are worth to reflect upon are:

- What goals do I want to achieve by integrating a CTI platform into the security monitoring of my organization? Desirable goals are, for example, to reduce the number of false-positives or to better assess the dangers posed by a threat. In addition, successful integration can help to prevent future attacks.
- What techniques and processes are currently in use to repel cyber-attacks? A close look at the security monitoring processes and workflows as described in Chapter 3.2 for the application scenario of this thesis, is therefore essential.
- How could they be improved to better detect and assess existing threats?
- What are currently the biggest threats to my organization? To integrate CTI data successfully, I need to be aware of the kinds of threats that are relevant to my organization.
- What kind of CTI could be the most valuable to better protect my organization from cyber-attacks?
- Where do I want to store the collected CTI? Do I want to store it locally or do I keep all information in the platform?

The individual categories in which the requirements analysis has been subdivided illustrate that there are various ways in which CTI can be used. Therefore, an important decision to be taken during the first step is how to use CTI within the organization. Basically, there are three use cases which, however, can be combined according to one's preferences. On the one hand, the focus can be on the **CTI collection**. This is the case if an organization wants to enrich its security devices with further useful data. The CTI platform then serves as an additional source for threat information besides, for example, one's own alerting mechanisms or CERT notifications.

On the other hand, **CTI analysis** could be the main motivation for integrating a CTI platform into the security monitoring. This would be the case if an organization wishes to

further investigate the threat information it already possesses. In this way, it would also be possible to generate your own CTI. For example, the platform could be used to gain a better understanding of the threat information reported by security devices already deployed, such as a firewall or NIDS. Moreover, one can get more information on the actual threat potential of the reported incident. Finally, **CTI sharing** as well as the mutual exchange between different organizations can be a major concern. Since attack patterns are generally the same or even repeat, it can be of great value to learn about other organization's experiences with cyber-attacks and to share knowledge about potential threats.

Only when an organization is aware of the various possibilities offered by a CTI platform and an appropriate strategy has been decided (if necessary, by combining the various options), one or more platforms should be selected for the integration into the security monitoring. In addition, there are further decisive factors for the selection of a platform. The support of standards, such as STIX and TAXII, greatly simplifies the automation of CTI collection and distribution. Moreover, depending on the security devices into which one wants to integrate the CTI, one or another platform can be more suitable. Therefore, it is important to take a closer look at the already deployed security architecture and to think about which platform best complements it. Finally, resources, such as money and staff, as well can have an influence on the selection of the platform. This is especially the case with the commercial versions which offer greater functionalities and bigger data rates.

For the first step of this implementation concept, an exact organizational consideration and self-assessment is mandatory. Once attainable objectives for the integration of CTI into the security monitoring have been developed, an overall strategy on how to use the CTI can be derived according to the above discussed considerations. Finally, the results of the requirement analysis in Chapter 3.3 can help to come to a decision in favor of one or more platforms that are best suitable for the integration. As a result of the first step, therefore, a clear idea of the expectations associated with the integration of CTI into the security monitoring should exist. This general concept can then be the basis for the selection of one or more platforms as well as the framework for the selection of relevant CTI.

4.2 Step 2: Definition of processes and workflows

Before starting the implementation of the integration of CTI into the security monitoring, it is useful to define processes for collecting, analyzing and sharing threat information. Therefore, step 2 examines possible workflows for the three use cases of CTI collection, CTI analysis and CTI sharing. The presented processes are simplified versions of the actual workflows which also will slightly differ when applied to the individual organizations.

CTI collection

Figure 4.3 shows the workflow for the collection of CTI. The first step is to decide on a relevant source, such as an IP Blacklist, that should be integrated into the security monitoring of the respective network. Once decided on the source, one must decide into which security device, e.g. SIEM, NIDS, or firewall, the CTI should be integrated. This decision will be typically based on the kind of CTI data. For example, if the source entails NIDS rules, they will be distributed to the deployed NIDS while an IP Blacklist might be disseminated to the

SIEM for further correlation. If the downloaded CTI is not directly integrated into the security device but first stored locally, it should be examined whether the data entails duplicates, and if so, one should delete these duplicates. At the same time, the priority of the respective incident or indicator should be increased since it appears several times and therefore is likely to be a relevant threat. Afterwards, the question arises whether the downloaded data has the right format. Depending on the security device to which the CTI will be disseminated, a conversion of the format as well as a restructuring or extraction of data might be necessary. After eliminating the duplicates and transforming the data into the correct format, the CTI can be integrated into the security device.

