computers

MDPI

Article

Interoperability Challenges in the Cybersecurity Information Sharing Ecosystem

Konstantinos Rantos [1,*], Arnolnt Spyros [2], Alexandros Papanikolaou [2], Antonios Kritsas [1], Christos Ilioudis [3] and Vasilios Katos [4]

1 Department of Computer Science, International Hellenic University, Agios Loukas, 654 04 Kavala, Greece; ankrits@teiemt.gr
2 Innovative Secure Technologies, 60 Monastiriou, 546 27 Thessaloniki, Greece; a.spyros@innosec.gr (A.S.); a.papanikolaou@innosec.gr (A.P.)
3 Department of Information and Electronic Engineering, International Hellenic University, Sindos, 574 00 Thessaloniki, Greece; iliou@ihu.gr
4 Department of Computing and Informatics, Bournemouth University, Poole BH12 5BB, UK; vkatos@bournemouth.ac.uk
* Correspondence: krantos@cs.ihu.gr

Received: 4 February 2020; Accepted: 29 February 2020; Published: 6 March 2020

Abstract: Threat intelligence helps businesses and organisations make the right decisions in their fight against cyber threats, and strategically design their digital defences for an optimised and up-to-date security situation. Combined with advanced security analysis, threat intelligence helps reduce the time between the detection of an attack and its containment. This is achieved by continuously providing information, accompanied by data, on existing and emerging cyber threats and vulnerabilities affecting corporate networks. This paper addresses challenges that organisations are bound to face when they decide to invest in effective and interoperable cybersecurity information sharing and categorises them in a layered model. Based on this, it provides an evaluation of existing sources that share cybersecurity information. The aim of this research is to help organisations improve their cyber threat information exchange capabilities, to enhance their security posture and be more prepared against emerging threats.

Keywords: cyber threat information; cyber threat intelligence; interoperability; cybersecurity; evaluation

1. Introduction

Information has undoubtedly become one of the most valuable assets for organisations, and whose dependence on it is constantly rising. At the same time, the frequency and ferocity of cyberattacks is also increasing, posing a great threat to business environments. According to a study conducted jointly by Ponemon Institute and Accenture, the average cost of cybercrime to the organisations in 2018 has risen to $13.0M [1]. Moreover, 79% of Chief Information Security Officers in the Banking sector believe that cybercriminals have become more sophisticated [2].

In this constant battle, organisations have to retain visibility of emerging and evolving threats and defend themselves against a wide range of adversaries with various levels of motivations, capabilities and access to resources. These adversaries typically range from amateur hackers to well-organised and highly capable teams that have direct access to vulnerabilities and exploits, and therefore become advanced, and sometimes persistent, threats to organisations. The impact of such sophisticated, dynamic and automated cyberattacks can be devastating.

The community has recognised the need of being able to share Cyber Threat Information (CTI) in a timely and reliable manner to enhance its ability to identify any malicious activity or sources

and mitigate attacks in a timely manner, prior to damaging organisations' assets. NIST [3] defines CTI as "any information that can help an organization identify, assess, monitor, and respond to cyber threats". In a survey conducted by SANS regarding the evolution of cyber threat information [4], 72% of survey respondents mentioned that in 2018 they had produced or consumed such information for their network defence. The respective percentage for 2017 was 60% [4]. This demonstrates that information sharing is becoming part of the organisations' strategy and the number of them that join the sharing community is rising. The types of information that can be produced and shared among communities include, among others, security appliances log entries and alerts, measurable and observable actions, security bulletins and advisories, identified vulnerabilities, news, reports and intelligent information.

The number of CTI sources is therefore increasing, as do cyber threat intelligence platforms capable of consuming information from threat intelligence feeds, analysing, evaluating and classifying it prior to sharing threat information with the community. Threat intelligence, according to the authors of [5], means evidence-based knowledge representing threats that can inform decision-making. Cyber Threat Information and Intelligence (CTII) facilitate situational awareness of the threat landscape, a deeper understanding of threat actors and their Tactics, Techniques and Procedures (TTPs), and greater agility to defend against evolving threats.

Nowadays, organisations have the ability to participate in such threat-sharing communities or intelligence groups, and analyse and evaluate CTII via their security operations team. Depending on the types of security appliances they are using, they can incorporate CTII into their security solutions, such as Unified Threat Management (UTM), Intrusion Detection/Prevention Systems (IDS/IPS) and Security Information and Event Management (SIEM). IT security vendors and service providers also collect such information from their clients worldwide, analyse it and feed it back to the deployed security appliances, thus gaining from the experience of their community. Analysis and evaluation of such information is considered essential, as sometimes the information that is shared is not properly filtered or checked.

Using unreliable sources poses risks to the organisations, as the information may not be accurate or complete. To confront this undesired situation, organisations go beyond closed groups and use multiple sources instead. As such, CTII sharing takes place among multiple actors, such as government agencies and organisations, private sector organisations and industry-focused groups. One of the most challenging issues in this process is achieving consensus regarding how this information should be shared among interested parties and the threat intelligence community. This requires having a common understanding on what information is shared, how it is shared and whether its sharing is law-abiding.

This paper addresses the interoperability challenges that the community faces when adopting specific sharing solutions. These challenges mainly stem from the adoption of multiple technical standards, strategies and policies among stakeholders, together with legal restrictions concerning information sharing. The aim of this work is to highlight points that impede the wide adoption of cyber security information sharing and the much-needed automation of the sharing process [4]. Security analysts would benefit from such automation, as they could devote more time on analysing collected interoperable data, as opposed to devoting their efforts on the collection process itself. Interoperability is also a means to widen the CTII sharing spectrum and engage more stakeholders in the exchange process, thus developing a global shield against emerging threats, with significant benefits for the community.

The rest of this paper is structured as follows. Section 2 provides some background information about the exchange of cyber security information. Section 3 identifies the interoperability barriers in CTII sharing and defines a model that comprises four interoperability layers. Section 4 provides an analysis of a number of CTII sources with respect to their policy, semantic and syntactic interoperability characteristics. Section 5 provides the conclusions and an outlook on future work.

2. Background

Cybersecurity Information Exchange (CIE) concerns sharing CTII with third parties to help organisations enhance their security posture and protect themselves from cyber threats. The main motive is to create new knowledge or services about cyber threats, as well as to make cyber-defence systems more effective and efficient.

CTI can be sometimes directly imported to security appliances, e.g., IP addresses reported as malicious can be directly imported to firewalls and IDS/IPS, to speed up the organisation's reaction to a potential threat. This, however, may come at the cost of introducing many false positives as, after an in-depth analysis and evaluation, these IPs may prove to be non-malicious. Therefore, prior to becoming a valuable asset for the organisation, cyber threat information has to be properly collected from various sources, correlated, analysed and evaluated to add significant value to the raw and/or unevaluated data, thus producing the so-called "Cyber Threat Intelligence". Analysed information significantly reduces false positives and establishes trust on the various sources it comes from. The downside of this process is the delay introduced to information sharing, due to the time-consuming analysis of, e.g., big data as well as other techniques, like threat sandboxing, which is used for monitoring the behaviour of suspicious targets.

