ST. PÖLTEN UNIVERSITY
OF APPLIED SCIENCES

Informatik
& Security

fh
st.pölten

Collaborative Sharing of Information among CTI Platforms

A Case Study using Taranis AI and MISP

Master thesis

for attainment of the academic degree of

Master of Science in Engineering (MSc.)

submitted by

Bc. Peter Leitmann

52304915

in the

University Course Cyber Security and Resilience at St. Pölten University of Applied Sciences

Supervision

Advisor: FH-Prof. Dipl.-Ing. Dr. Peter Kieseberg

Assistance: Dr. Dr. Florian Skopik

Assistance: BSc. Benjamin Akhras

Declaration

Title: Collaborative Sharing of Information among CTI Platforms

Type of thesis: Master thesis

Author: Bc. Peter Leitmann

Student number: 52304915

I hereby affirm that

- I have written this thesis independently, have not used any sources or aids other than those indicated, and have not made use of any unauthorized assistance.
- I have not previously submitted this thesis topic to an assessor for evaluation or in any form as an examination paper, either in Austria or abroad.
- this thesis corresponds with the thesis assessed by the assessor.

I hereby declare that

☒ I have used a Large Language Model (LLM) to proofread the thesis.

☐ I have used a Large Language Model (LLM) to generate portions of the content of the thesis. I affirm that I have cited each generated sentence/paragraph with the original source. The LLM used is indicated by a footnote at the appropriate place.

☐ no Large Language Model (LLM) has been used for this work.

Acknowledgment

I would like to express my sincere thanks to FH-Prof. Dr. Dipl.-Ing. Peter Kieseberg, Dr. Dr. Florian Skopik, and BSc. Benjamin Akhras for their insightful remarks, expertise, patience, and guidance during the writing of this thesis. I would also like to convey my heartfelt gratitude to my family and close friends for their unwavering support throughout this journey. Last but not least, my gratitude also goes to colleagues, who I had the pleasure to work with, in the course of my thesis at the Austrian Institute of Technology, Vienna.

Abstract

This thesis focuses on the problematics of information sharing and collaboration mainly within the area of strategic Cyber Threat Intelligence (CTI). This thesis delivers a case study of two open-source CTI platforms – the Threat Information Sharing Platform (MISP) and Taranis AI. The research focused on what is the viability of using a platform highly specialized on technical CTI, like MISP, for sharing strategic CTI. In the course of study it was analyzed the base use cases to achieve collaboration and, accordingly, a software design was created. Subsequently, the design was proved by the implementation of a Proof of Concept solution integrated in the Taranis AI platform. The solution was also validated in a simulated scenario with users.

The results of the research claim that even though MISP is a platform for mainly technical CTI, it has enough flexibility to support collaboration with strategic CTI. On the other hand, the research in this domain is not complete, and this thesis points out various problematic areas that need more exploration and research. For instance, conflict resolution strategies that support the workflows of CTI analysts, and what are the aspects of platforms that are missing, in order to better support sharing of strategic CTI. While the scope of this thesis was in the practical part narrowed down to the specific platforms, many of the concepts tackled are applicable in general to the whole domain of collaborative CTI environments.

Contents

1 Introduction **1**

1.1 Motivation 1

1.2 Problem Statement 2

1.3 Objectives and Scope of the Thesis 3

1.4 Research Methodology 4

1.5 Structure of the Thesis 4

2 Background and Related Work **7**

2.1 Foundations of Cyber Threat Intelligence and OSINT 8

2.1.1 Overview of CTI and OSINT platforms 10

2.1.2 Role of MISP in Threat Intelligence Sharing 10

2.1.3 Key Challenges in Collaborative Threat Intelligence sharing 11

2.1.4 Trust Establishment and Federated Collaboration 12

2.1.5 Information Sharing from Legislative Perspective 13

2.2 Case Study Platforms Description 13

2.2.1 MISP in Detail 13

2.2.2 Taranis AI in Detail 15

3 Use Case and Process Analysis **17**

3.1 Use Case Analysis 17

3.1.1 Use Case 1: CTI Analyst Shares a Story 17

3.1.2 Use Case 2: CTI Analyst Ingests a Story 18

3.1.3 Use Case 3: CTI Analyst Updates an Owned Shared Story 18

3.1.4 Use Case 4: Propose Changes to Story Metadata 19

3.1.5 Use Case 5: Propose Changes to the Story's News Items 19

3.1.6 Use Case 6: Transitively Share Information from One Collaboration Group to Another 19

3.1.7 Use Case 7: Connect to a specific collaboration group . . . 20
3.2 System Interactions Analysis . . . 20
3.2.1 Share Information with a Partner . . . 20
3.2.2 Contribute Back to the Originator . . . 21

4 Design and Architecture . . . 23
4.1 Requirements Analysis . . . 23
4.1.1 Functional Requirements . . . 23
4.1.2 Non-Functional Requirements . . . 24
4.2 High-Level Architecture Overview . . . 25
4.2.1 Hub-and-Spoke as a Federated Backbone . . . 25
4.2.2 Repository Pattern and Shared Knowledge Base . . . 26
4.2.3 Architectural Decisions and Trade-Offs . . . 26
4.3 Subsystems Overview . . . 27
4.4 Data Flow and Interactions Scenarios . . . 28
4.5 Nonfunctional Properties . . . 29
4.5.1 Interoperability . . . 29
4.5.2 Identity and Access Management . . . 30
4.6 Security Mechanisms . . . 30

5 Mapping Information Sharing Models to MISP . . . 31
5.1 Mapping Taranis AI Data Model to MISP . . . 31
5.1.1 Information Carrying Properties . . . 31
5.1.2 Logical Mapping of Taranis AI data to MISP . . . 33
5.1.3 MISP Attributes . . . 33
5.2 Functional Mapping of Taranis AI to MISP . . . 33
5.2.1 MISP Events . . . 34
5.2.2 MISP Objects and Attributes . . . 34
5.2.3 Missing link – Extended events . . . 34

6 Implementation . . . 35
6.1 Implementation Overview . . . 35
6.1.1 Bidirectional Communication . . . 35

6.1.2 Conflict Handling System . . . 38

7 Evaluation and Discussion . . . 47

7.1 Evaluation Session Context . . . 47

7.2 Results . . . 48

7.2.1 User Evaluation Results . . . 48

7.2.2 Research Results . . . 49

7.3 Limitations and Discussion of Current Work . . . 51

7.3.1 TLP Protocol . . . 51

7.3.2 Conflict Resolution . . . 51

7.3.3 MISP Integration Degree . . . 52

7.3.4 Sharing Circles Limitations . . . 52

7.3.5 Limited Collaboration Ability . . . 52

7.3.6 Future work . . . 52

8 Conclusion . . . 55

List of Figures . . . 58

Acronyms . . . 61

Bibliography . . . 64

A Appendix . . . 69

A.1 MISP Object Mapping Taranis AI news item . . . 69

A.2 MISP Object Mapping Taranis AI story . . . 71

A.3 Assignment for Users for Practical Evaluation . . . 73

A.4 Evaluation User Session with Users – Plots . . . 81

A.5 List of References to the PoC Implementation . . . 82

1. Introduction

1.1. Motivation

With the rise of the World Wide Web (WWW) and the internet in recent decades, the world's malicious actors have developed comparatively and adjusted to the most sophisticated protection mechanisms. Therefore, to keep up with the threat landscape, Cyber Threat Intelligence (CTI) played a huge role [1].

CTI is usually referred to as evidence-based knowledge concerning the threat landscape, taking advantage of artifacts collected during or after a cyberattack. This information is crucial for informed decisions for defence [2].

While CTI has been an established way to fight cyber threats for the past years, the continuous improvement and increasing complexity of single attackers and Advanced Persistent Threats (APTs) create a desire for solutions that can support CTI further [1].

Since a huge amount of information has become freely available, new ways of collecting it have emerged. The field of open information is called Open-source Information (OSINF). OSINF is the cornerstone and prerequisite of Open-source Intelligence (Open-source Intelligence (OSINT)). OSINF encompasses all the data that is freely available, not only tied to information online but also to information available in newspapers, governmental reports, maps, etc. [3].

When OSINF is collected and is used for intelligence purposes, then the whole process becomes OSINT. It enables more possibilities and reach of CTI to create a much better and cohesive understanding of the vicious threat situation. Although OSINT raises ethical questions, it could be handling personally sensitive information or audio and video recordings, and they may have specific copyright rules. Undeniably, for CTI, OSINT can be an invaluable source of information to increase awareness on a global scale of the current situation at the cybersecurity forefront. As the CTI, in general, is performed for non-malicious reasons, the usage of OSINT is more justified, however, it remains a problematic topic [3], [4].

Modern CTI heavily relies on two distinguished information categories – Tactics, Techniques, and Procedures (TTPs) and Indicators of Compromise (IoCs). IoCs are usually forensically gained information

gathered from devices compromised by attackers, like IP addresses, file hashes, etc. They highly contribute to the speed of understanding what happened during a crime act. The unfavourable character of IoCs is, they may have a limited lifespan because the footprints can be changed during an attack with reasonable effort, masquerading what truly happened. Therefore, the collections of IoCs shared today tend to consist of the simple ones, leaving the crucial ones unseen [5], [6].

TTPs are, in a broader sense, behaviours and manoeuvres attackers make and roughly describe how they think as well as their goals through unified names. This knowledge provides a lot of insight and room to identify attackers and APTs promptly and help act accordingly to counter current or possible future attacks. Even though it is much better to rely on TTPs than IoCs, it is difficult to gain such because they usually must be derived from unstructured data like penetration test reports which is complicated, and domain expertise is needed [7], [8].

As stated before, information about the attacker's behaviour is important, and because the attackers often do not target just a single victim, the exchange of this information is also very relevant and can help many. For this reason, the necessity to have established ways to communicate between OSINT platforms, that accumulate this information, can become critical in the future.

One of the prerequisites of such a mechanism is timeliness. Without an appropriate exchange of information in time, the shared insights are becoming less relevant, which is desired to avoid.

Another aspect to consider is sharing clusters of information ("stories") that would aggregate more sources of information on a topic and enable analysts to work more efficiently in a certain context. Working with contextualized information can be invaluable, however, creates higher complexity towards the sharing mechanism, as clustered information needs to be supported by the sending and receiving parties. Moreover, it creates another layer of complexity in handling classified parts of information that should not be disclosed and shared with other parties.

1.2. Problem Statement

Sharing information among OSINT platforms in a one-to-multiple setting is a non-trivial task. One of the cornerstones of such a mechanism is to understand what solutions are efficient for this given task.

Today's world knows numerous solutions that offer sharing of CTI information (e.g. OpenCTI, and Threat Intelligence Sharing Platform (MISP)), but there has always been a lack of platforms that focus on sharing of the strategic CTI [9]. Strategic CTI comes with complications, because of the unstructured character of the information.

When information is received, it must also be interpreted in an understandable way, which would allow no room for speculation on delivered information. Furthermore, sharing information has various implications. What some organizations, consider valuable, some organizations might find it distracting and irrelevant. As a result of that, effective mechanisms that allow the distribution of strategic CTI can be considered crucial. Another aspect to consider is collaboration. According to researches, it can produce higher quality results, as the process incorporates more entities [10].

The aim of this thesis is to explore the viability of usage of a CTI platform like MISP for sharing and collaboration with the strategic information.

Sharing of CTI information is one aspect, but to reach the point that information is shared, information needs to be collected and understood. One of the platforms which allow the collection and analysis of strategic CTI is the OSINT platform Taranis AI[1].

Due to that, the goal of this thesis is to understand the applicability of integration of these two platforms, and explore limitations that stem from using a platform for technical CTI to share and collaborate with strategic CTI.

1.3. Objectives and Scope of the Thesis

The objectives of this case study are to review the current state of collaboration in the information domain and CTI. For brevity of this thesis and the complexity that the topic, the scope of this thesis is narrowed to a case study of the two open-source CTI projects MISP and Taranis AI. Building upon that, the case study should provide an analysis of what collaboration requirements are needed. Furthermore, a suitable solution focusing on the main requirements identified should be designed and practically implemented.

The results should contain answered research questions and a validation of the practical output should be carried out.

This thesis recognizes three main research questions:

- RQ1 – To what extent does MISP (as a central node) and Taranis AI enable an interoperable collaboration on strategic CTI. Can MISP platform provide enough flexibility to its standard technical CTI focus?
- RQ2 – What are the challenges of resolving conflicts, and how can MISP and Taranis AI platform support a more smooth operation?
- RQ3 – What are the key characteristics that can significantly improve experience of CTI analysts

[1] https://taranis.ai

during collaboration?

1.4. Research Methodology

This thesis uses the design-science approach that focuses on designing a system for information sharing within the domain of CTI focusing on strategic oriented intelligence. Initially a review of the challenges and current state was carried out. Accordingly, using the open-source CTI platforms MISP and Taranis AI[2] were used for the implementation of the Proof of Concept (PoC) to demonstrate what can collaboration with CTI look like.

Afterwards, design and architecture was delimited and a requirements analysis was carried out on what are the characteristics that CTI collaborative environments require. Even though many of the concepts discussed are applicable technology-agnostic, the scope of the thesis is to elaborate on effective collaboration of multiple organizations in one sharing circle. According to the requirements a system design was described. Due to the scope of this thesis, it already considers the aforementioned candidates for the practical implementation.

The PoC was realized using the results of the architectural analysis making the software (SW) meet the identified requirements for CTI collaboration.

The implementation was afterwards evaluated with two real-world simulations with users, consisting mainly of scientists, system engineers. The evaluation was framed to a realistic like scenario helping the participants understand the workflows in the context of the domain. However, the main goal of the evaluation sessions was testing the capabilities of the designed and implemented software from a functional point of view and not the capabilities of the participants. Therefore, only results from the software perspective were discussed.

1.5. Structure of the Thesis

The structure of the thesis is as follows. In the chapter 1, the introduction of the thesis, definition of the scope and explaining the motivation of the thesis is set. It also includes a description of the methodology and approach used in the course of this thesis.

Chapter 2 focuses on the definition of important background knowledge in the domain of information sharing and CTI. Furthermore, it introduces the main open-source tooling used for the PoC implementation which supported the theoretical results of the research. It also sheds light on the challenges that encompass

[2]https://taranis.ai

the domain of information sharing, and the significant parameters of collaborative environments, like trust and willingness to share information.

Chapter 3 documents the use cases that are in the scope of this thesis supported with diagrams. Chapter 4 describes the precise requirements for a collaborative environment and introduces the design, architecture, and processes for the PoC implementation. Furthermore, chapter describes the mapping of data between the two open-source SW MISP and Taranis AI.

In the chapter 6, the realization of the PoC is described with the emphasis on the user perspective. The evaluation and validation of the practical output is described in chapter 7, and conclusion of the whole thesis is in chapter 8.

2. Background and Related Work

Sharing of Threat Intelligence (TI) can provide additional value. Either in terms of understanding the threat landscape more efficiently, explorations of outages of systems interconnected and foresee coming issues. Ultimately, sharing of TI can result in not only financial savings but can contribute to a better situational awareness or security maturity [11].

Another aspect to consider is complying to legal requirements [11], [12]. The successor of Network and Information Security Directive (NIS), Network and Information Security Directive 2 (NIS2), is requiring even richer information sharing and extends the scope of organizations that need to either report security incidents to authorities, or describes additional voluntary sharing of relevant information [12].

From another perspective, in the last decades the development of cyber threats has progressed immensely. Therefore, the natural evolution of threat defenders has achieved noticeable progress as well [13].

The look into the real operations of TI analysts has uncovered the true nature and aspects of the work of TI teams. TI teams work towards understanding of the cyber threat landscape and ingesting the Threat Information (TINF) and translating it into actionable tasks, TI [14]. This process resulting in TI and Situational Awareness (SA) was, historically, mostly technology oriented. That, however, has been deemed insufficient for the collaborative nature, as it was heavily directed to the objective information interpretation side of the information, rather than the strength of interpretation of information of individuals working on the assessments. As a result of that, incorporating cognitive task analysis indicated a stronger human-supported system as it helped to convey and store the threat situation from the certain analyst's perspective.

In the past, the nature of the job of TI analysts was perceived as individualistic. That, quite on the contrary, was proven not to be so conclusive.

Ahrend et. al. conducted a study (survey) conducted on the TI team members and the team strongly indicated the opposite, and that the true strength lies in the collaborative nature of the TI team and the so-called Threat and Defence Knowledge (TDK). In case of TDK, it was highlighted to be a difficult task to efficiently share it among other members of a single team, let alone the intentional dissemination of information among interested stakeholders of the organizations.

The TDK, according to the study results, was often shared during informal occasions, therefore not following or promoting a systematic approach which would support the (re-)usability of the information hastily and not making the information open to a broader audience. This type of collaboration was preferred, as the informal exchange of TDK was perceived as simple and good enough.

The need for a new solution that would encourage an easier, more structured way to track and share the TDK is real, and therefore circumventing situations when valuable information is passed along but parties of interest stay uninformed [15]. The fact that in the today's world is the threat landscape so dynamic, supportive mechanisms to share effectively is needed where one of the key attributes is also sharing anonymously [16].

2.1. Foundations of Cyber Threat Intelligence and OSINT

CTI is an evidence-based knowledge, associated with a specific threat and can stem from aggregation, transformation, analysis, interpretation, and enrichment of the context [17], [18].

It is important to recognize that CTI goes beyond mere TINF, providing deeper insights and context for effective cybersecurity decision-making. With TINF, it is understood that it can be any tangible artifact that is presented just as raw data of an event.

Commonly that are [17]:

- Indicators – technical indicators and traces that suggest attack (e.g. Intrusion Prevention Systems (IPS), file hashes, etc.).
- TTPs – high-level descriptions of behaviours of attackers.
- Security alerts – usually human-readable notes about vulnerabilities and exploits.
- Threat intelligence reports – can be extensive documents about APTs, including the TTPs and IoCs. However, e.g. Structured Threat Information eXpression (STIX) recognizes threat intelligence reports too, and in that format they are also computer readable reports.
- Tool configurations – best practices for processing threat information.

Depending on the intent, the sources for CTI can vary. In general, the information can come from two distinguished environments – internal and external. In the case of internal sources, they commonly consist of outputs of security systems like firewalls, Intrusion Detection Systems (IDS) or IPS and offer mainly structured data to work with [2].

When it comes to the information coming from external parties, it can include similar (structured) information, but can also offer more. The biggest difference is the number of the contributors that share knowledge that could be potentially leveraged commercial sources – threat feeds, Software as a Service (SaaS). Fur-

thermore, in today's technology society, the information shared everywhere on the Internet, e.g. using social media or forums, offers a vast room for the exploration of unstructured data. Intelligence, which comes from open sources, as previously stated, is called OSINT (includes also open-source threat feeds, Dshield and ZeuS Tracker) [2].

Given the processing, good CTI is more specific and actionable. Whereas the source information could be referred to as OSINF. On the other hand, it also introduces some difficulties. The creation of intelligence out of the information can be a peculiar and tedious process. For (high) quality results, a number of aspects need to be considered: time, context, use case, stakeholders, confidentiality, etc [2], [19].

Generally, there are several types of CTI [2].

- Strategic – TI reports, briefings or conversations.
- Tactical – also referred to as TTPs, explained above
- Technical – Technical Threat Intelligence (TTI) is the intelligence that feeds detection and similar systems.
- Operational – intelligence that reaches security managers allowing them to plan and prepare for potential threats.