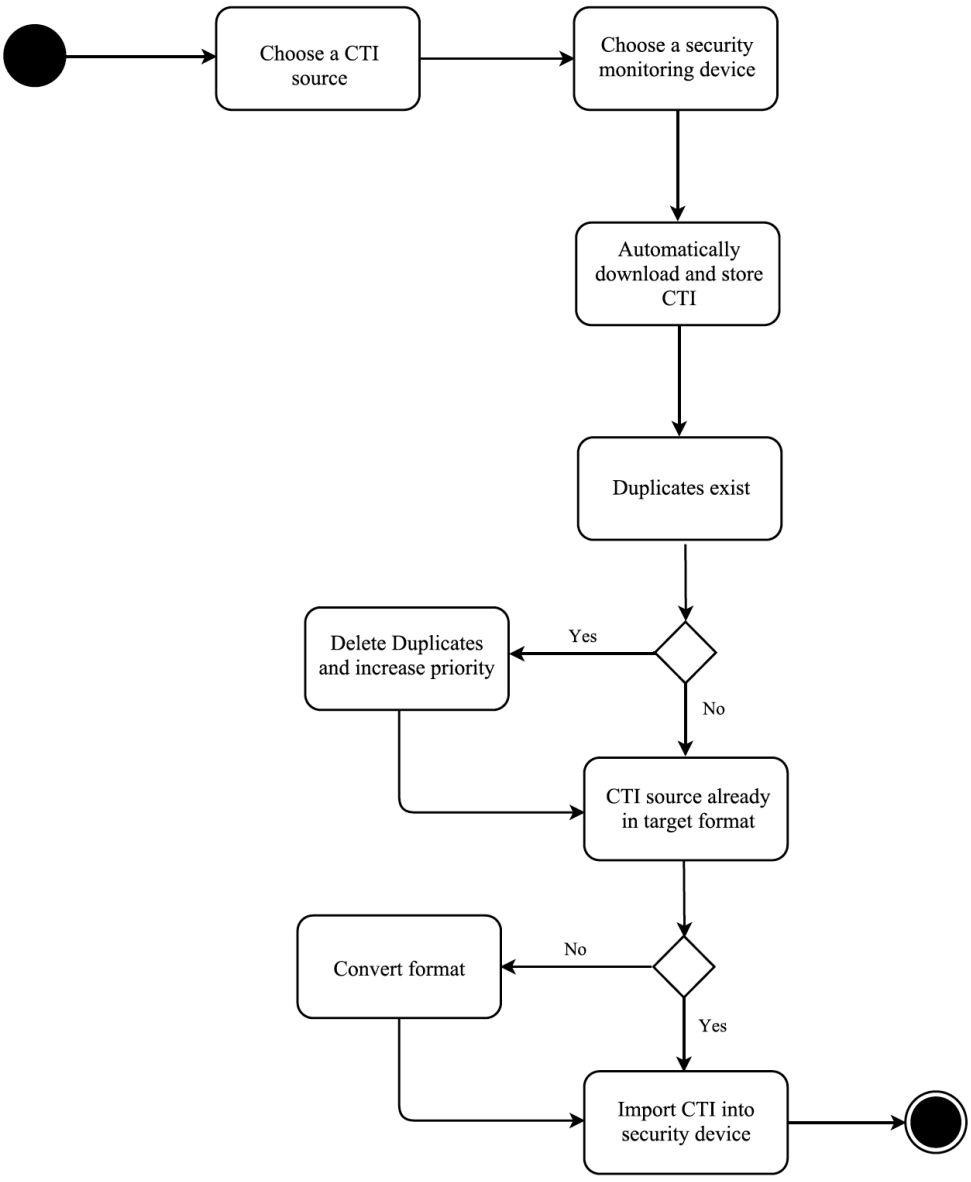

Figure 4.3: Process of CTI collection

CTI analysis

Figure 4.4 shows the workflow for the analysis of CTI. The starting point would be, for example, that an employee of the Abuse-Response-Team receives a DFN-CERT alert and wants to further investigate the respective incident. The employee now has two options: he can either do a keyword search or, depending on the format of the given threat information (as well as the available functionalities of the chosen platform), the employee can upload the file which contains the information and use the automatic analysis functionalities of the platform. In the latter case, it might be necessary to convert the file format of the given threat information into a format that is supported by the platform. Afterwards, the question arises whether the uploaded files or the manually investigated incident involves relevant information or not. If relevant, the data can be stored and imported into the matching security device. If not, no further action is needed.