The cooperation among organisations in the EU and the publication of detected security incidents are also encouraged, although there are also cases where they are obligatory, as stated in the Directive on security of Network and Information Systems (NIS) [6]. An implementation guidance for meeting the requirements of the NIS Directive is provided by ETSI [7]. In this context, a Computer Security Incident Response Team (CSIRT, also known as Computer Emergency Response Team – CERT) network has been established which counts more than 400 members willing to share incidents and risk-related information (an inventory of EU's CSIRT Network members is maintained by ENISA [8]). The network, however, is not restricted only to CSIRTs, but also includes commercial organisations, EU Institutions, Law Enforcement Agencies, Private & Public Sector organisations, National and Military Agencies.

Similar initiatives in the US that promote the exchange of information include the US Department of Homeland Security (DHS) Cyber Information Sharing and Collaboration Program (CISCP) [9] and the Automated Indicator Sharing (AIS) [10].

Examples of such shareable information include the following.

- Indicators of Compromise (IOCs), i.e., system artefacts or observables that contain patterns, such as malicious Internet Protocol addresses (IPs) or hashes of files containing malware, which can help identify suspicious or malicious activity.
- Tactics, Techniques and Procedures (TTPs), i.e., (detailed) description of the behaviour of an actor that can assist operational activities, such as details of exploits and malware delivery mechanisms. TTPs include tradecrafts, i.e., behaviour used to conduct a malicious activity, infrastructure used to deliver malicious content or maintain command and control capabilities, and attackers' intentions [11].
- Security alerts, i.e., notification, usually human-readable regarding security issues, such as vulnerabilities.
- Threat intelligence reports, i.e., collections of threat intelligence for various topics, such as threat actors, malware and attack techniques.
- Recommended security tool configurations, regarding automated collection and processing as well as healing of identified security issues.
- Vulnerabilities, i.e., known weaknesses in software implementations or procedures, and corresponding mitigations, such as security patches. Although in the information security literature, vulnerabilities are not threats, they are considered as an important component in the CTII sharing ecosystem.

CERT-UK together with the UK's Centre for the Protection of National Infrastructure define four subtypes of threat intelligence in a whitepaper published by MWR Infosecurity [12]. A similar

categorisation is also adopted by ENISA [13]. Based on the two schemes, stakeholders exchange information of one of the following types.

- Strategic reports providing high level threat analyses that is consumed at board level or by other senior decision-makers.
- Operational details providing information about impending attacks against an organisation.
- Advisories which include vulnerabilities, exploits patches and patch status, or high-level tactical patterns of activity on a host, service, network or internet level, tools and methodologies.
- Indicators of compromise, which include IP addresses, DNS names, URLs, specific values of format-specific fields (e.g., email headers), artefacts (e.g., hashes, registry and keys) related to malware and sequences of low-level events (e.g., syscalls and packets) linked to malicious behaviour.
- Low-level, i.e., network flow records and full packet captures, application logs, including typical IDS alerts, samples of executable files, documents and email messages.

Depending on the type of information that is being shared, organisations have the capability for prompt reaction, as a response to newly-emerged threats or vulnerabilities, for properly adapting their security defences against changes in their situational awareness at a more tactical level and make more strategic decisions on the allocation of their resources for upcoming threats.

The authors of [5] classify and distinguish existing threat intelligence types, focusing on technical threat intelligence issues, emerging researches, trends and standards. They also explain why there is a reluctance among organisations to share threat intelligence and provide sharing strategies that can help overcome policy interoperability issues. In [14], following a layered model, the authors propose a taxonomy to help classify existing threat-sharing standards and analyse interoperability. However, the five layers proposed by the authors do not address legal or semantic issues that are analysed in this paper.

So far, although CTII sharing is identified in the literature as an important issue, it is treated as a problem that can be solved with the adoption of some technical standards [3,14,15]. To the best of the authors' knowledge, no work has been published so far for a holistic approach to the interoperability problem in the CTII sharing.

3. Interoperable CTII Sharing

CTII information is generated and shared among devices and organisations that typically have well-established procedures to appropriately handle personal and classified information found within. When CTII is about to be shared, especially with external entities, several interoperability and security issues have to be confronted, which can be categorised into the four layers depicted in Figure 1 and that are further analysed in the following sections.

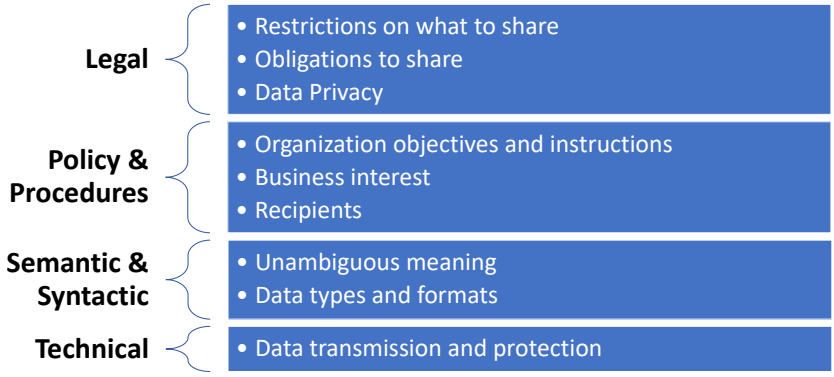

Figure 1. Cyber Threat Information and Intelligence (CTII) exchange interoperability layers.

3.1. Legal Interoperability

Legal interoperability is about ensuring that the legal frameworks under which organisations operate and provide services are aligned and do not impede the sharing of CTII. Legal frameworks open up the potential for improving the collaboration among public–private organisations and may encourage or require the sharing of cyber threat information. Examples of such frameworks are the EU's NIS Directive [6] and the US Cybersecurity Information Sharing Act (CISA) of 2015.

Legal constraints may also prohibit or restrict the uncontrolled sharing of CTII. Examples of the latter are any personally identifiable information (PII) that may be part of the shared sightings, such as usernames of entities that have been identified as sources of malicious activity and restrictions that stem from the corresponding telecommunications privacy legal framework. One of the main legal restrictions arises from the EU's General Data Protection Regulation (GDPR) [16] and relates to any PII shared with external entities without the user's consent. CTII sharing with external entities should not impact privacy and sharing parties must take measures to properly anonymise or pseudonymise any records of data that could otherwise be used to identify individuals.

Liability is another key factor that has to be considered when sharing information. Organisations have to take into account that they may be liable for any damage caused to stakeholders or entities that appear to be threat actors, by sharing information that may prove inaccurate or false. For instance, what would happen if an IDS/IPS falsely identified some malicious activity coming from a popular website and this was immediately reported to sharing parties?