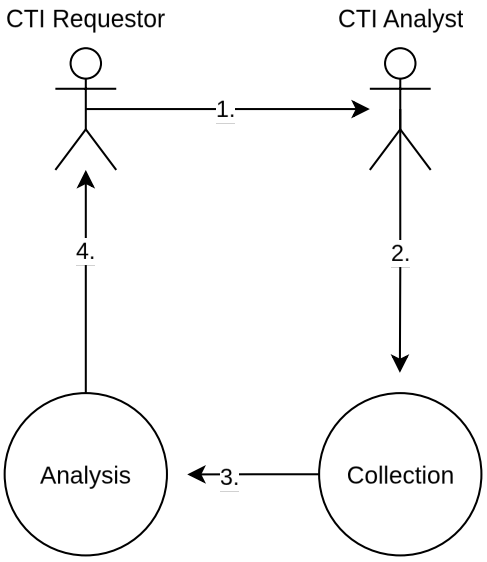


Figure 2.1.: Cyber Threat Intelligence Lifecycle

Overall, CTI is something that requires triage, analysis, and processing of information. It can also be described in the form of a lifecycle. It starts with a request for information. CTI analysts carry out the data collection, analysis, and deliver the intelligence product to the requestor (see fig. 2.1) [20].

2.1.1. Overview of CTI and OSINT platforms

As stated before, intelligence is a tangible information digest processed by an entity. The entity can be a human being or a (computer) system. In this regard, the high potential of advantages gained by the implementation of diverse text processing mechanisms (e.g. Large Language Models (LLM)) is recognized [21]–[23].

According to the review [22] incorporating several studies and evaluating the existing options of CTI platforms, a high number of CTI platforms available was identified.

The platforms come under various licensing standards. There are commercial solutions that are closed source as well as open-source.

A list of the most popular closed source CTI platforms:

- Anomaly STAXX
- Soltra Edge
- ThreatQ
- ThreatConnect
- IBM X-Force

A list of the most popular open source CTI platforms:

- MISP
- CRITs
- OpenCTI
- CIF

The platforms generally support the creation of intelligence and the subsequent access to the processed intelligence. The way how the platforms do it varies, however, the most sighted features resulted in the ability to import or export data in at least one of the popular data format (e.g. Extensible Markup Language (XML), JavaScript Object Notation (JSON) or Comma Separated Values (CSV))[16], [22].

It is, however, not easy to manage a high volume of CTI. It requires to find a suitable platform for a specific organization where it satisfies also the requirements of the human side in the long run [20].

2.1.2. Role of MISP in Threat Intelligence Sharing

MISP has evolved from an initially military project recognizing the tremendous importance of the fight against terrorism, thus supporting CTI in various sectors. MISP has evolved into a well-established platform driven by the community and other big organizations. It is an open-source project supported by the

Computer Incident Response Center Luxembourg (CIRCL) and the European Union (MISP 2024). Due to the European origins, the platform is used in various countries (e.g. Finland, Germany, Turkey, or Denmark) [24], [25].

MISP is, from the more technical point of view, software that acts as a standalone centralized server with the option to interconnect multiple instances of MISP and allow flow of TI based on rules that are applied to the participants. The power of MISP lies not only in its ability to solely connect different parties, but also helps to convey important knowledge using automatic correlation of related attributes, e.g. hash signatures and other similar IoCs [26].

From the data model perspective, the data is flat and. It relies on minimum viable data, unlike more complex data formats, like STIX, where, during the processing and information dissemination, a loss of information due to the higher data complexity can occur. Some parsing systems can have difficulties understanding it to such an extent, that some parts stay neglected or rejected by the systems.

CTI platforms play a key role to efficiently share information in a timely manner. The MISP platform is under constant development and brings new features that make understanding of shared information easier, e.g. event reports and others. [24], [27], [28]

2.1.3. Key Challenges in Collaborative Threat Intelligence sharing

Even though the significance of CTI sharing is now broadly accepted, the reality accommodates numerous obstacles in the course of it. Some researches emphasize the issue of trust establishment [28], [29].

Creation of trust has been recognized as an initial hurdle in establishing a partnership for CTI sharing, however, the later actions of the partners can as well undermine trust. One example can be not contributing back and not reciprocating the received value. On the opposite spectrum lies an argument that the more current issue in information sharing lies in the complex implications, which sharing of information – especially sensitive, encompasses. On one end, there is the fear of the subjects sharing information and becoming exposed, on the opposite side lies the vast amount of privacy implications that need to be taken into account by organizations [30].

Moreover, the problems organizations spotted span further. Along the process, various technical and operational requirements were recognized. It included the issues with the timeliness or lifetime of shared information, or the human resources needed to understand what information is relevant and of a good quality.

Another aspect of information sharing is the understanding of how much of the information should be shared, as sharing too much or too little is just as bad. Diving even deeper, the data shared can be relevant to

a certain group of stakeholders, and that the data reaches correct partners can be articulated as an important characteristic. That can not only help the efficiency of teams to work towards their own goals, it can also help the teams which are short-staffed and have difficulties handling bigger amounts of information. Therefore, the ability to filter on the side of the receiving part based on tags helps the receiver to process more relevant content [31].

From a more technical perspective, in 2014 it has been observed that another problem was holding back the ability to share information efficiently – interoperability. There were 53 protocols identified as the means of transport of CTI. Although the later prevalence of the STIX format has helped, it remains an open issue that organizations and the CTI face [28], [31].

2.1.4. Trust Establishment and Federated Collaboration

Trust is a natural aspect needed for effective interactions and activities for humans. Solutions that help manage trust in real world can be observed in many places, one example can be online shops or service providers. In the case of aforementioned instances, the trust indicators are fairly dependent on the customers' activity and active contribution to the rating system [32].

As recognized before, trust establishment and its management is a complex topic when it comes to services. One attempt to address this problematic was the Business Process Execution Language (BPEL). Even though has BPEL during its existence gained interest, the direct applicability deemed unsatisfactory for its rigid and hard-structured characteristic [32], [33].

Therefore, a more holistic approach should be taken into account, so that it counts in the multidimensional aspects of trust in CTI sharing and even more in collaborative environments [32].

- Social trust in collaborations – subjective, based on previous interactions, valid in a certain subject scope.
- Personalized trust inference – understanding the subjectivity, each entity perceives trust differently and basis it upon different requirements, therefore it should be allowed to personalize them.
- Multi-faced trust – considering multiple set of aspects like, previous or current experience with entities, certain skills or service features can influence trust in different directions.
- Compositional trust – trust should not only an act between two entities, but also be recognized on a bigger scale in multi-entities environments.

Trust to sharing partners can, however, be as well as neglected, or can be considered as a need for CTI sharing as long confidentiality is guaranteed. The trust aspect as well directly applies to the information quality and its actionability (in cases of dealing with intelligence) [34].

2.1.5. Information Sharing from Legislative Perspective

As stated in the beginning of this chapter, information sharing is a topic that is getting more and more recognized legislatively. Specifically, according to the European directive NIS2, the states are responsible to assure that organizations share cybersecurity relevant information between each other [12].
Another in European Union, the Cyber Solidarity Act officially entered into effect on the 4th of February. This regulation aims to harmonize cooperations of European Authorities. This is not a regulation that forces regular companies to follow rigorous processes. It serves as a framework to achieve a more efficient defence against bigger cyberattacks. In the context of information sharing, plays European Union Agency for Cybersecurity (ENISA) a role of central information sharing institution. ENISA can not only provide information, but also can help with creation of Information Sharing and Analysis Centers (ISACs), which are meant to share vital information in the fight against cyberattacks [35], [36].

2.2. Case Study Platforms Description

To establish a foundation for this case study, this section describes the stepping stones for the chosen CTI platforms – MISP and Taranis AI.

2.2.1. MISP in Detail

In the section 2.1.2 was the platform's role already explained. For the purpose of this thesis and the creation of a collaborative setting using MISP, it is important to understand its building blocks.
As MISP is a very complex software, there are many aspects that will not be covered in this thesis, but the most relevant parts for the collaborative system in the scope of this thesis are described below.

MISP Attributes

MISP attributes can be considered as the main building block. These attributes are particularly simple. There are numerous types of them and in the end are a simple type-value combination.
Brief selection of the attribute types[1]:

- filename – name of a file
- domain – a domain name used in the malware
- text – free text option
- boolean – binary value

[1] https://www.circl.lu/doc/misp/categories-and-types

- link – link to an external information
- email – an email address

To help to stay systemic, these attributes also fall into their categories[2] (e.g. Payload delivery, Payload type, etc.)

MISP Objects

Objects are a layer above attributes and they provide the ability to group attributes which are contextually linked[3]. Objects have again their types, and some of these types are predefined, but it is also possible to extend the set of objects to the needs of individual organizations.

A selection of objects[4]:

- Domain/IP – describes a domain and an Internet Protocol (IP) address
- File – describes metadata of a file
- Organization – describes an organization

Conceptually, the objects are based on templates. As stated before, the set of objects a MISP platform can work with is not finite. Each MISP instance can introduce their own objects by adding their templates, or if the objects do not need to be interactive in the UI, they can exist even without the templates.

MISP Events

To bring the attributes and objects together, they are concentrated into events. Besides that, events have some metadata to help represent a real-world event, for example[5]:

- Threat Level
- Analysis (stage)
- Description.

MISP Distribution Levels

Because MISP is a system designed for sharing, it offers mechanism, that helps to limit what kind of parties can reach certain information. MISP historically offers 4 main distribution settings[6]:

- Your organization only – information reaches only users of the creator's organization.

[2] https://www.circl.lu/doc/misp/categories-and-types/#categories
[3] https://www.misp-project.org/2021/03/17/MISP-Objects-101.html/
[4] https://github.com/MISP/misp-objects
[5] https://www.circl.lu/doc/misp/using-the-system/#creating-an-event
[6] https://www.circl.lu/doc/misp/sharing/#distribution-settings

- This community only – information reaches the organizations of the one instance.
- Connected communities – information reaches the community one hop away from the community that creates it.
- All communities – information is shared freely across connected communities without restrictions.

MISP also offers a special and a more granular distribution setting – **Sharing groups**. Such groups are useful when a subset of organization need to collaborate on a case[7].

2.2.2. Taranis AI in Detail

Taranis AI[8] is a platform with the main use case of handling soft CTI (strategic and tactical intelligence[9]) [37]. Among others, Taranis AI provides CTI analysts a mean to analyse the threat landscape from a high level to a more detailed view, due to the filtering capabilities. The data provided for analysis consists of news articles and intelligence feeds information, which concern (non-)cybersecurity or other specific domains. Because OSINT can be an overloading domain, one of the main limitations of analysts are often time and the information they have available [17]. Therefore, a number of mechanisms tries to ease the load, for instance, by grouping of relevant news items together or adding keywords (tags) for better context building [38].

Taranis AI News Item

News items are the main building block for real-world events in Taranis AI. Each news item usually a piece of information taken from the open internet (e.g. a news article) or (open or closed) intelligence feeds containing potentially interesting information for analysts.
A news item can exist only in one story, which can also help the analysts not to review the same articles over and over again. As visualized in the fig. 2.2, news items can be enriched with news item attributes.

Taranis AI Story

The context is built within the stories in Taranis AI. Stories are sets of news items grouped by context. A story is one layer above the news items and provides a space for additional information. Stories are dynamic, their content can change frequently, based on the analysts findings (e.g. metadata enrichments or collection of the news items when new arise). Stories contain:

[7] https://www.circl.lu/doc/misp/using-the-system/#create-and-manage-sharing-groups
[8] https://taranis.ai
[9] https://www.paloaltonetworks.com/cyberpedia/what-is-cyberthreat-intelligence-cti

- Comments – analysis remarks.
- Summaries – brief summaries of the whole context.
- Tags – extracted keywords from the story.
- Attributes – key value pairs enriching the context of the story.
- News items – news articles in the form of news items are grouped by a story to a single context.
- For more, see fig. 2.2.

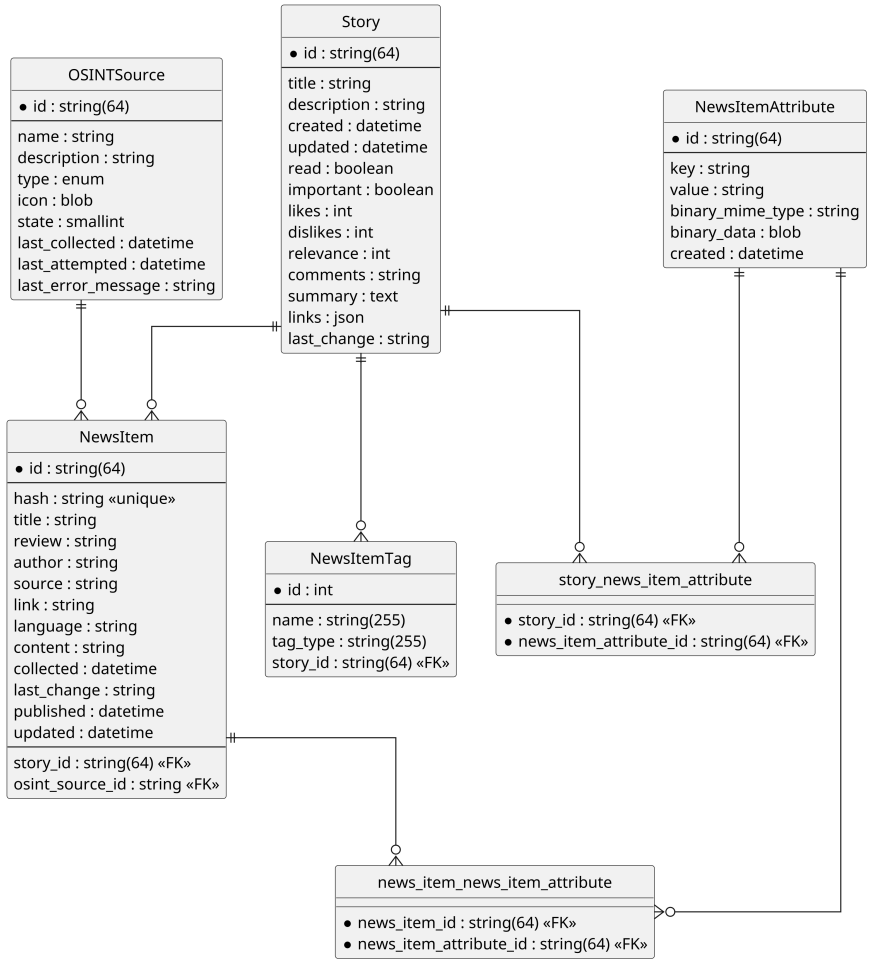


Figure 2.2.: Entity Relationship Diagram: Taranis AI internal data model scheme.

3. Use Case and Process Analysis

Before a SW is implemented and even designed, a requirements analysis is carried out. A part of this analysis is the use case analysis, where the users (actors) are the initiators of various system actions (use cases).
This chapter includes also sequence diagrams to better demonstrate the interactions on the system level.

3.1. Use Case Analysis

This is the user oriented section, where it is visualized what are the key capabilities of the collaboration environment. The main actors are:

- CTI analysts
- CTI platform administrators.

3.1.1. Use Case 1: CTI Analyst Shares a Story

This use case describes the workflow of a CTI analyst sharing a story with someone else (see fig. 3.1). The process of sharing necessarily involves the assessment of news items and stories. An assessment can be complemented by adjustments to enhance the completeness of the shared message.

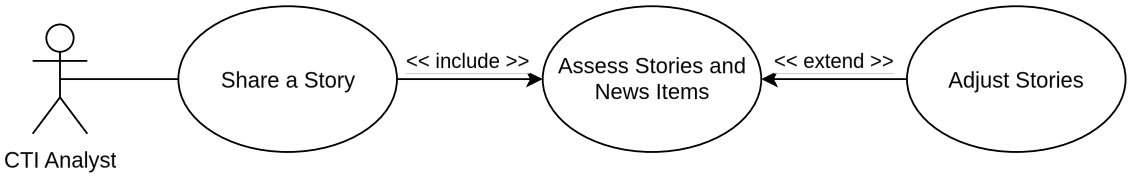


Figure 3.1.: Use case diagram: CTI Analyst shares a story after an assessment complemented by fine-grained adjustments.

3.1.2. Use Case 2: CTI Analyst Ingests a Story

A CTI analyst has a number of sources of information. This use case shows that an analyst ingests data (see fig. 3.2), but because some sources or partners provide overlapping data, there might need to occur a manual conflict resolution.

Furthermore, if an analyst is ingesting a story in a collaborative environment, it is also possible that the story has change requests from other collaborators. This is the latest stage, when an analyst reviews the proposed changes. This order helps analysts to work with the most up-to-date data and save time.

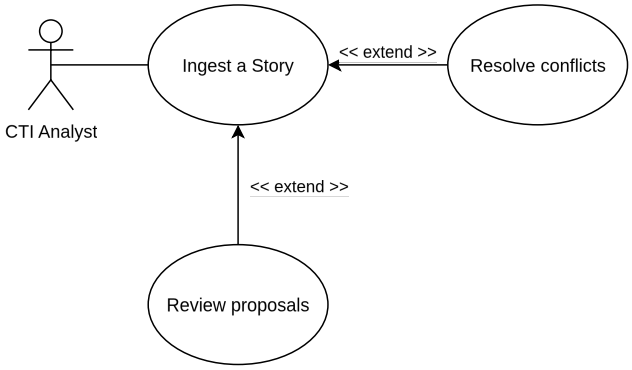


Figure 3.2.: Use case diagram: CTI Analyst ingests a story, verifies potential upstream change requests and resolves conflicts.

3.1.3. Use Case 3: CTI Analyst Updates an Owned Shared Story

A CTI analyst wants to update the story that they shared, because more information has been gained about that event, and because it is of interest of the collaborators, they should get the new information propagated to them (see fig. 3.3).

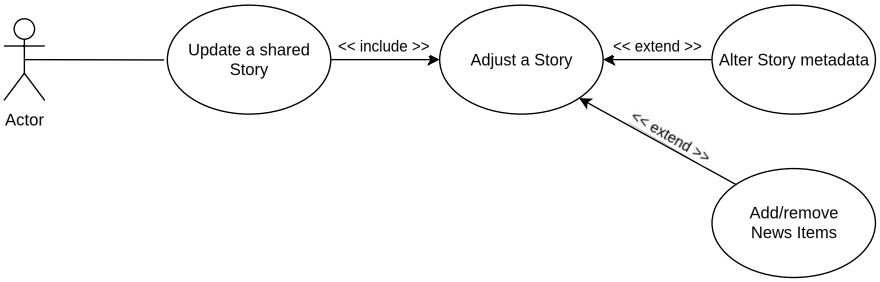


Figure 3.3.: Use case diagram: CTI Analyst updates a story with story metadata changes or news item changes.

3.1.4. Use Case 4: Propose Changes to Story Metadata

A CTI analyst assessed information. Some information is relevant to the story, which a partner had shared with them. Because the story's origin, the analyst proposes changes to it (see fig. 3.4) to add clarifying comments and suggestions to make the story more accurate (according to their latest analysis) and provide additional (or correcting) information to the partners.

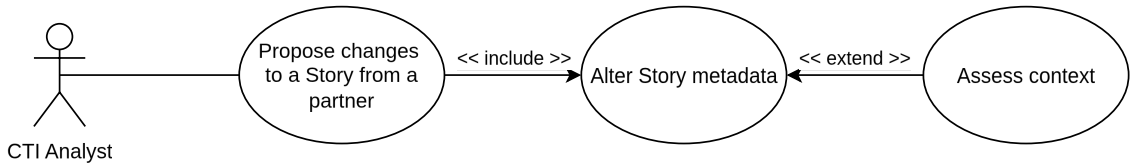


Figure 3.4.: Use case diagram: CTI Analyst proposes changes to a story originating from a partner.