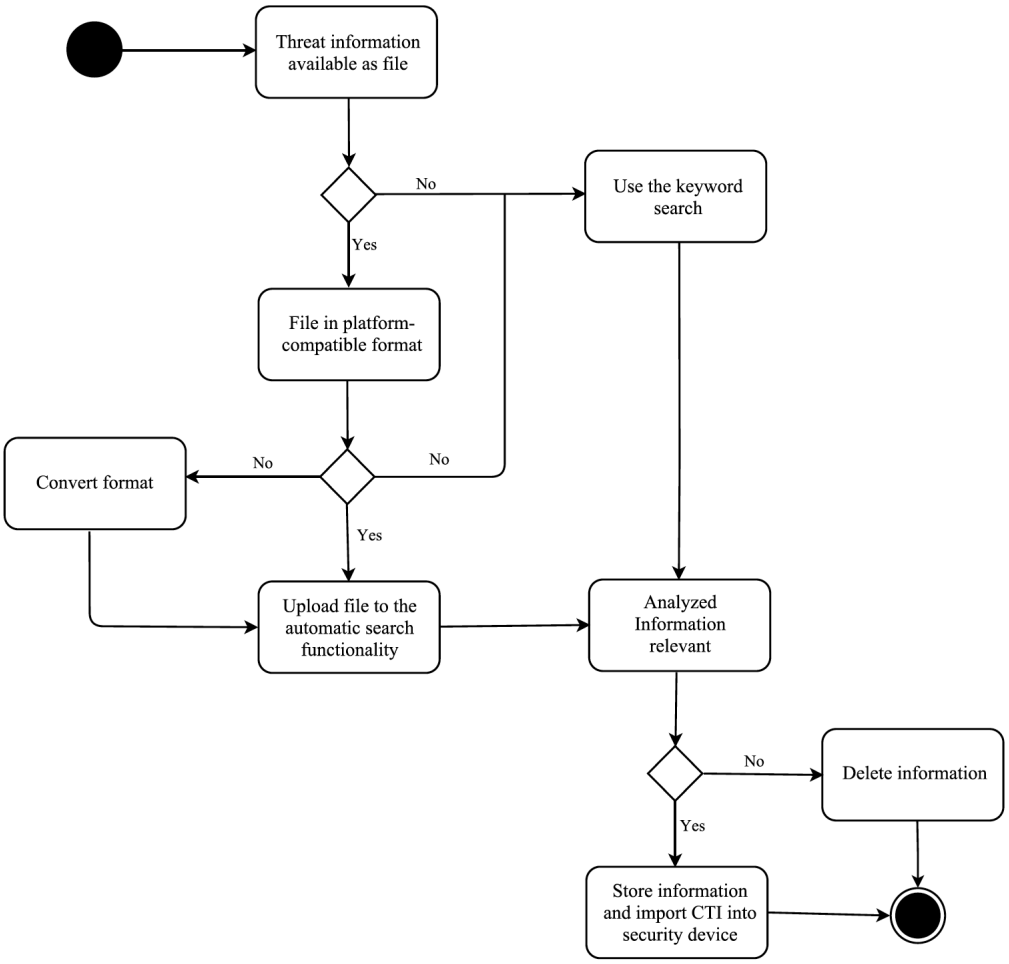

Figure 4.4: Process of CTI analysis

CTI sharing

Figure 4.5 shows the workflow for the distribution of internally aggregated CTI. There are multiple sources for the aggregation of internal CTI. Employees who are responsible for the security monitoring of a university network can gather such internal CTI from a NIDS, firewall or other deployed security solutions. Information on experienced exploit kits, malware infections and other attacks can be collected and enriched with further context. However, as discussed in section 2.2, the possibility of gathering internal CTI depends on given resources as well as the organization's analyzing capabilities and processing skills. Moreover, every organization needs to weigh up which information it wants to share in a community and which it regards as confidential.

The first and second steps are to collect internal data from deployed security devices and enrich it with further context. Afterwards, the question arises whether the aggregated data is confidential or can be shared with other users of the CTI platform. If the information is not intended for the public, it can be stored locally for the own usage. Otherwise, it can be structured in such a way that it fits the requirements for CTI sharing of the respective platform (e.g. brought to a supported format) and afterwards can be uploaded to the platform. Moreover, depending on the platform's functionalities, it might be possible to restrict the visibility of the uploaded data to certain users.

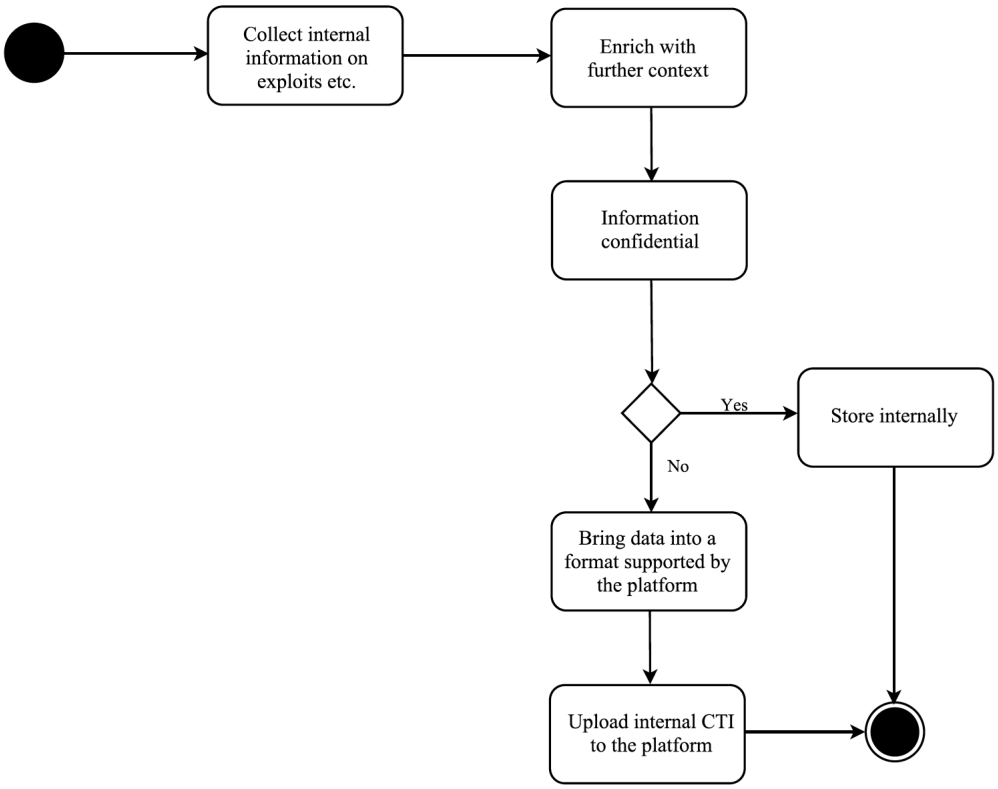

Figure 4.5: Process for the management and sharing of internal CTI

4.3 Step 3: Automation framework for the integration of CTI into the security monitoring of university networks

Since the manual processing of the above described work steps is very time-consuming, solutions for the automation of individual sub-processes of the described workflows are presented in this third step. For this purpose, various approaches for the technical implementation of the individual work steps are discussed. The implementation of these, however, depends on step 1 and 2 of this concept and will vary between different organizations.

Retrieving CTI

The CTI platform combines various CTI, such as open-source feeds, commercial feeds, and CTI created and shared by users, and thus serves as a central source for retrieving CTI. As can be seen in Table 3.4 of the requirements analysis, all platforms offer functionalities to automatically export threat data. Depending on the platform and the security device into which the CTI is to be integrated, this could be as easy as clicking a checkbox to enable the connection between the different systems. Afterwards, the only thing to do is to subscribe to those sources that are relevant to the respective organization. This would be the case, if the platform already offers direct integration with the security device, e.g. if both tools are provided by the same vendor. However, if this kind of direct integration is not available, there are other ways to retrieve the CTI automatically. For this purpose, all investigated platforms offer an API that can be accessed by its users. Clients can automate querying and retrieve information on DNS, IP addresses, malware, etc. via HTTP or, if supported, TAXII requests. The response to the request will usually be a XML, JSON or STIX file with the requested information.