Legal interoperability is even more important when data sharing spans multiple countries or legal domains. Therefore, restrictions imposed by each such domain must be carefully reviewed prior to information dissemination. The organisation has to establish boundaries on any sharing activities so that these will not infringe any legal restrictions.

Legal interoperability requires visibility control over CTII. The organisation has to have the means to perform deep inspection on what is practically going to be shared to identify any information that may cross the well-established boundaries regarding information sharing, and this should be reflected to the policies and procedures adopted by the organisation.

3.2. Policies and Procedures for Interoperability

Organisations' information sharing policies and procedures are formal statements that reflect the organisation's objectives and detailed instructions to achieve these objectives respectively. They are typically part of the organisation's information security policy and has to be endorsed by the organisation's top management.

Among the issues that the organisation has to consider at the Policy layer when deciding to share CTII, are the 5 Ws and 1 H that have to be answered prior to start sharing, as depicted in Figure 2. The policy should clearly address them, while well-established procedures and appropriate robust measures will properly support them to avoid policy violations, such as the leakage of classified or sensitive information. Therefore, to ensure that the security and privacy posture of the organisation is not negatively affected by the introduction of CTII sharing, the service's risk analysis should address the threats that are associated with the exchange of CTII with third parties and the consumption of the received information. The deployed security measures that will protect CTII and ensure that this is handled only by authenticated and authorised entities will minimise the associated risks.

The adopted procedures should facilitate information sharing according to the organisation's strategy and policy and should not be an impediment to the accurate, fast and secure exchange of high quality information. Organisations' policies may also be driven by security-related standards that address the sensitive issue of information sharing, such as the ISO 27010:2015 [17], which provides guidance on managing information within sharing communities and therefore facilitates trust establishment among stakeholders. Trust establishment with stakeholders is the foundation in CTII sharing and should be mainly addressed in the context of the "Who to share it with" question. A methodology that can be used for evaluating CTII sources based on quantitative parameters is

proposed in [18]. The methodology is based on a weighted evaluation method that allows each entity to adapt it to its own needs and priorities.

Figure 2. The 5 Ws and 1 H of CIE policy.

In the context of formulating its policy and procedures, an organisation may consider adopting a cybersecurity sharing framework for managing cyber threat information, which set provisions for information sharing and exchange. There are several frameworks and programs adopted and supported by governments or the industry, like the US Department of Homeland Security's (DHS) Automated Indicator Sharing (AIS) [10] and Cyber Information Sharing and Collaboration Program (CISCP) [9], which is more focused on critical infrastructure sectors, the standardised Cybersecurity Information Exchange Framework (CYBEX) [19], the Malware Information Sharing Platform (MISP—https://www.misp-project.org/), as well as proposed frameworks, like the Cybersecurity Information Exchange with Privacy (CYBEX-P) which also addresses data privacy and classified data management issues [20].

Data sanitisation [21] is one of the solutions that the organisations should consider utilising to ensure that no classified or sensitive information is disclosed to unauthorised entities while sharing CTII with external entities. In this context, the organisation should carefully review and classify the information that their security appliances use and report (e.g., logs from IDS/IPS), and ensure that no information violating the organisation's policy is accidentally shared using CIE tools. For example, the CIE solution should not disclose the internal IP that has been reported by the IDS/IPS as a target of malicious activity.

Sharing of information, if not public, should be driven by agreements among the sharing parties, which address policy restrictions. Such agreements should address, among others, the following issues [22].

- The type of information that satisfies the organisations' business needs.
- The form of information exchange that is going to take place, such as emails, bulletins, documents and automated sharing.
- The confidentiality requirements for the exchanged information and the dissemination restrictions.
- Parties authorised to access, process and use the information.
- Technical standards used for the exchange of information to satisfy the syntactic and semantic interoperability.
- Language issues for cross-border dissemination.
- Communication protocols and access to services.

One of the means that security organisations utilise to control access to the CTII they produce and process is licensing agreements. Several agreement schemes can be adopted by the organisation depending on its business model and needs, as well as the types of services provided to consumers and third parties. Schemes that enforce restricted use include, but are not limited to, commercial licenses, academic or research and personal use. Other licenses may be related to information reuse options and further dissemination of this information to third parties or other communities. For example, several vendors build their own communities, where information collected by their deployed security appliances is properly analysed, evaluated and disseminated back to their customers.

Policies can be technically supported by the use of specific protocols, such as the Traffic Light Protocol (TLP) [23] and the Information Exchange Policy (IEP) [24] that can semantically support the unambiguous transfer of policy rules among sharing parties, as explained in the following section.

3.3. Semantic and Syntactic Interoperability

Semantics are introduced to convey the necessary meaning for syntactically correct messages. Although in the CTII sharing process sources may disseminate unstructured information that hinder data processing (such as information found on social media, news or even on CERT and CSIRTs), several standards have been introduced for properly exchanging CTII among stakeholders. Compliance with standards facilitates the automated sharing of information as well as the ingestion, analysis and integration. Although the existence of a small set of standards facilitates interoperability, adopting a wide variety of competing non-interoperable ones, introduces significant overhead and can eventually hinder the exchange of information. The most prominent standards in the CTII sharing are the following.

- Structured Threat Information Expression (STIX) [25]: This OASIS adopted standard is an information model and serialization solution used to exchange CTII. STIX 2.X is based on JSON, while its predecessors, i.e., STIX 1.X were XML-based. STIX 1.X together with the CybOX 2.X [26], a standardised language for encoding and communicating information about cyber observables, were integrated into STIX 2.0.
- Malware Information Sharing and Threat Intelligence Sharing Platform (MISP) [27]: A data model composed of "events", which usually represent threats or incidents, which in turn are composed of a list of "attributes", such as IP addresses and domain names. Objects in MISP allow combinations of attributes and "galaxies", which enable a deeper analysis and categorisation of events.
- Common Attack Pattern Enumeration and Classification (CAPEC) [28]: It provides a comprehensive dictionary of known patterns of attack employed by adversaries to exploit known weaknesses.
- Malware Attribute Enumeration and Characterization (MAEC) [29]: A structured language used for encoding and sharing information about malware based upon attributes such as behaviours, artefacts and relationships between malware samples. It facilitates correlation, integration, and automation of malware analysis and reduces potential duplication of malware during analysis.
- IETF's Managed Incident Lightweight Exchange (MILE) set of standards: MILE working group develops standards to support computer and network security incident management focusing on the definition of data formats for representing incident and indicator data, and on standardising how application layer protocols such as HTTPS and XMPP are utilised for sharing the structured data. Among the standards adopted by MIME are the following.
  - IODEF v2 (Incident Object Description Exchange Format) [30,31]: It is an XML-based data representation of security incident reports and indicators. It is designed so that to be compatible with IDMEF (Intrusion Detection Message Exchange Format) and capable to include IDMEF message into incident objects.
  - Intrusion Detection Message Exchange Format (IDMEF) [32]: An XML-based data format that is used by intrusion detection and response systems to report alerts about suspicious events.