3.1.5. Use Case 5: Propose Changes to the Story's News Items

In a collaborative environment, a CTI analyst has much more news items about a certain event, thus can much better describe the latest (known) state. The analysts adjust the news items representing in the story (an event) and shares this back (see fig. 3.5). This can have two results – suggested additions or deletions.

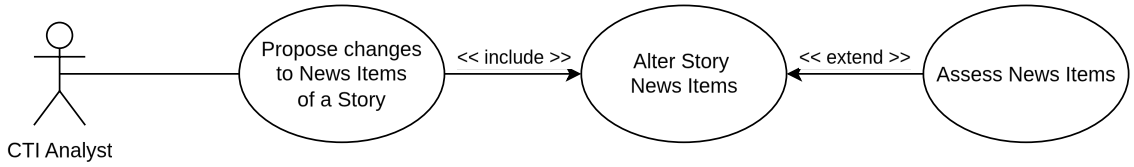


Figure 3.5.: Use case diagram: CTI Analyst proposes changes on the news item level to a story originating from a partner.

3.1.6. Use Case 6: Transitively Share Information from One Collaboration Group to Another

The CTI analyst works in more distinct collaboration groups. One group reports events that are highly relevant for the other working group. The CTI analyst considers the information sharing constraints and shares the information to the other interested group.

3.1.7. Use Case 7: Connect to a specific collaboration group

A CTI analysis centre wants to join forces and cooperate with one or more partners. For this, a group needs to be established with proper access control (see fig. 3.6). The group members can be managed by the administrator of the group. The end CTI platforms can decide, whether they contribute, consume, or both.

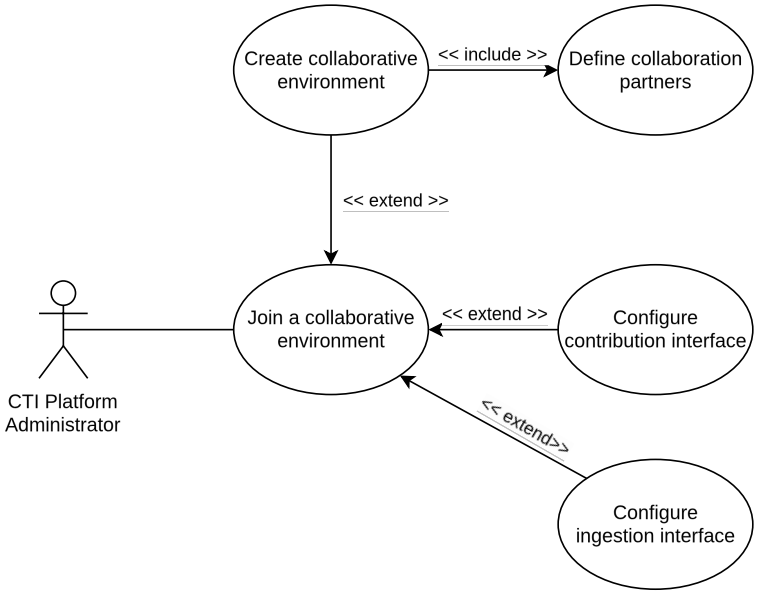


Figure 3.6.: Use case diagram: CTI Analyst proposes changes on the news item level to a story originating from a partner.

3.2. System Interactions Analysis

The following section with diagrams describes example scenarios to reach goals of collaboration (share information with a partner and the partner contributes back).
From the standard way MISP operates [39], it can be derived how a sharing setting can look like. As initially stated in the chapter 1, for the course of this thesis, a more simple scenario is considered to verify the applicability and usability of the suggested solution.
In the middle of it, as a central hub, lies the MISP platform. Via this, the organizations using Taranis AI can become connected share or ingest information, and also collaborate (see fig. 3.7).

3.2.1. Share Information with a Partner

The fig. 3.8 describes the process how one CTI analyst can send stories to a partner. The analyst, after conducting the analysis, sends the data to the joint MISP instance. That is queried by the Taranis AI platform

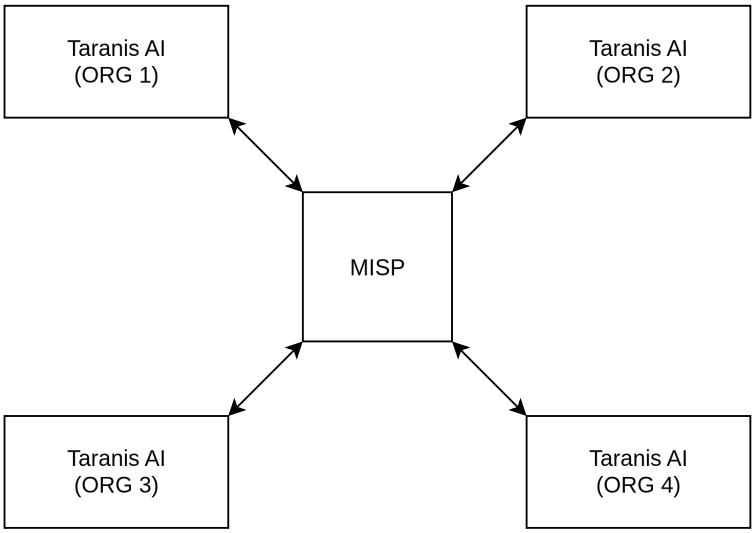


Figure 3.7.: Organizations using Taranis AI connect through MISP and collaborate.

to ingest new stories. Then the analyst from another organization makes an assessment in the current context. The fig. 3.8 reflect the cases of sharing new stories and providing updates to formerly shared stories.

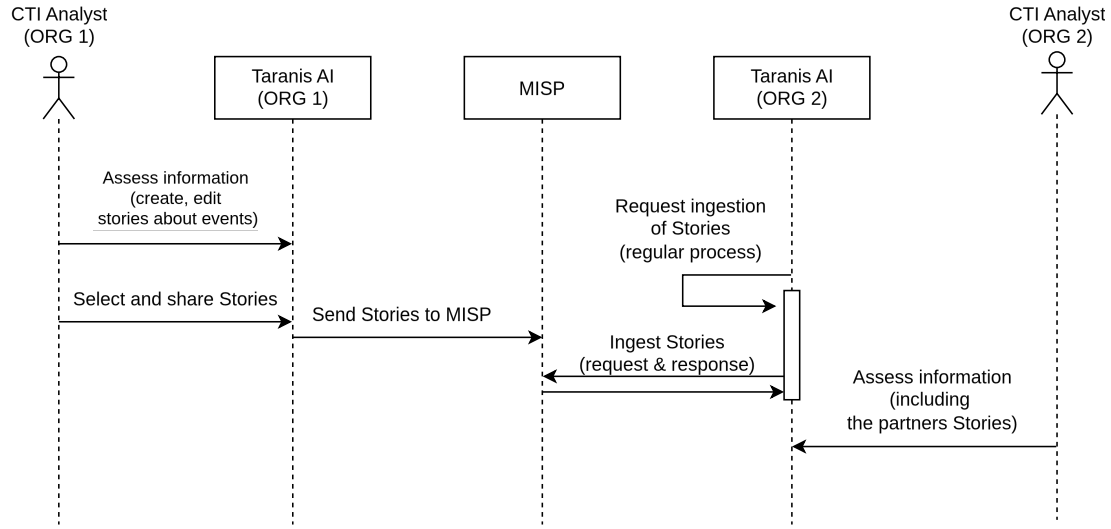


Figure 3.8.: Sequence diagram displays analysts from two organizations sending and receiving data.

3.2.2. Contribute Back to the Originator

The fig. 3.9 demonstrates the step of an initiated collaboration. A CTI analyst from a partner organization considers a story incomplete, therefore, proposes changes to the originator. The information is propagated back to MISP where it reviewed, approved, or rejected.

Then, on a regular basis, the originator's Taranis AI platform receives the stories, including the changes introduced by the partner.

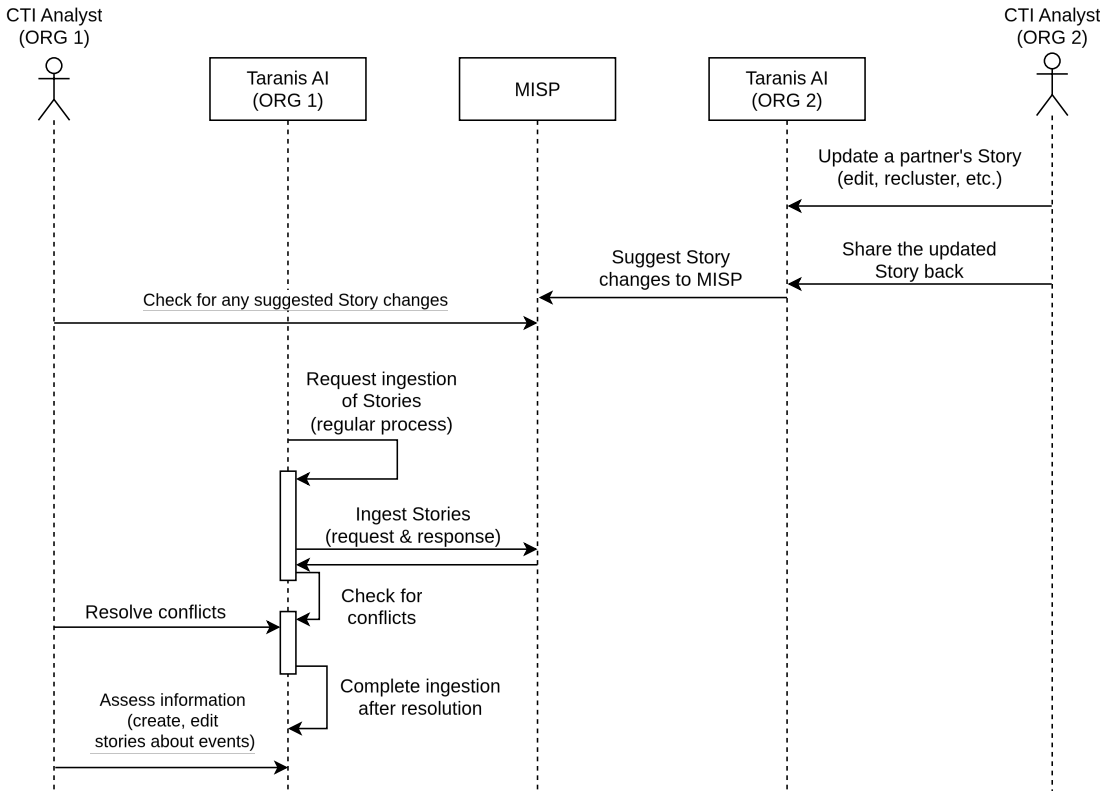


Figure 3.9.: Collaborating organization proposes a change request to the originator.

4. Design and Architecture

This chapter presents the design and architecture of the PoC implementation of a module which extends an open-source OSINT platform Taranis AI. The PoC implements a communication interface between an open-source CTI platform MISP and the aforementioned OSINT platform.

4.1. Requirements Analysis

This section describes in detail what are the key design and architecture motivators, including the functional and non-functional aspects of the SW.

4.1.1. Functional Requirements

Ingestion of information from various CTI partners

An organization can have more partners, even geographically far away. In case, these partners do not want to interact directly with each other, the communication channels must be split.

- **Architecture impact:** Ingestion component allows the usage of different channels of communication.

Outbound sharing of CTI to various partners

An important aspect of information sharing is the ability to share information privately with different partners. Such as in case of different collaboration contracts, or based on different information relevance or needs.

- **Architecture impact:** The exporting component is responsible for exporting the data in the correct format and making it available to the partner. Additionally, allow multiple outbound information channels of communication.

Allows many-to-many collaboration – with transitive sharing

Even though, the information flows in from different information channels, it is possible that some partners would find the information relevant to other partners. In that case, if no sharing constraints apply (e.g. Traffic Light Protocol (TLP)), data can be further shared to other partners – so-called Transitive sharing.

- **Architecture impact:** Allow sharing of information without originators check. This way the responsibility of the correct data transitivity is delegated on the end user (security analyst).

Collaboration during analysis with various partners with minimal delay

Timeliness is an important aspect during investigations, especially of security related events. Thus, the need to minimalize the delays between when one collaborator sends the information to the other collaborator is tangible.

- **Architecture impact:** Allow configuration to minimize delays of information exchange between collaborators.

Safe data manipulation and updates

The integrity of the data in a collaborative environment is not only essential, but also sets the ground for trustworthy partnerships.

- **Architecture impact:** Assure the data on the collaborators' ends are changed only intentionally.

Conflict detection and resolution mechanism

In a collaborative environment it is natural that the information diverts from the source. On an attempt of unification (merging) the versions into one, conflicts must be resolved manually, if the diversion cannot be solved automatically.

- **Architecture impact:** Introduce a conflict detection and resolution mechanism that will contribute to the safe data manipulation and updates.

4.1.2. Non-Functional Requirements

Interoperability

Building upon solutions that are trusted can on one hand increase the acceptability of a solution by the community. On the other hand, it can also help with the interoperability and ease the initial hurdle to integrate systems into organizations' workflows that already use some of the systems.

- **Architecture impact:** Create interfaces to reuse commonly used technologies, like MISP, to support the collaboration.

Scalability

It is important that SW is scalable and adjustable to the needs of a particular organization. This becomes even more significant in the currently always more commonly used cloud resources, because a dynamic adaptability to fluctuating resource needs can immensely influence the costs.

- **Architecture impact:** Extend the microservice architecture of the OSINT tool Taranis AI.

Extensibility

The shared data can change in time. Losing the ability to collaborate, because the data model on which the collaboration occurs should be avoided. Changing the templates should not be a breaking change, rather a deployment decision. Having an option to extend or change the data types should be prioritized.

- **Architecture impact:** Data on which the collaboration occurs is sophistically mapped to selectable templates that the senders and recipients both possess.

4.2. High-Level Architecture Overview

Considering the sharing system distributed characteristic — the collaborating partners are far apart — there are a few common ways the system can be designed. Among the general options are the Layered Architecture (separates different responsibilities into hierarchical layers), the Peer-to-Peer (direct communication of endpoints) or Client-Server (distinguished responsibilities of clients and servers) architecture, etc. However, because the

4.2.1. Hub-and-Spoke as a Federated Backbone

Considering the main components of the sharing system (MISP and Taranis AI), the system maps best to the sharing model of Hub and Spoke [40]. As the architecture should allow federation between multiple hubs, the hub-and-spoke model can be extended by more of the autonomous hub-and-spoke models. This directly mirrors the intended connection of MISP and Taranis AI. MISP acts as a hub and allows the participating organizations to connect to it. The spokes, Taranis AI instances, connect to one hub or more (see fig. 4.1). This creates a federated system.

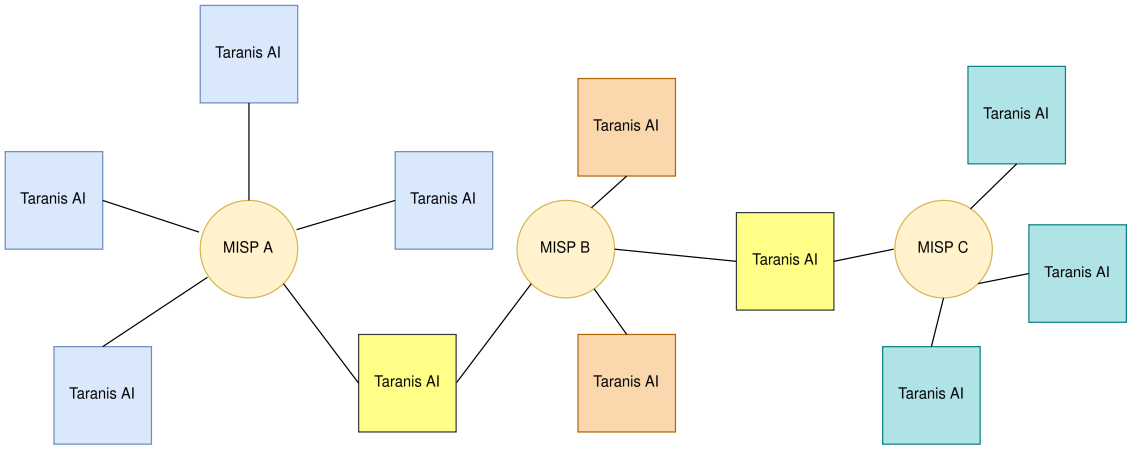


Figure 4.1.: Federated Hub-and-Spoke Architecture

In reality, the system is a bit more complex, because the hubs of MISP instances can be also interconnected between each other and share information directly.

4.2.2. Repository Pattern and Shared Knowledge Base

From the perspective of the hub-and-spoke model, the architecture type that fits the best for this sharing system is the Repository Architecture.
Within this architecture, Taranis AI and MISP have special roles. MISP acts as a database saving all the data and Taranis AI contributes to this database by sending their data there. This information can be later on retrieved, similarly to how databases act in systems.

4.2.3. Architectural Decisions and Trade-Offs

The final design was influenced by the previously discussed requirements in the section 4.1.

Implications of the Star Topology

Architecture styles strongly influence the SW's ability to perform. The following section highlights some advantages and disadvantages, and justifies the chosen architecture.
There are different names of the chosen style, especially across different fields of application – Star Topology (Information Technology (IT) networking), Hub-and-Spoke model (information sharing), Data-Centred Architecture (software architecture). All of them stem from the common perspective where a central system joints multiple systems in one point.

Delays

Distributed systems with this architecture tend to be burdened by delays in communication, as the information from the originator to recipient can take longer to propagate. The speed of flow and delay times can be mitigated to some degree with implementation and configuration, but delays usually cannot be evicted from the system. Typically, finding balance in the propagation speeds is preferred.

Push-and-Pull Approach

With a system of this design, there are two main ways to control the information propagation. Depending on the character of the traffic between the nodes, can be propagation of information be done via *push* or *pull*. The *push* approach benefits lower "latency" (delays) and can be exceptional, when the resources connecting the nodes are available all the time. Such approach can be favoured in cases of real-time communication, where low response times are required.

The *pull* approach works the opposite way. The nodes query the other systems on demand. This approach is useful when the nodes themselves do not require active input for their functioning. The nodes themselves request information when external input is favourable. This can save the bandwidth and also does not require availability of the communication mediums.

Considering the needs of this information sharing system, a close to real-time communication could be beneficial, but it is not required. The endpoints of the network can be very independent, and the information that is shared with others can highly vary, so that a combination of *pull* and *push* mechanisms is suitable.

Modularity

A data centric approach has also the potential to offer more modularity and interoperability. A system with a reliance on an arbitrary central storage-like node enables the community to use individually chosen systems on the receiving sides (e.g. replacing Taranis AI as an end system with a custom solution).

4.3. Subsystems Overview

The facilitation of communication between the OSINT platform and the CTI platform is split into two main modules (MISP Collector and MISP Connector) on the side of the OSINT platform. On the side of the CTI platform support to the sharing naturally happens as well, and without proper configuration (e.g. Application Platform Interface (API) access keys or authorizations), the communication would not be possible.

OSINT Platform Subsystems

MISP Collector

The collector module has the responsibility to query (pull) information from the CTI platform. It is its re-

sponsibility to convert the data from the format into the format understandable by the OSINT. The converted data is then passed through to the application where more processing occurs.

MISP Connector

MISP Collector is responsible for the outward sharing of the data and delivering it from the OSINT platform to the CTI platform. It, essentially, converts the internal data to the recipients compatible data representation and delivers it. This system also decides how is information displayed in the shared repository.