Besides the URL via which the API is available, using the API usually will require an API key as well as some sort of authorization information, such as a key and password or username and password combination. A HTTP client can be used to access the platform API and to pull the data from the associated server. Therefore, the CTI platform should provide an API documentation which helps to build the correct requests. The queries can, for example, be made via a Python or Bash script which runs as a daemon and repeats the queries to the API within a specified time frame. The same is true for a TAXII request. Moreover, by adding logging calls to the script, one could track major events that happen when the script is running. Activities such as the successful or failed login and the successful or failed download of CTI information can be captured by logging messages to a separate file. A configuration file can be used to set the framework parameters for retrieving the data. For example, the API key and, if required, a user name and password can be specified here.

Format conversion

A format conversion might be necessary in different cases and will be applied in form of a restructuring, an extraction, or an actual format conversion. A restructuring must be applied if the received data comes in the wrong representation. For example, a timestamp information, which includes the date and time of the creation of the data it refers to, could come in US date format when needed in German date format. Therefore, at first the respective data must be checked and then, if necessary, must be restructured to fit the requested format.

The prior checking of the data is also important, for example, to recognize whether data categorized as IPv6 actually corresponds to an IPv6 address. An extraction, on the other hand, could be necessary if only certain data of the retrieved file is relevant. In this case, those columns with the relevant data must get extracted and stored in a separate file for further processing. Moreover, only certain data of a specific category could be relevant. For example, only country-specific IP addresses could be interesting for an organization. In this case, extraction of the specific data to a separate file will be useful. In addition, a complete format conversion might be necessary as well. This would be the case if the retrieved data is, for example, in XML format and the further processing requires the CSV format.

For the above described data cleansing process, a script or program could be developed that carries out these tasks according to the ETL (extract, transform, and load) process. The ETL process, which is usually associated with data warehousing, foresees three steps for the automatically processing of data from multiple sources. First, the extraction process reads the data from various sources and extracts the relevant data. Second, the transformation process converts the beforehand extracted data into the desired form, e.g. by using rules or lookup tables. Third, the loading process writes the data into the target storage. Tools for the individual steps of the ETL process can be found online [VG10]. For the CTI conversion, it is conceivable to combine and adapt several tools from the Internet or to write a program from scratch, which fulfills the functions discussed above.

Storing CTI

Depending on how the CTI is being used, there are various ways for storing the data. If an organization does not want to store the collected CTI locally, the CTI platform or other cloud-based storages can be used. Moreover, it is thinkable that the CTI is being distributed to the security devices directly. However, a more advanced solution for the usage and further processing of a great amount of CTI is to store the data locally, e.g. in a database. This way the CTI can also be enriched with certain metadata, such as

- Date and time when the CTI was collected
- Label that indicates whether it is internal or external CTI
- Source of the CTI
- Priority of the CTI that indicates how important it is to exploit the included indicators for the security monitoring
- In case of internal CTI a label that indicates the confidentially of the data and whether it can be shared or not
- The type of CTI, e.g. IP addresses, URLs, NIDS rules, TTPs,...

Building a local repository of CTI requires a lot of capacity and thinking. For example, before storing the data, a program should check if there are duplicates. If duplicates exist it should then delete the duplicate and increase the priority of the respective indicator. Therefore, policies should be defined that describe rules for when data should be archived, deleted or adjusted. Rules for updating the database on a regular basis as well need to

be defined. Moreover, updating, deleting or overriding information as well as increasing or decreasing the priority of certain CTI should as well be possible through manual input. As with the retrieval process, the process of storing the CTI could also be written into a log file to facilitate a better traceability of the processes. The implementation of a database for storing the CTI is optional and as well can be implemented at a later stage, e.g. when an organization has made its first experiences with CTI and wants to expand its usage.

Distribution of data to security devices

After the CTI has been retrieved and, if necessary, stored, the actual integration of the CTI into the individual security monitoring devices follows. For further correlation and alerting purposes, SIEMs will be the most suitable integration point. Another security monitoring solution that is suitable to alert on specific CTI would be the deployed NIDS solution. For blocking purposes, the firewall would be the right choice for the integration of CTI. The various possibilities for the embedding of data now depend on different factors, such as the deployed security monitoring solutions and the decisions made in the previous steps. Integration of CTI that is provided by the same vendor as the deployed SIEM solution, for example, will not require major implementation efforts other than activating the integration via the web interface of the SIEM solution. However, if direct integration is not available, there are other options.

As discussed above, CTI can be retrieved by scripting against the API of the supplying platform. After successfully downloading the CTI, there are various possibilities of how to further proceed. The CTI could be directly delivered to the security device by scripting against its API. Another possibility is to store the CTI, as described in the previous section, and then deliver it automatically (again by scripting against the API). Furthermore, it would be possible to select required CTI manually and as well manually load it to the security devices. However, this can be very time-consuming and therefore will hardly be applicable to an organization.

It is important to note that, depending on the pre-processing of the retrieved CTI, various results can be achieved. For example, downloading CTI from a platform and then feeding it into a SIEM or firewall without prior data cleansing risks creating a huge number of false-positives and could easily lead to an overload of the systems. Therefore, the selection of the CTI to be integrated should be well thought-out (see step 1 of this concept) and, if necessary, a prior data cleansing should take place. As with data storage, predefined rules can help to control and improve the process of CTI selection and its distribution across individual security monitoring devices.