A comparison of the above CTII sharing technologies with respect to the data format and the types of data they can support is provided in Table 1.

Table 1. CTII technologies characteristics.

Technology	Format	Type of Data
STIX1.x	XML	TTP, Threat Actor, Incident, Exploit Target, Course of Action, Report, Package
STIX2.x	JSON	Attack Pattern, Campaign, Course of Action, Identity, Indicator, Intrusion Set, Malware, Observed Data, Report, Threat Actor, Tool, Vulnerability, Sighting
MISP	JSON	Sighting, Threat Actor, Incident, Network Activity, Antivirus Detection, Hashes, Malware Sample, External Analysis, Traffic Light Protocol
CAPEC	XML	Attack Pattern, Cross Site Forgery, SQL Injection, Buffer Overflow
MAEC	JSON	Malware, Cyber Observable, Entity Association, Incident Management
IODEFv2	XML/HTML	Incident Management, Hashes, Indicator
IDMEF	XML	IP Addresses, Malware, Alert, (Login, Date of Creation, Classification)

The latest version of STIX (2.x) is defined using JSON schemas, thus rendering it easier to parse and expand than its XML-based predecessor (STIX 1.x). CAPEC is supported in both STIX 1.x, as an instance type, and STIX 2.x, as an external reference for the Attack Pattern object type.

STIX and MAEC were designed based on similar use cases. Despite their similarities, they differ in the level of description. MAEC is intended to provide a comprehensive, structured way of capturing detailed information regarding malware samples, whereas STIX is meant to capture a broad spectrum of cyber threat information, including basic information on malware in a more conceptual framework. An analysis of the types of malware-related information that can be captured by MAEC, STIX and MAEC embedded in STIX is presented in [33].

STIX 2.x does not provide support for MAEC content. As a result, STIX 2.x cannot describe a malware in as much detail as STIX 1.x [34]. Nevertheless, STIX 2.1 will include a new object type called malware-analysis which could possibly include MAEC content [35].

Some of the aforementioned standards, such as STIX 2.0 and MISP, also support data markings which can facilitate enforcement of policies regarding the sharing of information to achieve technical and semantic interoperability, such as the following.

- Statements, e.g., copyright, terms of use, applied to the shared content.
- Traffic Light Protocol (TLP v1.0) [23]: It is a set of designations used to ensure that sensitive information is shared with the appropriate audience by providing four options as shown in Figure 3. Although optimised for human readability and person-to-person sharing and not for automated sharing exchanges, TLP can help restrict information sharing only with specific entities or platforms and avoid any further unnecessary or unauthorised dissemination thereof. TLP is supported by STIX 2.0 and MISP protocols.

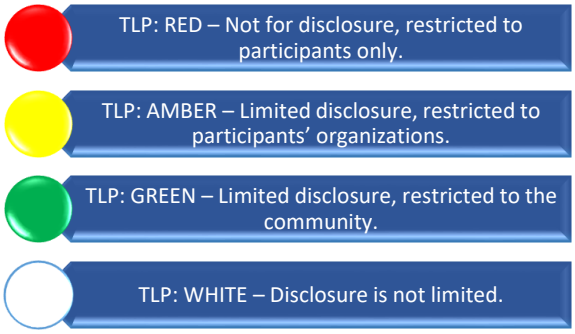

Figure 3. The Traffic Light Protocol (TLP).

TLP, however, cannot support fine-grained policies. Considering this limitation, FIRST developed Information Exchange Policy (IEP) [24], a JSON-based framework that content producers can use to specify restrictions for threat intelligence regarding the following:

- Handling: It can be used to ensure the confidentiality of the information being shared.
- Action: It can be used to define the permitted actions or uses of the information received that can be carried out by a recipient.
- Sharing: It can be used to define any permitted redistribution of information that is received.
- Licensing: It can be used to define any applicable agreements, licenses or terms of use that govern the information being shared.

Some of the aforementioned standards were developed to cover the needs for sharing cyber threat-related information, like incidents, patterns and indicators, and not vulnerabilities, which can be considered as a type of information that has its own needs and sets of standards for dissemination, such as the ones listed below.

- Common Vulnerability Enumeration (CVE): A de facto standard for listing vulnerability information where each CVE entry should include:
  - A CVE ID number,
  - A description of the vulnerability,
  - The anticipated impact in the form of a CVSS (Common Vulnerability Scoring System) value, and
  - References to vulnerabilities reports and/or advisories.
- Common Vulnerability Reporting Framework (CVRF) [36]: An XML-based standard that supports creation, update and interoperable exchange of security advisories as structured information on products, vulnerabilities and the status of impact and remediation among interested parties. According to MITRE, CVRF allows interested parties to download the entire CVE list at once, or download CVEs for a specific year, whereas each CVE can record when it was initially published by the CVE Team, and when it was most recently modified.
- Open Vulnerability and Assessment Language (OVAL): A language that can be used for representing system information, expressing specific machine states, and reporting the results of an assessment.

3.4. Technical Interoperability

Technical interoperability is much related to the implementation of the necessary tools and APIs to support the automated exchange of information, which includes both consumption and delivery, as well as the support of the underlying communication protocols used for conveying CTII information. It typically involves many layers of the TCP/IP stack in formulating and transferring these messages, yet the anticipated approach towards interoperable solutions is the adoption of well-established standardised protocols that can be easily supported by the communicating parties.

Protocols that have been proposed for the transmission of CTII include the following (please note that in this paper we do not consider typical TCP/IP protocols including those found at the application layer).

- Trusted Automated Exchange of Intelligence Information (TAXII) [37]: It is an application layer protocol that is used to exchange CTII, represented in STIX, over HTTPS.
- Resource-Oriented Lightweight Information Exchange (ROLIE) [38]: It defines a REST style approach for security information publication, discovery and sharing. Using ROLIE, security incidents, attack indicators, software vulnerabilities, configuration checklists, and other security automation information are provided as web-addressable resources. IETF's MILE working

group has drafted two extensions to ROLIE for CSIRTs needs [39] and for the exchange of vulnerabilities [40].

Other technical issues that have to be considered during sharing are related to the protection of information against unauthorised disclosure or modification, the mechanisms that are used thereof and the corresponding cryptographic algorithms.

4. The CTII Landscape

The CTII sharing community comprises a number of sources that disseminate information based on their respective strategy and policy. In this section, an analysis of CTII sources with respect to their interoperability characteristics is presented; it focuses on the semantic standards adopted by the parties, as well as the Licensing options and demonstrates the diversified approaches followed by them. Entities that want to join the CTII community or, most importantly, utilise the information provided by its members, have to consider these diversities to appropriately consume from the community or contribute to it.