Conflict Detection and Resolution Systems

The ability to resolve conflicts is very closely tied to the conflict detection ability. These systems (modules) are responsible for the change tracking of local information and in the case of incoming conflicting changes allow the user to manually resolve the conflicts.

CTI Platform Subsystems

CTI Platform The CTI platforms responsibility is to manage the correct system accesses and storage of the data, so it could be queried by partners. The platform is also responsible to allow working out some divergencies in terms of conflicting data, when such would be sent to the platform by the connected partners.

4.4. Data Flow and Interactions Scenarios

The distributed architecture allows the connection of more OSINT platforms together by the CTI platform remaining as the intermediate point. To make the knowledge exchange even more sufficient and open to other stakeholders, there is a potential of connecting more CTI platforms together directly and synchronize on the layer of the intermediate nodes as well.

The general flow is, however, simple, the selected data are sent to the CTI platform using the Connector module. Afterwards, the data can be consumed by partners into their own OSINT platforms (see fig. 3.8).

During the outward communication (sending), the Connector transforms the internal data to match the compatible format of the CTI platform.

When the data reaches the CTI platform, it is up to the platforms' settings, as to what data will be accessible to what partners. Thanks to the abilities of MISP, the data can, in theory, be shared between several instances and connect numerous partners together.

Furthermore, the ingestion of the data of partners is handled by the Collector module. The Collector collects the data present in at the CTI platform, and transforms the data to be compatible with the OSINT platform.

When data is ingested, then changes are made locally, the backpropagation of the changes to the source of

information varies depending on what changes are made (in MISP creation of proposals or extended events) and whether the contributor is from the owning organization, or is an external contributor.
When it comes to ingesting, there are multiple reasons why there would be conflicts. If more partners use the same source, the system would run in duplicated information. As a result, conflicts of duplicated (conflicting) news items are one level and reason of conflicts.
Another level of conflicts can occur when the partners started collaborating on one piece of information (one event context) in more detail. There, the changes are much more minor, but can add a lot of value (e.g. analysts comments, etc.).
If the same information can come from two sources, it can even come from more sources and this way can create an increased number of conflicts. This has strong effects on the conflict resolution workflow, as the resolution of conflicts of those items directly influence the rest of the conflicts. In other words, accepting a news item that comes from multiple sources and waits for conflict resolution (due to a different reason of a conflict), will result in a conflict with the freshly ingested news item.
All in all, the complexity of the system increases the moment a true collaboration is initiated – that is when the conflicts occur on the different levels and manual inspection and resolution of the data must be done.

4.5. Nonfunctional Properties

This section explains the involvement of non-functional properties, like interoperability, user/organization based permissions, authentication.

4.5.1. Interoperability

Interoperability of the design is well-supported thanks to the distributed design and the incorporation of a popular connecting entity – MISP. Therefore, due to the open characteristic of MISP and the PoC, this solution allows anyone to build upon it and introduce their own systems and interfaces to communicate with MISP or other solutions.
The system is compatible with the MISP data model. In case third parties would want to use their own solutions for the intermediate node and connect with organizations that use Taranis AI, they would need to adhere to the data model and its representation on the MISP level.

4.5.2. Identity and Access Management

Identity and Access Management (IAM) is done on more layers. The first and most basic support of IAM is present on the user level in the OSINT platform. The PoC features Role-based Access Control (RBAC), and allows the OSINT platform to assign each user of the platform to allow or forbid the usage of the Connectors, which are set up.
The second step in the IAM is delegated mainly to the administrator of the MISP instance and complements the PoC's (first) layer. The current design allows an organization level granularity. Meaning, the organizations, per se, can send information to the MISP instance under their organization's access key.
The intended use is, that the MISP administrator creates a dedicated user, creates or reuses a suitable role with secure permissions (following the least privilege principle), and assigns it to the user. Subsequently, the administrator crafts an API key for the user for later use in the CTI platform. This way the users sending information from the CTI platform will be represented in MISP by a single user, and therefore the granularity lies on the organization level.

Internal/External Parties Management

Additionally, the sole purpose of sharing is that the right information arrives or is accessible for the right parties. Users that belong to the creator organization have special privileges (e.g. edit permissions of their organization events, etc.).
So that the information is accessible for the correct parties, correct *distribution levels* need to be applied. The distribution of events across the interconnected system can then either span to all parties (when least restrictions are applied), or it can limit it to a single sharing group.

4.6. Security Mechanisms

This section the discusses selected security mechanisms that are responsible to support security and compliance. One of them, for increased privacy and operational security, the design allows usage of the SSL mode. That way, during the establishment of the connection with the CTI platform, the certificate validity is checked. It also supports self-signed or other certificates. For extra security, it allows submitting a client-side certificate (Mutual Transport Layer Security (mTLS) concept). Moreover, the Hypertext Transfer Protocol Secure (HTTPS) adapter can be enforced, for the whole communication.
Lastly, other properties, which can contribute to an increased security, can be configured – e.g. proxy settings, Hypertext Transfer Protocol (HTTP) headers or request timeouts.

5. Mapping Information Sharing Models to MISP

To effectively use MISP as an intermediate node for distribution, there need to be a structured way to handle it. It is similar to how communication protocols standardize communication on a big scale, a similar mindset needs to be applied for this system integration.
With the structured information, MISP can be used for the distribution between different instances of Taranis AI. For the purpose of the plain information distribution and collaboration, the data needs to be functional – support the collaborative nature of the sometimes quickly changing data, where more parties can contribute their opinions.
Another aspect to consider is the meaningful Taranis AI data representation outside Taranis AI. This particular characteristic significantly increases the potential of wider adoption, since the contribution could be done directly via the MISP, which, as previously mentioned, is used in the industry.

5.1. Mapping Taranis AI Data Model to MISP

To achieve this small scale standardization, so that information is easily readable and understood in two different systems, data models were mapped. Usually not all data carries information relevant for external system. With this in mind, an only the most important data (information carrying properties) are selected, whereas the parts with the potential of information noise will be omitted for the intersystem standardization.

5.1.1. Information Carrying Properties

In the section 2.2 was the structure of the data in Taranis AI outlined. Considering the fig. 2.2, it is visible there are many properties. For a meaningful sharing process it is good to focus on sharing the core and most relevant properties.

Stories and News Items Context

From the fig. 2.2 diagram, the fundamental properties that build up the correct context are:

- identifier (ID),
- title – title of a story;
- summary – summary of a story (generated or manually entered);
- description – manually added description;
- comments – manual comments from analysis;
- news item tags – keywords representing the context (generated or manually entered)
- news item attributes – key-value pairs of additional metadata (e.g. tracking internal processing, etc.),
- links – e.g. Uniform Resource Locator (URL) list of online resources.

From the news items, the important properties are:

- ID
- title – title of a news item;
- link – exact place where the information (an article) comes from;
- content – main content of an article;
- author – author of an article;
- hash – unique hash for checking of duplicate articles;
- language – language set on the source of the article;
- source – URL of the OSINT source's target (e.g. intelligence feed URL);
- review – review of an article;
- story ID – ID of the current umbrella story.

On the counterside, sharing information that is platform specific (e.g. read, likes, dislikes) is not useful to share and creates unnecessary noise. The main reason why some are platform specific is by design. Consider an example case: Whether a story is marked as read is irrelevant to the recipient. (i) The partner shared it with an intention so it was most likely "read". (ii) It could also lead to misunderstandings and incorrect states, due to different organization workflows and rules depending on e.g. when items are marked as read. Another example are properties such as likes and dislikes. They are tied to the number of users (analysts) of a platform. Each user can give either one like or one dislike. Because the number of analysts is not standard, it is hard to clearly interpret the meaning of number of total likes/dislikes. As described in the chapter 2, data that can be clearly interpreted is an important aspect for data sharing.

5.1.2. Logical Mapping of Taranis AI data to MISP

To follow up with the introduction (see chapter 5), the visual appeal could be also called *logical*, since it should be fairly straightforward to grasp the same understanding from looking at the information in MISP alone too.

The core properties for mapping are derived from the section 5.1.1. For simplification of PoC, the final data model includes also fields that are not information carriers for regular users (e.g. `osint_source_id`), but rather elements that allow some workarounds.

Building upon the knowledge from section 2.2.1, there are various options and combinations of how the data could be mapped to the MISP sharing system. Usage of MISP Attributes, Objects, Events, and Reports deemed as best as the combination of these provides an easy transition from the way analysts see in Taranis AI and then in MISP.

MISP Events

The name itself (MISP events) suggests that it represents an event which occurred. That is very similar to what function of stories is in Taranis AI. However, for the whole sharing process, the general metadata fields of MISP events are neither sufficient to describe a whole story, nor the metadata fields of a story known in Taranis AI. What ultimately these MISP events allow is the grouping of various other data types. As a result of that, the MISP events can be partially mirroring the Taranis AI story data model.

MISP Objects

MISP objects, as mentioned in section 2.2.1, cluster MISP attributes (the most basic pieces of information) together to build a greater context. MISP objects, for their nature, offer a good fit for representing more complex data models, like news items.

5.1.3. MISP Attributes

The attributes are a very simple property. It represents a simple value and supports different data types, so can directly be mapped to singular variables that are present in Taranis AI (hash, title, content, etc.).

5.2. Functional Mapping of Taranis AI to MISP

This section describes the functional aspects of the mapping between the Taranis AI's data model and MISP. While the logical mapping in section 5.1.2 is more simple and considers mainly the User Interface (UI)

from the user's perspective, the functional mapping is dedicated to understanding of what functionalities are natively used for collaboration.
This type of mapping also complements the validation of the logical mapping, because it is important that it is possible to accomplish the main goal — to collaborate — and not just to be able to see information someone shared.

5.2.1. MISP Events

Even though events encompass numerous features, for the process of sharing, the setting of distribution level is the most prominent. This distribution level is then advised to be inherited to all the attached properties (e.g. MISP attributes or objects). Other functionalities useful for the collaboration are ownership management using the organization metadata, contributors listing, and title.

5.2.2. MISP Objects and Attributes

Because MISP objects consist of attributes, both will be described together. Objects provide the basis for collaboration. They allow handling a complex piece of information as one entity. Attributes, on the other hand allow the power of collaboration with their ability to receive proposal changes from other partners than their originator.

5.2.3. Missing link – Extended events

With the previous mapping of functionalities to the requirement use cases, it is possible to fulfil the use case from fig. 3.1 and fig. 3.2, but only partially the use case from the fig. 3.9. That is because the changes in news items in a story would require proposals of additional news item objects to the event. Unfortunately, that is a technical limitation and it is not possible to propose fully fledged MISP objects to events.
As a result, another functionality of MISP needs to be used – extended events. With this feature, it is possible to also create new objects and in a specific way (by extending the original event) creating a link between the new object and the original event. Extended events are also called extension events, because they serve as an extension.

6. Implementation

The following chapter focuses on the specifics of the PoC implementation. The design and other requirements were translated into a piece of software, extending the open-source OSINT platform *Taranis AI*[1] integrating MISP.

6.1. Implementation Overview

Since the goal of the thesis was to extend the SW Taranis AI, naturally, but also for integration simplicity, the technology stack of the core sharing modules was influenced by the already involved technologies.
Taranis AI was at that moment built with two main frameworks for web applications – Flask (Python backend framework) and Vue.js (JavaScript frontend framework). The platform was also complemented by a message queue broker RabbitMQ[2] and Celery[3], which is a framework for processing asynchronous high load tasks (microservice worker pattern).
The modules comprising the application are split into isolated microservices containerized with the enterprise standard Docker.
For the central (intermediate) point that should joint the communicating partners was chosen the MISP platform. To integrate MISP the Python library PyMISP[4] was used.

6.1.1. Bidirectional Communication

The cornerstones of the inward and outward communication were built into the two components – MISP Collector and MISP Connector. Even though it was first considered to couple (integrate) them more tightly together (into one singular module) it was not the best approach from the application design standpoint.
Firstly, the use cases of the Collector and Connector are each other opposites, and by keeping them in one functional module, the modularity of the software would decrease and complexity of the one joint

[1] https://taranis.ai
[2] https://www.rabbitmq.com
[3] https://docs.celeryq.dev
[4] https://pymisp.readthedocs.io/

component would rise. Secondly, the MISP Collector was nothing conceptually new to the Taranis AI platform. The variety of Collector types (Collectors) are one of the core components and are responsible to ingests external data from different source types (RSS feeds, websites, Mastodon channels and more).
Therefore, the readiness of the platform to handle an additional source of information — MISP — was high. Taking advantage of the supportive built-in mechanisms for processing external information, it was not a huge challenge to add this new source type.
The part that was more complicated to integrate in the new system was the MISP Connector module. This was a new feature for the software on the conceptual level. While the feature to share intelligence reports was present, the ability to share information in the course of the analyses of security analysts' was missing.

MISP Collector

The MISP Collector was implemented as a new module added to the existing set of Collectors of other source types. This way, the different Collectors stay separated. The MISP Collector inherits from the `BaseCollector` base class. This base class is by generally reused in other Collectors responsible for incoming information.
The collection starts with the entry method `collect`, shown in Listing 1.

```
1 def collect(self, source: dict, manual: bool = False) -> None:
2     self.parse_parameters(source.get("parameters", ""))
3     return self.misp_collector(source)
```

Listing 1: Entry method of MISP Collector

The method takes three parameters, out of which one is optional (`manual`). First, the parsing of the parameters common for all Collectors and parameters special for MISP Collector are parsed for the execution.
The module is able to connect to the MISP instance, filter for relevant and compatible events and transform them into a compatible format to process it in Taranis AI.

MISP Connector

MISP Connector was decided to become a dedicated module of Taranis AI as well. Conceptually, this new Connector module is responsible for exporting data in specific forms, in this case the MISP friendly format

and send the data to a MISP instance. Most importantly, it allows seamless support for various organization, where MISP Proposals and MISP Extended Events are used.

To make the experience of users using MISP better, each MISP event has an automatically generated MISP Report. This report is created from the story's description and comment (see fig. 6.1). This report is always kept up-to-date with the latest story information and there is always only one report per event.



Figure 6.1.: MISP Event report with a description and summary properties extracted from a Taranis AI story.

The connector allows the user to fulfil the use cases from figs. 3.1 and 3.3. To share or update a story, the same user interaction with the Connector is needed. In the *Assess* section, has the analyst stories that can be shared. The stories can be one by one sent to MISP using the story action button "*Share to Connector*"(see fig. 6.2).

Then, a popup is displayed, and the analyst chooses what preconfigured Connector will be used (see fig. 6.3). A story that was shared can be worked upon, and in case of an update, the analysts can propagate this update with the exact same steps as mentioned before.

Building upon the information from section 5.1.2, as a result, in MISP this story creates a MISP Event with the distribution settings and other metadata. The information of the story is carried by custom news item and story objects. Each event has one story object with the story data and at least one news item object present.

If the story is not owned by the organization, the steps in Taranis AI are the same, but the way it is displayed in MISP is different. If there is a change proposed to a certain property of a story, it is, in most of the cases, displayed as an attribute change proposal (see fig. 6.4).

When the user wants to change the news items, however, this is propagated differently. If a user suggests a news item from a story for a removal, a proposal is created to the *hash* attribute of the news item object with a comment: *This news item is not considered a good fit for the story. Please consider deletion and update of this event.* (see fig. 6.5). The owner of the story can then decide if the news item gets removed, or this

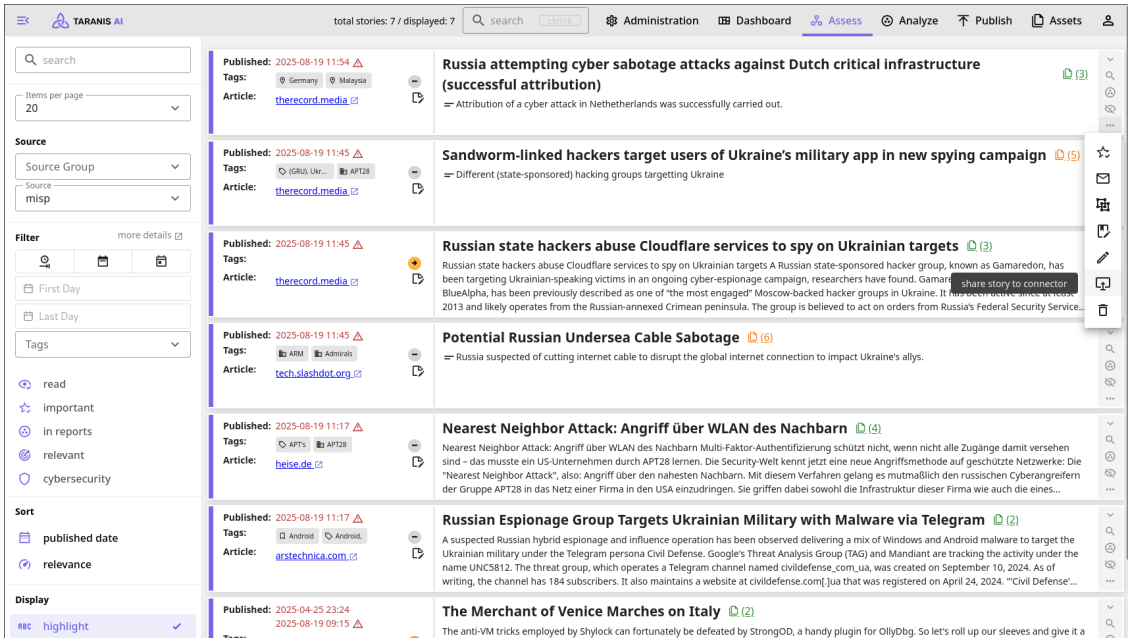


Figure 6.2.: Share a story to MISP – story action button.

proposal would be discarded.

When an additional item is proposed to this event, it is visualized differently. Building upon the knowledge from section 5.2.3, because it is not possible to propose a new object to an event directly, extended events are used (see fig. 6.6). The extended event extends the original event, and contains only the news item objects suggested for addition. The fact that the original event is being extended is also visible at the top of the MISP event (see fig. 6.7).

An event can be extended by more extension events. The implementation consists of a few mechanisms that help with a better usability. Every MISP organization can create one extension event per event. This one event can contain one or more news items objects. Then, is this extension event from an organization updated accordingly if some news items are removed or added. There is also a deduplication mechanism, so all news item objects in the extension events stay unique.

6.1.2. Conflict Handling System

Not even this collaborative setting is able to avoid conflicts that require human input. Because of that, a system which can recognize and handle conflicts was implemented. To recognize a conflict or states when data should not be overridden, rules had to be defined.