5 Implementation

To demonstrate how automation of the process displayed in Figure 4.1 and 4.2 can be achieved, the automatically retrieving of certain CTI from a CTI platform as well as the storage of the downloaded CTI has been prototypically implemented. For demonstration purposes, the OTX platform was chosen. Figure 5.1 shows the screenshot of a pulse of the OTX platform concerning a certain malware named "Rurktar". As can be seen from the pulse description, Rurktar is a trojan spy that is still under development, but could become dangerous as soon as it is fully developed. On the basis of the summary one can see that currently seven indicators are connected with the Trojan. It refers to YARA signatures, IPv4 addresses and SHA256 hashes.

To implement the process of receiving CTI from the OTX platform, a Python script (see Figure 5.2) was written that uses the OTX-Python-SDK provided by AlienVault. Through the DirectConnect API, users of the OTX platform gain access to all pulses that they have subscribed to and can automatically download the related indicators. The script logs into the platform by using the API key that one receives after registering at the platform. It takes the required API key information from a configuration file (config.ini) in line 18 of the script. The script then downloads the indicators that are related to the before selected pulse by using its Pulse_ID: *59760ec4a87db71c15caeedc* (see line 26 of the script). This Pulse_ID was selected as an example and could be replaced by any other valid ID. Afterwards the script first outputs each indicator on the terminal (line 29 and 30) and then creates a CSV file and writes the indicators to that file together with the following information, which is representing the columns of the CSV:

- **Downloaded on**: The date and time on which the indicators have been retrieved via the OTX platform
- **Indicator**: The actual indicator information
- **Type**: The type of the indicator
- **Creation of Indicator in OTX**: The date and time on which the indicators have been added to the respective pulse
- **Expiration of Indicator**: The date and time when the indicator will expire
- **Pulse_ID**: The Pulse_ID of the pulse to which the indicator belongs

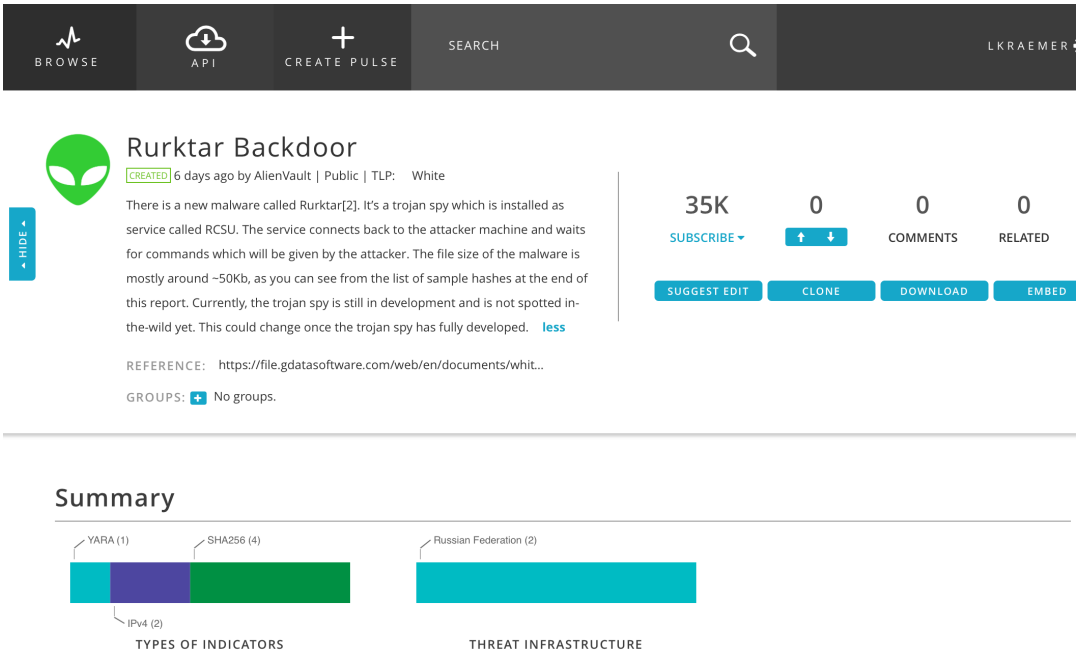

Figure 5.1: Screenshot of a selected pulse of the OTX platform

```python
#!/usr/bin/env python
#
# Retrieve IOCs from Open Threat Exchange
# https://otx.alienvault.com

# Use OTX-Python-SDK
from OTXv2 import OTXv2

import time
import dateutil
import ConfigParser
import csv

#Load the configuration file
with open("config.ini") as f:
    sample_config = f.read()
config = ConfigParser.RawConfigParser(allow_no_value=True)
config.readfp(open('config.ini'))

#Get API key from configuration file
OTX_KEY = config.get('otx', 'API_key')

#Get Indicators from Pulse by Pulse ID
otx = OTXv2(OTX_KEY)
pulse_ID = "59760ec4a87db71c15caeedc"
indicators = otx.get_pulse_indicators(pulse_ID)

#Output Indicators on the Terminal
for indicator in indicators:
    print indicator["indicator"] + indicator["type"]

#Create and safe csv-file named "otx_indicators_pulseID.csv" and write Indicators in it
with open('otx_indicators_'+pulse_ID+'.csv','wb') as csvfile:
    fieldnames = ['Downloaded on', 'Indicator', 'Type', 'Creation of Indicator in OTX', 'Expiration of Indicator', 'Pulse_ID']
    writer = csv.DictWriter(csvfile, fieldnames=fieldnames)

    writer.writeheader()

    for indicator in indicators:

        writer.writerow({'Downloaded on': str(time.strftime("%c")), 'Indicator': str(indicator["indicator"]),
        'Type': str(indicator["type"]), 'Creation of Indicator in OTX': str(indicator["created"]),
        'Expiration of Indicator': str(indicator["expiration"]), 'Pulse_ID': pulse_ID})
```