The authors have studied a large number of threat intelligence sources, including those that provide high-level information, such as CERTs and conducted an analysis on interoperability-related characteristics of 32 of them. These sources are listed in Appendix A, whereas a listing of the characteristics that were used in the analysis is provided in Appendix B. Sources providing only vulnerabilities, such as the NIST NVD, MITRE and CVEDetails, have been excluded from this study, as the focus was more on threat sharing information, i.e., indicators, sightings, hashes, as opposed to vulnerability sharing. In [41] there is an extended list and analysis of those sources that provide vulnerabilities. Moreover, although an analysis of multiple National CERTs could be conducted, the authors have intentionally omitted this as the majority of National CERTs adopt similar, if not the same approaches in information sharing. As a result, only Finland's National CERT is included in the analysis as an example.

The 32 sources that are analysed in this paper include original data providers, data aggregators, intelligence platforms and report providers, the number of which is shown in Figure 4. The analysis demonstrates the diversities in the CTII sharing landscape, where the majority of the sources choose to share information based on generic standards. JSON and CSV were chosen by 50% an 31% of sources, respectively, whereas another 59% of them chose plaintext, as shown in Figure 5. Unfortunately, only a small set of sources have chosen to invest on the set of CTII-oriented standards, like STIX, MISP and OpenIOC.

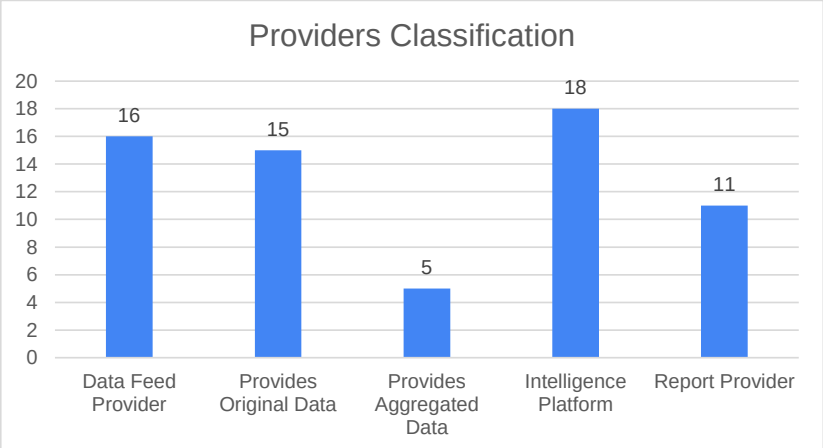

Figure 4. Providers' classification.

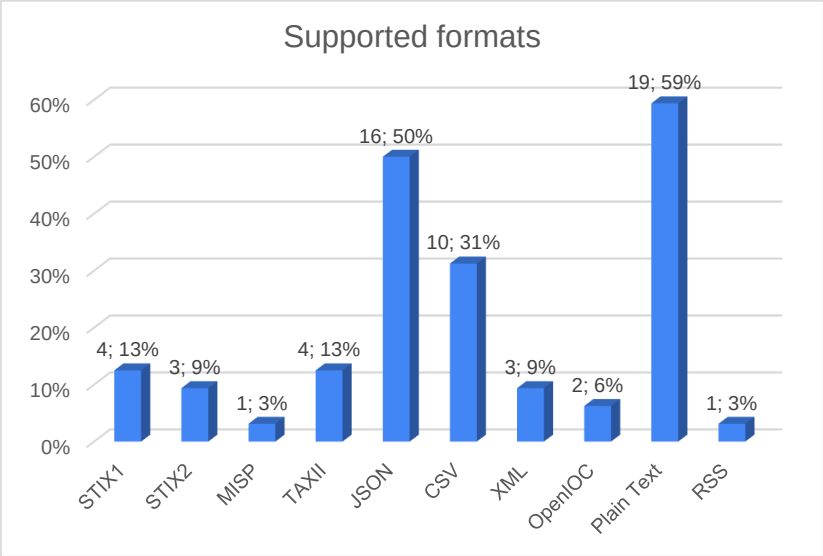

Figure 5. CTII sources supported formats.

The sources that have adopted the various formats are shown in Table 2.

Table 2. Sources supporting semantic/syntactic formats.

Semantic/Syntactic Formats	Sources
STIX1	CRITs, MISP Platform, OTX AlienVault, US CERT AIS
STIX2	Abuse.ch, Anomali STAXX, MISP Platform
MISP	MISP Platform
TAXII	Anomali STAXX, CRITs, OTX AlienVault, US CERT AIS
JSON	Block List Project, CINSscore, Dshield, Fortinet, Google Safebrowsing, Hybrid Analysis, Malc0de, MalShare, MISP Platform, OpenPhish, OTX AlienVault, PhishTank, Proofpoint, Shadowserver, ThreatMiner, VirusTotal
CSV	Abuse.ch, Autoshun, MISP Platform, OpenPhish, OTX AlienVault, PhishTank, Proofpoint, Shadowserver, Spamhaus, ThreatMiner
XML	Fortinet, Hybrid Analysis, PhishTank
OpenIOC	MISP Platform, OTX AlienVault
Plain Text	Abuse.ch, Bambenek, Bitdefender (Advanced Threat Intelligence), Block List Project, BruteForceBlocker, CERT-EU, CINSscore, Comodo Site Inspector, DNS8, Dshield, ESET, Fortinet, Malc0de, MalShare, National CERTs (CERT-FI), OpenPhish, Spamhaus, TalosIntelligence, Trustwave
RSS	CERT-EU

Moreover, the majority of the sources provide information only in one or two formats, as shown in Figure 6, which are generic JSON and CSV formats and not CTII sharing related ones. This generally hinders interoperability.

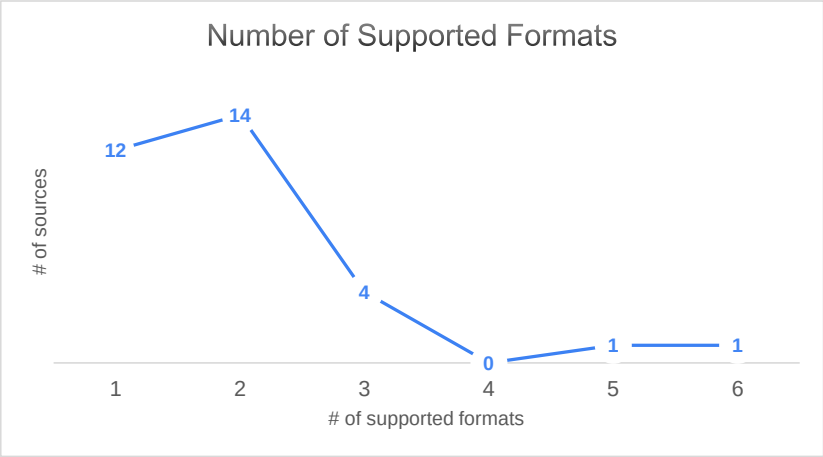

Figure 6. Formats supported by cybersecurity sources.