The conflict handling system does not reuse the main database. Even though, for the scope of this thesis, the system does not count with conflicts with thousands of conflicts, the computation power for certain

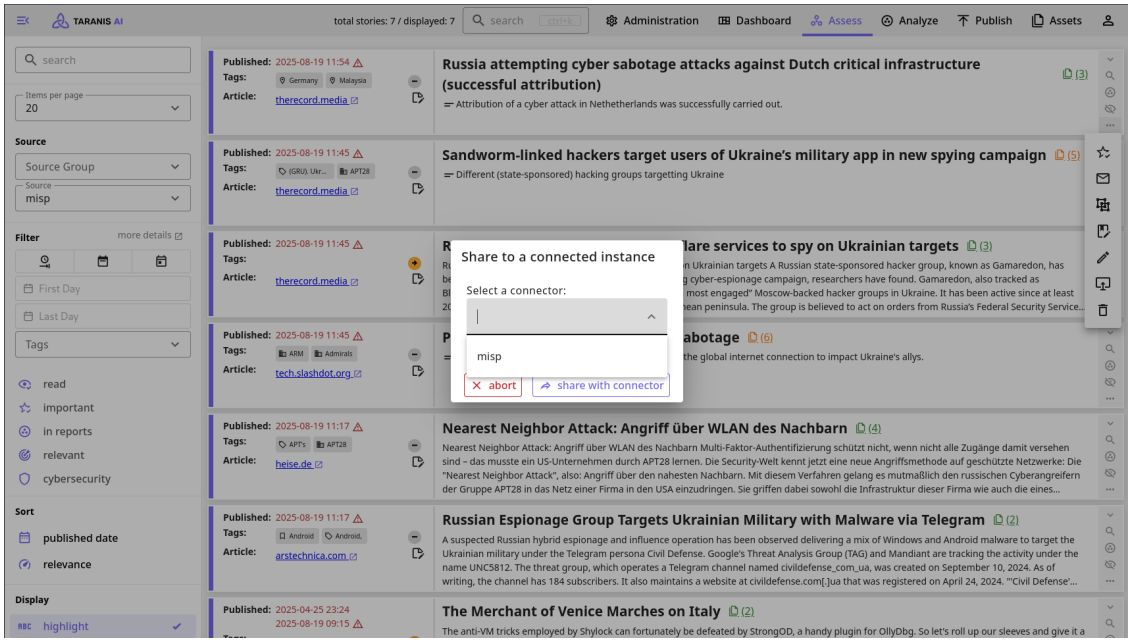


Figure 6.3.: Share a story to MISP popup, choose a Connector.

operations (e.g. recalculation of conflicts) require a fast memory. Because of that an in-memory storage was chosen. That is because after a conflict is resolved and the internal state changes, more incoming items might be effected by this change (e.g. in case of overlap of news items in more incoming stories), all news items conflicts need to be reevaluated against the latest internal state. Due to that, two data python data classes[5] were implemented for the conflict storage – one for the conflicts that occur on the Story level, and one for the conflicts that occur on the News Item level. The access to the conflict handling feature have users, that are write permissions to the *Assess* section with news item assigned.

Conflict Resolution for Story Conflicts

To inspect conflicts, the user uses the conflict resolution section. The conflict resolution view is split into two parts. At the top, the story conflicts can be resolved. This part features collapsable cards that provide a view to see two text fields implemented with the library Mergely[6]. The left side shows the internal state of the story, and on the right side is the external state ingested from MISP.

The left and right view show the data in the raw JSON format. The text needs to be sorted, otherwise it would be hard to compare the two sides. Any differences are highlighted with colours. Red text means deletions and blue addition. A user can scroll over the content or take advantage of the button "*Next Diff*"

[5] https://docs.python.org/3/library/dataclasses.html
[6] https://github.com/wickedest/Mergely

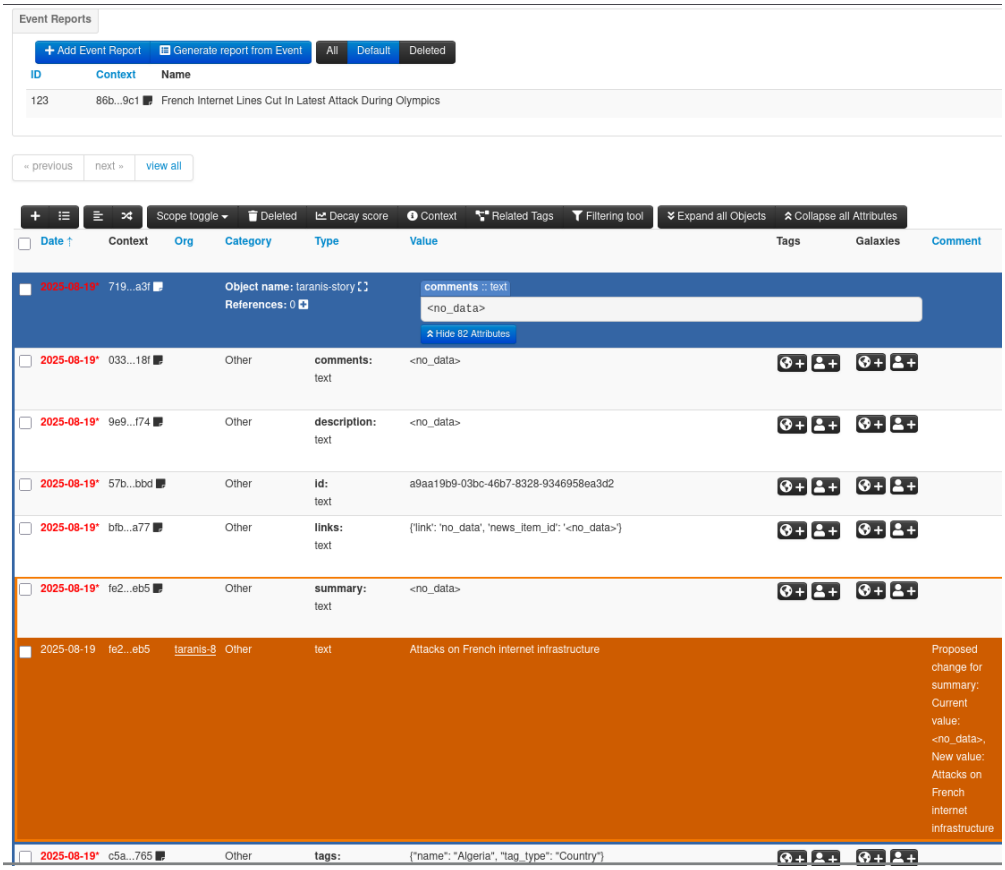


Figure 6.4.: An attribute change proposal was submitted.

to jump to the next difference.

By going through this view, the user can manually alter the content to the state desired. The determining side, that is used for submission is the right side, so the user, effectively, needs to move things from left, to right, if removal should not happen, or alter text on the right side to the desired state.

After the user finishes the review, the button "*Get right side (Incoming)*" is clicked, to validate the proper JSON format, and it further provides a view to once again review the final version of the diff. If everything looks good, the user can finally finish and press the button "*Submit Resolution*". After that the story vanishes and an update to the story is carried out. It is still possible, that the story incoming has conflicts with the news items. In that case, the story ends up in the news item conflicts below.

Additionally, this story conflict view features a few visual and functional aids to improve the experience. At the top, a violet circular box shows the total number of proposals that are present in the story conflicts. Each collapsable card features a dynamic status button that that distinguishes between two states. If a story in conflict has no proposals in MISP, it is disabled, grey and shows "*No Proposals*". If, however, the story has open one or more proposals, it shows "*View Event with Proposal(s)*" and on click it redirects the user

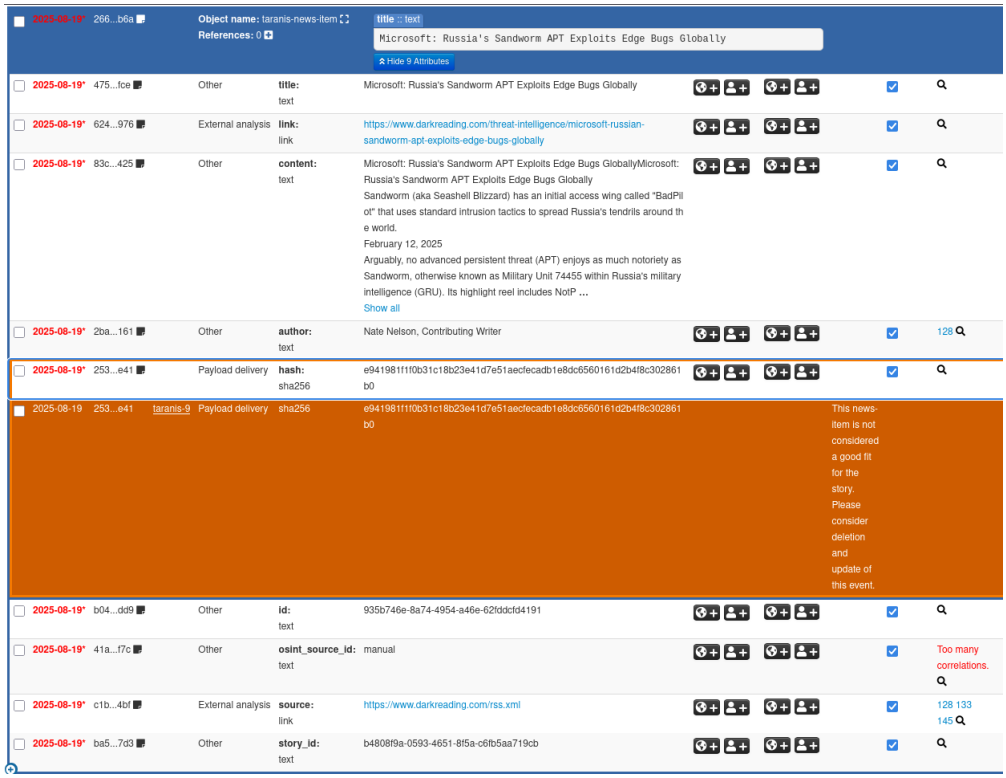


Figure 6.5.: A MISP proposal to remove a news item.

to the specific MISP event. To support the user, above the two text blocks are labels describing what side contains the local and external data (see figs. 6.8 to 6.10).

Conflict Resolution for News Item Conflicts

News item conflicts can be resolved in the section below the story conflicts. In this section, each incoming story with a conflict is displayed in a visually separated grey square. This grey square contains two columns of card-style blocks. Each card represents a single story. Here, in the left column is the incoming story card. The second column occupy the story cards that are present locally and in conflict with the story on the left. Each card with a story follows the same base layout. The story ID is at the top and below in a prominent font size is the story title that has a deep link to the internal story detail view. Below is a bullet point list of news item titles obtained in the story.

The internal story cards feature additional text and labels that help to add more context about them (see fig. 6.11). Next to the story ID, is the count of items and the value of the *relevance* field of the story. To the right of it are rendered labels conditioned to the story's characteristic. Label *MISP* informs the user that the story is present in MISP (or was at some point). Label *Has conflicting update* informs about this story being

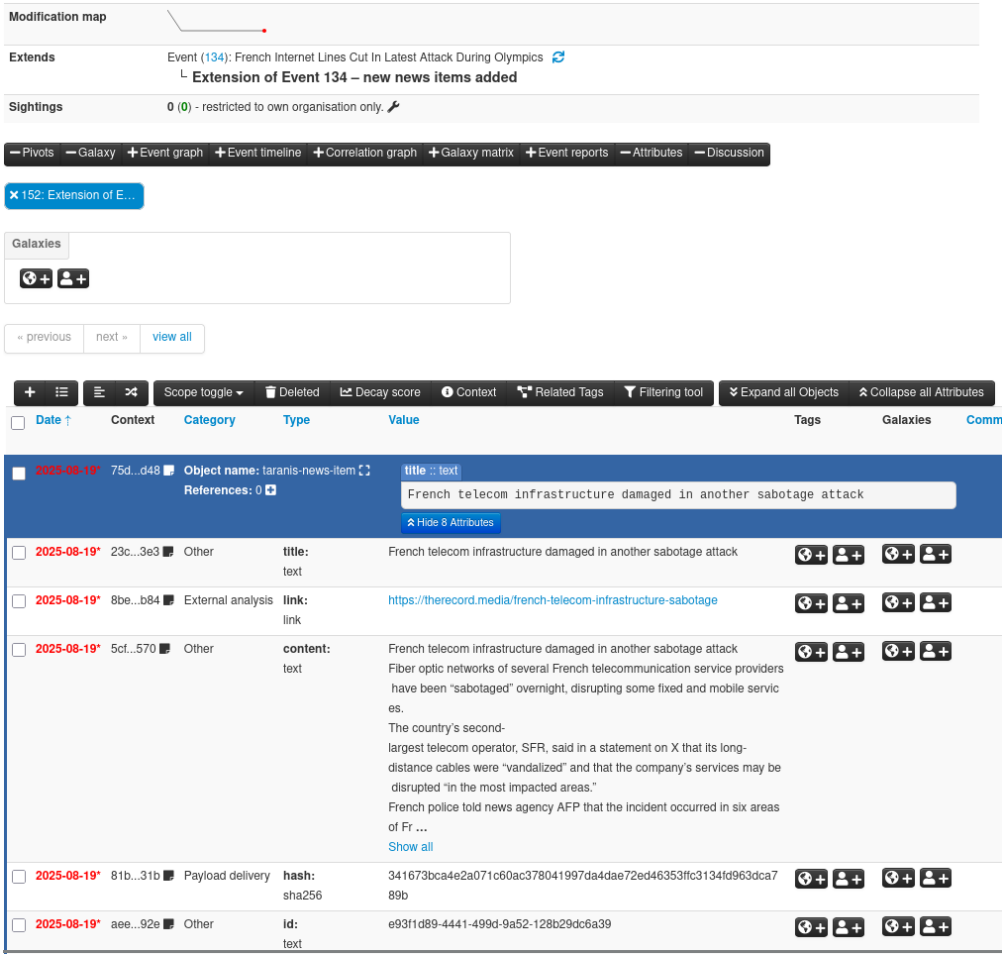


Figure 6.6.: A MISP extension event suggests that one news item should be added.

also in conflict in the story conflict section above. Finally, the label *Appears in multiple conflicts* informs the user about it being in conflict with more incoming stories.

The colours of text of news items title support the readability of the conflicts. The titles in the second column with the internal news items can be in two colours – red and black. The black colour is neutral, and means that the item does not cause a conflict. The red colour means that the news item is responsible for the news item conflict, because it comes also in the incoming story. The items that are coloured red (as conflict cause) in the right column, are coloured grey on the left one – in the incoming story. The incoming story features two other title colour options – black and blue. Black means, that the news item is completely new and is not present in any of the stories from the right column. The blue colouring informs the user that the news item is incoming within more incoming stories in conflict. All in all, this colour coding is useful for the analyst to more effectively understand the conflict (see fig. 6.12).

After a thorough research of the implications that conflict resolution flexibility has, two options to resolve

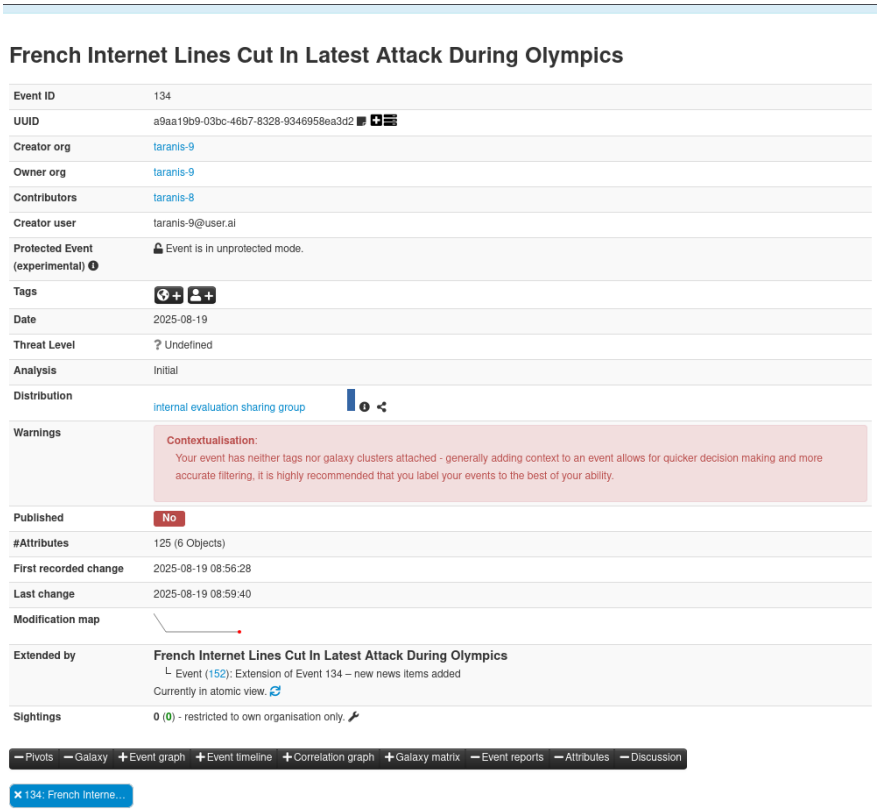


Figure 6.7.: A MISP event is extended.

conflict were derived. Each news item conflict therefore features two buttons – "*Ingest Incoming Story & Ungroup Local Stories*" and "*Keep Internal & Ingest Unique News Items*" button (see fig. 6.12). The former allows the user to ingest the incoming story fully, with all the items and with all the metadata (that is not directly visible in the conflict view, but can be accessed directly in MISP). It is useful if the internal stories do not provide a considerable value on their own, therefore reusing the clusters from partners is helpful. After a thorough research of the implications that flexibility causes in conflict resolution (described in detail in the chapter 7), two options to resolve conflict were derived. Each news item conflict therefore features two buttons – "*Ingest Incoming Story & Ungroup Local Stories*" and "*Keep Internal & Ingest Unique News Items*" button (see fig. 6.12). The former allows the user to ingest the incoming story fully, with all the items and with all the metadata (that is not directly visible in the conflict view, but can be accessed directly in MISP). It is useful if the internal stories do not provide a considerable value on their own, therefore reusing the clusters from partners is helpful.

In case the incoming story is in conflict with internal stories with high value (e.g. after a thorough analysis), the analyst can choose the button on the right (see fig. 6.12) to keep the stories on the right side. Only the unique news items, which are present in the incoming story, are ingested.

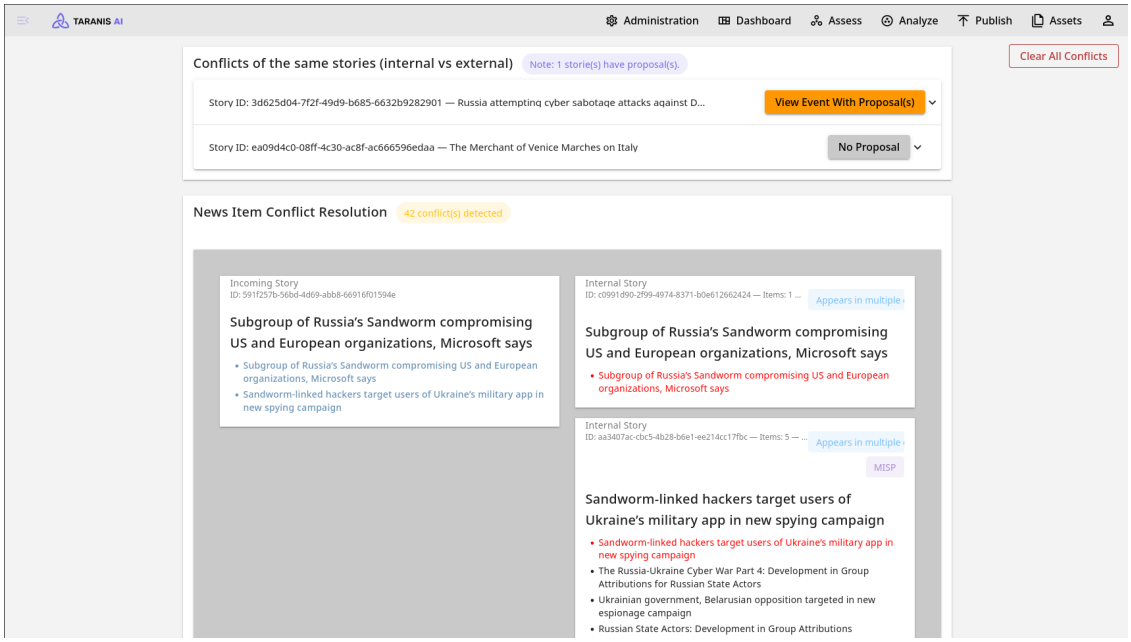


Figure 6.8.: Conflict view – overview.

