DYNAMO

D4.1

Initial prototypes of the cyber-threat intelligence gathering, extraction, sharing components and AI-based solutions

Project number	101069601
Project acronym	DYNAMO
Project title	Dynamic Resilience Assessment Method including combined Business Continuity Management and Cyber Threat Intelligence solution for Critical Sectors
Start date of the project	1st October, 2022
Duration	36 months
Programme	HORIZON-CL3-2021-CS-01

Deliverable type	Other
Deliverable reference number	CL3-CS1-01-101069601 / D4.1 / V2.0
Work package contributing to the deliverable	WP4
Due date	January 2024 – M16
Actual submission date	23rd October 2024

Responsible organisation	CERTH
Editor	Ioannis Chalkias
Dissemination level	PU
Revision	2.0

Abstract	The document is reporting the current state of the initial prototypes of the cyber-threat intelligence gathering, extraction, sharing components and AI-based solutions of DYNAMO while reflecting on the whole DYNAMO architecture and their collaboration with the rest of the platform.

Funded by the European Union

"Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Commission. Neither the European Union nor the granting authority can be held responsible for them."

Keywords	CTI, prototypes, detection, forecasting, knowledge graphs, prediction, anonymisation, information-sharing, situational awareness, IDS, business continuity, actionable, trust, correlation, enrichment

Editor

CERTH

Contributors (ordered according to beneficiary numbers)

CERTH

LAU

VST

RHEA

IRTSX

Document control			
Version	**Date**	**Author(s)**	**Change(s)**
0.1	16.10.2023	CERTH	Table of Contents
0.2	31.10.2023	CERTH, RHEA	Update with feedback from RHEA, CERTH plus further drafting.
0.3	10.01.2024	CERTH, RHEA, VST, LAU, IRTSX	Addition of content.
0.4	24.01.2024	CERTH	Minimal updates following internal review.
0.5	29.01.2024	CERTH	Minimal updates following external review
1.0	30.01.2024	Fraunhofer	Creation of V1.0 from V0.5.
1.1	15.07.2024	CERTH, RHEA, VST, LAU, IRTSX	Updates based on the Midterm Review feedback
1.2	10.10.2024	CERTH, RHEA, VST, LAU, IRTSX	Updates following review of the document
2.0	18.10.2024	Fraunhofer	Creation of V2.0 from V1.1

Disclaimer

Executive Summary

This document provides a detailed analysis of the implementation status of the initial prototypes of the cyber-threat intelligence gathering, extraction, sharing components and AI-based solutions that will be implemented for the DYNAMO platform, more specifically for its CTI Framework.

In Section 2, the detailed information on the topics that render the development of the prototypes significant are reported.

In Section 3, a brief overview of the architecture of DYNAMO and a discussion on how the prototypes will collaborate within the context of the architecture is given.

In Section 4 detailed information for each developed tool regarding its operation, high level architecture, supported by visual information providing evidence of the technical progress of the prototype (i.e., diagrams, screenshots etc) is included.

The tools are developed for each task of Work Package 4. T4.1 (Cyber-threat intelligence gathering and extraction); T4.2 (Cyber-threat intelligence sharing and orchestration); T4.3 (Advanced AI-based analysis & correlation); T4.4: (AI-based Predictive Analytics).

The document also aims at addressing the comments from the Midterm Review feed in order to provide a better understanding of the CTI Framework, its components and its collaboration with BCM. In the revised version of the document the following comments have been addressed:

- Collaboration with the BCM Frawework (Sections 3.1.1 and 4.1.2)
- The use of network data for incident detection/forecasting and near real time response (Sections 4.5, 4.6)
- Allignment of DYNAMO with Regulatory Frameworks and and NIS2 (Section 2.2)
- Added information regarding the use of knowledge graphs and ontologies (Sections 2.3.1, 2.3.2, 2.3.3)
- Discussion for the provision of data for the tools (Included in the sections of the respective tools)
- Additional technical information for each tool (as it is also explained in the text, the work is aligned with the deliverable D2.2 where the main technical information for the DYNAMO platform is documented
- Addition of footnotes and references to increase the relationship between the work and the relevant literature
- Addition of a table with the status of each tool at the beginning of the project and the aimed status at the end of the project (Section 4.1.1)

Table of Content

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Meaning
AA	Attribute Authority
ABE	Attribute-based Encryption
AI	Artificial Intelligence
BCM	Business Continuity Management
BOW	Bag of Words
CA	Certificate Authority

CVE	Common Vulnerability and Exposure
CVSS	Common Vulnerability Scoring System
CPE	Common Platform Enumeration
CERT	Computer Emergency Response Team
CI	Critical Infrastructure
CAF	Cyber Attack Forecasting
CTI	Cyber Threat Intelligence
CKG	Cyber Knowledge Graph
COTS	Commercial off-the-shelf
DAT	Data Anonymisation Tool
DP	Destination Port
E-EWS	ECHO Early Warning System
ENN	Edited Nearest Neighbours
GA	Genetic Algorithm
GDPR	General Data Protection Regulation
GUI	Graphical User Interfaces
HIPAA	Health Insurance Portability and Accountability Act
IoC	Indicator of Compromise
IDS	Intrusion Detection System
IPS	Intrusion Prevention System
LSTM	Long Short Temporary Memory
ML	Machine Learning
MISP	Malware Information Sharing Platform
MI	Mutual Information
NER	Named Entitiy Recognition
NLP	Natural Language Processing
NIS2	Network and Information Security 2

NS3	Network Simulator 3
OWL	Web Ontology Language
PoC	Proof of Concept
RDF	Resource Description Framework
RDFS	RDF Schema
RFE	Recursive Feature Elimination
SaaS	Software as a Service
SIEM	Security Information and Event Management
SSO	Single Sign-On
SVM	Support vector machine
SMOTE	Synthetic Minority Over-Sampling Technique
TTP	Tactics Techniques and Procedures
WP	Work Package

Chapter 1 Introduction

1.1 Context

Threat intelligence is considered as information that allows end-users to protect an infrastructure from attacks or obtain information about recent threats and campaigns. This is achieved by obtaining information that is produce by other similar entities that share the same needs or have experienced relevant incidents. The end-users can also assist in increasing the robustness of their industry/community by producing and sharing Cyber Threat Intelligence (CTI) with others. Based on this, the DYNAMO platform aims to offer to the users a selection of CTI tools that can ingest and process CTI coming from relevant sources and/or trusted users, in a timely manner for defending against present or future threats and attacks. For the DYNAMO platform, the effective utilisation of CTI by the user is a key factor in achieving business continuity and increasing resilience within the premises of an organisation.

1.2 Purpose and Scope

The purpose of this deliverable is to summarise the initial development of the tools of DYNAMO's CTI Platform in Work Package (WP) 4 (Milestone 5) and review the design of the DYNAMO Platform's architecture (Milestone 6). The tools that are being developed will support the DYNAMO Platform in achieving the DYNAMO objectives of addressing the specific needs and requirements of the end users and raising their awareness by increasing their collaboration and information sharing. This document contains information about all the prototypes designed and developed in the context of the objectives of each respective task. T4.1 aims to achieve the gathering and extracting of CTI from a variety of sources by developing tools that are able to obtain this information not only from the online but also from internal sources in order to identify their CTI requirements and match them with the existing information. T4.2 is the task responsible for the orchestration and sharing of CTI in order for the users of the DYNAMO platform to share trusted and relevant information with the use of appropriate channels of information sharing. T4.3 receives the information collected for from T4.1 and uses it in order to correlate the information with other existing threats and remediations while also providing useful knowledge graphs for the user. T4.4 will provide the analysis that is required to achieve predictions for upcoming attacks by offering tools that rely on the CTI that is processed and provided by the tools of the other tasks.

1.3 Relation to other work in the project

The activities of WP4 represent a software development activity that requires relationships with other tasks and activities carried out within the DYNAMO project. The developed tools will assist in raising the awareness of the user by providing tailored threat information to the tools of the Business Continuity Management (BCM) Framework (WP3) as parts of the efforts of the Platform to provide a systemic resilience assessment. The tools will also provide material for the training and simulation activities that will take place in WP3. The prototypes of the CTI Framework will be integrated and refined in the DYNAMO platform under the efforts of WP5, along with the tools of WP3 and other products from previous projects (e.g. the H2020 ECHO project[1]):

- The work on this deliverable takes information from the deliverables:

[1] https://echonetwork.eu

    - D2.1: Use cases definition and user requirements
    - D2.2: DYNAMO system specification and reference architecture
    - D3.1: Business continuity processes, human factors, assessment methodology, and DYNAMO AI-solution prototype plan.
- The work on this deliverable will be evaluated by the efforts in Work Package 6 “Cross-sector pilot demonstration, evaluation & training”.
- D4.2 “Final version of the cyber-threat intelligence gathering, extraction, sharing components and AI-based solutions” will receive D4.1 as initial input to conclude the development in WP4.

1.4 Document structure

The document is structured in the following manner:

- Section 1 – Introduction, describing the purpose and structure of the document as well as the relation with other work packages
- Section 2 – Description of the concepts that the development efforts in WP4 are addressing
- Section 3 – Description of the internal architecture of the CTI Framework and its relationship with the rest of the platform’s architecture
- Section 4 – Description and analysis of the early versions of the CTI tools of WP4, also providing diagrams, mockups, screenshots, etc
- Section 5 – Conclusions, summarising the findings, results and recommendations
- Section 6 – Bibliography section.

Chapter 2 Overview and Concept

2.1 Intelligence Gathering and Sharing

Contemporary organisations continuously add to their infrastructures increased number of diverse cyber(physical) and IoT devices and adopt the use of new tools with extended and diverse capabilities. This unavoidable evolution introduces new attack vectors that are reshaping the threat landscape [1] and advanced threat actors with enhanced skills and resources that affect the operation in the private and public sector. The detection and deterrence of these advanced attacks require innovative approaches in organising and orchestrating cybersecurity. CTI is defined as "evidence-based knowledge, including context, mechanisms, indicators, implications, and actionable advice about an existing or emerging menace or hazard to assets that can be used to inform decisions regarding the subject's response to that menace or hazard" [2]. CTI consumers collect and analyse data regarding security threats, threat actors, exploits, malware, vulnerabilities and Indicators of Compromise (IoC) in order to produce actionable threat information for the organisation or even share it with other organisations that share relevant cybersecurity objectives. This information increases the proactive cybersecurity posture of an organisation. It includes strategies for limiting the effect of cyberattacks, allowing the continuation of operations to a reasonable extent.

The dissemination of information regarding potential risks and threats creates a series of challenges on how companies can obtain and employ threat information to improve their defence posture against threats and attacks [3]. Those challenges can include the following:

- Threat Data Overload
- Data Quality of shared feeds
- Trust in the content of the CTI tokens
- Actionability of information
- Timeliness of CTI
- Privacy-related regulations
- Internal policies and confidentiality of information

Artificial Intelligence (AI) and automations have been introduced to the CTI production and sharing mechanisms in order to address those challenges, create situational awareness among stakeholders, and be alerted about a threat as immediately as possible. With the existence of multiple sources of CTI, an organisation is receiving overwhelming amount of data that needs to be processed and analysed. With a major portion being false, irrelevant, or fake, it is inevitable that manual analysis of data cannot be effective or sometimes even possible. On many occasions, organisations cannot afford the time that is required to undergo that process in a manual manner.

One of the aims of utilising AI solutions in CTI is to decrease the human error rate during the processing of information, and also the subjective relevance filtering by automating some of the processes [4]. In the CTI paradigm a critical part of the information comes from unstructured text data, such as threat reports, social media posts, news articles, and hacker forums. Therefore, Natural Language Processing (NLP) techniques are used to automatically examine and process textual data for the identification of relevant threats, IoCs, threat actors and their Tactics Techniques and Procedures (TTP) [5]. Identifying the relationships between entities (cyber-actors) and events creates a more complete operational picture for the security teams. Modelling CTI by using NLP and creating taxonomies and ontologies can be key factors to identifying attack groups, potential attacks and threats for an organisation or even an industry [6].

Protecting a whole industry against threats and attacks requires mutual trust and overcoming human, cultural and organisational aspects [7]. On many occasions, the shared information can be a subject of confidentiality, expose the defensive measures of an organisation, or challenge privacy-preserving legislation. These factors can reduce the actionability of the shared information and also limit the

sharing of threat information with external entities. Building a trusted network of partners and providing them the means of sharing information that respect those limits is one of the mail elements can enable a common culture for information sharing.

DYNAMO is aiming to develop tools and create a platform that helps its user to obtain relevant, enriched and actionable information in order to achieve increased situational awareness that not only will protect against threats and attacks but will also inform an organisation about recent trends by allowing the sharing of trusted information within a trust environment of entities that share common cybersecurity needs. The tools will be able to offer the necessary automation for information gathering and analysis and build ontologies and taxonomies to correlate the information that is received in order to protect against current and future threats.

2.2 Regulatory requirements and trust environments for CTI

Due to the nature of threat information, its handling and sharing are subject to various requirements and regulations. Key considerations include the General Data Protection Regulation (GDPR)[2], the second Network and Information Security (NIS2) Directive[3], and the upcoming AI Act[4] Regulation. Additionally, there may be impacts of industry-specific provisions, national regulations, or laws depending on the sector (e.g., the Health Insurance Portability and Accountability Act (HIPAA)5 in healthcare). At the end of this section, we discuss a little about the specifics of the trust environment in the context of DYNAMO.

2.2.1 General Data Protection (GDPR)

The GDPR aims to reduce legal uncertainty and limit interpretations by establishing clear rules and conditions for the processing and sharing of personal data. Organisations must ensure they process only the necessary amount of personal data to achieve their purpose. GDPR defines specific roles and responsibilities for data processors and data controllers.

According to GDPR, personal data includes any information related to an identified or identifiable natural person, such as name, ID number, location data, or factors specific to their physical, physiological, genetic, mental, economic, cultural, or social identity. IP addresses may be considered personal data as they can enable identification.

In the context of threat intelligence sharing through tools like the ECHO Early Warning System (E-EWS) and Malware Information Sharing Platform (MISP), the focus is primarily on exchanging threat and malicious actor information rather than sharing personal data. However, since these data often contain personal Data (e.g. IP etc) or personal information, those sharing information must be aware of the purpose of data processing and the roles of relevant parties.

Under GDPR, security teams can process and share information in line with the principles of purpose limitation and necessity to ensure network and data security. Therefore, under GDPR, security teams can process and share information in line with the principles of "purpose limitation" and "necessity to ensure network and data security".

When information is shared via an information sharing platform (e.g. EWS or MISP), in most cases, personal data has not been obtained directly from the data subject but comes from threat analyses. In such cases, GDPR's Article 14 triggers the application of the transparency principle, requiring data controllers to inform data subjects.

[2] https://eur-lex.europa.eu/eli/reg/2016/679/oj
[3] https://digital-strategy.ec.europa.eu/en/policies/nis2-directive
[4] https://eur-lex.europa.eu/eli/reg/2024/1689/oj
[5] https://www.hhs.gov/hipaa/index.html

The GDPR permits the sharing of information when it aligns with the principles of purpose and necessity and is essential for ensuring network and data security.

In conclusion, cybersecurity actors can use DYNAMO's tools to share information within the framework of data protection legislation, to guarantee that their activities are designed to serve the general interest and security while processing personal data only to the necessary extent.

2.2.2 EU's AI Act

The EU's artificial intelligence regulation is the world's first concrete initiative to regulate artificial intelligence. Its goal is to make Europe a global centre for reliable artificial intelligence by establishing uniform regulations that guide the development, marketing, and use of artificial intelligence in the EU. Implementing the AI Act to guarantee that artificial intelligence systems used in the EU are safe and respect fundamental rights and values. In addition, the EU, must use investment and innovation, and improve governance and supervision to encourage the development of a single EU market in the field of artificial intelligence.

The AI Act introduces a risk-based approach to AI regulation, focusing on applications that could potentially harm people. This includes situations where AI systems are used in critical infrastructure. Before deploying AI systems, assessments must be conducted, including requirements related to any level of risk. Critical infrastructures examined in the DYNAMO project are derived from the healthcare, energy, and maritime sector.

In the DYNAMO context, the requirements of the AI Act must be considered already during the construction phase, as AI is used in it. The first step is to identify all AI models in use and their roles as developers or procurers and to list the identified AI models in a database.

The next step is to assess the risk rating of the AI models in question. The EU AI regulation distinguishes different risk categories. Models of unacceptable risk are prohibited. High-risk models are allowed but must meet several recommendations and be evaluated for compliance before the model can be submitted. These models must also be registered in the database established by the EU. Use of high-risk AI models require an appropriate risk management system, event tracking capability, and human oversight or ownership.

For low-risk models (such as chatbots), transparency is required, i.e. the user must know that he is interacting with artificial intelligence. The AI Act allows the free use of low-risk AI. In the context of DYNAMO, these could include AI-powered spam filters and gamified training environments.

In cyber threat prevention, it should be considered that, following biometric identification, they are high-risk and subject to strict requirements. It may be that scraping biometric data from social media would be prohibited. Also, tools for predicting who will commit crimes would be prohibited. Exceptions may occur after this when they are necessary to prevent the threat of terrorism or to detect, locate, identify, or prosecute a criminal or suspect. Their use requires the permission of a judicial body or other independent body and appropriate deadlines, geographical coverage, and the database that is the subject of the search. However, the AI Act proposes measures to promote trust in AI, such as security operations centres using AI. One would therefore expect that the use of artificial intelligence to improve the common good and safety would even be encouraged.

2.2.3 NIS2 Directive

The NIS2 directive, which is an update to the original EU NIS (Network and Information Security) directive, defines requirements to improve the general level of cybersecurity in the EU. Organisations sharing open-source intelligence and threat information (including the DYNAMO end-users) must consider the requirements set by NIS2. New obligations must be included from October 2024.

NIS2 emphasises the importance of implementing appropriate and proportionate technical and organisational measures to manage risks to the security of network and information systems. The

directive requires e.g. management accountability, rapid reporting of significant cybersecurity incidents (within 24 hours of incident detection), appropriate security policy, disruption management, continuity management, risk management, ensuring supply chain security, national cybersecurity strategies to be implemented by member states, regular assessment of security measures and increasing awareness and training.

NIS2 encourages information sharing and cooperation between actors in the same sector, between different sectors and with relevant national and EU authorities. DYNAMO's aim is to provide a platform that will be valuable for sharing information by also ensuring compliance with regulations.

2.2.4 Trust environment

Threat actors often target specific sectors, such as healthcare, energy, or maritime. Once they find a successful technique against one organisation, they typically try the same methods on similar organisations. Therefore, sharing cyber threat information in a trusted environment benefits all parties by reducing the likelihood of similar organisations falling victim to the same threats. The DYNAMO platform supports all phases of cyber resilience. One of its key capabilities is storing and sharing cyber threat information among DYNAMO partners, with a strong emphasis on protecting shared data. To facilitate this, the project develops a trusted environment for users of the DYNAMO platform. Within this trusted environment, CTI is exchanged primarily through EWS. MISP is utilised for CTI exchange between DYNAMO partners and their stakeholders. MISP aims to prevent and identify targeted cyber-attacks using Indicators of Compromise (IoCs). It provides a unified structure for threat information and automatically consolidates similar data, making it easier to store and share information among organisations facing similar threats. MISP also features a REST API for adding functionalities and external data sources. TheHive is an open-source Security Incident Response platform for managing information security incidents. It can synchronise with MISP instances to investigate MISP events, and research results can be exported and shared as MISP events within trusted environments. Cortex, developed by TheHive Project, is designed to analyse IoCs and works specifically with MISP and TheHive. When sharing cyber threat information, it is crucial to balance transparency, privacy, and confidentiality.