The analysis has also demonstrated that 15 out of the 32 analysed sources provide open, publicly available information (see Figure 7). Moreover, information reuse for personal use is allowed by 13 of them.

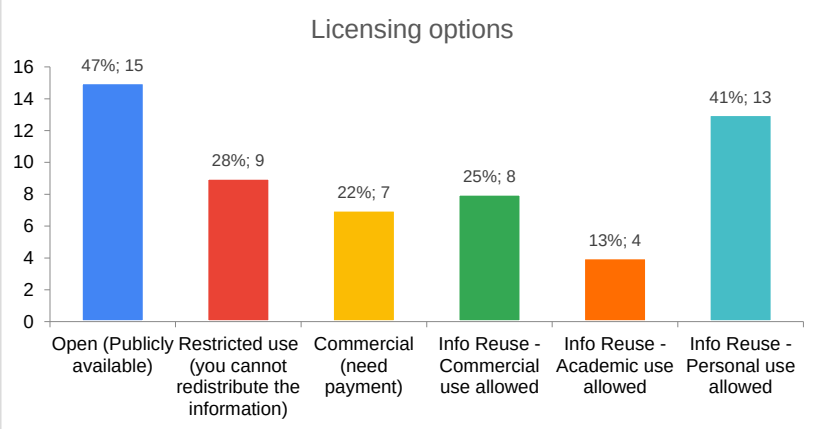

Figure 7. Licensing options for the studied sources.

Another interesting characteristic that the analysis revealed is the number of sources that provide an API, such as REST, that can be used to easily consume the provided information, as shown in Figure 8. These sources are listed in Table 3.

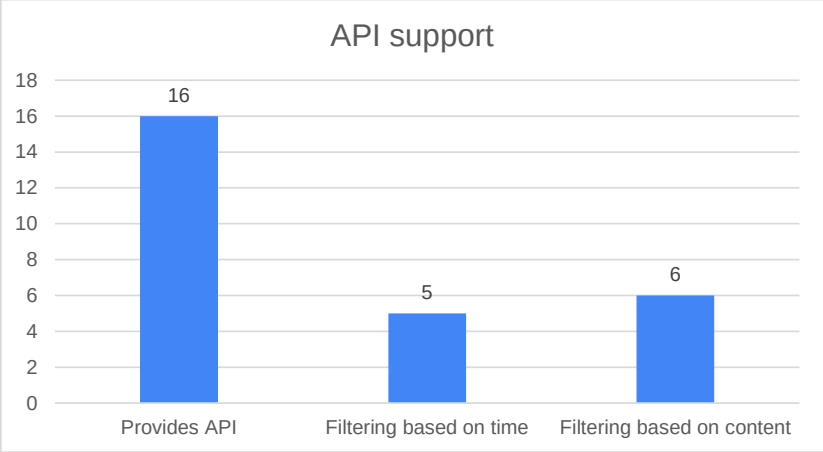

Figure 8. Sources supporting API.

Table 3. Sources supporting API.

API Support	Sources
Provides API	Abuse.ch, Autoshun, Bitdefender (Advanced Threat Intelligence, CRITs, Dshield, ESET, Fortinet, Google Safebrowsing, Hybrid Analysis, Malshare, MISP Platform, OTX AlienVault, PhishTank, Threat Miner, Virus Total

5. Conclusions

The need for sharing information about cyber threats has been established by the cyber security community as a means for timely and effective protection against the aforementioned threats. The main requirements are sharing to be performed using a structured format (so as to be machine readable and facilitate automated actions) and have the ability to provide as much context as possible for a given threat. Nevertheless, interoperability among cyber threat sources is not related only to technical issues. This paper has proposed a layered interoperability model which addresses all those factors that can affect the exchange of cybersecurity information among stakeholders. The research conducted based on this model, has proven that there is significant diversity in the way cyber security information gets shared by the involved sources. This approach directly impacts interoperability both from the perspectives of sharing it and/or consuming it, and also makes it harder for security analysts to effectively extract knowledge.

Future work could involve the development of an interoperability maturity model that will guide stakeholders towards the development of interoperable CTII sharing solutions, or the adaptation of their existing ones. Improving the interoperability of cyber security information sharing will facilitate more effective protection against cyber threats in the future.

Author Contributions: Conceptualisation, K.R.; Data curation, A.S. and A.K.; Methodology, K.R., A.P. and C.I.; Writing—original draft, K.R. and A.P.; Writing—review & editing, C.I. and V.K. All authors have read and agreed to the published version of the manuscript.

Funding: This work has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement Nos. 740723 and 830943).

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A

The appendix provides a listing (in alphabetical order) of the sources that have been analysed in this work.

Table A1. CTII sources.

Source	URL
Abuse.ch	http://abuse.ch/
Anomali STAXX	https://www.anomali.com/community/staxx
Autoshun	https://www.autoshun.org
Bambenek	https://www.bambenekconsulting.com/
Block List Project	https://blocklist.site/
Bitdefender (Advanced Threat Intelligence)	https://www.bitdefender.com/
BruteForceBlocker	http://danger.rulez.sk/index.php/bruteforceblocker/
CERT-EU	https://cert.europa.eu/cert/filteredition/en/CERT-LatestNews.html/
http://cinsscore.com/	http://cinsscore.com/
Collaborative Research Into Threats (CRITs)	https://crits.github.io/
Comodo Site Inspector	http://siteinspector.comodo.com/

DNS8	https://www.layer8.pt/products/dns8/
DShield	https://www.dshield.org/
ESET	https://www.eset.com
Fortinet	https://www.fortinet.com/
Google Safebrowsing	https://safebrowsing.google.com/
Hybrid Analysis	https://www.hybrid-analysis.com/
Malc0de	http://malc0de.com/
Malshare	https://malshare.com/
MISP Platform	https://www.misp-project.org/
National Certs (NCSC-FI example)	https://www.cybersecurityintelligence.com/national-cyber-security-centre-finland-ncsc-fi-1916.html
OpenPhish	https://openphish.com
OTX AlienVault	https://otx.alienvault.com/
PhishTank	https://www.phishtank.com/
Proofpoint	https://www.proofpoint.com/us/daily-ruleset-update-summary
Shadowserver	https://www.shadowserver.org/
Spamhaus	https://www.spamhaus.org/
TalosIntelligence	https://talosintelligence.com
Threat Miner	https://www.threatminer.org/
Trustwave (SpiderLabs Blog)	https://www.trustwave.com
US DHS - Automated Indicator Sharing	https://www.cisa.gov/automated-indicator-sharing-ais
Virus Total	https://www.virustotal.com

Appendix B

This appendix provides a listing of the characteristics that have been used for the analysis that was made for the sources identified in Appendix A, which adopts the approach followed in [18]. The analysis did not consider vulnerability sharing sources. It rather focused on the resources that share indicators and sightings.