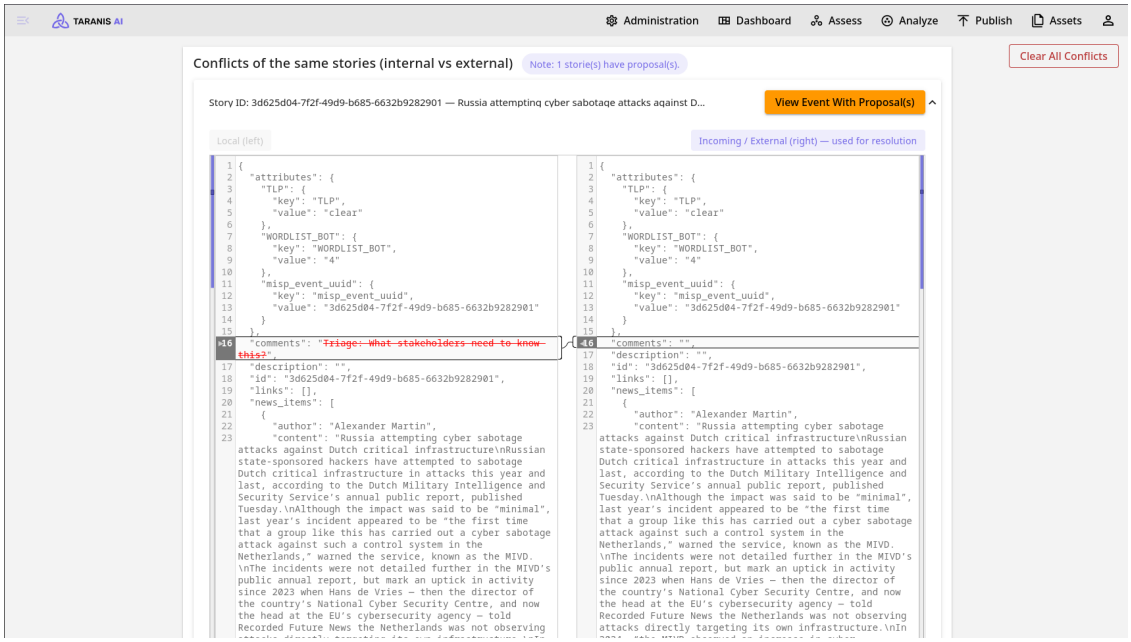


Figure 6.9.: Conflict view – Story conflict (I.)

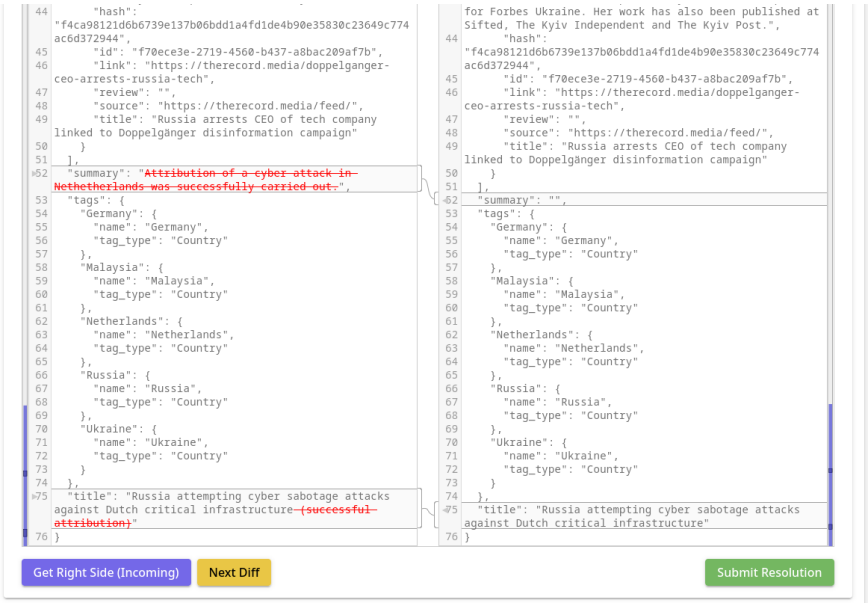


Figure 6.10.: Conflict view – Story conflict (II.)



Figure 6.11.: News item conflicts – column 2, internal story (detail with labels).

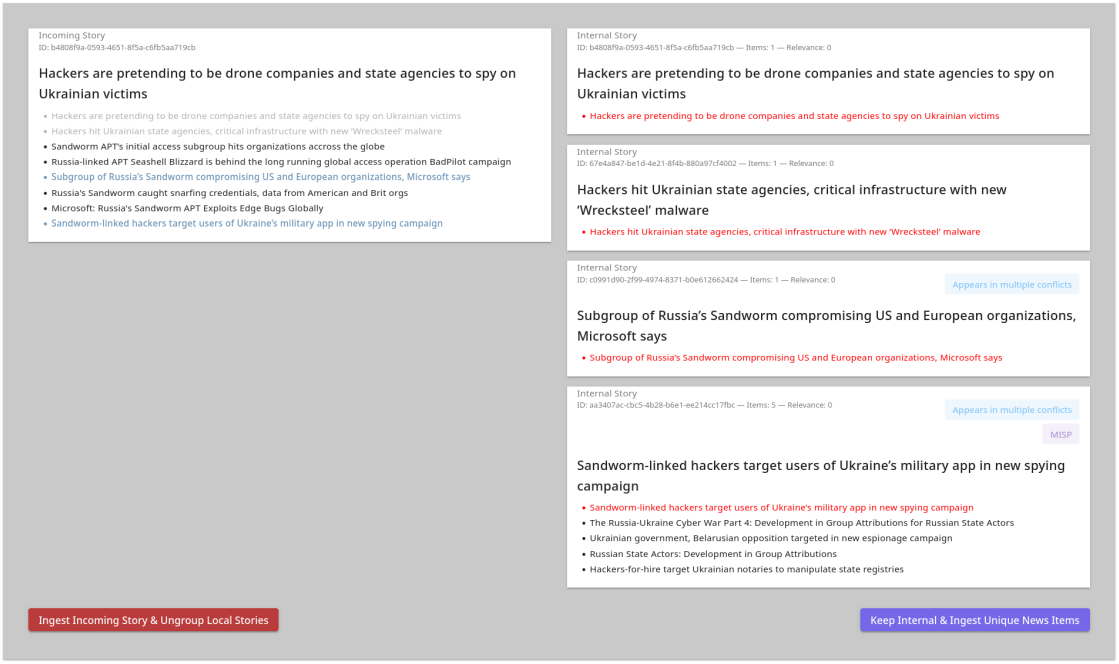


Figure 6.12.: Example of a news item conflicts with multiple internally conflicting stories.

7. Evaluation and Discussion

To evaluate the practical results of the thesis, two evaluation sessions with distinct participants were carried out. Each scenario-based exercise was accompanied by a hand out for the participants, which is attached in the appendix A.3. It served as a guide, and helped to explain the goals of the exercise and how to use the systems that were interacted with.

7.1. Evaluation Session Context

Both sessions took 120 minutes. In the first session 10 people participated and in the second session 11 people participated. The main goal was to gather a better understanding, what impact can collaboration have for CTI analysts. From the participants, 20 filled out a survey consisted of 7 engineers, 4 scientists, 4 project managers, 2 administrative workers, 1 software developer, and 1 Doctoral and 1 Master student.
The role-played scenario by the users was, that each participant was an individual organization that focuses on CTI production, and together they would explore cross-organization collaboration. To realize that, participants had access to their own instance of the OSINT platform Taranis AI. Each instance was pre-populated with OSINF data. Across all instances, 3 different sets of datasets were used, to ensure, some commonalities in the datasets, but also to allow some participants access to data, that was not accessible everyone. At the start, each participant started with analysis of their own data. The end of the first part resulted in participants (and each organization) to have their own understanding, and the results of it were sent to the shared MISP instance.
In the second part the users explored the collaborative character of CTI analysis. The participants were ingesting results from other participants. This exposed them to new stories in their Taranis AI instance. Alongside the new stories ingested, many participants faced conflicts in their incoming data and progressed to resolve them.
All in all, this collaborative environment allowed them to, potentially, gain more insights from other participants, and use this to their own advantage. Either by getting new insights, or finding similarities and new information clusters (stories). After each part, participants filled out a survey about their experience. The

survey was pseudo-anonymous and only metaquestions about participants' profile were asked, e.g. years of experience, main professional role, etc.

7.2. Results

This section describes the tangible results that stem from the practical evaluation of the implementation. Moreover, this section answers the research questions this thesis identified. The practical output, providing a resource for validation of the research and design, is referenced in the appendix A.5.

7.2.1. User Evaluation Results

The most important result that can be derived from the two sessions is that the practical output of this thesis is functional and participants were able to fulfil the use cases defined in the chapter 3. Furthermore, the practical sessions resulted in rich feedback on the usability of the solution.
According to the feedback, it is deduced, that the implementation is not perfect, but the field of research is valid. The participants reported positive feedback on the idea of collaboration, and acknowledged, they were able to gain information thanks to the analyses shared by other participants.
The downsides reported concerned the most the deficiency of the conflict resolution view. The User Experience (UX) was reportedly considerably affected when the volume of conflicts (reaching more than 70 conflicting incoming stories) reached higher numbers. It was recognized, that the ability to select multiple conflicts to resolve in bulk, or the ability to sort conflicts would improve the analysts experience and efficiency.

User Feedback Interpretation

According to the 20 participants who answered the survey, results can be interpreted. The weight of the interpretations can be debatable, as some answers are inconclusive as either data is insufficient, sample size is limited or the answers could go against each other. A subset of plots showcasing the results of the thesis are in the appendix A.4.
So even though, the overall experience when transitioning (from individual analysis) to collaboration was *disruptive* to 43% of participants and [neutral] to another 43% of participants, 86% *noticed enhancements* in their understanding of the threat landscape (see fig. A.1f).
Moreover, on one hand, 40% of participants perceived the information received from others as *somewhat useful*, and, on the other hand 40% perceived it as *somewhat unuseful* (see fig. A.1d). Despite this overall

result, comparing the results of first, and second group, 10 participants (7 – *somewhat useful* and 3 neutral) out of 11 claimed considerably more optimistic experience. Therefore, these results could suggest, that the second group performed better on the level of creating threat intelligence. This suggestion also correlates with the number of people that took part and have more than 5 years of professional experience. During the first session, only 2 participants had more than 5 years of experience, in contrast with 8 in the second testing group. However, to really understand the reasons for worse intelligence shared, more research would be needed.

From the point, how the received clusters were used, the participants claim that 24% of them *built mostly upon clusters from peers*. 33% of them claim to *built mostly upon their own clusters*, and 43% of them used an approach that combined a balanced approach of both (see fig. A.1c). These findings are evenly distributed, and do not draw any conclusive suggestions, rather than recognizing the fact, that user workflows vary case by case.

Additionally, the user interface for conflict resolution was perceived as rather problematic, where a considerably big group of participants (50%) perceived it as ineffective (see pie2).

Conversely, 57% claimed confidence, that if they were CTI analysts, they could believe that a tool like this could be part of their professional routine (see fig. A.2a).

To conclude these numbers, it could be said that the implemented prototype worked, yet there would be much more work to it to achieve an enjoyable experience. Last but not least, the results also suggest, that a tool enabling collaboration of strategic CTI could complement the workflows of analysts to gain a better situational awareness.

7.2.2. Research Results

Condensing the knowledge gained in the course of the research of this research, the two research questions are answered as follows:

- **RQ1 – To what extent does MISP (as a central node) and Taranis AI enable an interoperable collaboration on strategic CTI. Can MISP platform provide enough flexibility to its standard technical CTI focus?**

  Considering the practical PoC implementation it is concluded, that the combination of MISP and Taranis AI are feasible candidates providing a functional solution for strategic CTI collaboration. The chapter 5 describes what is the feasible mapping of functional and logical mapping of the different data models to enable collaboration and to convey the most crucial parts.

  Additionally, the usage of the MISP platform implicitly allows an interoperable environment, because

anyone with the access to MISP can adopt a custom solution to enter, ingest or contribute to shared CTI.

- **RQ2 – What are the challenges of resolving conflicts, and how can MISP and Taranis AI platform support a more smooth operation?**

  Conflicts and their complexity during collaboration can have a significant impact.

  The conflicts can be caused even just by sending overlapping data from one party to another. Therefore, the opportunity and overflow of conflicts seems like a real complication. The conflicts stemming from co-authoring, require a high flexibility of the resolution option. Therefore, the granular difference checker UI is suitable. However, the higher flexibility is not always an advantage. Especially, considering the user feedback from the survey, with the higher number of conflicts the process resolving them can become tiring and error-prone. This is the reason the section 6.1.2 proposes a more simplified resolution UI for news item conflicts. With this approach, the cognitive load on the analyst is decreased as only two options to resolve a conflict are available (in comparison to the story conflict UI). This approach circumvents another issue — incoherent resolutions. If more flexibility is allowed, it is more likely to end up in mixed states by accident, and thus decrease the value of receiving information. When removing a single news item from a story would make the whole story with its metadata hard to interpret and create confusion, it is not good to remove it.

  In the MISP platform conflicting data can appear. Mainly, the owners of the events make sure what proposals from others are accepted, and what is not. Even though, the current PoC employs the usage of the extended events. If new objects could be proposed to events, and extended events were replaced by object proposals, control, and understanding over the data and conflicts would be better.

- **RQ3 – What are the key characteristics that can significantly improve experience of CTI analysts during collaboration?**

  This thesis explored several aspects that can improve experience of CTI analysts working in a collaborative environment. From the collected feedback discussed in section 7.2.1, it can be concluded, that from the analyst view, it is important when the tool tries to minimize the amount of overhead work during day-to-day tasks of collaboration. In specific, during conflict resolution, allowing the analysts to focus on the differences. Another aspect of significance is the ability to adjust the view to the context the analyst is interested in, e.g. conflict sorting feature in a specific timeframe, or ignore and mark as read feature.

  Furthermore, in the light of problematics of context-switching, minimizing the amount of context switches when analysts need to change tools can also positively impact the experience. Especially in

a setting when more tools are used to achieve a goal (exactly the way this thesis does with MISP and Taranis AI), developers might be prone to take shortcuts and keeping things only good enough. This can result into some details not being accessible where analysts need it.

Last but not least, the ability to exchange ideas and discuss threats with the collaborators can have a significant impact on the efficiency of the experience.

7.3. Limitations and Discussion of Current Work

The thesis scope only discussed the fundamental baseline of collaboration in the information sharing domain. Further research on various topics is needed.

7.3.1. TLP Protocol

Even though TLP protocol is a cornerstone in information sharing, in this scope, it was not directly considered. TLP has significance in regulating, what information can reach what entities. Once information is delivered, the handler should then comply with the TLP.

Similarly, in this case study. Even though, Taranis AI uses TLP, there are no support mechanisms that would support the analyst while collaborating (e.g. showing TLP level in collaborative stories), or does not forbid or warn the user when information with highly restrictive TLP is shared. Improvements in this regard have potential to decrease the cognitive load of analysts, or support analysts during days with high overload.

7.3.2. Conflict Resolution

The section 6.1.2 described how the conflict resolution was implemented. As described, the story conflicts resolution section adopts a view that uses a simple comparison of two texts next to each other adding highlighting to visualize differences. Even though the rendered contents are after filtration, the user needs to work with views long texts and sees parts without changes. To have an option to allow the user to only see the differences in a condensed, more user-friendly way could improve the experience.

Furthermore, news items conflicts provide only two options to resolve conflicts (see section 6.1.2). Considering a scenario – there is a news item conflict with an internal story that has news items A, B, C and D. The incoming story comes with news items A, B, C, but news item D is not included. The analyst realizes the news item D is incorrectly assigned to the story internally, and wants to remove it. If the conflict resolution was more flexible, a user could just remove the news item D from the internal cluster to resolve that conflict.

This could potentially damage the context of the story. It could still be, that the story references the news item D in other properties (e.g. description, summary, etc.).
To avoid cases like these, the conflict resolution was implemented in a more rigid way. On the other hand, according to the user evaluation, this was perceived as limiting and would require a more detailed analysis of this area to design a better solution.

7.3.3. MISP Integration Degree

In this thesis, it was fundamental to explore whether it is viable to use a CTI plarform that is mainly used for sharing more technical CTI like MISP as a platform for sharing strategic CTI. As a result of that, only features that were requirements to allow collaboration were implemented. Researching, what are other features that could improve collaboration is needed. There is a vast room for it, e.g. in the direction of implementing a chat-feature, potentially by integrating the Event Discussion Thread[1].

7.3.4. Sharing Circles Limitations

The PoC implementation considers a simplified setting where one group of partners collaborate. This is a scenario where a single organization pushes only to a single distribution level (or a sharing group) that MISP recognizes[2]. Once one Taranis AI instance would push to multiple MISP instances or different distribution levels, the system would have issues. Either technical limitations, that are present in the PoC (e.g. unable to share an event to two different instances), or sharing an event to different sharing groups (which is not possible on MISP level unless the even is “duplicated”).

7.3.5. Limited Collaboration Ability

The PoC is unable to allow partner organizations to contribute additional tags, attributes, and links to originators. It was scoped out from the PoC implementation as it was not foreseen as a hard requirement.

7.3.6. Future work

This thesis demonstrates what it can look like when a threat sharing platform built for technical CTI is used for sharing and collaboration with strategic CTI. Given the findings, there is a certain level of applicability and usability. What it, however, shows too is that for a software to be production ready, much more needs to be done.

[1] https://www.circl.lu/doc/misp/using-the-system/#general-event-information
[2] https://www.circl.lu/doc/misp/using-the-system/#creating-an-event

There are several areas that need more research. While it is natural, that a collaborative system encounters data conflicts, during this thesis it has been one of the most problematic areas. There is definitely a room for research of what are the suitable methods to be offered to CTI analysts, so that their work can be efficient and supports their natural workflows and needs.

Moreover, this thesis has mostly discussed what are the opportunities and needs on the side of the software the analysts use for strategic CTI (Taranis AI). As a result of that, there is a certain room for research on the area, *if* and *how* could CTI platforms for technical CTI, like MISP, support such a scenario, that also strategic CTI would be handled by them.

8. Conclusion

With the rise of WWW has the world and the threats, which the world faces, changed. The world became much more globalized and interconnected. This has brought many advantages but also disadvantages in terms of threats.

Nowadays, organizations commonly use shared databases of malicious software to fight against computer threats. The fact that organizations can profit when they share threat information among themselves is also known in today's world. However, this is a problematic topic with not only security but also business continuity, as sharing critical information about organization's breaches is very sensitive.

One of the key and rising aspects of the fight against cyber-attacks is the usage of OSINT. OSINT has become a powerful tool, especially, since there is so much information on today's internet that is free to access and provide the base OSINF.

Currently, not only nation states, but also organizations take advantage of OSINT. It is common to subscribe to OSINT information feeds. However, in the steps before producing intelligence feeds stand the CTI analysts and work towards actionable intelligence that will help face trending threats.

This CTI analysis is a complex task and involves a lot of research. In the year 2025 in Europe, with the new NIS2, there is a strong incentive to share threat information. This thesis addressed the area of how could the open-source platforms MISP and Taranis AI work together to allow not only sharing, but also collaboration on cyber threat intelligence.