MISP's goal is to prevent and identify targeted cyber-attacks, for example with the help of IoCs. MISP gives threat information a unified structure and automatically combines similar information, making it easier to store and share information despite the large amount of information among organisations that are likely to face similar threats. MISP has a REST API programming interface for adding functionalities and external data sources. TheHive is an open-source Security Incident Response platform, i.e. a platform for managing information security incidents. It can be synchronised with MISP instances to investigate MISP events. The research results can be exported and distributed as a MISP event among the trust environments. Cortex is a software developed by TheHive Project for analysing IoCs, which is intended specifically for use with MISP and TheHive.

Organisations should share CTI to help others understand threats better in a trusted environment, but avoid sharing identifiable or sensitive information, such as company names, internal IP addresses, personnel names, customer identifiers, or business-related information. The shared information should focus solely on threats, not on the organisation's own business, infrastructure, employees, or customers. To maintain privacy and confidentiality, it is important to review shared threat information. An approval chain can be used to reduce risks and ensure confidentiality and accuracy before sharing information with other organisations.

2.3 Knowledge Graphs

Knowledge Graphs are used to model CTI concepts and the relations between them, which is structured according to an ontology. This knowledge graph is built by mapping the entities provided by each data source onto ontology concepts and by extracting additional relations between them.

2.3.1 RDF Graphs

There are many modelling approaches for knowledge graphs, e.g., labelled property graphs, hypergraphs, RDF graphs or even relational databases. The conceptual framework adopted for this project was the one standardised by the W3C Semantic Web standards[6], i.e., RDF graphs. These standardise both a modelling framework (RDFS and OWL, see below) and a graph query language (SPARQL). The main reasons for this are:

- The framework requires minimal design time "schema" restrictions. This is particularly suitable for the cyber intelligence domain, where data comes from heterogeneous data sources and a rigid schema would introduce unnecessary constraints.
- It provides formal semantics, which allow the use of inference/reasoning. This was one of the exploration areas initially proposed for the WP 4.2 of this project.

RDF graphs model knowledge as a set of statements about resources (RDF triples). Such statements can be either a relation between two resources (Object Properties, e.g., "Cyberthreat A *is associated with* Threat Actor B") or that a resource has a property with a given value (Data Properties, e.g., "Vulnerability V *has exploit score* N").

RDF triples provide a way to express arbitrary statements about resources, but these statements do not constitute knowledge unless they have semantics associated with them. Such semantics can be achieved by restricting the statements to a given vocabulary, where each vocabulary term has a well-defined meaning. Such vocabulary constitutes an ontology.

W3C includes multiple standards to define ontologies, but in the scope of this project only the RDFS and OWL are being used. These standards have the particularity that the ontology itself is defined using RDF triples. The same language (RDF) is used to define both the data and its "schema".

2.3.2 Ontology

The use of RDF to define ontologies makes it possible for the ontology to be extensible in "runtime", i.e., not restricted to a vocabulary set in design time. However, there is an initial ontology comprised of the following groups of classes:

- Cybersecurity, which represent cyber threat intelligence concepts. This group is aligned with the STIX standard entities.
- IT assents, which models IT infrastructure objects. Examples of these include "Server", "Software", "IP Address", "Software Artifact", etc.
- Domain Specific objects. An ontology for the space domain was explored in another project, and while there is potential for domain specific classes, no one was identified yet at the time this document was written.

The relations that exist between objects of the "Cybersecurity" and "IT Assets" group are essential ones because they allow understanding how CTI is related to the organisation.

[6] See https://www.w3.org/2001/sw/wiki/Main_Page

2.3.3 Building the Knowledge Graph

The following techniques are used to build the knowledge graph:

- Merging entity data from multiple sources; some sources provide only partial information about entities, which is merged into single entities.
- Explicit relations on the source; some sources already provide relationships between entities.
- Rule-based enrichment; nodes and relations can be fetched on the fly, as the knowledge graph is built. For example, when an “IP address” node is processed, an enriching element can be triggered to fetch additional information for that IP address from the “ipinfo” data source, thus adding new graph relations.
- Entity extraction from text; text associated to knowledge graph nodes can reference concepts in the ontology (e.g., IP-addresses, hashes and Common Vulnerability and Exposure (CVE)). These references lead to new relations in the knowledge graph. Two entity extraction methods are provided:
    - Rule based, where entity references are defined explicitly as regular expressions.
    - Named entity recognition (NER), where entity references are extracted using a machine learning (ML) model trained with a labelled corpus of text. This allows entity extraction not achievable with rule-based techniques, as the cost of a possible loss of precision.
- Use of the inferencing/reasoning mechanisms available in the semantic web standards, and that are all provided by the graph database[7]. An example is the use of the owl:InverseFunctionalProperty entailment rule to merge data from multiple sources, where the same entity can be identified in several ways in each source.

Having information structured as a knowledge graph makes it possible for users to explore the relationships in an interactive way, adding insights/situational awareness that would be harder to achieve by exploring the entities in isolation.

2.4 AI-Driven predictive analysis of CTI

Panigrahi et al. [8] state that the Intrusion Detection Systems (IDS) fall into one of two classes based on the categorisation methods they use:

- Signature-based intrusion detection systems. In order to identify known attacks, they frequently use pattern matching algorithms that compare the signatures of previous invasions in a threat signature database.
- Anomaly-based intrusion detection systems. These systems evaluate an entity's typical behaviour using ML, statistical, or knowledge-based techniques. Any significant deviation from this usual behaviour is seen by the system as abnormal and hence as an intrusion. According to Umer et al. [9], ML enables the identification of entirely new threats as well as updates to established attacks.

IDSs may be further divided into groups based on the kinds of algorithms they use. To create intelligent solutions, ML improves the ability to extract knowledge from massive amounts of data. With little assistance from humans, a ML system may provide data-driven conclusions, forecasts, and assessments based on what it has discovered from prior data. ML algorithms may be divided into the following groups according to their training methods:

- Supervised learning. The pre-labelled data is used to train these algorithms. According to the outcome of the relevant topic, these labels describe each index [10]. In order to easily classify a newly tested index to a known category, training aims to generate patterns that correspond

[7] https://graphdb.ontotext.com/documentation/10.5/reasoning.html

to each known category-label. The number of unique labels that are present throughout the training phase restricts the categories that may be created.

- Unsupervised learning. There is no user monitoring of these algorithms. It also enables the model to function independently, revealing information and patterns that were previously undiscovered. It is comparable to the process of learning that takes place in the human brain upon acquiring added information. It performs admirably even with unlabelled data.
- Reinforcement Learning. Reinforcement learning algorithms are trained by putting their knowledge to the test and rewarding right responses with positive reinforcement or punishing bad responses with negative reinforcement.

One of the most important methods for detecting intrusions is the near real-time extraction of statistical information from network data. To operate correctly and detect unusual behaviour, IDSs use network flows that are created based on source/destination IP, source/destination port, protocol, and timestamp [11], [12]. A network flow is an assemblage of IP packets that pass through a monitoring point during a certain period and have certain features [13].

The calibre of the dataset used to train the system directly affects how effective a ML-based intrusion detection system is. Four primary factors may be used to determine the quality of a dataset:

- Up-to-dateness: The dataset must be updated with the most recent list of known cyberattacks due to the ongoing evolution of cyberattacks in terms of both deployment and complexity.
- Real network traffic: Including real network traffic rather than just simulated traffic helps create realistic traffic patterns.
- Volume and variety of traffic: More types of network traffic included in a dataset lead to more accurate traffic patterns being produced. Furthermore, a dataset's size is important since a larger dataset has more indexes, which improves the system's accuracy.
- Accessible to the general public. A dataset that is accessible to the public can also be inspected by the public.

Ullah and Mahmoud [14] proposed a hybrid model anomaly detection approach based on flow-based anomaly detection for the classification of the CICIDS2017 and UNSW-15 datasets. They used Recursive Feature Elimination (RFE) to identify significant features, Synthetic Minority Over-Sampling Technique (SMOTE) to oversample, and Edited Nearest Neighbours (ENN) to clean the dataset. A decision tree classifier was used to determine if the network flows at level 1 were normal or abnormal, which allowed the network flows to be forwarded to level 2 (multi-classification) in order to identify the type of attack. Level 2 results for recall, precision, and F-score were measured at 100% for the CICIDS2017 dataset and 97% for the UNSW-15 dataset, while level 1 results for specificity, recall, precision, and F-score were measured at 100% for the CICIDS2017 dataset and 99% for the UNSW-15 dataset.

Vijayanand and Devaraj [15] created a special IDS for wireless mesh networks that uses many support vector machine classifiers and evolutionary algorithm-based feature selection. To attain higher accuracy, they pick certain features using the Support Vector Machine (SVM) classifier and Genetic Algorithm (GA)-based feature selection. The system is assessed using a WMN-generated intrusion dataset, and it is simulated in the Network Simulator 3 (NS3) tool using the standard intrusion dataset. Additionally, they validate the system using the CICIDS2017 and ADFA-LD intrusion datasets. After a comparative analysis between the proposed system and Mutual Information (MI)-based feature selection, it is determined that GA-based feature selection with SVM classifier shows higher performance metrics, including lower computational complexity and a larger accuracy of over 99%.

Zhang et al. [16] reported a neural-network-based anomaly detection model. They developed an IDS using the LeNet-5 convolutional neural network and utilised the Long Short-Term Memory (LSTM) network for feature extraction. The experiments employed the CTU and CICIDS2017 datasets for multi- and binary classification. They employed CNN, LSTM, and a combination of the two in binary-classification and multi-classification investigations, yielding effective classification outcomes. There

was about a 99% accuracy rate. They also looked at the flows that were essential for efficient anomaly detection and classification.

Ahmad et al. [17] primarily focused on creating an ensemble for feature selection using several evaluation techniques in order to produce an intrusion detection system. They particularly recommended a number of feature extraction and selection strategies, and they developed an IDS model with Random Forest as their learning algorithm. Evaluations were carried out using a range of evaluation datasets, including KDDCup'99, NSLKDD, UNSW-NB15, and CICIDS2017, in order to demonstrate the efficacy of the proposed model. The final high-performance measures, which attained 99.88% accuracy, demonstrated that the specific subset of characteristics is promising [19] when compared to other techniques, as per the data.

Ullah and Mahmoud [18] presented an anomaly detection sequence model based on a LSTM recurrent neural network (RNN). Two bidirectional LSTM models were used to process embedded sequences to create the proposed system. CICIDS2017 was used for the testing trials, and the model outputs were meant for multi-classification.

C Zheng et al. [19] looked on mapping ML methods to programmable network devices within the framework of IDSs especially made for enterprise networks. Furthermore, the functionality, resources, scalability, and throughput of recently developed and state-of-the-art in-network ML algorithms are assessed and contrasted. They employed six datasets for intrusion detection, including AWID3 and KDD99. For decision trees, their accuracy varied from 97.47% to 49.37%. Furthermore, the goal of researchers [20] was to identify the minimal set of classifier characteristics that could be used to any implementation version of 802.11. Their study also examined the effectiveness of ML algorithms in identifying several types of network attacks by utilising datasets from the AWID family. For their research, the authors selected 16 attributes that were essentially relevant to every frame type and subtype of 802.11, ensuring that they could be directly applied to a wide range of network setups. In particular, they avoided traits that showed recurring patterns since these may lead to biases and overfitting. Because it was anticipated that this collection of attributes would remain consistent across all frames, avoiding analytical bias, it was selected. The AWID3 dataset's cases were split into three groups by Chatzoglou et al.: Normal, Flood, and Impersonation. The Flood category includes attacks like Deauth, Disas, Assoc, and Kr00k, while the Impersonation category had attacks like RogueAP, EvilTwin, and Krack. In practice, the authors' average accuracy for ML and deep learning techniques combined was 99.96%. Their study demonstrated the value of the selected variables and the efficacy of deep learning and ML techniques in accurately detecting and classifying network threats. The study's findings back up the development of 802.11-based networks' intrusion detection systems.

In order to expand the operational capabilities of the IDSs, researchers have worked on including forecasting in deterring cyber-attacks. This requires the use of models that forecast future hostile behaviour based on system flaws and attacker behaviour by identifying attack patterns in previously collected information. Provided timely, the predictions of attacks can be used to disrupt potential cyber-attacks [21], [22]. These methods require the use of several ML techniques; more specifically deep learning methods have effectively used in time series forecasting [23] . The use of automations in attack forecasting can potentially decrease the element of surprise that the attackers exploit, the time spent of generating forecasting reports and also remove any biases in predictions [22].

Chapter 3 Architecture of the CTI Framework

3.1 The CTI Framework within the DYNAMO Architecture

The DYNAMO platform is designed to combine CTI and BCM, cumulating in a comprehensive situational awareness framework. To avoid overlapping with the content of the deliverable D2.2, the topic is extensively covered and analysed in D2.2.

The design of DYNAMO platform is supported by frameworks and tools to allow modular approach, scalability, high-availability, and high volume of information exchange. The analysis of the tools is provided in Section 4, CTI Framework's Components.

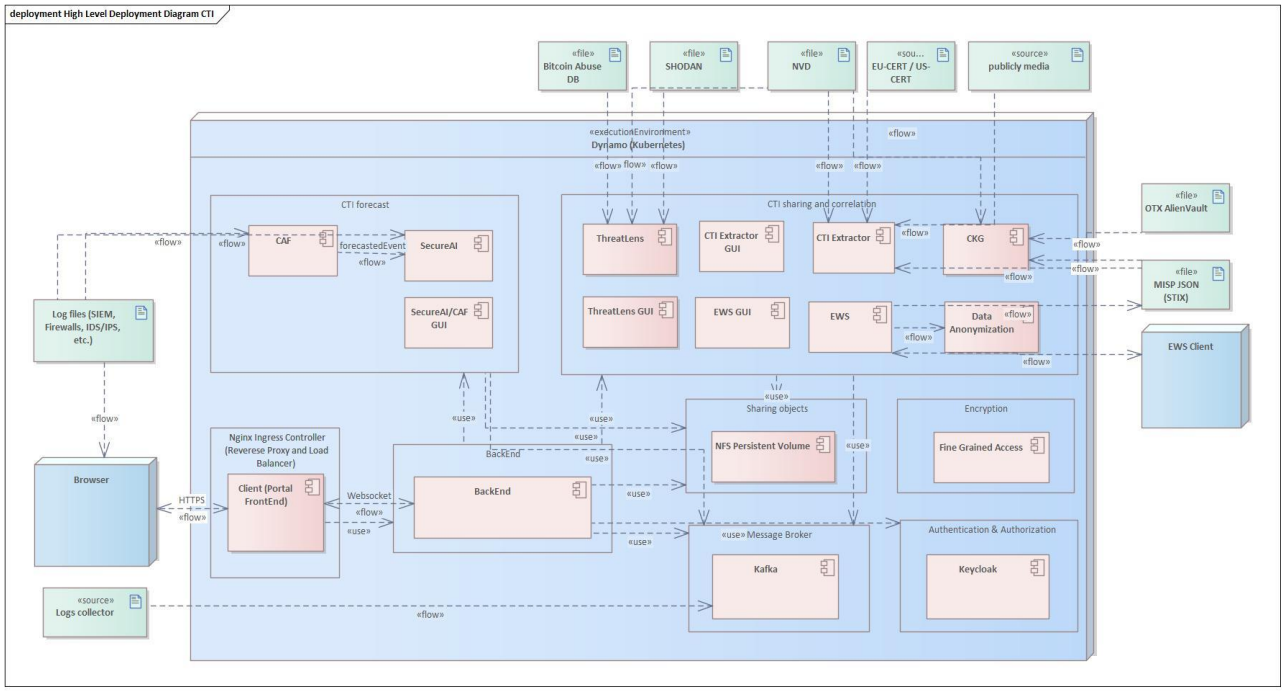


Figure 1: The DYNAMO Platform

The platform is a web portal, combining Software as a Service (SaaS) and On-premises application models. This ensures situational awareness through overview and dashboards and supports decision making over the risk scenarios. The portal provides access, through a common view and Single Sign-On (SSO) - Keycloak[8], to the BCM / CTI Components' Graphical User Interfaces (GUI), supporting a deeper analysis of the CTI and BCM information.

The message exchange between CTI and BCM components is ensured by web services and Kafka[9] publish-subscribe pattern to handle high-velocity, high-volume and fault-tolerant data streams.

The deployment of the platform is done on Kubernetes[10] environment, achieving the scalability and high availability of the system.

The CTI Framework provides detection and forecasting of threats, correlation with the use of identified relevant CTI, and sharing with external entities. All this is done after the anonymisation of information, provided by the Data anonymisation tool (DAT). Detection and forecasting of threats,

[8] https://www.keycloak.org/
[9] https://kafka.apache.org/
[10] https://kubernetes.io/

two tasks respectively performed by the SecureAI and the Cyber Attack Forecasting (CAF) components, is pushed to the Client GUI, to be visualised and acknowledged by the user. The user can then choose to share it with external entities and external sources, through E-EWS and/or MISP, or get in deeper analysis of its correlation using the components ThreatLens, CTI Extractor, Cyber Knowledge Graph (CKG), or a combination.

The CTI information is shared, through a message broker, with the BCM framework, which generates attack paths, risk and impact scenarios, and the mitigations plans, and pushes them to the Client GUI.

3.1.1 CTI and BCM information workflow

The BCM framework ingests the CTI information from Kafka topics and generates dashboards and impact graphs with containing information on the relevant risk situation.

In the BCM framework, a key role is played by the CaESAR tool. CaESAR component receives CTI information from defined Kafka topics, requests correlated information from the CTI Framework, generates visualisations on impact scenarios and provides information on resilience metrics. The impact scenarios are pushed to the Client GUI.

The following sequence diagram represents the flow of information and actions after an attack detection.