Table A2. CTII sources characteristics.

Generic Properties	Explanation
1. Type of Data	A CTII source may manage and share different types of data, ranging from vulnerabilities and exploits to sightings and courses of actions
1.1 Indicators	An indicator is a collection of cyber security relevant information containing patterns that can be used to detect suspicious or malicious cyber activity.
1.2 Sightings	A sighting is an observation that someone has shared with the community, without adding additional intelligence to it.
1.3 Courses of Action	A course of action is information concerning how to prevent or mitigate an event shared by, e.g., an indicator.
1.4 Vulnerabilities	A vulnerability is a weakness in a hardware or software appliance that could be exploited to breach the appliance.
2 Provider Classification	An indicator to assess the type of sharing an information sharing provider does, and how much original information can be expected from this provider.
2.1 Data Feed Provider	A data feed provider is the entity that produces cyber security information, or shares received information with minimal or no additional intelligence added to it.
2.1.1 Provides Original Data	A data feed provider that is the original provider of this information, which has been shared in one form or another by the source of, e.g., an incident.

2.1.2 Provides Aggregated Data	A feed provider that does not provide information originally shared with this provider, but shares information aggregated from other feeds.
2.2 Intelligence Platform	A provider that adds additional intelligence/ analysis to the information that is shared with the provider in one form or another.
2.3 Report Provider	A provider that provides, e.g., statistical information in form of reports rather than data feeds.
3 Licencing Options	An indicator to assess the licencing of data use and/or access to the data source API.
3.1 Open (Publicly available)	The data is freely available to collect and use.
3.2 Restricted use	Some restrictions apply as to how the data can be used (e.g. academic or commercial context).
3.3 Commercial	The data provider has commercial interest and provides the data for a fee.
3.4 Information Reuse	Specifies how the data provided by a data source can be reused. Options include commercial, academic or personal use.
3.4.1 Commercial use allowed	The data can be reused in a commercial context, offering services and collecting fees based on this data is allowed.
3.4.2 Academic use allowed	The data can be used in an academic context without restrictions, restrictions apply in other contexts.
3.4.3 Personal use allowed	The data can be used for personal use without restrictions, restrictions apply for other contexts.
4 Interoperability/ Standards	An indicator to assess the interoperability of a tool provider with state-of-the art cybersecurity threat exchange standards and relevant tools/libraries.
4.1 STIX1	Supports the STIX1 threat expression standard.
4.2 STIX2	Supports the STIX2 threat expression standard.
4.3 MISP	Supports the MISP threat exchange protocol.
4.4 TAXII	Supports the TAXII threat exchange protocol standard.
4.5 JSON	Supports the JSON protocol data exchange format.
4.6 CSV	Supports the CSV data expression standard.
4.7 XML	Supports the XML-based threat exchange.
4.8 OpenIOC	Supports the OpenIOC cybersecurity artefact description standard.
4.9 Plain Text	Supports plain text data expression.
4.10 RSS	Supports the RSS feed standard.
5 Advanced API	Indicates whether a data source supports or enables relevant advanced API features.
5.1 Supports API	The source implements an API, such as REST.
5.2 Filtering based on time	Supports filtering based on time for data access. Relevant for data collection to only collect entries since last access.
5.3 Filtering based on content	Supports filtering based on content. Relevant for context specific data collection.

References

1. Bissell, K.; LaSalle, R.M.; Dal Cin, P. The cost of cybercrime—Ninth annual cost of cybercrime study. Technical report, Accenture, 2019. Independently conducted by Ponemon Institute LLC and jointly developed by Accenture. Available online: https://www.accenture.com/_acnmedia/pdf-96/accenture-2019-cost-of-cybercrime-study-final.pdf (accessed on 3 March 2020).
2. Kellermann, T.; Young, B. Modern Bank Heists: The Bank Robbery Shifts to Cyberspace. Technical report, Carbon Black, OPTIV, 2019. Available online: https://www.carbonblack.com/resources/threat-research/modern-bank-heists-the-bank-robbery-shifts-to-cyberspace/ (accessed on 3 March 2020).

3. Johnson, C.S.; Badger, M.L.; Waltermire, D.A.; Snyder, J.; Skorupka, C. *Guide to Cyber Threat Information Sharing*; Special Publication (SP) 800-150; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2016. [CrossRef]
4. Brown, R.; Lee, R.M. The Evolution of Cyber Threat Intelligence (CTI): 2019 SANS CTI Survey, 2019. SANS Institute. Available online: https://www.sans.org/reading-room/whitepapers/threats/paper/38790 (accessed on 3 March 2020).
5. Tounsi, W.; Rais, H. A survey on technical threat intelligence in the age of sophisticated cyber attacks. *Comput. Secur.* **2018**, *72*, 212–233. [CrossRef]
6. European Parliament and Council. Directive (EU) 2016/1148 concerning measures for a high common level of security of network and information systems across the Union. Official Journal of the European Union, 2016. Available online: https://eur-lex.europa.eu/eli/dir/2016/1148/oj (accessed on 3 March 2020).
7. European Telecommunications Standards Institute (ETSI). *CYBER; Implementation of the Network and Information Security (NIS) Directive*; TR 103 456; European Telecommunications Standards Institute: Sophia Antipolis, France, 2017.
8. ENISA. CSIRTs in Europe. Available online: https://www.enisa.europa.eu/topics/csirts-in-europe/csirt-inventory (accessed on 3 March 2020).
9. US Department of Homeland Security. Cyber Information Sharing and Collaboration Program (CISCP). Available online: https://www.cisa.gov/ciscp (accessed on 3 March 2020).
10. US Department of Homeland Security. Automated Indicator Sharing (AIS). Available online: https://www.cisa.gov/automated-indicator-sharing-ais (accessed on 3 March 2020).
11. ODNI. A Guide to Cyber Attribution. Office of the Director of National Intelligence, 2018. Available online: https://www.dni.gov/files/CTIIC/documents/ODNI_A_Guide_to_Cyber_Attribution.pdf (accessed on 3 March 2020).
12. Chismon, D.; Ruks, M. Threat Intelligence: Collecting, Analysing, Evaluating. MWR InfoSecurity, Whitepaper, 2015. Available online: https://www.foo.be/docs/informations-sharing/Threat-Intelligence-Whitepaper.pdf (accessed on 3 March 2020).
13. ENISA. *Actionable Information for Security Incident Response*; Technical Report; European Union Agency for Network and Information Security: Helion, Greece, 2014. Available online: https://www.enisa.europa.eu/publications/actionable-information-for-security/at_download/fullReport (accessed on 3 March 2020).
14. Burger, E.W.; Goodman, M.D.; Kampanakis, P.; Zhu, K.A. Taxonomy Model for Cyber Threat Intelligence Information Exchange Technologies. In Proceedings of the 2014 ACM Workshop on Information Sharing & Collaborative Security, Vienna, Austria, 23–25 October 2014; pp. 51–60. [CrossRef]
15. Skopik, F.; Settanni, G.; Fiedler, R. A problem shared is a problem halved: A survey on the dimensions of collective cyber defense through security information sharing. *Comput. Secur.* **2016**, *60*, 154–176. [CrossRef]
16. European Parliament and Council. Regulation (EU) 2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Off. J. Eur. Union L* **2016**, *119*, 1–88. Available online: http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32016R0679 (accessed on 3 March 2020).
17. ISO Central Secretary. *ISO/IEC 27010:2015: Information Technology—Security Techniques—Information Security Management for Inter-Sector and Inter-Organizational Communications*; Standard; International Organization for Standardization: Geneva, Switzerland, 2015.
18. Schaberreiter, T.; Kupfersberger, V.; Rantos, K.; Spyros, A.; Papanikolaou, A.; Ilioudis, C.; Quirchmayr, G. A Quantitative Evaluation of Trust in the Quality of Cyber Threat Intelligence Sources. In Proceedings of the 14th International Conference on Availability, Reliability and Security, Canterbury, UK, 26–29 August 2019; pp. 83:1–83:10. [CrossRef]
19. Rutkowski, A.; Kadobayashi, Y.; Furey, I.; Rajnovic, D.; Martin, R.; Takahashi, T.; Schultz, C.; Reid, G.; Schudel, G.; Hird, M.; Adegbite, S. CYBEX – The Cybersecurity Information Exchange Framework (X. 1500). *Comput. Commun. Rev.* **2010**, *40*, 59–64. [CrossRef]
20. Sadique, F.; Bakhshaliyev, K.; Springer, J.; Sengupta, S. A System Architecture of Cybersecurity Information Exchange with Privacy (CYBEX-P). In Proceedings of the 2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 7–9 January 2019; pp. 493–498. [CrossRef]