In the course of this thesis, a requirements analysis was carried out that explored the needed use cases to achieve collaboration. Important information properties of the Taranis AI data model were mapped to the data model of MISP. In the scope of this thesis, a PoC implementation was also programmed.

The PoC implemented in the course of this thesis was also officially released in the version 1.2.4[1] of Taranis AI. Moreover, two groups of users tested the PoC implementation and contributed essential feedback. As a result of that, it can be concluded that the MISP platformis able to support collaboration among the Taranis AI instances, however, only within a limited scope of this thesis.

[1] https://github.com/taranis-ai/taranis-ai/releases/tag/1.2.4

All in all, this thesis describes the minimum baseline for collaboration on strategic CTI while usage of two open-source platforms, but also highlights that some areas, like conflict handling and resolution circles represent complex topics that need further research.

List of Figures

2.1 Cyber Threat Intelligence Lifecycle 9
2.2 Entity Relationship Diagram: Taranis AI internal data model scheme. 16

3.1 Use case diagram: CTI Analyst shares a story after an assessment complemented by fine-grained adjustments. 17
3.2 Use case diagram: CTI Analyst ingests a story, verifies potential upstream change requests and resolves conflicts. 18
3.3 Use case diagram: CTI Analyst updates a story with story metadata changes or news item changes. 18
3.4 Use case diagram: CTI Analyst proposes changes to a story originating from a partner. . . . 19
3.5 Use case diagram: CTI Analyst proposes changes on the news item level to a story originating from a partner. 19
3.6 Use case diagram: CTI Analyst proposes changes on the news item level to a story originating from a partner. 20
3.7 Organizations using Taranis AI connect through MISP and collaborate. 21
3.8 Sequence diagram displays analysts from two organizations sending and receiving data. . . . 21
3.9 Collaborating organization proposes a change request to the originator. 22

4.1 Federated Hub-and-Spoke Architecture 26

6.1 MISP Event report with a description and summary properties extracted from a Taranis AI story. 37
6.2 Share a story to MISP – story action button. 38
6.3 Share a story to MISP popup, choose a Connector. 39
6.4 An attribute change proposal was submitted. 40
6.5 A MISP proposal to remove a news item. 41
6.6 A MISP extension event suggests that one news item should be added. 42

6.7 A MISP event is extended. 43
6.8 Conflict view – overview. 44
6.9 Conflict view – Story conflict (I.) 44
6.10 Conflict view – Story conflict (II.) 45
6.11 News item conflicts – column 2, internal story (detail with labels). 45
6.12 Example of a news item conflicts with multiple internally conflicting stories. 46

List of Listings

1 Entry method of MISP Collector 36

Acronyms

API	Application Platform Interface
APTs	Advanced Persistent Threats
BPEL	Business Process Execution Language
CIRCL	Computer Incident Response Center Luxembourg
CSV	Comma Separated Values
CTI	Cyber Threat Intelligence
ENISA	European Union Agency for Cybersecurity
HTTP	Hypertext Transfer Protocol
HTTPS	Hypertext Transfer Protocol Secure
IAM	Identity and Access Management
ID	identifier
IDS	Intrusion Detection Systems
IoCs	Indicators of Compromise
IP	Internet Protocol
IPS	Intrusion Prevention Systems
ISACs	Information Sharing and Analysis Centers
IT	Information Technology
JSON	JavaScript Object Notation

LLM	Large Language Models
MISP	Threat Intelligence Sharing Platform
mTLS	Mutual Transport Layer Security
NIS	Network and Information Security Directive
NIS2	Network and Information Security Directive 2
OSINF	Open-source Information
OSINT	Open-source Intelligence
PoC	Proof of Concept
RBAC	Role-based Access Control
SA	Situational Awareness
SaaS	Software as a Service
STIX	Structured Threat Information eXpression
SW	software
TDK	Threat and Defence Knowledge
TI	Threat Intelligence
TINF	Threat Information
TLP	Traffic Light Protocol
TTI	Technical Threat Intelligence
TTPs	Tactics, Techniques, and Procedures
UI	User Interface
URL	Uniform Resource Locator

UX	User Experience
WWW	World Wide Web
XML	Extensible Markup Language

Bibliography

[1] Yinghai Zhou, Yi Tang, Ming Yi, Chuanyu Xi, and Hai Lu, "Cti view: Apt threat intelligence analysis system," *Security and Communication Networks*, vol. 2022, no. 1, p. 9 875 199, 2022. DOI: https://doi.org/10.1155/2022/9875199. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1155/2022/9875199. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1155/2022/9875199.

[2] Wiem Tounsi, "What is cyber threat intelligence and how is it evolving?" In *Cyber-Vigilance and Digital Trust: Cyber Security in the Era of Cloud Computing and IoT*. 2019, pp. 1–49. DOI: 10.1002/9781119618393.ch1.

[3] Quirine Eijkman and Daan Weggemans, "Open source intelligence and privacy dilemmas: Is it time to reassess state accountability," *Sec. & Hum. Rts.*, vol. 23, p. 285, 2012.

[4] Bert-Jaap Koops, Jaap-Henk Hoepman, and Ronald Leenes, "Open-source intelligence and privacy by design," *Computer Law & Security Review*, vol. 29, no. 6, pp. 676–688, 2013, ISSN: 0267-3649. DOI: https://doi.org/10.1016/j.clsr.2013.09.005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0267364913001672.

[5] Onur Catakoglu, Marco Balduzzi, and Davide Balzarotti, "Automatic extraction of indicators of compromise for web applications," in *Proceedings of the 25th International Conference on World Wide Web*, ser. WWW '16, Montréal, Québec, Canada: International World Wide Web Conferences Steering Committee, 2016, pp. 333–343, ISBN: 97814-50341431. DOI: 10.1145/2872427.2883056. [Online]. Available: https://doi.org/10.1145/2872427.2883056.

[6] Antonio Villalón-Huerta, Ismael Ripoll-Ripoll, and Hector Marco-Gisbert, "Key requirements for the detection and sharing of behavioral indicators of compromise," *Electronics*, vol. 11, no. 3, 2022, ISSN: 2079-9292. DOI: 10.3390/electronics11030416. [Online]. Available: https://www.mdpi.com/2079-9292/11/3/416.

[7] Fernando Maymí, Robert Bixler, Randolph Jones, and Scott Lathrop, "Towards a definition of cyberspace tactics, techniques and procedures," in *2017 IEEE International Conference on Big Data (Big Data)*, 2017, pp. 4674–4679. DOI: 10.1109/BigData.2017.8258514.

[8] Yizhe You, Jun Jiang, Zhengwei Jiang, Peian Yang, Baoxu Liu, Huamin Feng, Xuren Wang, and Ning Li, "Tim: Threat context-enhanced ttp intelligence mining on unstructured threat data," *Cybersecurity*,

vol. 5, no. 1, p. 3, Feb. 2022, ISSN: 2523-3246. DOI: 10.1186/s42400-021-00106-5. [Online]. Available: https://doi.org/10.1186/s42400-021-00106-5.

[9] Randy Borum, John Felker, Sean Kern, Kristen Dennesen, and Tonya Feyes, "Strategic cyber intelligence," *Information and Computer Security*, vol. 23, no. 3, pp. 317–332, Jul. 2015, ISSN: 2056-4961. DOI: 10.1108/ICS-09-2014-0064. eprint: https://www.emerald.com/ics/article-pdf/23/3/317/730073/ics-09-2014-0064.pdf. [Online]. Available: https://doi.org/10.1108/ICS-09-2014-0064.

[10] Pedro Santos, Rafael Abreu, Manuel J. C. S. Reis, Carlos Serôdio, and Frederico Branco, "A systematic review of cyber threat intelligence: The effectiveness of technologies, strategies, and collaborations in combating modern threats," *Sensors*, vol. 25, no. 14, 2025, ISSN: 1424-8220. DOI: 10.3390/s25144272. [Online]. Available: https://www.mdpi.com/1424-8220/25/14/4272.

[11] Florian Skopik, Giuseppe Settanni, and Roman Fiedler, "A problem shared is a problem halved: A survey on the dimensions of collective cyber defense through security information sharing," *Computers & Security*, vol. 60, pp. 154–176, 2016, ISSN: 0167-4048. DOI: https://doi.org/10.1016/j.cose.2016.04.003. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167404816300347.

[12] *Directive - 2022/2555 - EN - EUR-Lex — eur-lex.europa.eu*, https://eur-lex.europa.eu/eli/dir/2022/2555/oj/eng, [Accessed 07-09-2025].

[13] Zulaikha Amirah Johari, Aziatul Waznah Ghazali, Yusarina Mat Isa, Nur Aima Shafie, and Soliha Sanusi, *Digital disruption and cybersecurity threats: Redefining the role of internal auditing*, 2023. DOI: 10.15405/epsbs.2023.11.65.

[14] Sara Qamar, Zahid Anwar, Mohammad Ashiqur Rahman, Ehab Al-Shaer, and Bei-Tseng Chu, "Data-driven analytics for cyber-threat intelligence and information sharing," *Computers & Security*, vol. 67, pp. 35–58, 2017, ISSN: 0167-4048. DOI: https://doi.org/10.1016/j.cose.2017.02.005. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167404817300287.

[15] Jan M. Ahrend, Marina Jirotka, and Kevin Jones, "On the collaborative practices of cyber threat intelligence analysts to develop and utilize tacit threat and defence knowledge," in *2016 International Conference On Cyber Situational Awareness, Data Analytics And Assessment (CyberSA)*, 2016, pp. 1–10. DOI: 10.1109/CyberSA.2016.7503279.

[16] Wiem Tounsi and Helmi Rais, "A survey on technical threat intelligence in the age of sophisticated cyber attacks," *Computers & Security*, vol. 72, pp. 212–233, 2018, ISSN: 0167-4048. DOI: https://doi.org/10.1016/j.cose.2017.09.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167404817301839.

[17] Christopher Johnson, Mark Badger, David Waltermire, Julie Snyder, and Clem Skorupka, *Guide to cyber threat information sharing*, en, Oct. 2016. DOI: https://doi.org/10.6028/NIST.SP.800-150.

[18] Henry Dalziel, *How to define and build an effective cyber threat intelligence capability*. Syngress, 2014.

[19] Bank of England, *CBEST Intelligence-Led Testing*, https://www.bankofengland.co.uk/-/media/boe/files/financial-stability/financial-sector-continuity/understanding-cyber-threat-intelligence-operations.pdf, [Accessed 08-06-2025].

[20] Sarah Brown, Joep Gommers, and Oscar Serrano, "From cyber security information sharing to threat management," in *Proceedings of the 2nd ACM Workshop on Information Sharing and Collaborative Security*, ser. WISCS '15, Denver, Colorado, USA: Association for Computing Machinery, 2015, pp. 43–49, ISBN: 9781450338226. DOI: 10.1145/2808128.2808133. [Online]. Available: https://doi.org/10.1145/2808128.2808133.

[21] Thomas D. Wagner, Esther Palomar, Khaled Mahbub, and Ali E. Abdallah, "A novel trust taxonomy for shared cyber threat intelligence," *Security and Communication Networks*, vol. 2018, no. 1, p. 9 634 507, 2018. DOI: https://doi.org/10.1155/2018/9634507. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1155/2018/9634507. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1155/2018/9634507.

[22] Alessandra de Melo e Silva, João José Costa Gondim, Robson de Oliveira Albuquerque, and Luis Javier García Villalba, "A methodology to evaluate standards and platforms within cyber threat intelligence," *Future Internet*, vol. 12, no. 6, 2020, ISSN: 1999-5903. DOI: 10.3390/fi12060108. [Online]. Available: https://www.mdpi.com/1999-5903/12/6/108.

[23] Mauro Conti, Tooska Dargahi, and Ali Dehghantanha, "Cyber threat intelligence: Challenges and opportunities," in *Cyber Threat Intelligence*, Ali Dehghantanha, Mauro Conti, and Tooska Dargahi, Eds. Cham: Springer International Publishing, 2018, pp. 1–6, ISBN: 978-3-319-73951-9. DOI: 10.

1007/978-3-319-73951-9_1. [Online]. Available: https://doi.org/10.1007/978-3-319-73951-9_1.

[24] Katja Henttonen and Jyri Rajamäki, "CTI sharing practices and MISP adoption in finland's critical infrastructure protection," *Proc. Eur. Conf. Inf. Warf. Secur.*, vol. 23, no. 1, Jun. 2024.

[25] MISP, *MISP Communities and MISP Feeds — misp-project.org*, https://www.misp-project.org/communities/, [Accessed 07-05-2025].

[26] MISP Contributors, *User guide of MISP intelligence sharing platform — circl.lu*, https://www.circl.lu/doc/misp/, [Accessed 07-03-2025].

[27] Cynthia Wagner, Alexandre Dulaunoy, Gérard Wagener, and Andras Iklody, "Misp: The design and implementation of a collaborative threat intelligence sharing platform," in *Proceedings of the 2016 ACM on workshop on information sharing and collaborative security*, 2016, pp. 49–56.

[28] Borce Stojkovski, Gabriele Lenzini, Vincent Koenig, and Salvador Rivas, "What's in a cyber threat intelligence sharing platform? a mixed-methods user experience investigation of misp," in *Proceedings of the 37th Annual Computer Security Applications Conference*, 2021, pp. 385–398.

[29] Robert Dover, Michael S Goodman, and Claudia Hillebrand, Eds., *Routledge companion to intelligence studies*. London, England: Routledge, Sep. 2013.

[30] Adam Zibak and Andrew Simpson, "Cyber threat information sharing: Perceived benefits and barriers," in *Proceedings of the 14th International Conference on Availability, Reliability and Security*, ser. ARES '19, Canterbury, CA, United Kingdom: Association for Computing Machinery, 2019, ISBN: 9781450371643. DOI: 10.1145/3339252.3340528. [Online]. Available: https://doi.org/10.1145/3339252.3340528.

[31] Thomas D. Wagner, Khaled Mahbub, Esther Palomar, and Ali E. Abdallah, "Cyber threat intelligence sharing: Survey and research directions," *Computers & Security*, vol. 87, p. 101 589, 2019, ISSN: 0167-4048. DOI: https://doi.org/10.1016/j.cose.2019.101589. [Online]. Available: https://www.sciencedirect.com/science/article/pii/%5CS016740481830467X.

[32] Florian Skopik, Daniel Schall, and Schahram Dustdar, "Modeling and mining of dynamic trust in complex service-oriented systems," *Information Systems*, vol. 35, no. 7, pp. 735–757, 2010, ISSN: 0306-4379. DOI: https://doi.org/10.1016/j.is.2010.03.001. [On-

line]. Available: https://www.sciencedirect.com/science/article/pii/S0306437910000153.

[33] Hanine Tout, Azzam Mourad, Hamdi Yahyaoui, Chamseddine Talhi, and Hadi Otrok, "Towards a bpel model-driven approach for web services security," in *2012 Tenth Annual International Conference on Privacy, Security and Trust*, 2012, pp. 120–127. DOI: 10.1109/PST.2012.6297928.

[34] Vitor Jesus, Balraj Bains, and Victor Chang, "Sharing is caring: Hurdles and prospects of open, crowd-sourced cyber threat intelligence," *IEEE Transactions on Engineering Management*, vol. 71, pp. 6854–6873, 2024. DOI: 10.1109/TEM.2023.3279274.

[35] *Information Sharing and Analysis Centers (ISACs) | ENISA — enisa.europa.eu*, https://www.enisa.europa.eu/topics/cybersecurity-of-critical-sectors/information-sharing-and-analysis-centers-isacs, [Accessed 07-08-2025].

[36] *Regulation - EU - 2025/38 - EN - EUR-Lex — eur-lex.europa.eu*, https://eur-lex.europa.eu/eli/reg/2025/38/oj/eng, [Accessed 12-08-2025].

[37] Florian Skopik and Benjamin Akhras, *Taranis AI: Applying Natural Language Processing for Advanced Open-Source Intelligence Analysis — ercim-news.ercim.eu*, https://ercim-news.ercim.eu/en136/r-i/taranis-ai-applying-natural-language-processing-for-advanced-open-source-intelligence-analysis, [Accessed 09-08-2025].

[38] *Taranis AI — taranis.ai*, https://taranis.ai, [Accessed 07-09-2025].

[39] *CIRCL & BB; MISP - Open Source Threat Intelligence Platform — circl.lu*, https://www.circl.lu/services/misp-malware-information-sharing-platform/, [Accessed 08-08-2025].

[40] Okeoghene Odudu, "Indirect information exchange: The constituent elements of hub and spoke collusion," *European Competition Journal*, vol. 7, no. 2, pp. 205–242, 2011. DOI: 10.5235/174410511797248324. eprint: https://doi.org/10.5235/174410511797248324. [Online]. Available: https://doi.org/10.5235/174410511797248324.

A. Appendix

A.1. MISP Object Mapping Taranis AI news item

```
{
    "attributes": {
        "title": {
            "description": "Title of the Taranis news item",
            "misp-attribute": "text",
            "ui-priority": 0
        },
        "link": {
            "description": "URL address to the Taranis news item source.",
            "misp-attribute": "link",
            "ui-priority": 0
        },
        "content": {
            "description": "Content of the Taranis news item",
            "misp-attribute": "text",
            "multiple": true,
            "ui-priority": 0
        },
        "author": {
            "description": "Author(s) of the news item",
            "misp-attribute": "text",
            "multiple": true,
            "ui-priority": 0
        },
        "hash": {
            "description": "Hash of the news item",
```

```
            "misp-attribute": "sha256",
            "ui-priority": 0
        },
        "id": {
            "description": "Unique identifier of the Taranis item",
            "misp-attribute": "text",
            "ui-priority": 0
        },
        "language": {
            "description": "Language of the news item",
            "misp-attribute": "text",
            "ui-priority": 0
        },
        "osint_source_id": {
            "description": "OSINT source unique identifier",
            "misp-attribute": "text",
            "ui-priority": 0
        },
        "source": {
            "description": "Source feed or page that provided the news
            ↪  item",
            "misp-attribute": "link",
            "ui-priority": 0
        },
        "review": {
            "description": "Review or analysis notes of the news item",
            "misp-attribute": "text",
            "ui-priority": 0
        },
        "story_id": {
            "description": "Unique identifier for the story",
            "misp-attribute": "text",
            "ui-priority": 0
        }
    },
    "description": "An object describing a news item from Taranis AI",
```

```
    "meta-category": "misc",
    "name": "Taranis AI News Item",
    "uuid": "e4b7e5e3-2c4c-4edf-b9a3-2c2a32844e62",
    "version": 2
}
```

A.2. MISP Object Mapping Taranis AI story

```
{

    "attributes": {
        "comments": {
            "description": "User comments for the story",
            "misp-attribute": "text",
            "ui-priority": 0
        },
        "description": {
            "description": "Longer descriptive text of the story",
            "misp-attribute": "text",
            "ui-priority": 0
        },
        "id": {
            "description": "Unique identifier of the story",
            "misp-attribute": "text",
            "ui-priority": 0
        },
        "links": {
            "description": "One or more links associated with the story",
            "misp-attribute": "text",
            "multiple": true,
            "ui-priority": 0
        },
        "summary": {
            "description": "Short summary of the story",
```

```
27              "misp-attribute": "text",
28              "ui-priority": 0
29          },
30          "tags": {
31              "description": "Tags or labels for the story",
32              "misp-attribute": "text",
33              "multiple": true,
34              "ui-priority": 0
35          },
36          "title": {
37              "description": "Title of the story",
38              "misp-attribute": "text",
39              "ui-priority": 0
40          },
41          "attributes":{
42              "description": "Attributes of the story",
43              "misp-attribute": "text",
44              "multiple": true,
45              "ui-priority": 0
46          }
47      },
48      "name": "Taranis AI Story",
49      "uuid": "a2716df6-920b-4c4e-85f7-79a009476f3d",
50      "description": "An object describing a story item from Taranis or
        ↪  similar source",
51      "meta-category": "misc",
52      "version": 1
53  }
54
```

A.3. Assignment for Users for Practical Evaluation

AIT – Center for Digital Safety and Security Security and Communication Technologies Unit

Evaluation of a CTI Collaborative Environment

authored by: Bc. Peter Leitmann
coordinated by: Dr. Dr. Florian Skopik

Goal

- To evaluate **effects** and **feasibility** of CTI collaborative environments.