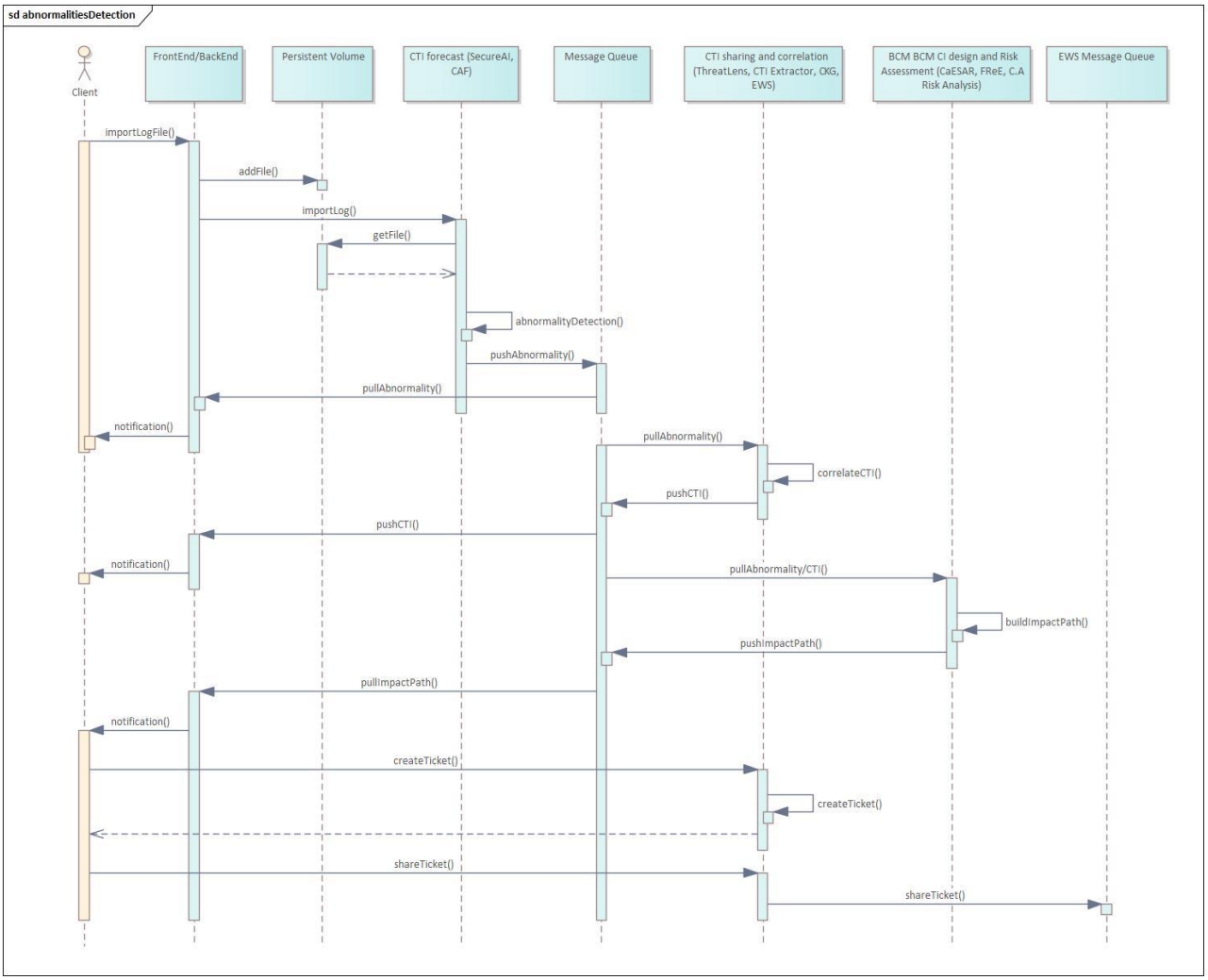


Figure 2: Sequence Diagram on Attack Detection in DYNAMO

The BCM framework supports generation of risk scenarios and mitigation plans, as well as evaluation of the risks after the mitigation actions. The Client GUI displays the mitigation plans, identified by RETA and Pianista components, over an identified risk situation. The CTI Frameworks contributes to the generation of the response plans with SecureAi, since the tool provides mitigation suggestions to be considered for utilisation by the repsonse plans, based on identified abnormalities. The CaESAR component interacts with RETA and Pianista, providing them the needed information (Critical Infrastructure (CI) attack paths, impact scenarios, CTI information from CTI Extractor, etc.) to generated mitigation plans.

The CaESAR tool supports the Computer Aided Risk Analysis tool to generate the risk scenarios and perform evaluation after the mitigation actions.

The following diagram describes a high-level flow of the mitigation request.

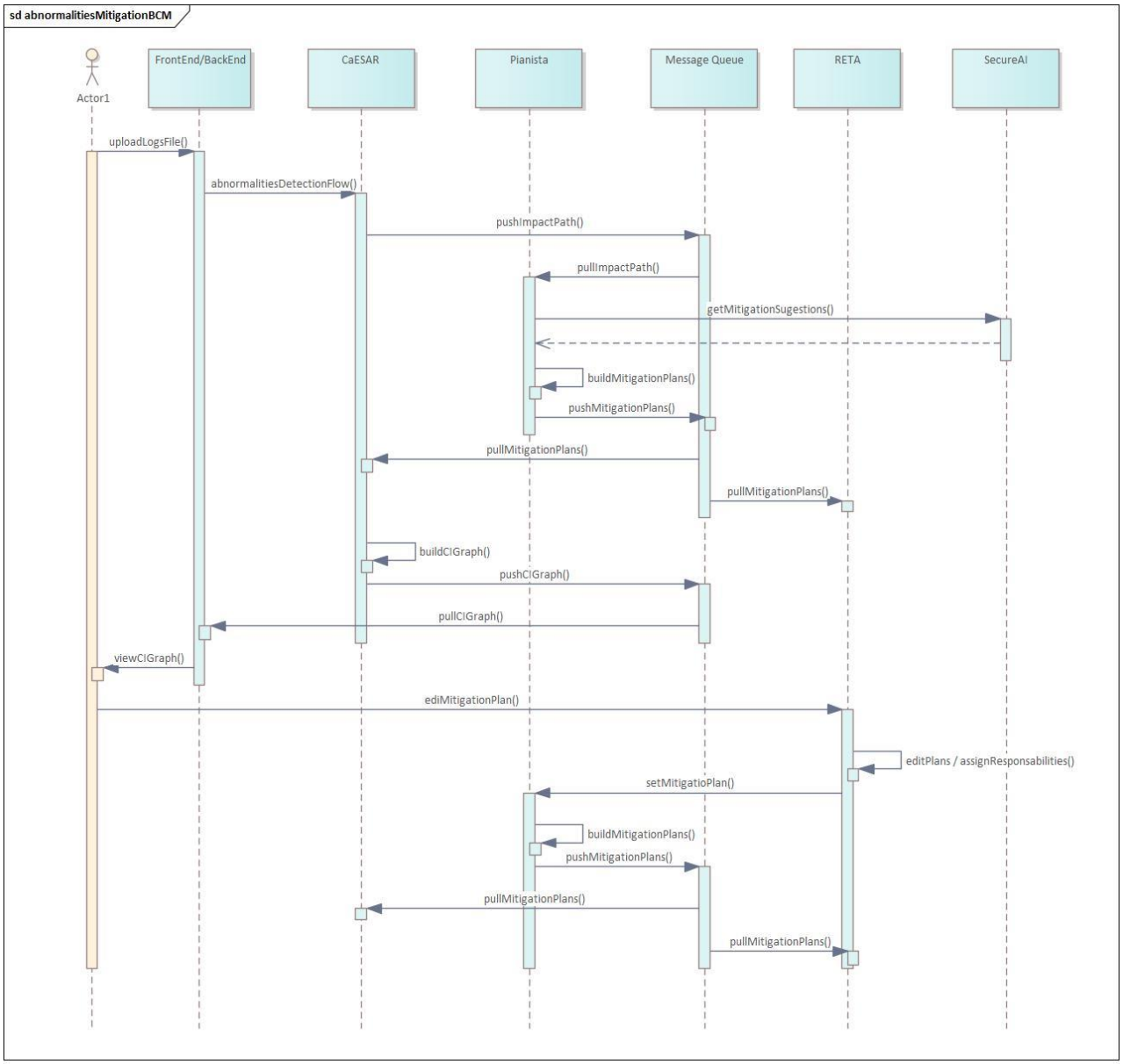


Figure 3: Mitigation Request in DYNAMO

3.2 Internal Architecture of the CTI Framework

The CTI Framework is consisted of multiple components, providing near-real-time detection and forecasting (SecureAI and CAF), correlation and enrichment over the identified CTI (ThreatLens, CTI Extractor and CKG) and sharing with external entities (E-EWS), after anonymisation of information with the use of DAT.

The CTI Framework is available to the Client with SSO access to components and supports situational awareness pushing warnings and information to the GUI of the current CTI.

The components inside the CTI framework can be used in the main flow of CTI identification, but also as supporting tools to get a deeper understanding of CTI and impacts, synchronised with external CTI sources. The high-level flow of information is described in the following steps:

- CAF, SecureAI, ThreatLens subscribe to Kafka message topics containing logs / source files and publish detected CTI information.
- CTI Extractor, CKG, ThreatLens subscribe to latest CTI information correlate, process, and share updated CTI in Kafka topics.
- CTI Extractor, CKG, ThreatLens support requests from BCM and provide deeper understanding of the CTI information. The request is implemented through webservices and to provide a normalised view of the correlated CTI information.
- SecureAI and CAF support requests from the Client, to provide historical analysis of the forecasted and detected incidents.
- SecureAI provides service to propose mitigation suggestions (measures) based on detected CTI to be included in the response plans. This feature is used by the Client GUI through push notifications but also on request by the BCM mitigation framework.
- E-EWS supports sharing with external entities of CTI information and mitigation plans, sanitising data with the help of DAT.

The following diagram presents a high-level overview of the CTI Framework that included the forecasting and detection interfaces.

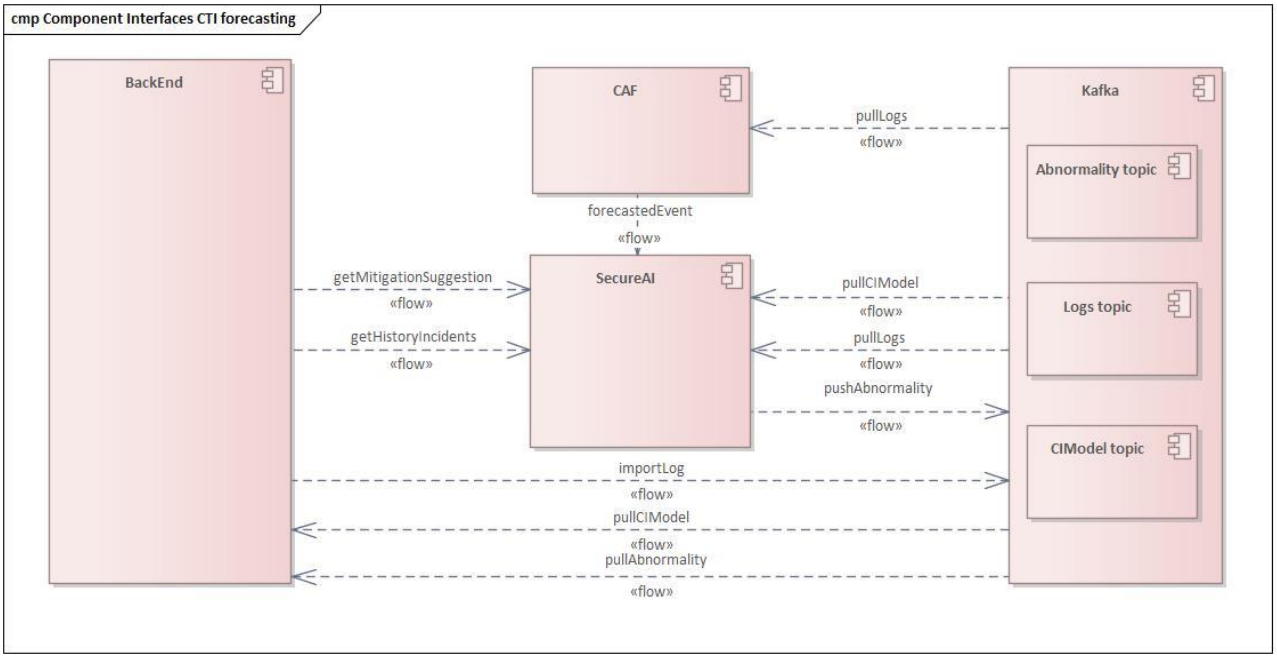


Figure 4: DYNAMO's CTI Framework - Forecasting and detection interfaces

The following diagram presents a high-level overview of the CTI Framework – correlation, enrichmenent and sharing interfaces.

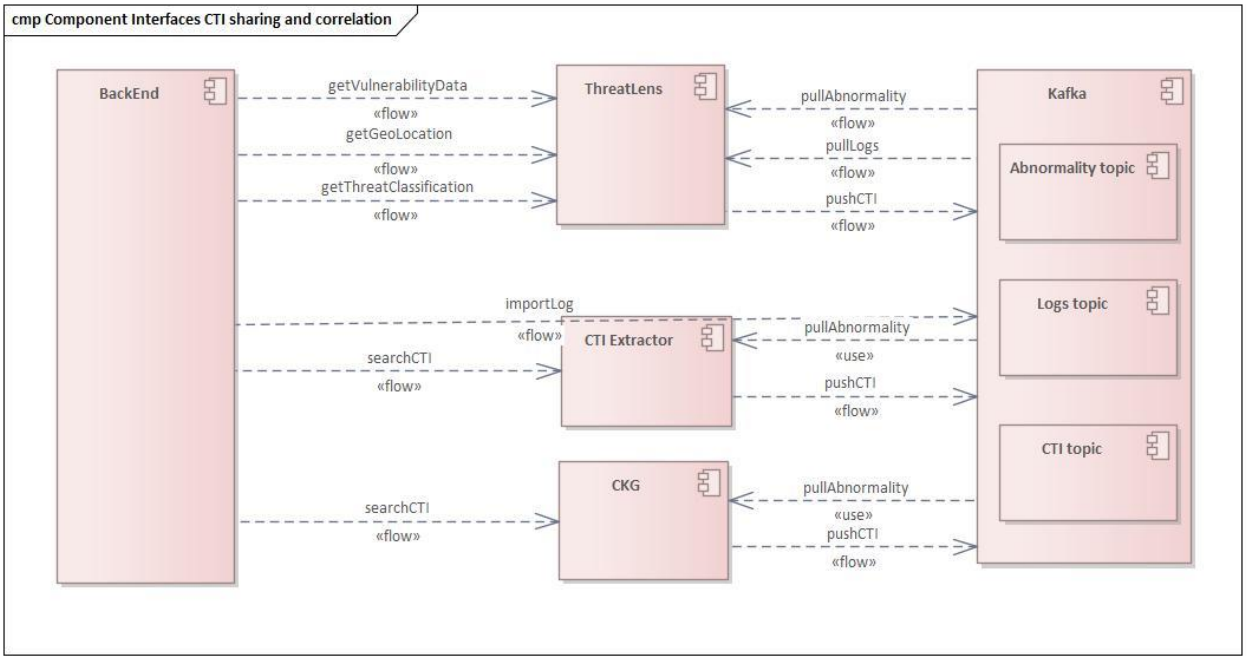


Figure 5: DYNAMO's CTI Framework – Correlation,enrichment and sharing Interfaces

3.3 CTI information workflow – Interconnection between the components

While the previous section provides a technical overview of the CTI's Framework, this section will provide an example of the way that the DYNAMO platform will utilise the CTI information and how this will be transferred with the use of it. DYNAMO does not simply aim to create a selection of independent tools. Its purpose is to compile a toolset that will assist the end-users with a comprehensive approach on collecting, analysing and utilising CTI and provide relevant and tailored information to the BCM framework. That will lead to a synergistic integration of CTI and BCM.

For that particular reason, the tools need to be able to collaborate with each other. This is depicted in the Figure 6 below.

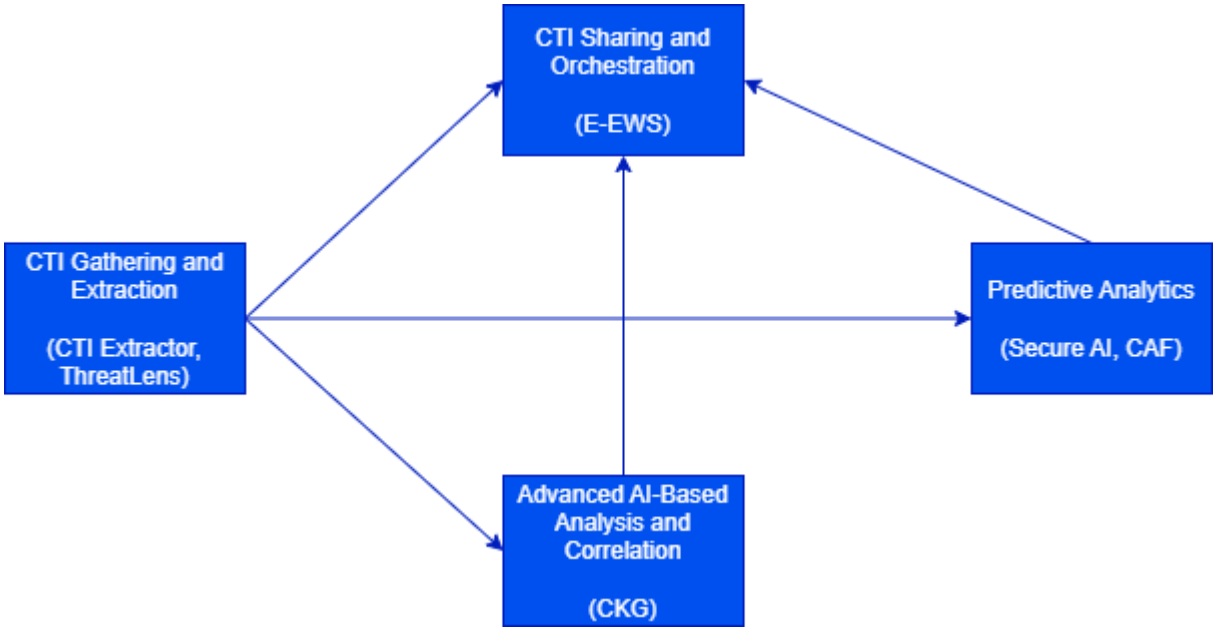


Figure 6: Information Workflow between the CTI Framework's tools.

CTI Extractor and ThreatLens are responsible for gathering and extracting CTI that is fed to the other tools for the following purposes:

- CKG will use this information in order to provide knowledge graphs using AI-Based analysis and correlation techniques
- E-EWS will use the extracted/enriched information to share it between the members of the organisation and also the members of other teams
- SecureAI and CAF will use the CTI to provide prediction and forecasting of possible upcoming attacks.

The Fine-Grained Access tool and DAT are tools that contribute to the general operation of the DYNAMO platform, playing integral role in various points in its architecture. Due to their support in the information-sharing activities of users, the tools are considered as part of T4.2.

Since the E-EWS is the main medium for sharing information, ThreatLens and CKG can use it in order to enrich information before sharing it with other entities.

An example of possible use of the CTI Framework is depicted in the following Figure 7.

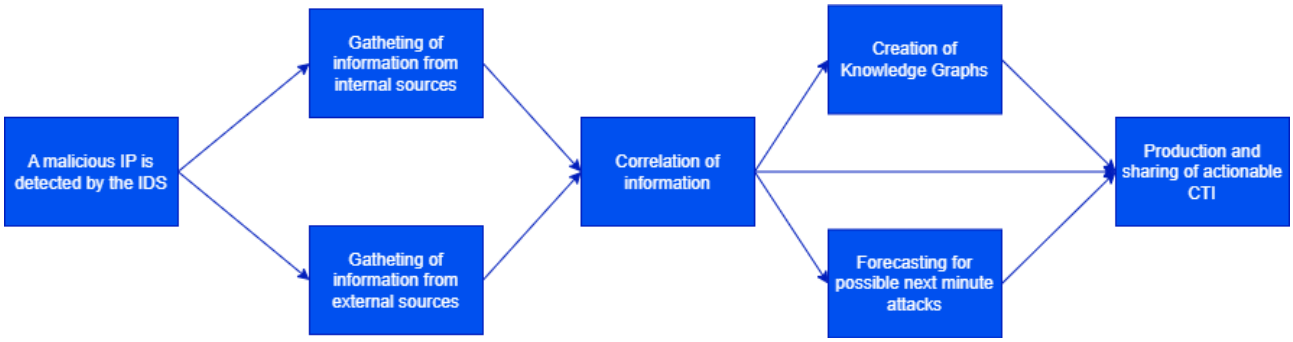


Figure 7: Example of the utilisation of the CTI Framework after an IDS alert.