21. Bishop, M.; Bhumiratana, B.; Crawford, R.; Levitt, K. How to sanitize data? In Proceedings of the 13th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, Modena, Italy, 14–16 June 2004; pp. 217–222. [CrossRef]
22. ISAO Standards Organization. *ISAO 300-2: Automated Cyber Threat Intelligence Sharing*; Standard; ISAO Standards Organization: San Antonio, TX, USA, 2019.
23. FIRST. Traffic Light Protocol (TLP), FIRST Standards Definitions and Usage Guidance—Version 1.0. Forum of Incident Response and Security Teams (FIRST). Available online: https://www.first.org/tlp/docs/tlp-v1.pdf (accessed on 3 March 2020).
24. FIRST. Information Exchange Policy Framework, Version 1.0. Forum of Incident Response and Security Teams (FIRST). Available online: https://www.first.org/iep/ (accessed on 3 March 2020).
25. OASIS. Structured Threat Information Expression (STIX). Available online: https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=cti (accessed on 3 March 2020).
26. MITRE. Cyber Observable eXpression. A Structured Language for Cyber Observables (CybOX). Available online: https://cybox.mitre.org/about/ (accessed on 3 March 2020).
27. Malware Information Sharing Platform (MISP). Available online: https://www.misp-project.org/ (accessed on 3 March 2020).
28. MITRE. Common Attack Pattern Enumeration and Classification (CAPEC). Available online: https://capec.mitre.org/index.html (accessed on 3 March 2020).
29. Malware Attribute Enumeration and Characterization (MAEC). Available online: https://maecproject.github.io/ (accessed on 3 March 2020).
30. Danyliw, R. The Incident Object Description Exchange Format Version 2. RFC 7970, 2016. Available online: https://tools.ietf.org/html/rfc7970 (accessed on 3 March 2020).
31. Kampanakis, P.; Suzuki, M. Incident Object Description Exchange Format Usage Guidance. RFC 8274, 2017. Available online: https://tools.ietf.org/html/rfc8274 (accessed on 3 March 2020).
32. Debar, H.; Curry, D.; Feinstein, B. The Intrusion Detection Message Exchange Format (IDMEF). RFC 4765, 2007. Available online: https://tools.ietf.org/html/rfc4765 (accessed on 3 March 2020).
33. Characterizing Malware with MAEC and STIX, Version 1.0, 2014. Available online: https://stixproject.github.io/about/Characterizing_Malware_MAEC_and_STIX_v1.0.pdf (accessed on 3 March 2020).
34. OASIS. *stix2-elevator Documentation*; Release 1.0.0; OASIS Open: Burlington, MA, USA, 2019
35. OASIS. stix2-elevator – Mappings from STIX 1.x to STIX 2.x. Available online: https://stix2-elevator.readthedocs.io/en/latest/stix-mappings.html (accessed on 3 March 2020).
36. Common Vulnerability Reporting Framework (CVRF). Version 1.2. OASIS Common Security Advisory Framework (CSAF). Available online: http://docs.oasis-open.org/csaf/csaf-cvrf/v1.2/csaf-cvrf-v1.2.html/ (accessed on 3 March 2020).
37. OASIS. TAXII^{TM} Version 2.0, Working Draft 02. OASIS standard, 2017. Available online: https://www.oasis-open.org/committees/cti/ (accessed on 3 March 2020).
38. Field, J.; Banghart, S.; Waltermire, D. Resource-Oriented Lightweight Information Exchange (ROLIE). RFC 8322, 2018. Available online: https://tools.ietf.org/html/rfc8322 (accessed on 3 March 2020).
39. Banghart, S.; Field, J. Definition of ROLIE CSIRT Extension. Internet-Draft draft-ietf-mile-rolie-csirt-05, IETF Secretariat, 2019. Available online: http://www.ietf.org/internet-drafts/draft-ietf-mile-rolie-csirt-05.txt (accessed on 3 March 2020).
40. Banghart, S. Definition of the ROLIE Vulnerability Extension. Internet-Draft draft-ietf-mile-rolie-vuln-02, IETF Secretariat, 2019. Available online: http://www.ietf.org/internet-drafts/draft-ietf-mile-rolie-vuln-02.txt (accessed on 3 March 2020).
41. ENISA. State of Vulnerabilities 2018/2019—Analysis of Events in the Life of Vulnerabilities. Technical report, European Union Agency for Network and Information Security, 2019. Available online: https://www.enisa.europa.eu/publications/technical-reports-on-cybersecurity-situation-the-state-of-cyber-security-vulnerabilities/ (accessed on 3 March 2020).