Figure 5.2: Script that automatically downloads CTI from the OTX platform

The resulting CSV file, which is shown in Figure 5.3, is named otx_indicators_59760ec4a87d b71c15caeedc.csv. In order for the file to be assigned to the corresponding pulse from which the indicators have been downloaded, it contains the Pulse_ID. It also contains the information listed above, such as Type or Expiration date. The file forms the basis for the further processing steps which have been descriped in Chapter 4 under the topic *Format conversion*. For example, to store the indicators separately by type, the YARA signatures can be written to a separate YARA file, the IPv4 addresses into an IPv4 file, and the SHA256 hashes to a particular file that contains only hashes. Moreover, one could add priorities to the indicators for easier evaluation of their threat potential. After the data cleansing and preparation, they can finally be fed into the appropriate security tools. For example, a NIDS can process the SHA256 hashes and thus prevent the opening of compromised files.

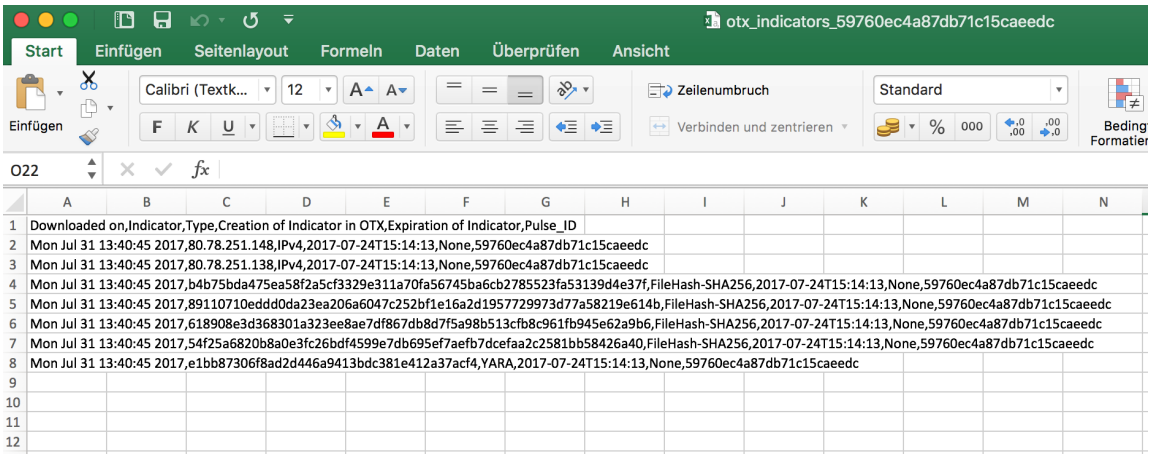

Figure 5.3: CSV file with collected indicators of the OTX pulse

Another possibility is to forward the IP addresses to the deployed SIEM for further correlation. To demonstrate how such an advanced processing of CTI data could look like, the following sections briefly describe how the downloaded IPv4 addresses from the OTX pulse can be integrated into QRadar, the SIEM solution used at the LRZ.

In QRadar lists of elements, such as IP addresses or other IOCs, can be written to so-called *reference sets*. In this way users can integrate externally retrieved CTI into QRadar, which uses this data to detect suspicious behavior by correlating it with events and incidents that occur on the network. Basically, there are two possibilities to create reference sets and write data to it: Users can either log in to the QRadar interface as administrator or they can use the API. In the first case, users can create a reference set to hold a list of IP addresses via the QRadar user interface and then import the elements to be added from a locally stored CSV file or manually add them to the set. In the latter case, users can access the API by sending HTTPS requests to a specific API endpoint, depending on the action they want to perform.

To manage reference data collections one must send its requests to the */reference_data endpoint*. The following HTTP methods specify the action that is to be performed by the request: GET, POST, PUT, or DELETE. Admin permissions are required for POST and

DELETE requests that are send to the /reference_data endpoint. To add or update an element in a reference set the *POST /reference_data/sets/name* request is used. Required parameters for this request are the name of the reference set to which one wants to add or update an element and the value of the element that is to be added or updated [IBM6b].

Once the IP addresses retrieved via the OTX pulse are written to a reference set in QRadar by using one of the above described options, rules can be created that make use of these indicators. For example, such a SIEM rule can be used to check whether an IP address that is accessing the network is entailed in the reference set. Moreover, the rule could be defined to send notifications to the administrator about communication from a potential malicious IP address and to initiate further analysis or to invoke direct actions.

6 Summary and Outlook

The aim of this work was to provide an initial overview of the emerging issue of CTI and to assess possibilities for the integration of CTI into the security monitoring of higher education networks. Since no thesis on this topic was written in conjunction with the LRZ before, it was an additional goal of this study to create the basis for possible future works. Therefore, Chapter 2 provided an introduction to the main goals as well as existing technologies and standards within the field of CTI. Subsequently, in Chapter 3, requirements for the use of CTI platforms in the context of the security monitoring of university networks were developed and an analysis of four different platforms was carried out. In Chapter 4, a multistage concept for the automated integration of CTI into the security monitoring of university networks was developed. For the thesis to be useful for other universities as well, it was tried to keep the requirements analysis and the concept as generic as possible. Therefore, the application scenario was only occasionally addressed, for example for demonstration purposes or to provide a more vivid description.