- Once a malicious IP is detected by the a deployed IDS on the CI or by the detection module of DYNAMO (SecureAI & CAF components), the CTI Extractor is used to analyse the received information, using its external source and correlating it with existing information on the IP.
- The collected information is used by the CTI Extractor and ThreatLens in order to obtain situational awareness on a possible threat that can be attributed to the malicious IP.
- This information can be used further with the following purposes:
    - Forecasting of possible next minute attacks with the use of CAF
    - Provision of predictive analytics by SecureAI
    - Creation of knowledge graphs for increased situational awareness with the use of CKG
    - CTI Extractor can be used to share this information with MISP, ThreatLens can enrich an E-EWS cyberticket before sharing it

The entire process is supported by DAT that is anonymising any data that require such actions prior to sharing information. As seen in the previous section, the results of the CTI Framework's operations will not only be used as part of the CTI practises but will also be utilised by the BCM Framework in order to obtain information that is critical for the operations of an organisation, increasing resilience, simulations and training content.

Chapter 4 CTI Framework's Components

4.1 Introduction to the CTI Framework

The chapter discusses an overview of the tools of the CTI Framework. To avoid overlapping with the technical content of the deliverable D2.2, the section focuses on the operation of the tools, their theoretical background and their utilisation within the DYNAMO platform.

4.1.1 Introduction to the needs behind choosing the tools of the CTI Framework

Tool Name	Tool's Purpose	Tool's state at the beginning of the project	DYNAMO's aim for the tool
	Why was the tool selected for DYNAMO	**Initial state of the tool at the beginning of DYNAMO**	**Tool's capabilities at the end of the development cycle**
CTI Extractor	• Uses internal and external sources and correlates them with existing threats and vulnerabilities. • Crawls the internet to identify and obtain relevant information • Utilises MISP to collect, correlate and share information	• MISP utilisation • Connection with external sources • Correlation capabilities • Crawling	• Advanced Correlation Mechanisms • Expansion of crawling capabilities • Improvements in the UI • Addition of geolocation • Correlation of network topologies to existing vulberabilities • Extension of the list of external sources
ThreatLens	OSINT analysis platform for providing insights about identified attacks, threats, and vulnerabilities, Visual analytics and Natural Language Processing capabilities for security-related incidents. Can function as a Security	• Plugin for EWS (Early Warning System) platform. • Capabilities for analysing data from various sources including E-EWS tickets, MISP and third-party databases • IP address, CVE and bitcoin address analysis • NLP capabilities for estimating	• Enhanced algorithms for vulnerability clustering and categorisation using NLP techniques. • Continued support and enrichment for of visual analytics toolset., NLP-based severity estimation, and security data analysis from external sources • Adaptation of the tool to work with DYNAMO's selected information sharing system (e.g.,

	Information and Event Management (SIEM) system and a supplementary security Knowledge Base	vulnerability severity and categorisation	EWS or MISP) and components (Kafka).
CKG	Complements situational awareness features provided by other tools.	• Property graph • No inference • Limited text search capabilities • No NLP • Static correlation	• RDF store • Use of inference • Use of NER • On demand correlation / virtual graphs.
SecureAI	AI-powered security inspection solution that can identify risks, threats, and abnormalities from various data sources (network, application, firewall logs) provides information and resources to safeguard against threats and harmful software. Contributes to the creation response plans (detection-based suggestions to the responsible BCM framework component) while also helps creating training material to help users prepare against cyberattacks.	• Attack detection • Mitigations suggestion • Network and packet inspection • Data gathering from internal and external sources • Deep Learning mechanisms • Visual analytics toolset	• Anomaly detection on gathered data from various internal sources (network, application, host-based logs) • Predictive analytics on input data based on AI-ML techniques. • Capability of historical data analysis • Capability to simulate attacks that can be used as training to build resilience and increase awareness against possible attacks. • Classifies attack types and proposes mitigation suggestions based on predefined playbooks • Dashboard for visualising results of the analysis and alerts with key findings. • Integration with DYNAMO's collaborating components.

CAF	Provides a forecasting for possible attacks in the coming seconds	• Selection of specific features • Identification of attacks • Identification of types of attacks • ML Based Models	• Improved Accuracy • Increase in the number of ports used to identify attacks • UI Improvements • Automatic processing for designated ports
Fine Grained Access Tool	Provide a secure sharing mechanism	• Allows to entire of the partial encryption of a document • Allows a differential access to information in a document based on user attributes • User identity is checked using certificates	• Define a multilevel sharing policy • Rely on an SSO to check user attributes
DAT	Provides an efficient way to detect and anonymise sensitive data in a CTI before it is shared with partners	• Designed to anonymise Database before sharing them. • Apply an appropriate anonymisation method to each sensitive data type • Implement state-of-art anonymisation techniques	• Designed for CTI anonymisation • Applicable to raw out JSON text • Apply NLP techniques to automatically recognise sensitive data • Specially designed to recognise cybersecurity sensitive data. • Apply anonymisation techniques according to recognised sensitive data type
EWS	Provides a cyber-incident ticket management platform, allowing sharing and coordination between operational units.	• Management and sharing of EWS (Early Warning System) tickets • History view • Reports generation	• Parsing of STIX format and mapping to ticket structure • Connection to message broker and ticket update based on CTI enrichment • Anonymisation through Data Anonymisation tool • Encryption through Fine Grained Access Tool

Table 1: CTI Framework Tools Purpose and Status

The characteristics of the tools can be matched to the challenges mentioned in section 2.1.

- CTI Extractor, ThreatLens and CKG to utilise timely and actionable CTI while avoiding threat data overload to increase awareness
- SecureAI and CAF produce actionable information in a timely manner to prepare and protect from threats and attacks
- E-EWS, Fine-Grained Access Tool and DAT provide the prerequisites for a trusted and secure environment that allows information sharing according to regulation and with accordance to organisational rules.

Their combined use and their collaboration with tools of the BCM Framework is providing an advantage to an organisation in order not only to protect its infrastructure and operation on a tactical level but also support the decision-making process and policies for the purpose of security and continuation of operations.

4.1.2 Collaboration with the BCM Framework

CTI is used by an organisation to provide answers to the What, When, Why, Who, How that are related to an incident. The purpose of using CTI it not only to respond to an incident but also to improve the protection of an infrastructure against potential threats. BCM aims to achieve coordinated response and recovery during incidents and minimise the impact of an attack to the infrastructure and the operations of the organisation.

Achieving a synergetic approach between for CTI and BCM, will improve incident response, achieve better coordination among responsible entities and provide more complete response plans for the protection against threats and attacks. BCM can provide to CTI the means to increase the granularity is required in order to find tailored answer to the questions mentioned above that protect against threats and identify vulnerabilities in a system by decreasing information overload. As it is stated in the deliverable D3.1, “the integrated BCM-CTI approach encourages the operational teams to engage and communicate proactively with decision makers. Establishing regular engagement is key to ensuring that in an emergency, communication remains clear and functional.” With the use of its tools, DYNAMO can offer effective and informative visualisations that can be used in order to bridge the communicational gap between an analyst/engineer and the decision-maker of an organisation. Also, with the use of the E-EWS and MISP, DYNAMO not only increases the communication and information-sharing among the members of an organisation but also sets up a strong prerequisite to achieve communication with external entities.

These answers can be utilised by the Business Continuity Plans that BCM is developing.

At this point there is no evident integration of BCM and CTI in any market solutions or other frameworks. DYNAMO, driven by the needs that are identified in the literature and also by the requirements of establishing a working cooperation between the CTI and the BCM Framework has implemented an approach where the ideas, the theoretical approaches and the challenges behind the two frameworks are fused and materialised by the collaboration of tools from the two frameworks. With this approach, DYNAMO is aiming to bridge this gap by providing a platform that allows the exchange of technical information to provide resilience and protection against threats and identification of vulnerabilities.

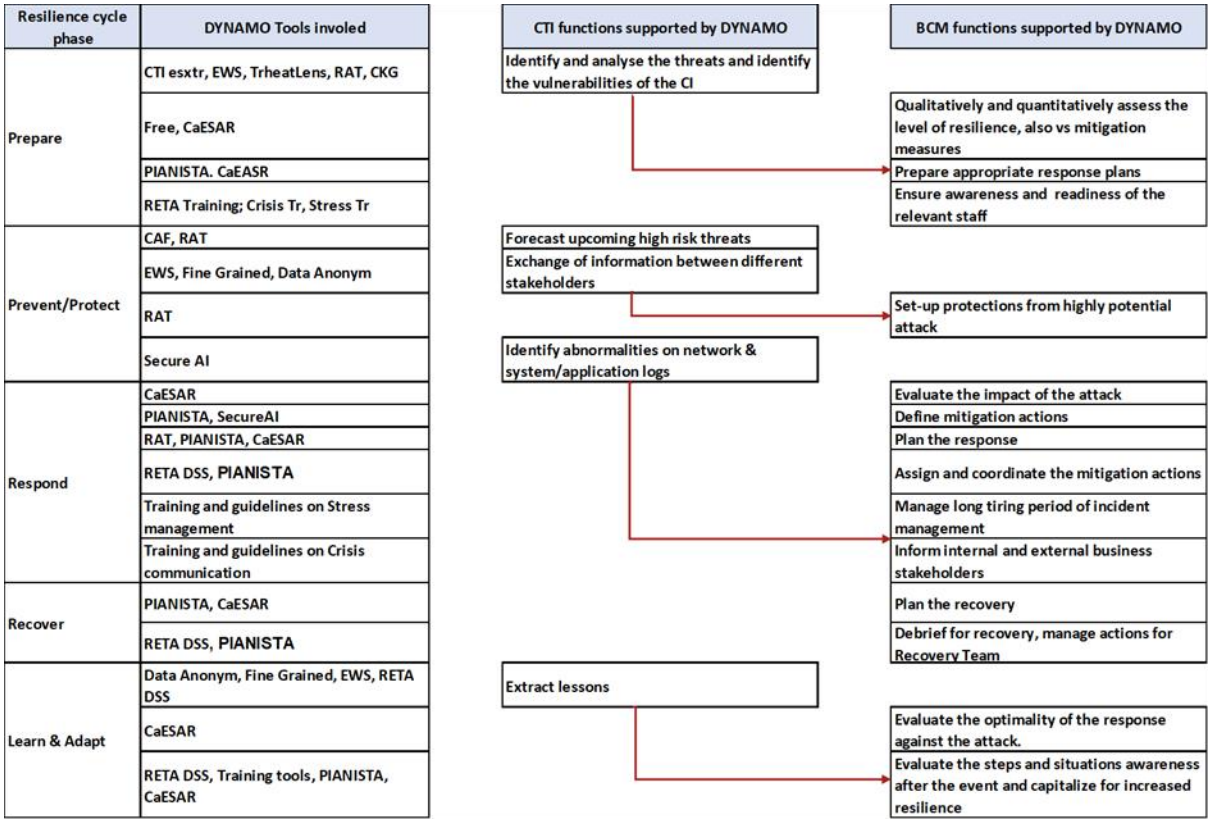


Figure 8: DYNAMO Tools mapped to resilience cycle and how CTI links to BCM functions. (source: D3.1)

Figure 8, taken from D3.1 summarises the integration of BCM and CTI tools across the resilience cycle; within the concept of the DYNAMO framework. The tactical activities of the CTI tools can provide support in the operational and strategic level that BCM is established while addressing each resilience phase.

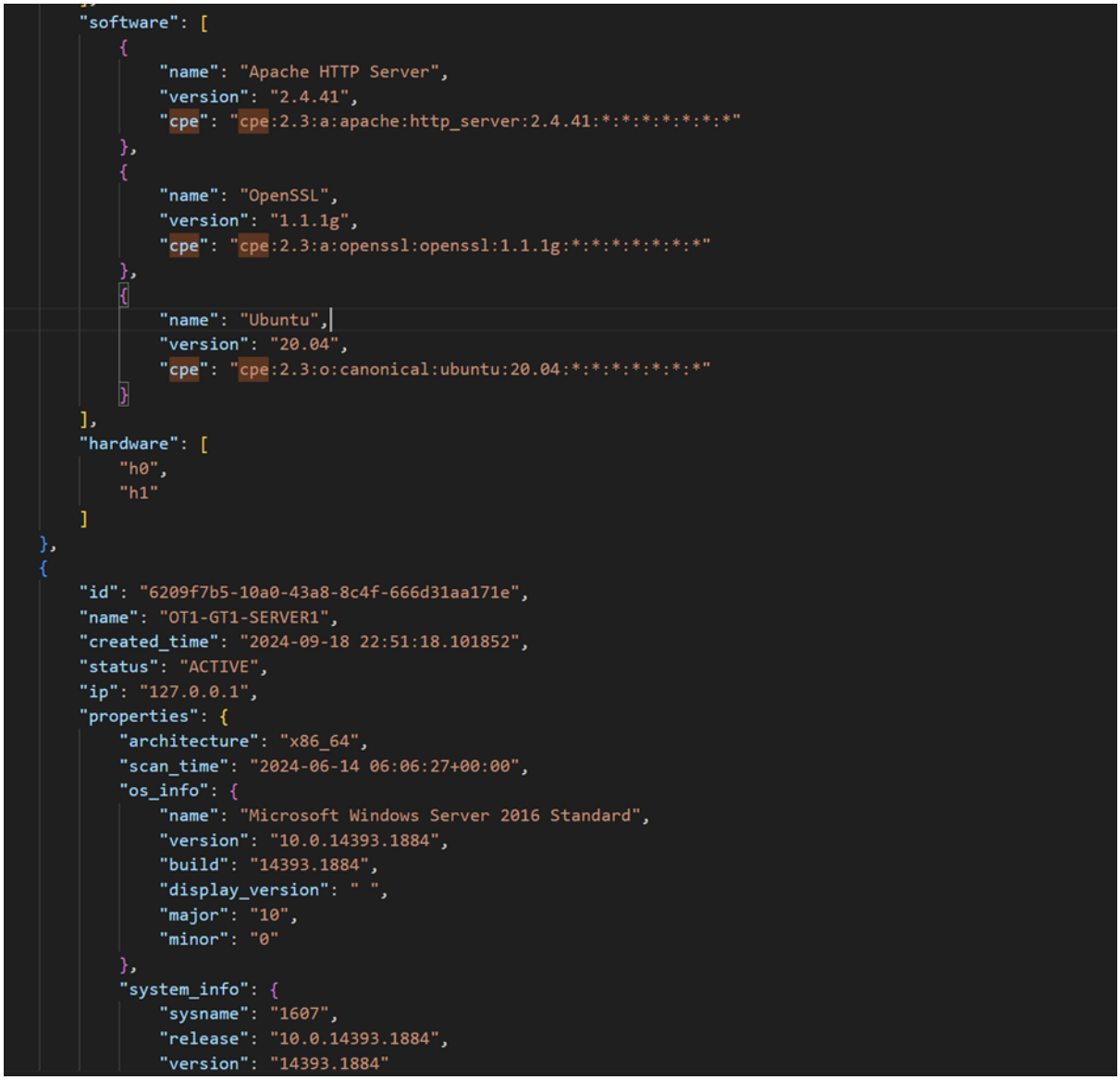
Two specified workflows that can be used to highlight the technical collaboration of tools are the following:

CaESAR (T3.3) – CTI Extractor (T4.1)

CaESAR is a tool that provides the topology of an infrastructure in order to simulate single/multi-point failures in a modelled CI and support in the preparation stage of the resilience cycle. It can perform offline and online analyses of impacts on networks to identify single point/multi-point failures or specific threats/failures in the network in real-time, based on data from a suitable integrated platform. It uses network, and threats modelled as described above along with probabilities for failures, delays, and repair times with certain variance to quantify and analyse resilience of the networks.

CTI Extractor is a tool utilised for the collection, extraction, analysis and correlation of CTI from both several external (i.e., online), as well as internal sources. The collected data are analysed in order to identify correlations between the information collected both from external as well as internal sources (e.g., correlation between CPEs and CVEs).

CTI Extractor will obtain a network topology (or any updates in a given topology) from CaESAR in order to use its correlation engine and identify any existing vulnerabilities related the network's infrastructure.



```
    "software": [
        {
            "name": "Apache HTTP Server",
            "version": "2.4.41",
            "cpe": "cpe:2.3:a:apache:http_server:2.4.41:*:*:*:*:*:*:*"
        },
        {
            "name": "OpenSSL",
            "version": "1.1.1g",
            "cpe": "cpe:2.3:a:openssl:openssl:1.1.1g:*:*:*:*:*:*:*"
        },
        {
            "name": "Ubuntu",
            "version": "20.04",
            "cpe": "cpe:2.3:o:canonical:ubuntu:20.04:*:*:*:*:*:*:*"
        }
    ],
    "hardware": [
        "h0",
        "h1"
    ]
},
{
    "id": "6209f7b5-10a0-43a8-8c4f-666d31aa171e",
    "name": "OT1-GT1-SERVER1",
    "created_time": "2024-09-18 22:51:18.101852",
    "status": "ACTIVE",
    "ip": "127.0.0.1",
    "properties": {
        "architecture": "x86_64",
        "scan_time": "2024-06-14 06:06:27+00:00",
        "os_info": {
            "name": "Microsoft Windows Server 2016 Standard",
            "version": "10.0.14393.1884",
            "build": "14393.1884",
            "display_version": " ",
            "major": "10",
            "minor": "0"
        },
        "system_info": {
            "sysname": "1607",
            "release": "10.0.14393.1884",
            "version": "14393.1884"
```


Figure 9: Sample Topology from CaESAR

By combining the use of the tools, the DYNAMO user will be able to obtain relevant information for their infrastructure that can be up to date, actionable and avoid information overload.

SecureAI (T4.4) – Pianista (T3.4)

SecureAI is an AI driven solution that acts as a cybersecurity Awareness tool that provides users and security experts with the knowledge to protect their systems and organisations from threats and malicious activities. The tool can detect data irregularities and identify future risks, security threats and provide useful insights. It can support in the prevent/protect and also the respond phase of the resilience cycle.

Pianista is a tool that supports the user at the respond and recover phase. It automatically generates solutions for problems, such that the solution is in the form of a set of actions (or ‘plan’) that must be followed to solve the problem. The tool takes as input representations of the world (in the form of a ‘domain’), its current state and the desired final state (both in the form of a ‘problem’). The plans are then generated using classical AI planning algorithms and heuristics.

By combining the use of the tools, the DYNAMO user will be able to achieve an improved response to a threat or an incident. SecureAI will provide tailored mitigation suggestions to Pianista to address the problem that appears in a critical topology.

SecureAI's mitigation suggestions can also be used as training material by RETA, BCM's training tool in order to increase the awareness of the DYNAMO users and also reflect on the human aspect that is part of BCM.

From these examples it can be seen that the communication between DYNAMO's two frameworks is bidirectional. BCM and CTI require information and data from each other in order to increase efficiency, specificity in their operation and to provide the user with tailored, working solutions to their problems.