Hardware Required

- a computer

Software Required

- Taranis AI
  - Each group has one instance accessible and assigned.
    - https://taranis-[group_number].[redacted].[redacted]
    - group_number: presented in the lab
      - username: *sent per request*
      - password: *sent per request*
- Threat Intelligence Sharing Platform (MISP)
  - Everyone reuses the same MISP instance with different users for simulation simplicity purposes.
    - https://[redacted].[redacted]
      - username: *sent per request*
      - password: *sent per request*

1 Background

This is a lab exercise to evaluate the usability and feasibility of collaborative Cyber Threat Intelligence (CTI) environments. The use of intelligence in general has deemed useful in many areas, especially in warfare or police investigations.

In today's world, investigations or analyses in the Internet environment takes place too. Several entities participate in such analyses. They either occur as a precaution against cyberattacks, in parallel during attacks, or are done after cyberattacks – so-called post-mortem analyses. These activities are often carried out by specialized personnel or whole Security Operation Center (SOC) teams of organizations, or so-called Computer Emergency Response Team (CERT) and Computer Security Incident Response Team (CSIRT) teams.

In the following activity, we will dive deep into Open-source Intelligence (OSINT). OSINT utilizes the disclosed information as new articles, social media accounts and many more. OSINT is the digested and actionable information gained from Open-source Information (OSINF) (the raw data is not intelligence) and developed for a certain stakeholder (e.g., client, organization, country). The important characteristic is the applicability and actionability of the individual information.

Your role – SOC Analyst

You will embody a SOC analyst. In the background, you have already gained a bit of an understanding what it takes to be one. Now, you will specialize in the area of OSINT and analyse the situation in the past year on a subject described later.

2 Practical part

2.1 Before you start

Everyone (done together) in the exercise will:

- check that you can log into Taranis AI and MISP
- In section Administration check the OSINT Source configuration
  - "misp" (URL should have prepended "=";
  - last 3 characters of API the key matching what I sent you; sharing group ID = 1,
  - organisation ID = to what was preshared;
  - refresh interval = "*/3 * * * *");
- in the administration check the Connector (might not be possible)
  - check URL has no "=";
  - sharing group;
  - organisation ID

2.2 Part I. – individual analysis

You have a common scenario ahead of you. You use the OSINT platform Taranis AI. You have access to data collected across the past year and on your own (or with your group mate). You will need to create an assessment of the landscape.

If you are new to the Taranis AI platform, check out the walk-through video[1] and focus on the Assess section and how news items are grouped into stories (see 3.1).

2.2.1 Your tasks

1. Assess the news items in Taranis AI regarding the topic – Ukraine war. The aim is to gain a better understanding of what has occurred in the timeframe according to your data.
2. Create at least 5 groups of news items (with other words - create at least 5 stories) with a maximum of 10 news items per story (keeping it around 5 is completely normal, too) that represent distinct and important events/aspects of the past year.
3. At the end, share the created stories to the MISP connector.
4. Take a short survey [URL redacted].

[1] https://taranis.ai

2.2.2 Tips for your searching!

When looking for relevant news items, you can leverage the following queries.

- Full Text Search without Tags:
  - broken undersea cable (9)
  - undersea cable (48)
  - ukraine russia disinformation (120)
- Full Text Search with Tags: Ukraine
  - sanction (135), sanctions (105)
  - refugees (26)
- Full Text Search with Tags: Ukraine, Russia
  - sanctions (84)
  - russia political influence (73)
  - cyber attack (478)
  - cyber war (470)
  - australia russia (44)
  - sandworm (35) (APT group)
  - europe russia (436)
  - europe russia influence (79)
  - transport (55)

2.3 Part II. – collaborative approach

Currently, the stories you see in your *Assess* tab should only be yours. To ingest data from partners, go to *Administration* tab (section), then click on the left side *OSINT Source* and click on the source called *misp*. Here, find the *URL* field, and remove the prefixed equal sign ("="). Before you submit the change with the button at the top, make sure to **enable the source**). Then, you will soon ingest stories from other organizations across the globe (other teams).

You can also manually start the ingestion (collection) by clicking the *collect* button on the right side of the main OSINT sources view.

2.3.1 Your tasks

1. Now the aim is to gain a better understanding of what has occurred, leveraging data shared by your partner organizations (other teams, participants). Notice, other partners have access to different feeds, which might but don't need to contain overlapping news items.
2. Go to the *Dashboard* in *Administration* or the "general" *Dashboard* accessible from the top bar. Click the *Connectors* card and check if there are any conflicts and resolve them (see 3.3) – you will need to continue doing this from time to time (at the moment, there is no supportive notification about new conflicts). [2]

[2] When handling the Story conflicts where you compare differences between the same story, you might encounter an issue of "jumping diffs". In that case – hover with your mouse over the diff which you want to review – the issue should not occur

3. **Goal**: Enhance your story clusters (stories) with the new partners' clusters, or enhance the clusters of the partners.

   How: Now go to the *Assess* tab again and check what new stories come from your cooperating organizations (teams). You can do this by setting a filter in *Assess*, to check items from the MISP source (*Filter by Source: misp*).

   Important: do not always use the *Filter by Source: misp*. At some cases, the clusters coming from others just group items you already have and you would not see these.

4. With the new information, adjust the existing and new clusters from others, share your changes again to connector, so the changes would get propagated to the MISP instance.

   You can focus on the addition of editing the stories – **add or remove news items, or add comments or summaries to stories**

5. **Check your MISP instance for any proposals and review/resolve** them[3], the collection happens every 3 minutes or manually retrigger the collection so you ingest the latest state.

6. Continue with repeating the steps 2, 3, 4, 5 for the rest of the excercise

7. Take a short survey [URL redacted].

[3] here you can check existing proposals (URL redacted)

3 Apendix

3.1 How story grouping works

Grouping means, using the merge button at the bottom bar after you select multiple stories. The first item that gets selected, will be the leading story. The other selected will be added to the leading story with all the metadata (attributes, tags, etc.) (see Figure 1)

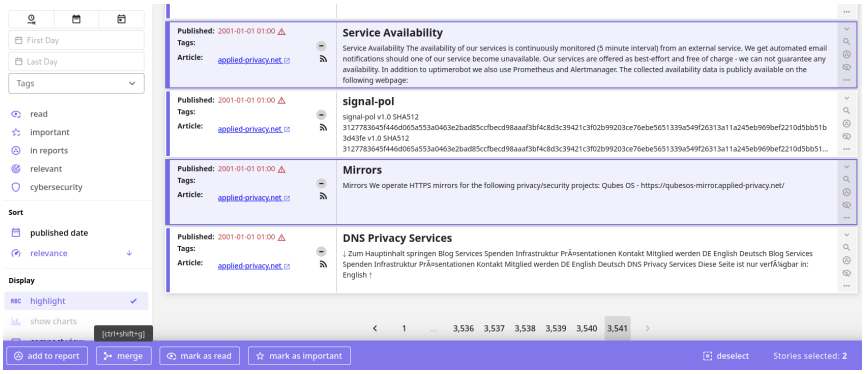


Figure 1: Story grouping - use the button *merge* at the bottom

3.2 How story ungrouping works

Ungrouping of stories can be done in two ways.

1. Ungroup a whole story at once. (see Figure 2
2. Ungroup (button *remove*) a single news item from a story (see Figure 3).

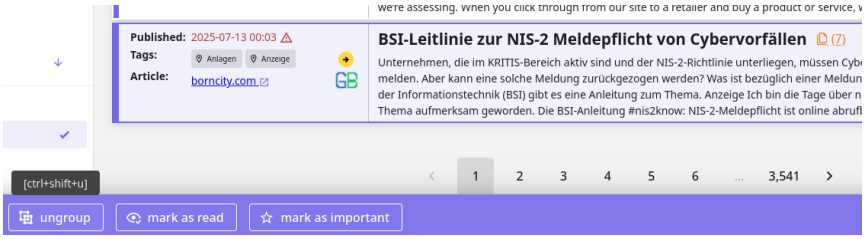


Figure 2: Ungroup whole story - select story, press *ungroup* button

3.2.1 Ungroup a whole story at once

This can be done by selecting a story in Assess view and clicking the button at the bottom bar – ungroup.

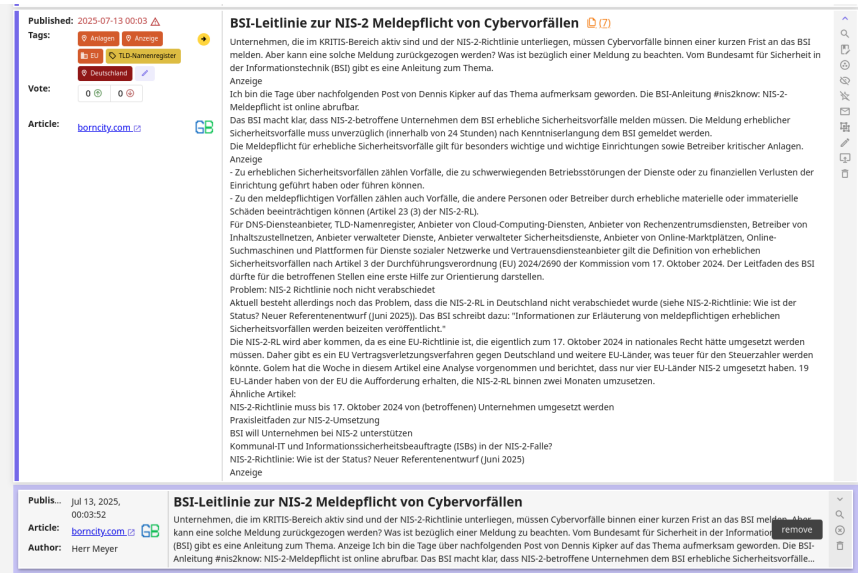


Figure 3: Ungroup one item from story - expand a story and press *remove* button next to the news item

3.3 Conflict resolution explanation

There are different levels where conflicts can occur.

- Story conflicts - these are the ones that are displayed first. They can be resolved using a text editor showing you the difference between the incoming and internal version of the story. On the left side is the internal version, on the right side resides the incoming story.
  - Left view: a story you have.
  - Right view: a story that is incoming.
  - What you want to ingest, you move to the right side of the comparison view. Then click the button *Get Right Side* and then click *Submit resolution* (see Figure 4).
- News item conflicts (see Figure 5) - happen when a story, which is incoming, contains a news item you already own
  - You can either choose to *keep the internal story cluster and ingest new news* items from the left side, or *keep the incoming story cluster and ungroup the internal stories* that are in conflict.

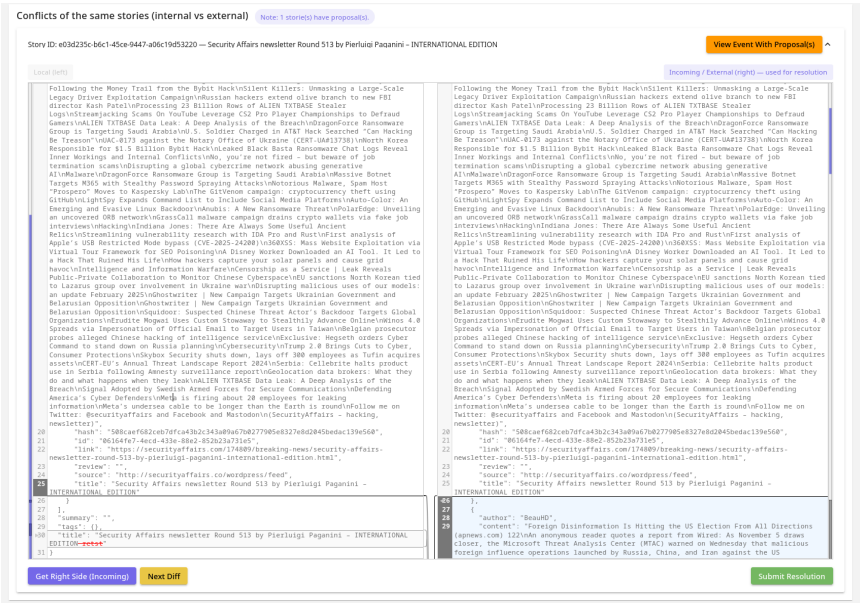


Figure 4: Story conflict view

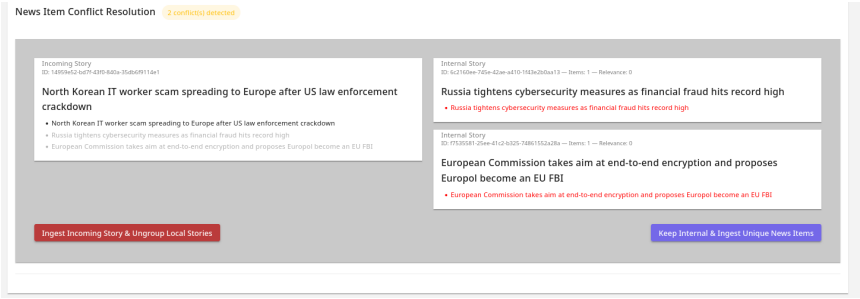
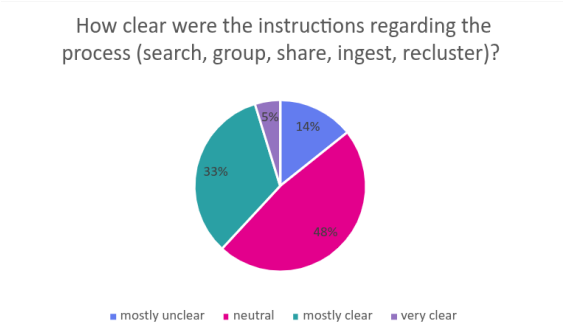


Figure 5: News item conflict view

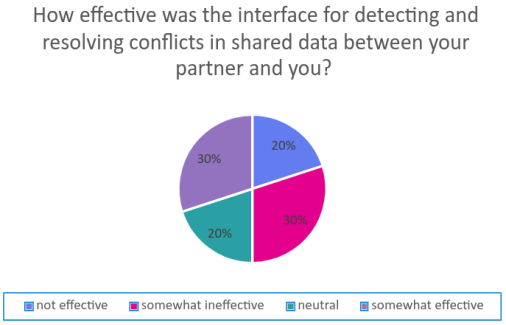
List of symbols and abbreviations

CERT	Computer Emergency Response Team
CSIRT	Computer Security Incident Response Team
CTI	Cyber Threat Intelligence
MISP	Threat Intelligence Sharing Platform
OSINF	Open-source Information
OSINT	Open-source Intelligence
SOC	Security Operation Center

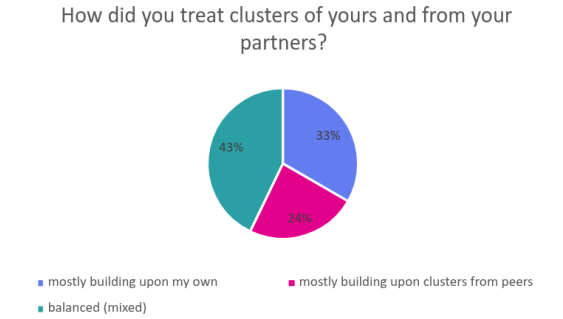
A.4. Evaluation User Session with Users – Plots



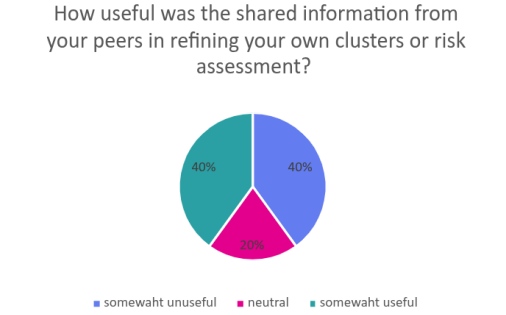
(a) How clear were the instructions regarding the process (search, group, share, ingest, re-cluster)?



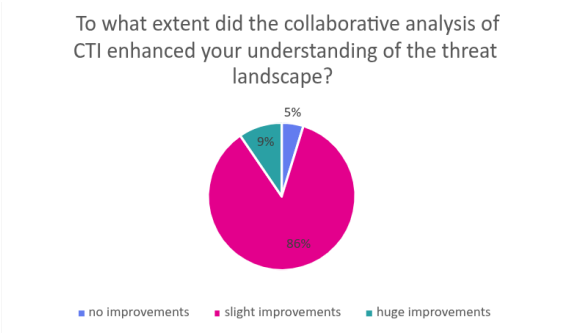
(b) How effective was the interface for detecting and resolving conflicts in shared data between your partner and you?



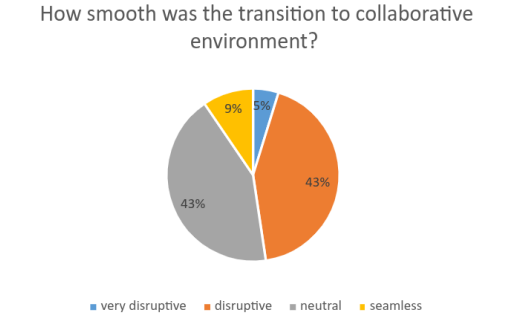
(c) How did you treat clusters of yours and from your partners?



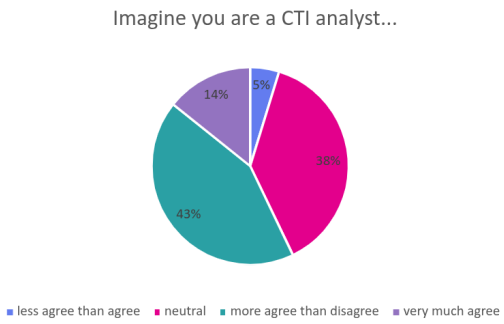
(d) How useful was the shared information from your peers in refining your own clusters or risk assessment?



(e) To what extent did the collaborative analysis of CTI enhanced your understanding of the threat landscape?



(f) How smooth was the transition to collaborative environment from an individual analysis?



(a) Imagine you are a CTI analyst, you strongly believe that you could use this software very effectively if it became part of my routine workflow in my profession as a security/CTI analyst.

A.5. List of References to the PoC Implementation

- `https://github.com/taranis-ai/taranis-ai/issues/436` – Dynamic tracking issue of topics concerning collaboration feature in Taranis AI
- `https://github.com/taranis-ai/taranis-ai/releases/tag/1.2.4` – Officially released version of Taranis AI that includes the output of this thesis.
- Pull requests merged in the course of this thesis:
  - `https://github.com/taranis-ai/taranis-ai/pull/513` – Reaching full collaboration with extended MISP events and conflict resolution improvements.
  - `https://github.com/taranis-ai/taranis-ai/pull/511` – Conflict handling enhancements and fixes.
  - `https://github.com/taranis-ai/taranis-ai/pull/481` – Introduction of conflict handling.
  - `https://github.com/taranis-ai/taranis-ai/pull/378` – Introduction of the new MISP Connector with basic functionality.