To exploit the full potential of CTI, a lot of thought must be put into the preprocessing. The success of using CTI in the case of network security at universities therefore depends on existing resources as well as the appropriate handling. The simple feeding of threat information to security controls without prior consideration and pre-processing might lead to a system overload or create false-positives and thus would even lead to a worsening of the security monitoring process. If, however, the individual steps described in Chapter 4 are applied carefully, the integration of CTI into the security monitoring can make a valuable contribution to the prevention of cyber-attacks. There are two main reasons for this. First, these platforms offer good functionalities for the automation of processes and can therefore reduce the workload of employees and provide faster processing. Second, CTI platforms offer near to real-time data and information on malware that is in development as opposed to, for example, DFN-CERT alerts that only report on malware that has already been spread.

Step 1 of the concept has been partially implemented in the course of this thesis. For example, some of the organizational characteristics and the security monitoring architecture of the LRZ have been discussed in Chapter 3.1.5 and 3.2. For future implementation endeavors, it would be necessary to conduct a detailed analysis of the threat landscape to determine what kind of CTI would be the most relevant to optimize the security monitoring currently deployed to protect the MWN. This is an inevitable step for the pre-processing of the CTI offered by the various platforms. Interviews with the employees of the Abuse-Response-Team could be conducted to get a better understanding of the current situation and to determine which CTI would be the most relevant for the integration into the security monitoring of the MWN. Step 2 could also be implemented based on these interviews by first determining the purpose of using CTI at the LRZ and then adapting the respective use cases accordingly. Step 3 was prototypically implemented by downloading CTI from one of the platforms by using a simple script query. A future work could develop and implement an advanced system

design which is based on the process depicted in Figure 4.1 and automatically integrates the workflows developed in step 2 of the concept.

As has been discussed in the previous chapters, sharing of threat data is a major concern in the field of CTI. Sharing experiences and indicators related to cyber-attacks with other organizations, especially with those within the same industry, can be very helpful. Therefore, with regard to the application scenario of this thesis, sharing CTI with other universities would be a promising step towards a more sophisticated security monitoring approach. However, this would require an overarching coordination between different organizations. As a first step towards such a general CTI exchange between universities, a regional cooperation, e.g. between Bavarian universities, could be formed.

The present thesis has dealt mainly with tactical CTI in form of IOCs. Another interesting topic which could not be addressed in more detail, is the processing and integration of long textual descriptions of adversary's TTPs. A future thesis could examine ways to automatically process strategic CTI, such as textual descriptions of new attack methods. It would be interesting to investigate how to map a new attacker's approach to the security monitoring. Therefore, one could examine ways to process long TTP descriptions and to translate them into machine readable rules or to derive settings for the deployed security devices.

List of Figures

List of Tables

Bibliography

[Ali7a] AlienVault: *Open Threat Exchange® User Guide*, 2017a. https://www.alienvault.com/documentation/resources/pdf/otx-user-guide.pdf. Accessed: June 23, 2017.

[Ali7b] AlienVault: *Open Threat Exchange®*, 2017b. https://www.alienvault.com/open-threat-exchange. Accessed: August 03, 2017.

[Bar14] Barnum, Sean: *Standardizing Cyber Threat Intelligence Information with the Structured Threat Information eXpression (STIXGuide™)*. MITRE, 2014. https://stixproject.github.io/about/STIX_Whitepaper_v1.1.pdf. Accessed: May 27, 2017.

[Bia13] Bianco, David: *The Pyramid of Pain*, 2013. http://detect-respond.blogspot.de/2013/03/the-pyramid-of-pain.html. Accessed: May 27, 2017.

[Bit15] Bitkom e.V.: *Spionage, Sabotage und Datendiebstahl. Wirtschaftsschutz im digitalen Zeitalter*, 2015.

[Bro16] Bromiley, Matt: *Threat Intelligence: What It Is, and How to Use It Effectively*, 2016.

[Bun16] Bundesministerium des Innern: *Cyber-Sicherheitsstrategie für Deutschland*, 2016. https://www.bmi.bund.de/cybersicherheitsstrategie/BMI_CyberSicherheitsStrategie.pdf. Accessed: April 29, 2017.

[Com16] Computer Incident Response Center Luxembourg (CIRCL): *Malware Information Sharing Platform (MISP) - A Threat Sharing Platform*, 2016. https://www.circl.lu/services/misp-malware-information-sharing-platform/. Accessed: June 24, 2017.

[CS17] Clemens Sauerwein, Christian Sillaber, Andrea Mussmann, Ruth Breu: *Threat Intelligence Sharing Platforms: An Exploratory Study of Software Vendors and Research Perspectives.* In Brenner, W. (editor): *Proceedings der 13. Internationalen Tagung Wirtschaftsinformatik (WI 2017)*, 2017.

[Dal15] Dalziel, Henry: *How to Define and Build an Effective Cyber Threat Intelligence Capability.* 2015.

[Der17] Der Tagesspiegel: *VDE-Unternehmen fordern nationale Cyber-Security-Strategie*, 2017. http://www.tagesspiegel.de/advertorials/ots/vde-verb-der-elektrotechnik-elektronik-vde-unternehmen-fordern-nationale-cyber-security-strategie/19707788.html. Accessed: April 29, 2017.

[EWB14] Eric W. Burger, Michael D. Goodman, Panos Kampanakis, Kevin A. Zhu: *Taxonomy Model for Cyber Threat Intelligence Information Exchange Technologies*. 2014.

[Far13] Farnham, Greg: *Tools and Standards for Cyber Threat Intelligence Projects*, 2013.

[FF16] Florian Flade, Lars-Marten Nagel: *Aufklärung von NSA-Cyberspionage unwahrscheinlich*, 2016. https://www.welt.de/politik/deutschland/article159223327/Aufklaerung-von-NSA-Cyberspionage-unwahrscheinlich.html. Accessed: April 29, 2017.

[Gar14] Gartner, Inc.: *Threat Intelligence: What is it, and How Can it Protect You from Today's Advanced Cyber-Attacks?*, 2014.