4.2 CTI Extractor

4.2.1 Tool overview and concepts

CTI extractor is a tool that is being developed for the collection, extraction, analysis and correlation of CTI from both several external (i.e., online), as well as internal sources. The external sources include sources such as vulnerability databases (e.g., the NVD[11]), Computer Emergency Response Team (CERT) feeds, databases with Proof of Concept (PoC) exploits, social media, forums, and relevant web pages from the Surface Web (e.g., websites relevant to cybersecurity) and the Dark Web (e.g., darknet market). The internal sources include web services, databases, computer systems, and application logs.

Data collection and analysis targets concerning the acquisition of IoCs:

- Hash Values
- IP Addresses
- Domain Names
- Network/Host Artifacts
- Tools
- TTPs

The IoCs are used to obtain CTI Information coming from the collection and analysis of data which needs to be converted into an actionable data format (CTI Extractor is using the STIX[12]2.1 format). The actionable data can be used to feed information to the rest of the DYNAMO platform (as seen in section 3.3) and as part of the information-sharing activities of CTI analysts (i.e. internal or external sharing).

CTI Extractor filters the collected data to avoid storing personal data leveraging rule-based techniques and extracts CTI from the collected sources using rule-based and ML-based techniques. Subsequently, the collected data are further analysed in order to identify possible correlations between the information collected both from external as well as internal sources (e.g., correlation between Common Platform Enumeration (CPE)s and CVEs). CTI Extractor utilises both simple (e.g., MISP correlation) and advanced (e.g., ML-based) correlations of threats. The collection process can be conducted either manually through a user-friendly GUI or automatically.

[11] https://nvd.nist.gov
[12] https://oasis-open.github.io/cti-documentation/stix/intro.html

4.2.2 Architecture

The CTI Extractor's architecture is depicted in Figure 10.

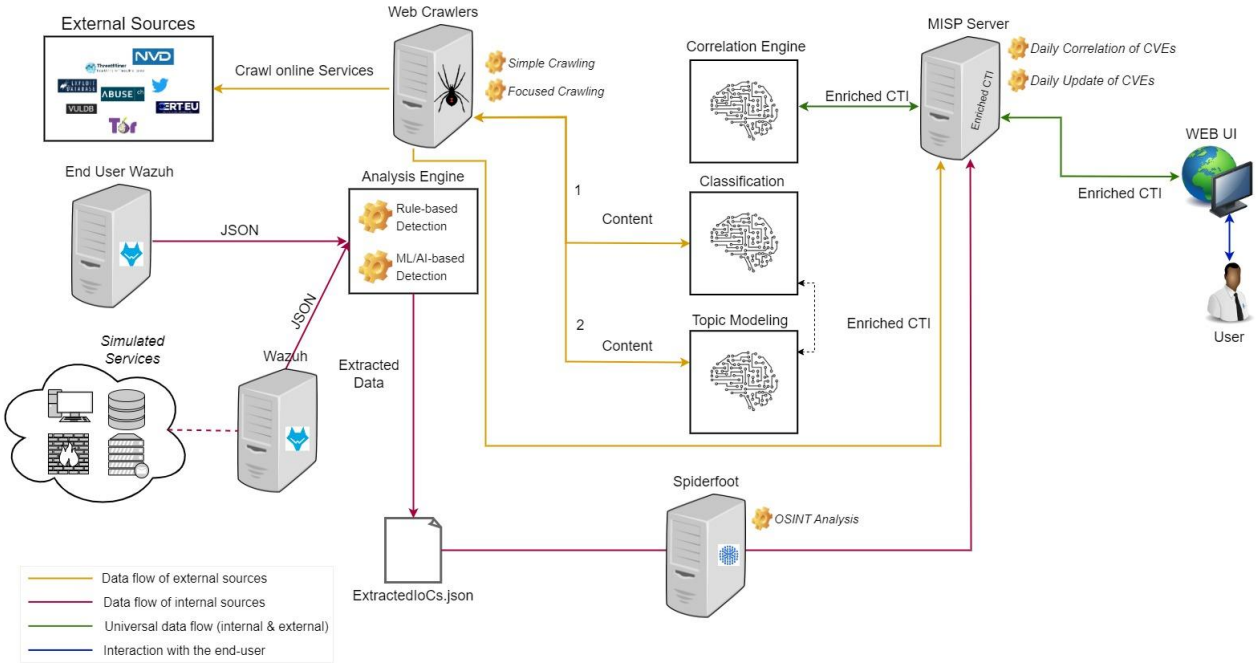


Figure 10: CTI Extractor's Architecture

For the collection of information from the aforementioned external sources, the tool is using Web Crawlers for several types of sources including Surface Deep and Dark Web sources in order to collect relevant information on threats, attacks, vulnerabilities etc. The options for crawling are the following:

- **Simple Crawling:** The user is targeting a URL to discover and index its content
- **Focused Crawling:** The crawler focuses on CTI related content, based on a confidence threshold

The simple crawler is comprised of three submodules, the *frontier*, the *fetcher*, and the *parser*. The architecture of the crawler in presented in Figure 11.

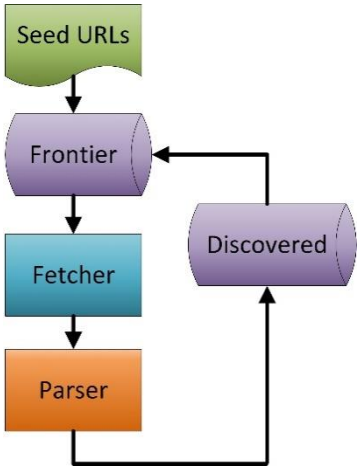


Figure 11: Simple Crawling Architecture

The frontier contains the list of URLs that are discovered and are to be downloaded. It initially includes the seed URLs (i.e., the URLs that are used as input to the crawler) and updates the list with new URLs.

The fetcher is utilised for downloading and parsing the content from the web pages. The Fetcher iteratively removes URLs from the frontier and downloads their content. To allow the crawler to seamlessly use the Surface and the Dark Web we utilise Privoxy[13], a proxy service that is responsible for forwarding each web pages to the respective dark web service. This process is repeated until one of the following goals have been achieved.

- The desired depth (i.e., a maximum distance between the seed and crawled web pages) has been reached
- A designated maximum time duration has passed
- There are no more web pages available for download.

The *parser* extracts the hyperlinks contained in each web page and feeds them to the frontier.

During the focused crawling, the crawler retrieves web pages that are relevant to the cyber-security domain and contain CTI-related information. Therefore, a filtering of the crawled web pages is required. The architecture of the focused crawler is added in Figure 12.

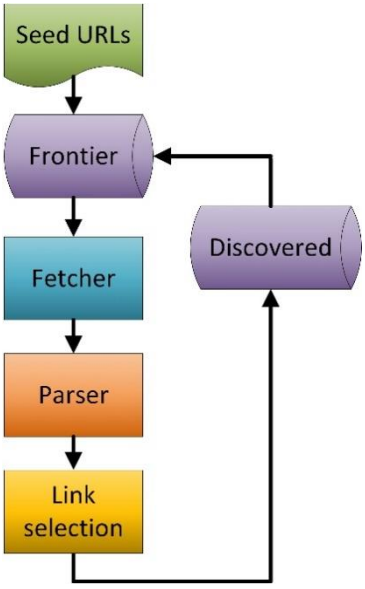


Figure 12. Focused crawling architecture

The training of the classifier is based on a manually labelled dataset that is created from six cyber security related websites, two technology-related and one well-known general news website. Each web page is cleaned off its boilerplate with the use of the Readability tool[14], so that only the main article is kept. To transform each document into a numerical feature vector, we use the Bag of Words (BoW) model [24] and the TF-IDF weighting scheme [25]. For the pre-processing, the module uses stemming, stop word removal, and replacement of CVEs. Finally, the extracted feature vectors are used as input for the initial configuration of the SVM algorithm's hyperparameters. After the configuration, the SVM classifier can be trained. The relevance of new instances of web documents as relevant to cyber-security is designed by the trained classifier.

The focused crawler will be further updated to support the automatic classification of documents to the DYNAMO domains (i.e., healthcare, energy, and maritime). For this, a second text classifier, similar to the one presented above, will classify the documents into four classes: documents that are relevant to the three domains, and documents that are not relevant to any of the aforementioned domains of interest. The domain classification functionality is placed after the CTI classification function, in order to select only those CTI-related documents that are relevant to the three domains.

[13] https://www.privoxy.org
[14] https://github.com/mozilla/readability

For this classifier, the documents were represented with the TF-IDF representation scheme. Due to their results in CTI classification, the linear SVM and the Random Forest algorithms have been used. We use multi-label classification[15] to identify the relevant content that can be detected for multiple domains. Similarly to the CTI classification, a dataset consisted of web documents is needed to be collected in order to be used for the training and evaluation of the classifier.

The infrastructure's services provide the data coming from the internal sources. These data include logs generated by servers, databases, security monitoring tools (e.g., IDS, Intrusion Prevention System (IPS)), and various other services which operate within the organisation. Data are also collected from honeypots to provide the tactics and techniques as well as the attack patterns that are used against the exposed services of an infrastructure. A Wazuh agent is used to analyse the data coming from these sources and provide alerts and notifications for detected malicious activities.

CTI Extractor's correlation functionalities enable further analysis of the stored CTI (e.g., correlation between CPEs and CVEs) in order to identify possible correlations. Through correlations between attributes and indicators from security related content (e.g. malware related information), information related to attack campaigns, advanced persistent threats etc. is produced and extracted. The outcome of this threat information analysis can be stored and shared in the case of mapping different TTPs used by the same threat actor in different attacks across a sector.

For simple correlation, the MISP correlation engine aims to identify relationships between attributes among the stored MISP events. In particular, MISP leverages a rule-based approach to examine if a certain value of an object's attribute (e.g., IPv4 address) exists in other events. The advanced correlation methods are using ML-based algorithms by analysing various features of the data. Text similarity methods are used on CERT feeds in order to identify different texts that refer to the same threat and/or relevant threats. A text similarity module takes as input a web document and gives a list of the most relevant documents to this initial output.

The events and the outcome of the extraction, correlation and analysis of the data are stored in a MISP server. The outcome can be visualised via the MISP's REST API or a potential dashboard of the DYNAMO's platform. By using MISP, the user can share any threat information that can be considered relevant to a threat. The data can also be stored in a local MongoDB instance.

The user can utilise CTI Extractor either from the tool's UI or automatically.

4.2.3 Mockups / Screenshots

Figure 13 illustrates the main dashboard of CTI Extractor and alerts that CTI Extractor can obtain from Wazuh. They included timelines, information on IPs, malware etc.

[15] https://en.wikipedia.org/wiki/Multi-label_classification

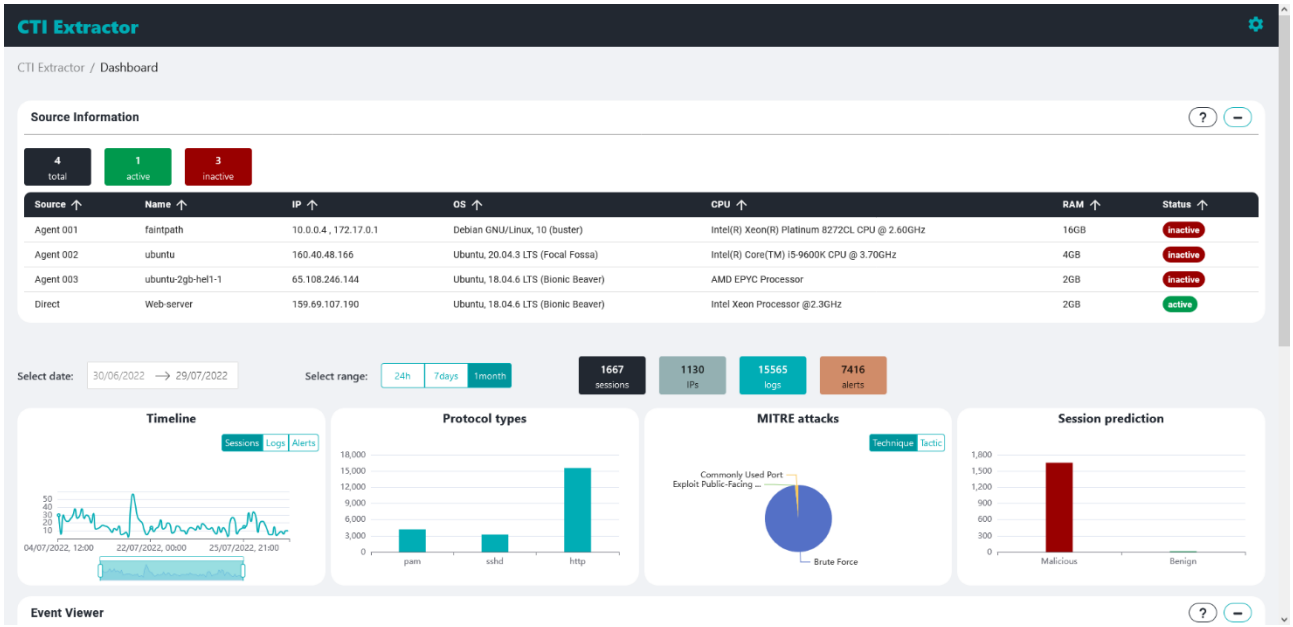


Figure 13: CTI Extractor's Dashboard

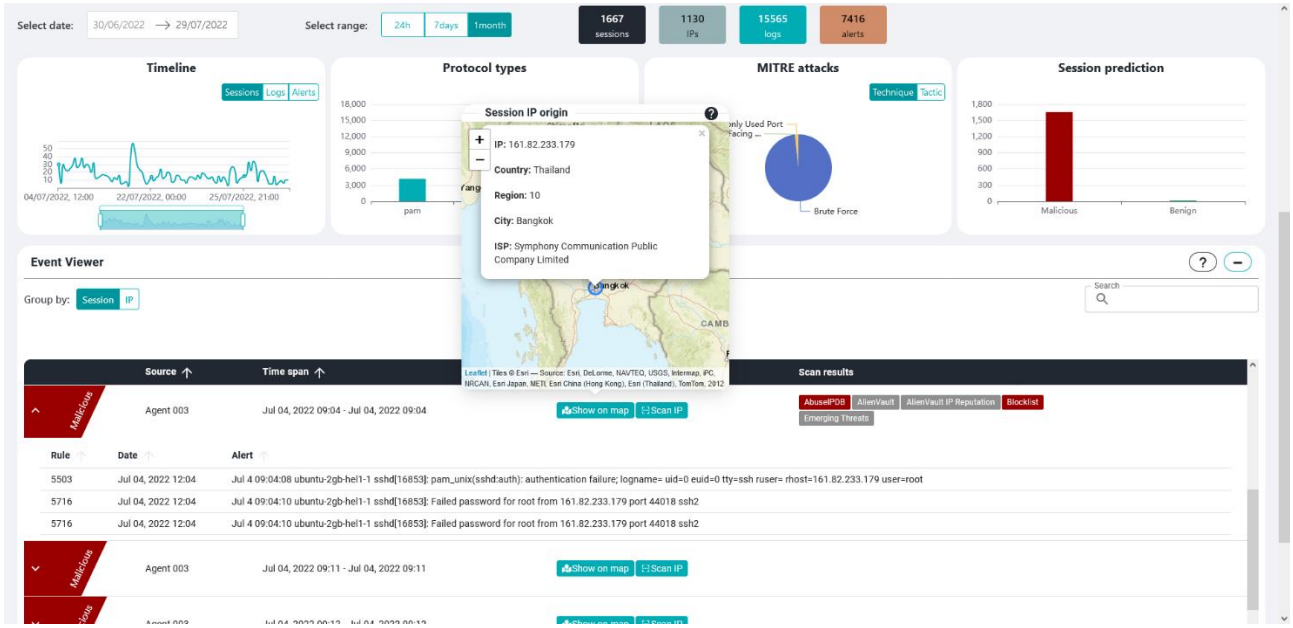


Figure 14: CTI Extractor's Dashboard with extra information from alerts

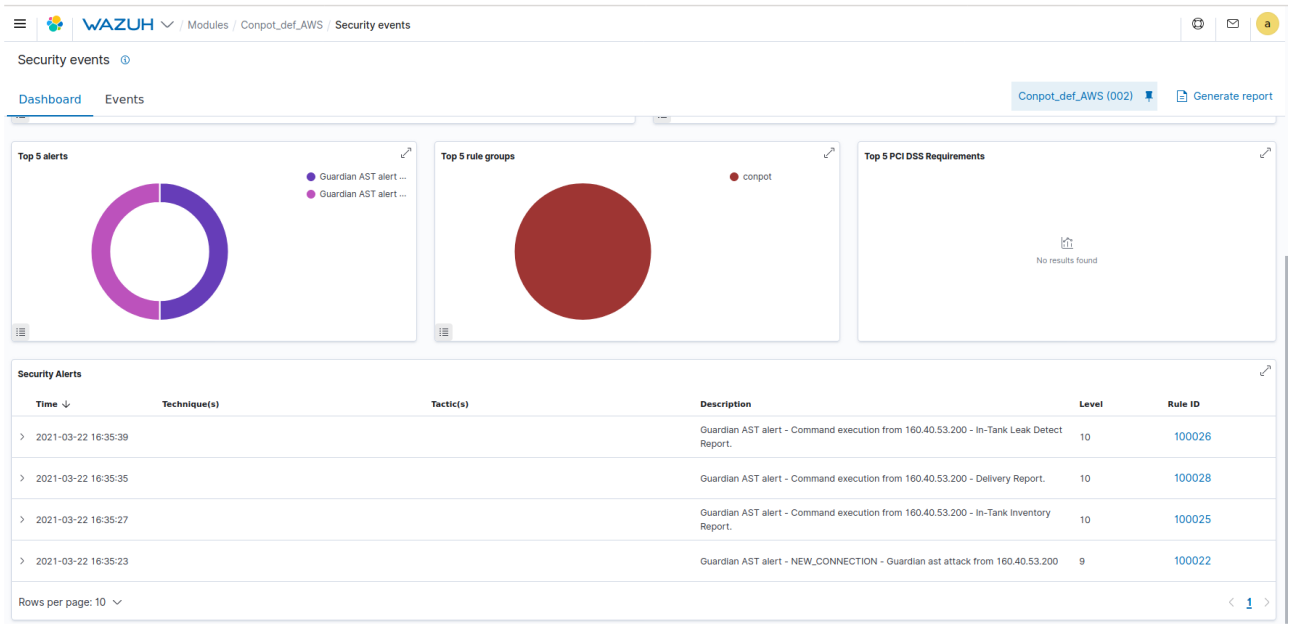


Figure 15: Wazuh Alerts to be used by CTI Extractor

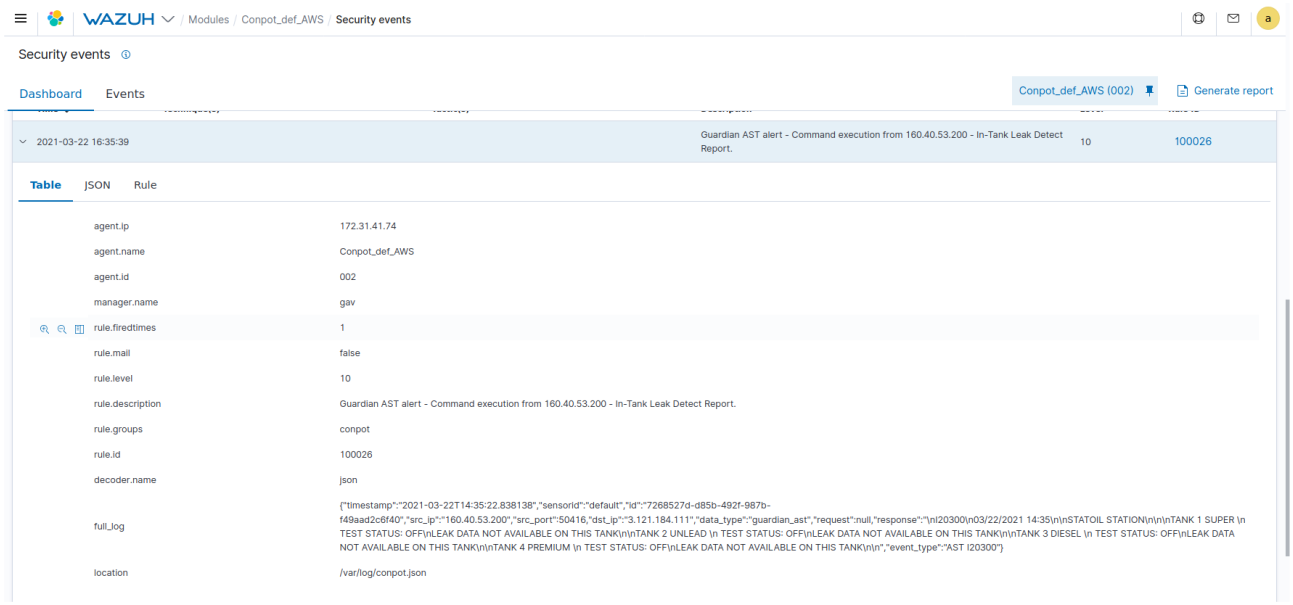


Figure 16: More information from Wazuh alerts to be used by CTI Extractor

4.2.4 Data

The plan for data utilisation for CTI extractor involves the following:

- Honeypot data coming from CERTH's honeypot
- Wazuh alerts coming from end user's Wazuh
- MISP data
- Data coming from X's API
- OSINT data (e.g. blacklisted IPs)

CERTH's deployed honeypots can be used to provide data to be mimic the internal sources that the tool requires in order to train its ML models for correlation. The T-Pot solution will be used as an all-

in-one multi honeypot platform which utilises a variety of different honeypots with some of them being relevant to the healthcare domain (e.g., dicompot, medpot). Each honeypot provides a different set of vulnerable services from different domains, increasing the range of the attracted attacks and the incoming data. Using the honeypots will provide the necessary data that is required for the training of the ML models without exposing any infrastructure to threats. The majority of the data that come from a honeypot is related to malicious actions. In order to have a more realistic training dataset, these data will be fused with normal traffic coming from the end users in order to replicate a more useful collection of data.

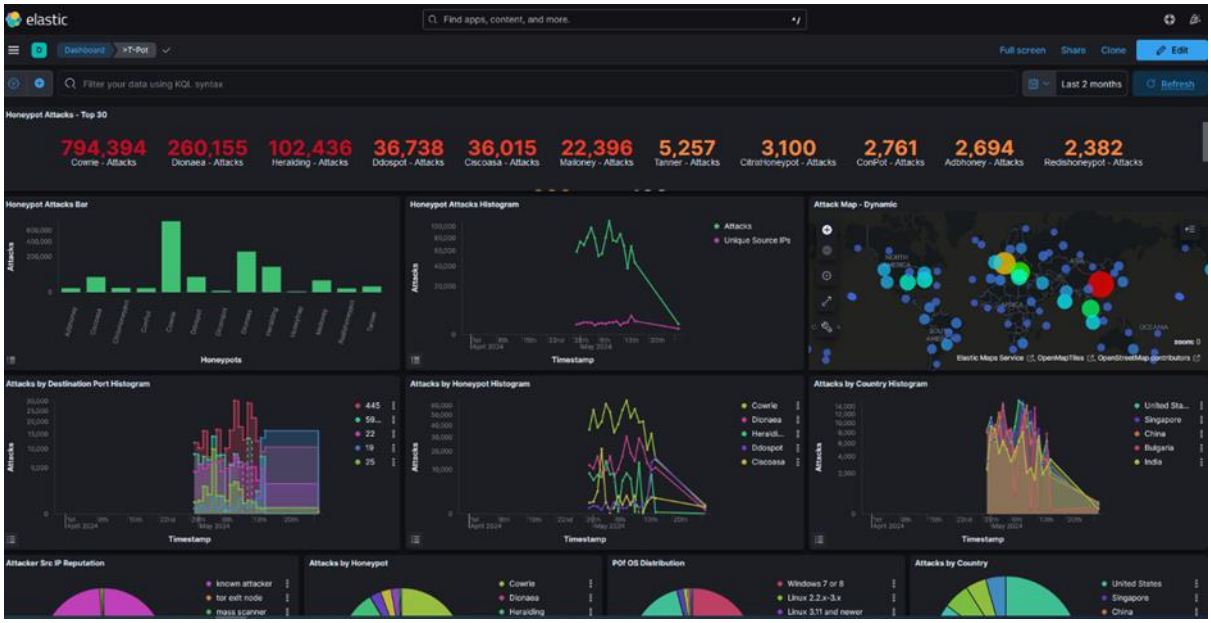


Figure 17: Sample Image from T-Pot's Dashboard

CTI Extractor will also use input from the end users Wazuh in order to correlate information from Wazuh's alerts to existing vulnerabilities.

```
"hits": {
    "hits": [{
            "_id": "T_DK4I4B54fz_2e7BfZc",
            "_index": "wazuh-alerts-4.x-2024.04.15",
            "_score": 0.0,
            "_source": {
                "@timestamp": "2024-04-15T08:05:18.216Z",
                "agent": {
                    "id": "001",
                    "ip": "192.168.10.15",
                    "name": "ot1-gt1-server1"
                },
                "data": {
                    "vulnerability": {
                        "..."
                    }
                },
                "decoder": {
                    "name": "json"
                },
                "id": "1713168318.102626370",
                "input": {
                    "type": "log"
                },
                "location": "vulnerability-detector",
                "manager": {
                    "name": "wazuh"
                },
                "rule": {
                    "description": "CVE-2021-24074 affects Windows Server 2016",
                    "firedtimes": 61,
                    "gdpr": "[...]",
                    "groups": "[...]",
                    "id": "23506",
                    "level": 13,
                    "mail": true,
                    "pci_dss": "[...]",
                    "tsc": "[...]"
                },
                "timestamp": "2024-04-15T08:05:18.216 0000"
            }
```


Figure 18: Alerts taken from Wazuh

The alerts will be utilised by CTI Extractor's correlation engine and the result of the correlation will be posted on MISP, where they can potentially also be shared according to the sharing policies of each organisation (Figures18, 19).

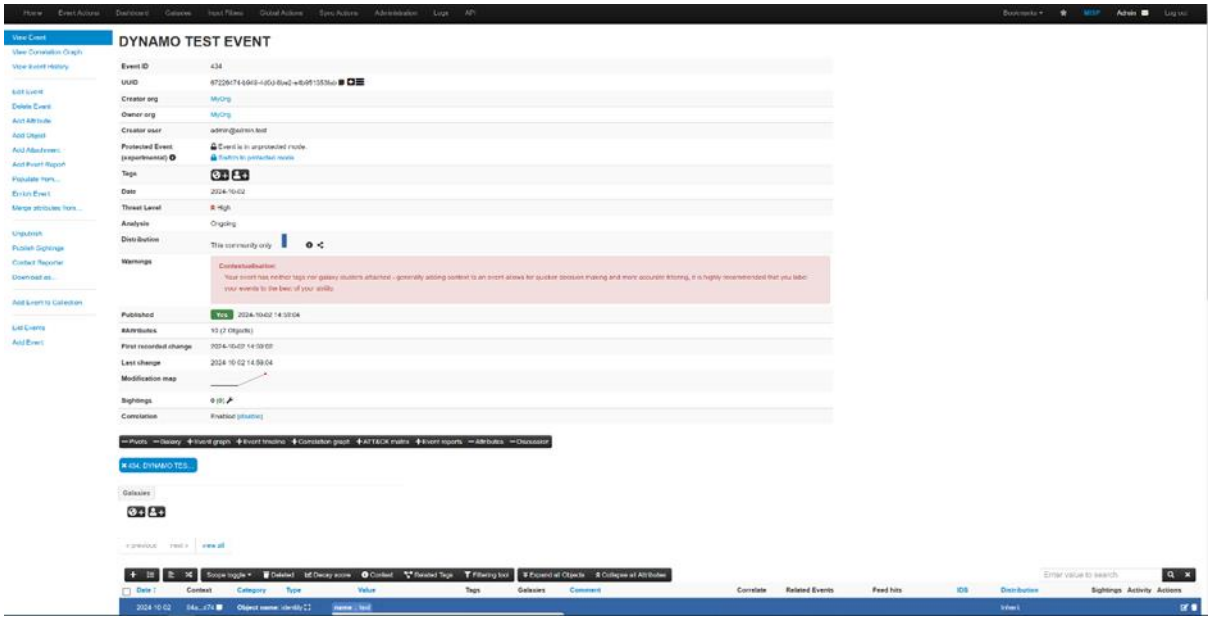


Figure 19: CTI Extractor's correlated event on MISP

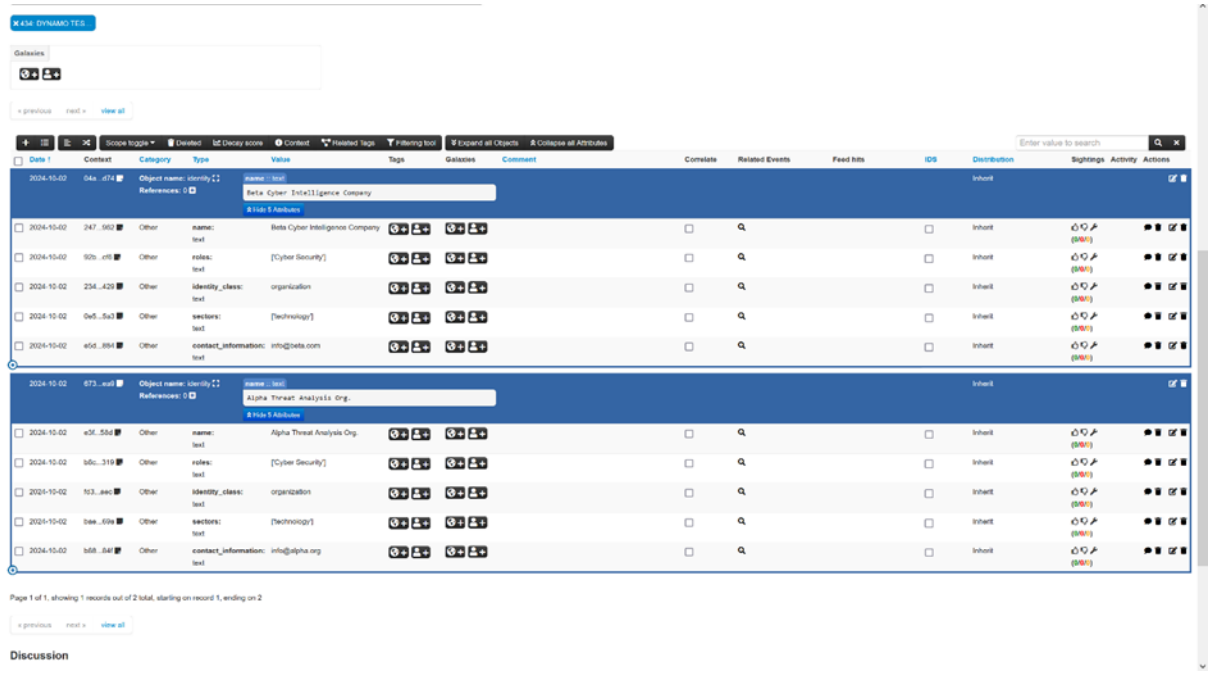


Figure 20: CTI Extractor's list of correlated objects

Additionally, For the improvement of the operation of the data crawlers, CERTH has obtained access to X's API and has created an extensive collection of relevant social media accounts to be used by the crawlers in order to obtain further data. The social medial crawler performs two actions:

(ii) The gathering of new posts from specific X accounts (prechosen by the user
(ii) The keyword-based queries containing CVE IDs (e.g., "CVE-2024-38819") to calculate the number of tweets and X accounts that are discussing them.

4.2.5 Summary and Next Steps

The CTI Extractor is an innovative tool for gathering, extracting, analysing and producing actionable CTI from multiple internal and external sources to raise awareness regarding possible threats and vulnerabilities. CTI extractor is a hardware independent tool that uses state of the art technologies and techniques, aiming to achieve quality results and create actionable threat information to assist

the operations of the DYNAMO platform. The tool can increase the relevance of the received information, by tailoring it to the specific needs of the user, when it is being correlated to a specific topology (i.e. collaboration with CaESAR) or when it is being correlated to specific alerts, received by other tools that a user is already operating (e.g. received alerts from Wazuh). This correlation can also assist in the decrease of the information overflow that is common in platforms like MISP while increasing the relevance of the shared information.

For the next stages of the development, the aims for CTI Extractor are the following:

- Creation of and analysis of actionable threat information for the information sharing purposes of the E-EWS and/or MISP (T4.2).
- Utilise the ML-based and NLP techniques for the analysis, correlation and classification of information to facilitate the creation/adaptation of dynamic taxonomies and ontologies to assist the Cyber Knowledge Graph Tool.
- Introduce dynamic taxonomies for further CTI Enrichment.

4.3 ThreatLens

4.3.1 Tool Overview and concepts

ThreatLens is an OSINT analysis tool created by CERTH as part of DYNAMO's Work Package 4. The tool utilises inputs/tickets from CTI-sharing platforms, such as MISP and DYNAMO's Early Warning System (EWS) and gathers various relevant information related to the collected tickets relevant to the nature and behaviour of cyber incidents (attacks happening on the system, open ports and services, vulnerabilities, etc.).

By analysing external third-party libraries, the tool is able to collect CTI relevant to the data originally included in the ticket and enhances the ticket with information that might be useful to an operator to analyse and gain insights on threats involving their system(s). Moreover, through its visual analytics, ThreatLens aims to provide security professionals as well as ordinary users with a comprehensive assessment of their system's security status. The dashboard comprises a map displaying the geolocation of IPs initiating attacks towards the target system(s), comprehensive details on relevant CVEs and CWEs, pertinent information on BTC wallets potentially associated with illicit actions, bar plots, time-series charts, and more.

ThreatLens also aids security professionals in assessing the criticality of identified vulnerabilities and serve as a valuable initial step in prioritising their efforts to address them. Through the utilisation of NLP techniques, the user can articulate in simple language a possible weakness, danger, or plan of action. ThreatLens' prediction and classification module automatically assess the seriousness of the input text given by the operator, based on the NVD Common Vulnerability Scoring System (CVSS) schema[16].

ThreatLens enables security teams to enhance their organisations' overall resilience by increasing their understanding on their system's security posture and identified events by analysing, assessing and characterising them. ThreatLens' capabilities include:

- Ticket Retrieval: ThreatLens receives tickets from EWS and MISP. These tickets are pushed to DYNAMO's kafka broker and after receiving them, ThreatLens collects security-related ticket information about IP addresses, CVEs, suspicious BTC addresses and other. Currently ThreatLens has been tested with tickets coming from EWS and MISP, but it can be adapted to any information sharing/ticketing-platform that supports security related events.
- Diagnostic Analytics: Using third-party databases, ThreatLens is able to acquire extra information about input threats using a combination of information sources.

[16] https://nvd.nist.gov/vuln-metrics/cvss

- Data Storage: ThreatLens stores all the input and extracted data into a database, giving the user the ability to visualise historical data and refer to if a similar threat appears in the future.
- Visual Analytics: All the gathered information is visualised in a robust and clear way to the user using several easy-to-understand plots and diagrams.
- Natural Language Processing (NLP): ThreatLens includes Machine Learning and NLP techniques in order to estimate the severity of a threat based only on a textual description. This feature serves the operator in case of uncharacterised or 0-day vulnerabilities that might be hard to evaluate and analyse.

In summary, ThreatLens is an OSINT (open-source intelligence coming for shared tickets) analysis and visualisation component that leverages AI/ML to provide insights to the user, while also enriching the information stored in a ticket. It offers an interactive interface for visual analytics of CVEs, identified attacks, BTC addresses, as well as a world map displaying the geolocation of attackers based on IP addresses. In addition to these features, ThreatLens provides users with detailed reports to help them stay informed about system security and emerging threats. One of ThreatLens' standout features is its application of NLP techniques to provide vulnerability estimation [26], while also by using unsupervised machine learing the tool is able to provide vulnerability categorisation [27] [28], allowing users to swiftly assess the severity of potential threats and take prompt action. More specifically, by simply describing a vulnerability, users can receive an estimated CVSS v2.0 vector, severity score and CWE category, enabling faster and more effective threat management. The dataset used for implementing the severity estimation & classification module was collected from NVD and was chronologically divided into training-set and testing-set.

4.3.2 Architecture

Figure 21 illustrates the structural diagram of ThreatLens. The core features are encapsulated in containers to provide transferability, scalability, and ease of deployment. The elements consist of:

- **Services**: This component is responsible for processing the API calls made by the user interface. It is responsible for managing, preprocessing, and delivering the findings to the dashboard, and it is connected to all the other components
- **Database**: PostgreSQL is included into the design to facilitate the storage of historical data.
- **Scrappers**: The scrappers are responsible for retrieving data from tickets (EWS, MISP) and other third-party libraries to ensure the data is complete
- **NLP** serves as the fundamental element of ThreatLens. Multiple models and other binary files are encapsulated in containers and can establish communication with the services through HTTP requests.

The I/O communications of ThreatLens intricately link various components through REST APIs and HTTP requests, enabling interaction between its sub-components, external sources, and docker-connected containers within its network. Tickets obtained from a ticketing platform (E-EWS and MISP) are used as inputs. ChainAbuse[17], IPinfo[18], and NVD are the third-party libraries that provide CTI, while user input is used to estimate the severity of vulnerabilities (raw text). The tool's outputs include a dashboard that displays information that is pertinent to the inputs, such as charts, bars, graphs, and raw text, but also enriched ticket pushed back to the information sharing platform. REST APIs, HTTP requests and DYNAMO's kafka broker are the communication interfaces that are implemented by ThreatLens. These interfaces are utilised both between the sub-components and external parties, as well as between the containers that are connected through the docker network. As far as the data formats are concerned, the tool anticipates receiving inputs in STIX 2.1 compliant JSON format, and it employs the same format for the output as well.