[HR16] Helmut Reiser, Stefan Metzger: *Das Münchner Wissenschaftsnetz (MWN) Konzepte, Dienste, Infrastruktur und Management*, 2016.

[IBM6a] IBM Corporation: *IBM X-Force Exchange Commercial API*, 2016a. https://public.dhe.ibm.com/common/ssi/ecm/wg/en/wgd03094usen/WGD03094USEN.PDF. Accessed: June 22, 2017.

[IBM17] IBM Corporation: *IBM X-Force Exchange*, 2017. https://exchange.xforce.ibmcloud.com. Accessed: August 03, 2017.

[IBM6b] IBM Corporation: *IBM Security QRadar. Version 7.2.8 API Guide*, 2014/16b. http://public.dhe.ibm.com/software/security/products/qradar/documents/7.2.8/en/b_qradar_api.pdf. Accessed: August 12, 2017.

[JF15] Jon Friedman, Mark Bouchard, CISSP: *Definitive Guide™ to Cyber Threat Intelligence*, 2015.

[Lei16] Leibniz-Rechenzentrum (LRZ) der Bayerischen Akademie der Wissenschaften: *Jahresbericht 2015*, 2016.

[Lei17] Leibniz-Rechenzentrum (LRZ) der Bayerischen Akademie der Wissenschaften: *Security and NAT Gateway for the Munich Scientific Network (MWN)*, 2017. https://www.lrz.de/services/netzdienste/secomat_en/. Accessed: June 15, 2017.

[MIS7a] MISP Community: *MISP - User Guide: A threat sharing platform*, 2017a. https://www.circl.lu/doc/misp/book.pdf. Accessed: June 23, 2017.

[MIS7b] MISP Community: *MISP - Open Source Threat Intelligence Platform & Open Standards For Threat Information Sharing*, 2017b. http://www.misp-project.org. Accessed: August 03, 2017.

[OAS7a] OASIS Cyber Threat Intelligence Technical Committee: *STIXGuide™ Version 2.0. Part 2: STIX Objects*, 2017a. http://docs.oasis-open.org/cti/stix/v2.0/csprd02/part2-stix-objects/stix-v2.0-csprd02-part2-stix-objects.pdf. Accessed: May 27, 2017.

[OAS7b] OASIS Cyber Threat Intelligence Technical Committee: *About STIX*, 2017b.

[OAS7c] OASIS Cyber Threat Intelligence Technical Committee: *Indicator for Malicious URL*, 2017c. https://oasis-open.github.io/cti-documentation/examples/indicator-for-malicious-url. Accessed: May 27, 2017.

[OAS7d] OASIS Cyber Threat Intelligence Technical Committee: *TAXII™ Version 2.0.*, 2017d. https://docs.google.com/document/d/1eyhS3-fOlRkDB6N39Md6KZbvbCe3CjQlampiZPg-5u4/edit#heading=h.ehzdxcsjrzgp. Accessed: May 27, 2017.

[Sha15] Shackleford, Dave: *Who's Using Cyberthreat Intelligence and How?*, 2015.

[SM11] Stefan Metzger, Wolfgang Hommel, Helmut Reiser: *Integriertes Management von Sicherheitsvorfällen.* 2011.

[SQ17] Sara Qamar, Zahid Anwar, Mohammad Ashiqur Rahman, Ehab Al-Shaer, Bei-Tseng Chu: *Data-driven analytics for cyber-threat intelligence and information sharing.* 2017.

[Ste15] Steinke, Michael: *Schwachstellenmanagement in Hochschulnetzen am Beispiel des Münchner Wissenschaftsnetzes*, 2015.

[Ste17] Stephenson, Peter: *IBM Security IBM X-Force Exchange*, 2017. https://www.scmagazine.com/ibm-security-ibm-x-force-exchange/article/629730/. Accessed: June 22, 2017.

[Süd16] Süddeutsche Zeitung: *Antisemitische Pamphlete: Hackerangriff auf Uni-Drucker*, 2016. http://www.sueddeutsche.de/news/politik/extremismus-antisemitische-pamphlete-hackerangriff-auf-uni-drucker-dpa.urn-newsml-dpa-com-20090101-160421-99-668212. Accessed: April 29, 2017.

[Thr17] ThreatConnect: *TC Open®*, 2017. https://www.threatconnect.com/free/. Accessed: August 03, 2017.

[TM14] Troy Mattern, John Felker, Randy Borum, George Bamford: *Operational Levels of Cyber Intelligence.* 2014.

[VDE16] VDE: *VDE: Networking the Future*, 2016. https://www.vde.com/en/about-us. Accessed: April 29, 2017.

[VG10] Vishal Gour, Dr. S.S.Sarangdevot, Govind Singh Tanwar, Anand Sharma: *Improve Performance of Extract, Transform and Load (ETL) in Data Warehouse*, 2010. http://www.enggjournals.com/ijcse/doc/IJCSE10-02-03-108.pdf. Accessed: July 23, 2017.

[VMw16] VMware: *University Challenge: Cyber Attacks in Higher Education*, 2016. http://www.comtact.co.uk/wp-content/uploads/2016/04/University-Challenge-Cyber-Attacks-in-Higher-Education-April-2016.pdf. Accessed: April 29, 2017.

[WH13] Wolfgang Hommel, Stefan Metzger, Helmut Reiser, Felix von Eye: *IT security concept documentation in higher education data centers: A template-based approach*. 2013.

[Wun7c] Wunder, John: *STIX 2.0 Finish Line*, 2017c. `https://www.mitre.org/capabilities/cybersecurity/overview/cybersecurity-blog/stix-20-finish-line`. Accessed: May 27, 2017.