[17] https://www.chainabuse.com
[18] https://ipinfo.io

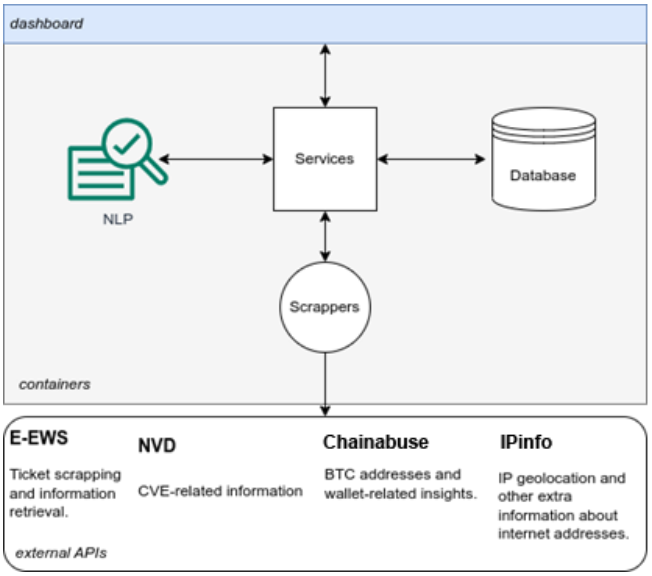


Figure 21: ThreatLens' Architecture

Within the DYNAMO project, ThreatLens interfaces with the DYNAMO's components that perform data gathering, analysis and sharing, aiming to improve the quality of CTI shared. The tool is being developed under WP4, and more especially in Task 4.1 Cyber-threat intelligence collection and extraction, as part of the information sharing and enrichment framework. This framework will be capable of collecting data from both internal and external sources, ultimately resulting in the identification of vulnerabilities and threats. In this sense, ThreatLens serves as a tool to support the workflow of analysing vulnerabilities and enhancing the CTI shared across the involved components and collaborating parties.

4.3.3 Mockups / Screenshots

The following figures (Figures 22–25) illustrate the information that ThreatLens provides by analysing inputs stored in a ticket coming from an information sharing platform. The tool can be used to enrich a cyberticket and collect relevant information on a threat or an attack. The results of the analysis is presented both to the operator but also are being stored as additional information to the retrieved ticket which can then be shared with other entities within or outside the organisation through EWS.

Figure 22: ThreatLens Attack geolocation map

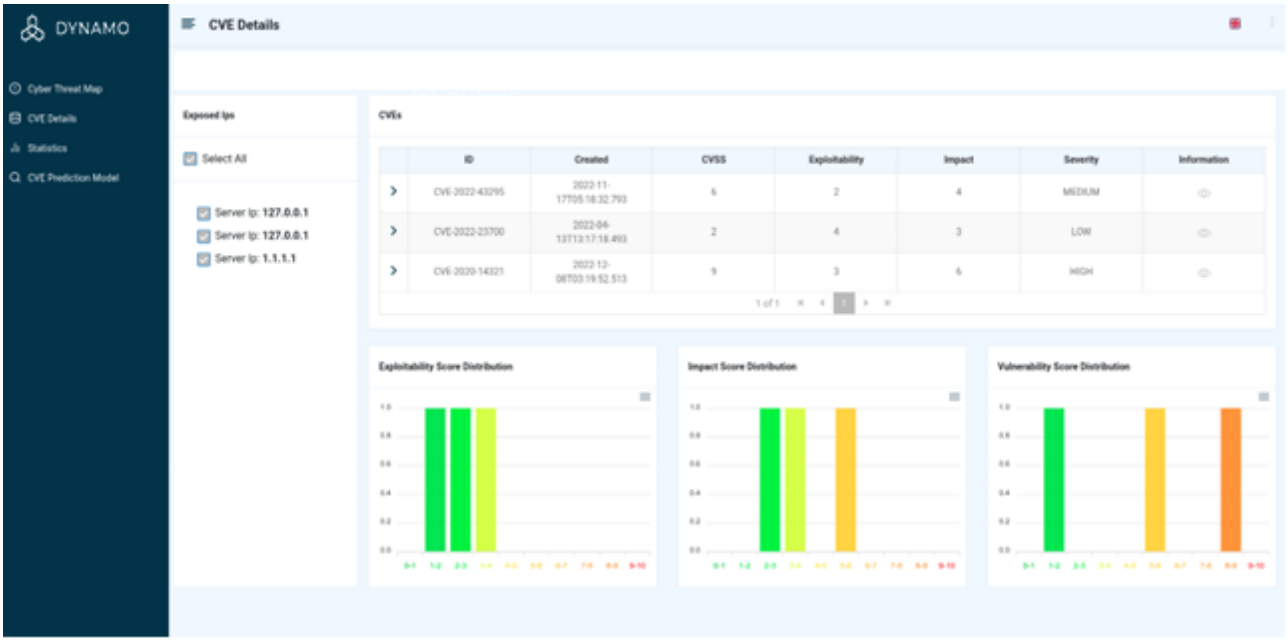


Figure 23: ThreatLens CVE information

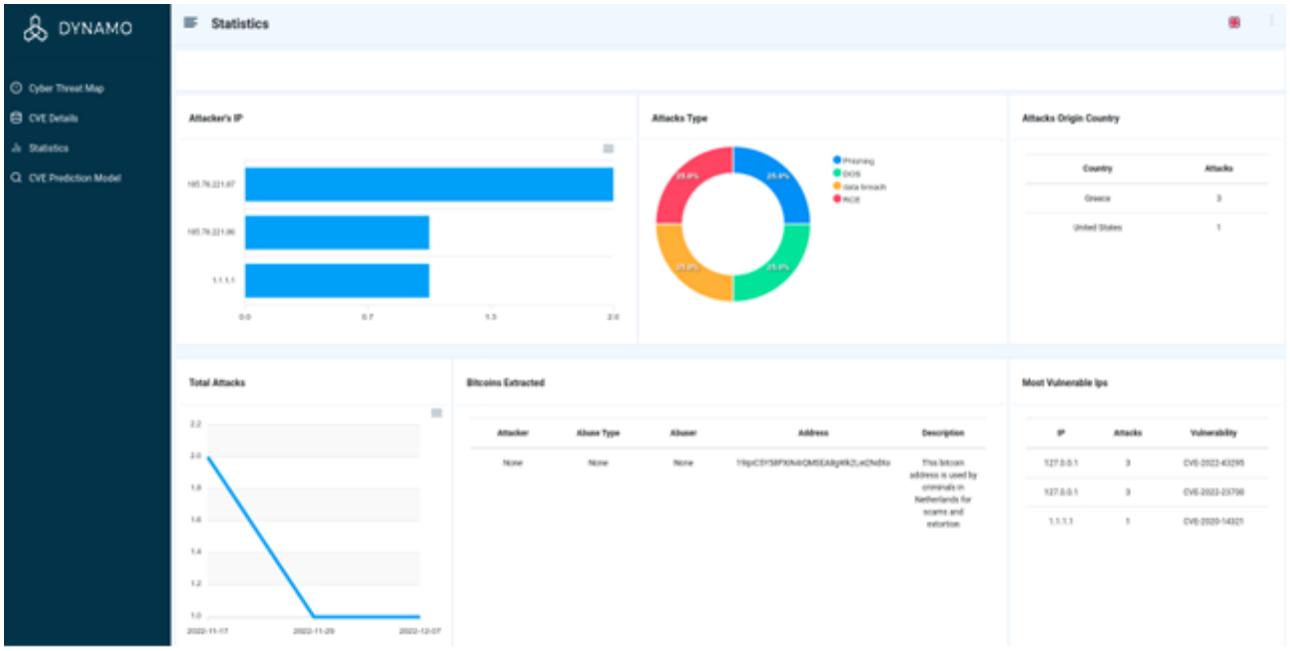


Figure 24: IP/CVE/BTC and other statistics relevant to identified attacks and vulnerabilities stored in tickets

Figure 25: NLP-based Vulnerability Prediction/Classification module of ThreatLens

4.3.4 Summary and Next Steps

ThreatLens is a technology that helps security professionals to analyse and visualise potential vulnerabilities and threats. Its role within DYNAMO is to collect the tickets shared across collaborating parties and enrich them with relevant CTI. Through the use of its visual analytics and natural language processing techniques, ThreatLens aims to give security analysts insights relevant to identified incidents on their critical infrastructure. In that frame, ThreatLens can play the role of a Knowledge Base and Information Extractor in addition to its primary function for the DYNAMO project to enrich information stored within a ticket with relevant CTI.
ThreatLens is being developed under the scope of Task 4.1 (M04-M30). The table below displays a brief timeline for the development of the task.

Time Period	Milestone Description
M04-M08	State of the Art analysis/ Research on available technologies
M08-M12	Design and Architecture of the system
M12-M18	High-Level Requirements and Early version of the system
M18-M24	Delivery of integration ready prototype
M24-M28	Testing and Refinements
M28-30	Documentation, Reporting and final Integration

Table 2: Timeline of ThreatLens' development under Task 4.1

4.4 CKG

4.4.1 Tool Overview and concepts

CKG is a cybersecurity intelligence management system that collects and correlates threat intelligence related information from various sources and maps it to the concepts of an ontology, which includes both cybersecurity concepts and related ones (e.g., IT infrastructure subject to cyber-attacks). This can then be queried and analysed by security analysts, increasing the situational awareness on the impact of cyber threats.

This tool can be used as an E-EWS extension, as it provides additional/complementary information to ticket information and allows enriching tickets and/or the reference library. It can also be used as a standalone search & analysis tool.

4.4.2 Architecture

Figure 26 illustrates the CKG's key building blocks.

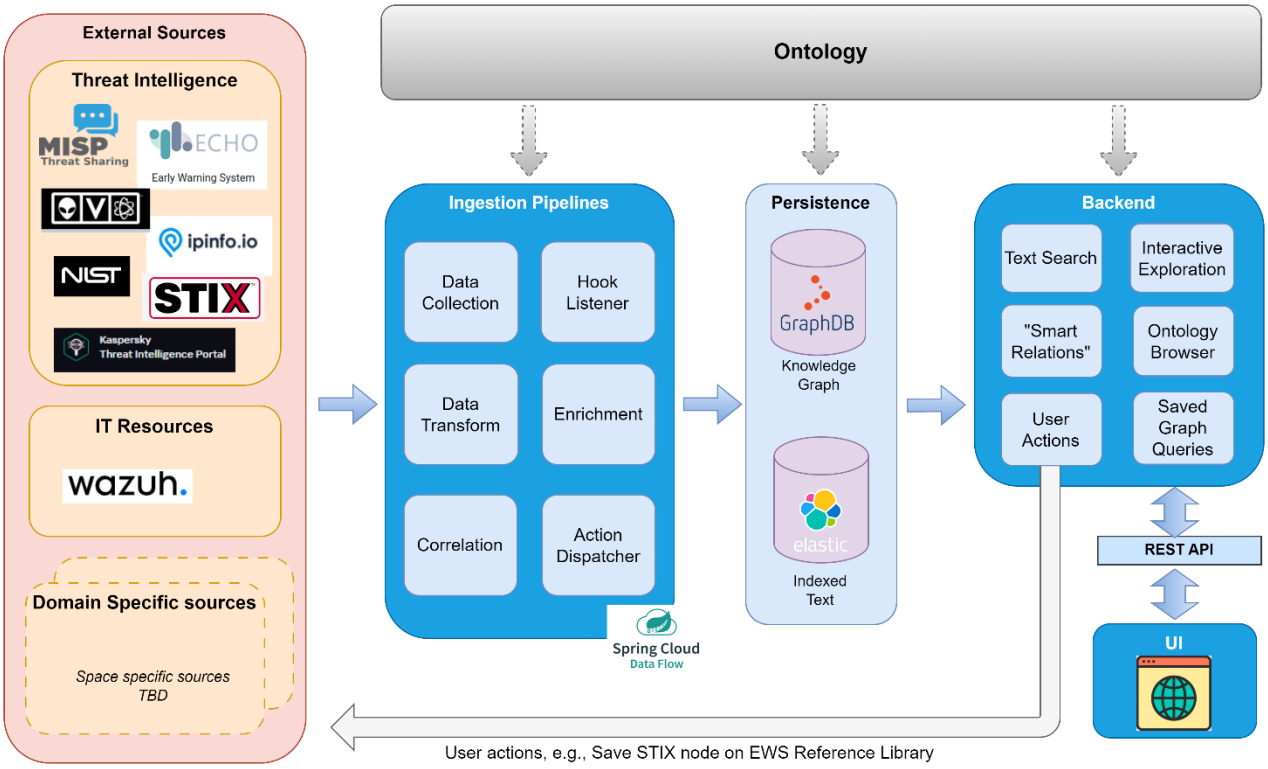


Figure 26: CKG Architecture

Ontology: The ontology used is the RDF ontology, as described in Section 2.3.2. The ontology provides the logical structure of the knowledge graph and so is transversal to all components. It models the domain concepts and the relations between them.

The ontology is modular, where each module represents the concepts specific to a domain. The current ontology has three modules: cyber threat intelligence, IT resources and space-specific entities. Cross domain relations are defined as an "aggregator" module.

The ontology is specified using the semantic web OWL framework[19]. This allows not only capturing the conceptual model, but also to take advantage of the inferencing capabilities it provides (Inferencing was not implemented yet; it is expected to be one of the key capabilities to be developed under the DYNAMO project).

External Sources: All data in the knowledge graph is retrieved from external sources of various natures:

- Threat intelligence sources. These include both public sources (OTX Alienvault[20], NVD, etc) and internal threat sharing platforms (MISP and E-EWS).
- IT resources, e.g., installed operating system and software. Linking IT resources with cyber threat intelligence is key, as it allows assessing the impact of cyber threats in the hosting organisation assets.

The process of building the knowledge graphs with the use of external sources is described in section 2.3.3.

Ingestion Pipelines: Ingestion pipelines retrieve data from external sources and, after a series of transformation, enrichment and correlation actions, store the retrieved knowledge in the graph database and text indexes.

Data ingestion was implemented using a micro-service approach, allowing pipelines to be defined in a modular way. Processing node orchestration is performed using Spring Cloud Dataflow, but the adopted micro-service approach allows replacing it by another platform with minimal changes in the core processing logic[21].

Data can be retrieved from the sources in many ways:

- Incremental collection, where all data from the sources is retrieved.
- Push-based notifications, where CKG receives data in real time by “listening” to a publish interface in the source. The E-EWS webhook interface is an example.
- On demand, where the source is queried in response to events. For example, an E-EWS ticket mentioning an IP address may trigger fetching additional information about that IP from the IP info source.

Pipelines include also “action handler” nodes that allow implementing custom user actions. For example, users can add threat intelligence nodes backed by STIX data to the E-EWS reference library. These actions are represented by pipeline nodes, similar to on-demand data retrieval ones.

Knowledge Graph: The knowledge graph persisted in a graph database, allowing graph traversals to be specified using a graph query language.

Currently CKG supports OrientDB[22] and Neo4j[23] (or any other encapsulated by Apache Tinkertop[24]). These databases fall under the category of “property graph databases”, where the graph is modelled as a set of nodes and edges that can have properties associated.

While property graphs provide all the modelling capabilities required by the ontology, they lack inferencing capabilities and have limitations when data falls under the “open world” assumption, i.e., when the same entities can be represented/identified in different ways by different sources. These limitations are overcome by having the graph modelled as and by using the RDF store[25].

[19] https://www.unb.ca/cic/datasets/ids-2017.html
[20] https://otx.alienvault.com
[21] We’re considering replacing Spring Cloud Dataflow by a more generic platform; In addition to be inefficient in terms of memory usage (each processing node runs on a separate JVM), it may not be suited for having nodes in languages other than Java, which may limit the addition of machine-learning features for which most of the libraries are written in Python.
[22] https://orientdb.org
[23] https://neo4j.com
[24] https://tinkerpop.apache.org
[25] https://www.ibm.com/docs/en/watson-explorer/11.0.0?topic=administration-rdf-stores

RDF stores represent knowledge as statements about the domain (RDF statements), which is compatible with the property graph abstraction and provides additional semantics. In RDF stores, even the ontology definition (in OWL) is represented as a set of statements ("knowledge about the knowledge") that have precise semantics. In particular, OWL statements can be used to infer new knowledge (statements) based on explicit knowledge.

As part of the DYNAMO activities, we plan to move the graph database from a property graph to an RDF store, more specifically, Ontotext GraphDB. This database includes an extensible inferencing engine (that includes most of the OWL inference rules) and optimisations that allow handling entity equivalence in an efficient way.

Indexed Text: Text searches are currently implemented using the Apache Lucene library, which is the one behind Commercial off-the-shelf (COTS) products such as Elasticsearch[26] and Solr[27]. This was due to licence cost concerns for Elasticsearch at the time it was implemented, and that are no longer applicable.

The CKG text search implementation will be totally replaced by Elasticsearch, thus simplifying the code base and benefiting of the security, scalability, reliability and fault tolerance offered by Elasticsearch out of the box. The latest versions of Elasticsearch also provide search capabilities not available in Lucene, namely dense vectors and K-nearest neighbour search that allow using transformer techniques for semantic search.

User Interface: CKG exposes all its functionality in a REST API, which is used by a web rich client application implemented in Angular. This application allows users to query and navigate the knowledge graph, thus exposing the collected data and their relations. Screenshots in section 4.3.2.1 provide an idea of the current interface.

Communication and I/O. CKG provides the following interfaces:

- Webhook interface to E-EWS, where E-EWS ticket changes are received in real time.
- REST API with all the functionality used in the UI.
- REST API for pipeline management operations – provided by Spring Cloud Dataflow.

CKG uses the following interfaces:

- REST APIs exposed by the data sources.
- E-EWS REST API to add contents to the reference library.
- Kafka topic on any pipeline node. This allows ingesting data directly to a pipeline by publishing in the corresponding Kafka topic.

CKG may also interface with any OpenId-compatible IDP interface (e.g., Keycloak) in case the single sign on functionality is enabled.

Data Persistence: CKG persists data in two repositories:

- Knowledge graph database, structured according to the classes defined in an ontology.
- Lucene indexes, for searchable text data.

Both stores are schema-less.

4.4.3 Mockups / Screenshots

Figure 27 shows the results for a search for the APT-27 threat actor[28]. It is known as "Lucky Mouse" and is related to actor "Emissary Panda" through MISP ticket MISP-1362. Nodes may have "actions" and "smart relations" associated, which depend on the type of node.

[26] https://www.elastic.co
[27] https://solr.apache.org
[28] https://attack.mitre.org/groups/G0027/

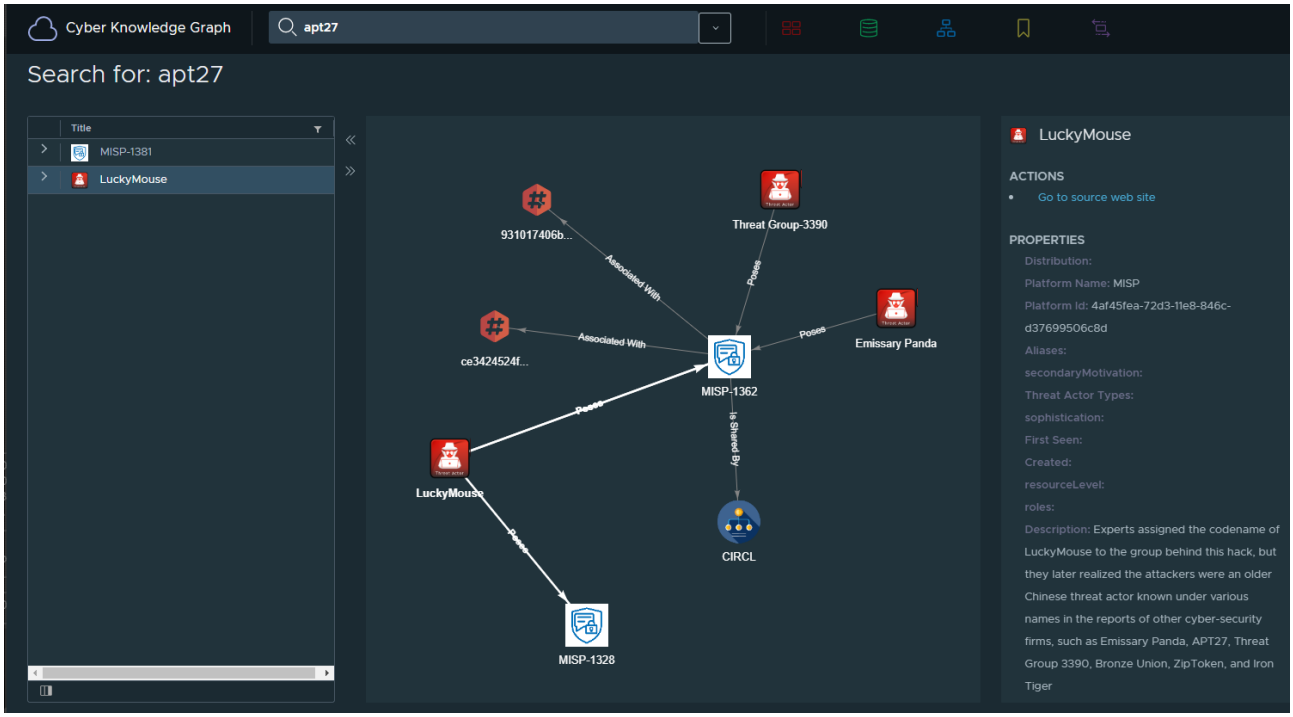


Figure 27: CKG's Search and Results Navigation

The ontology view shows the concepts captured by CKG, allowing to see their definitions and sample data from the knowledge graph.

Figure 28: CKG's ontology view

Bookmarks allow (advanced/admin) users to create graphs using custom graph database queries:

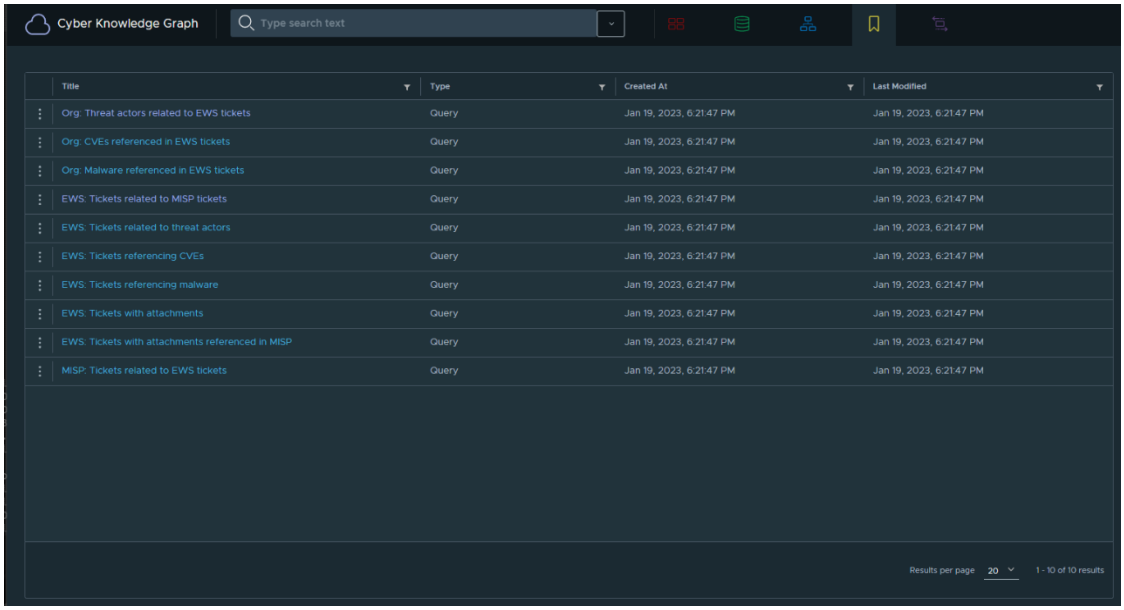


Figure 29: CKG's Bookmarks page

Data pipelines are managed by Spring Cloud Dataflow. Its UI supports creation, deployment, start and stop of pipelines.

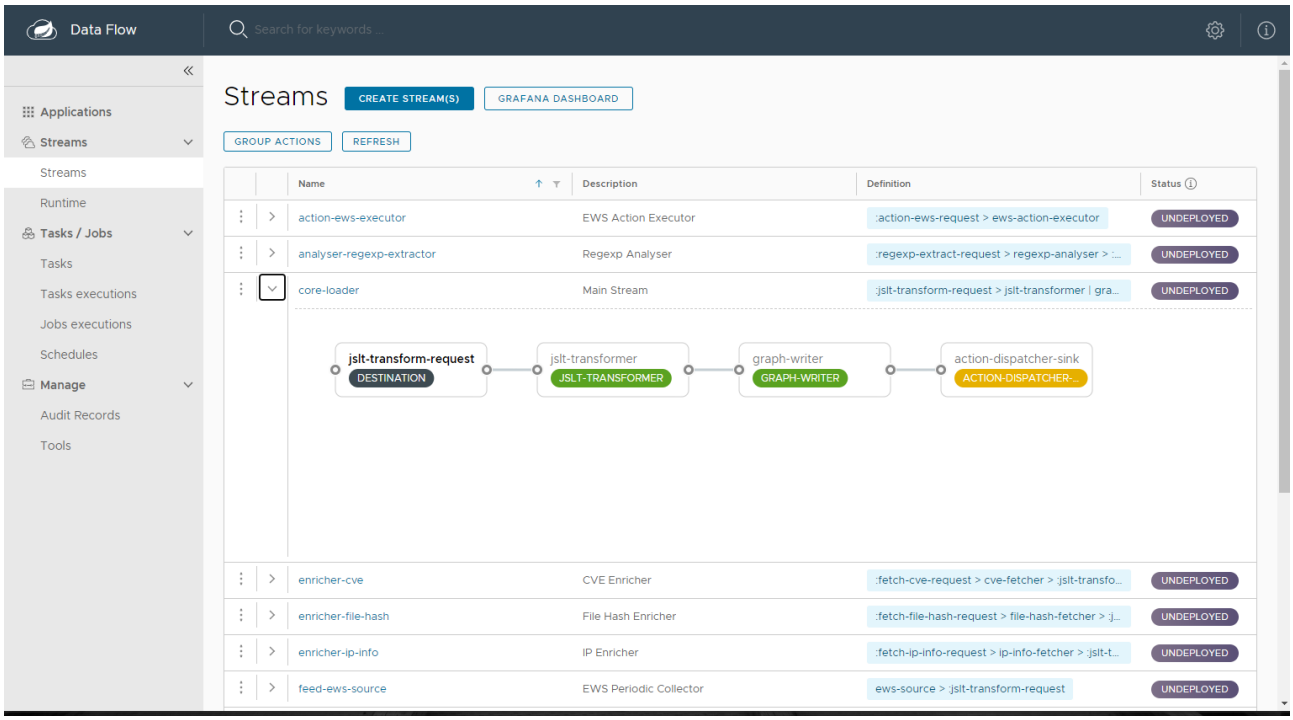


Figure 30: CKG's Pipeline Management with Spring Cloud Dataflow

CKG provides instrumentation data to a Prometheus timeseries database, allowing system and application level to be displayed in Grafana dashboards.

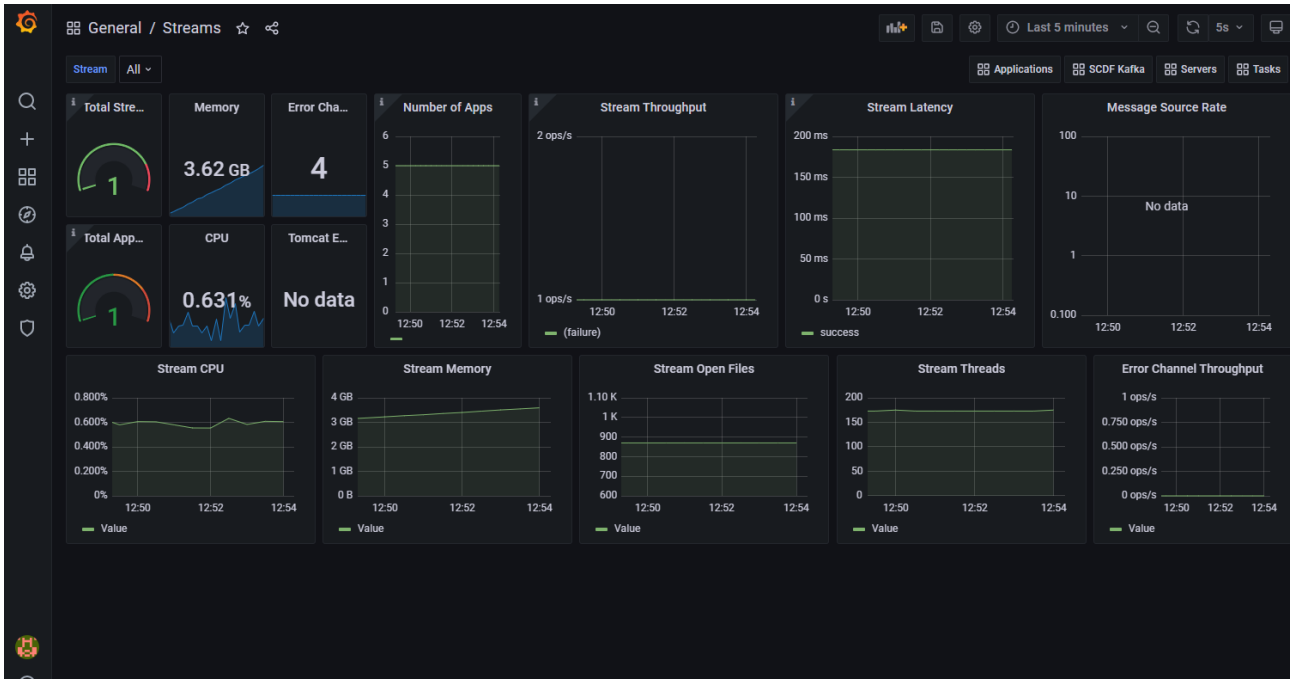


Figure 31: Grafana dashboard for a pipeline stream

4.4.4 Summary and Next Steps

Within DYNAMO, CKG will be used to further analyse and correlate the knowledge base gathered and extracted from different information sources. More specifically, it can be used:

- As a standalone search and exploration engine to improve situational awareness, relate threat intelligence with IT resources
- To enrich E-EWS information by expanding context from an E-EWS ticket and adding items to E-EWS reference library.

Another possibility we are exploring is the enrich of CKG with named entities, resulting from NER models trained on CTI data. In the coming term, the collaboration with CTI Extractor will be finalised in order to avoid overlapping in the topic of correlation.

4.5 CAF

4.5.1 Tool Overview and concepts

CAF is a tool that provides next-minute cyber-attacks forecasts, by utilising network traffic measurements and by including the type of attack (DoS, DDoS, Port Scan, SSH-P etc.). The tool is destination port (DP) oriented, and uses measurements only taken from the targeted DP to provide forecasts for possible upcoming cyber-attacks (next minute). Near real time forecasting is a decision that is based on the needs that are defined by DYNAMO's end users and their opting for near time responses.

The operation of the tool is described in the following steps:

- **Features' selection:** Taken from the network traffic flow in order to enhance the performance of forecasting
- **Fixed time-scale definition:** Applied at the network traffic measurements, to avoid the use of incomplete timestamps that appear in the network flow datasets

- **Cyberattack identification:** The attacks are identified and leveraged for each timestamp in order to provide forecasts for the near future
- **Forecasting:** An ML based model (LSTM) utilises the features that are extracted from the previous steps to forecast the type of any cyberattacks within the next minutes. The feasible predictions are designated as “non-attack” or “attack” activities.

4.5.2 Architecture

The high-level architecture of CAF is depicted in the following diagram (Figure 32).

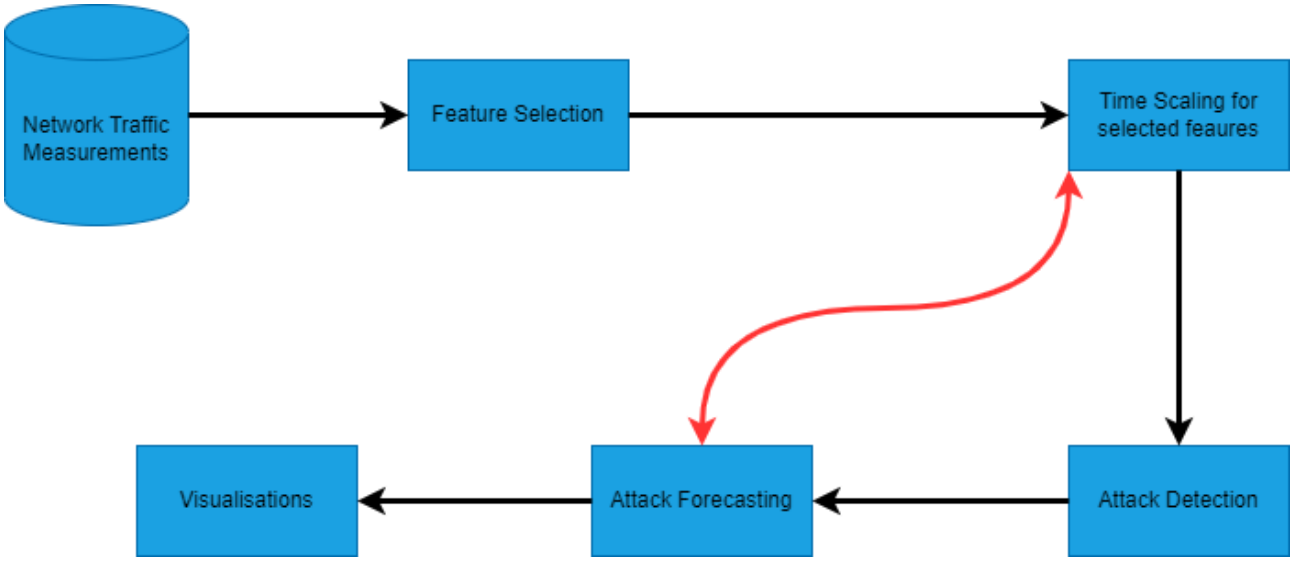


Figure 32: CAF's Architecture

- **Network Traffic Measurements-Feature Selection:** Network traffic measurements include distinctive features, such as *total packets in the forward direction*, *total packets in the backward direction* etc from the degree of correlation is identified (linear or non-linear). Therefore, the number of features is reduced in order to increase the efficiency of the tool and use less measurements.
- **Selection of timescale:** In order to provide a forecast for the *next time-step*, the time-steps are defined based on the available data. In most cases, the timestamps of network traffic measurements include the hour, the minute, and the second. In order to avoid the cases where the seconds are not provided, CAF is not including the seconds.
- **Attack Identification:** Identifying the type of cyber-attack at each minute (if any) is part of the process of forecasting future threats. During the development process, CAF is obtaining historical data from the CIC-IDS2017[29] dataset, a publicly available dataset provided and generated by the Canadian Institute for Cybersecurity to identify any DP oriented attacks with the use of the Random Forest algorithm.
- **Attack Forecasting:** After the selection of a DP, given the features in a given number of previous time steps, the LSTM model is able to forecast a possible attack for the coming minutes, along with the type of attack.
- **Visualisation:** With the proper visualisation, provided by the tool’s UI, the user is able to obtain information regarding imminent attacks.

[29] https://www.unb.ca/cic/datasets/ids-2017.html

4.5.3 Mockups / Screenshots

The Figures 33,34 and 35 illustrate information that can appear on CAF's UI. They include user friendly buttons for selecting a desired port, time and the duration of the potential forecasting, along with prediction on possible threats.

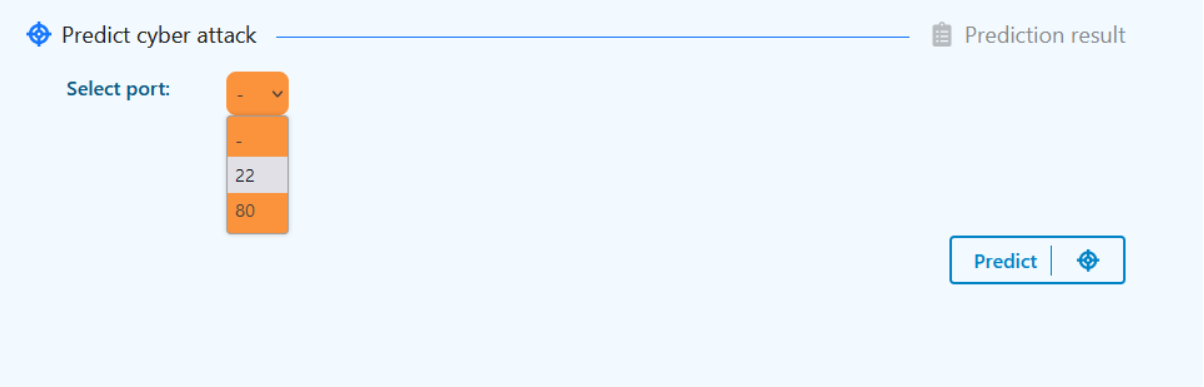


Figure 33: CAF's DP selection

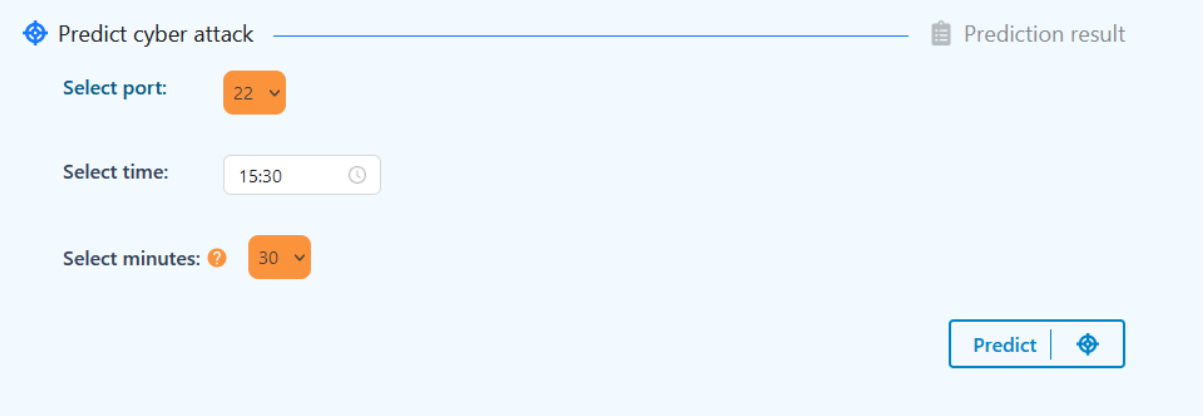


Figure 34: CAF's Selection of Time and Duration

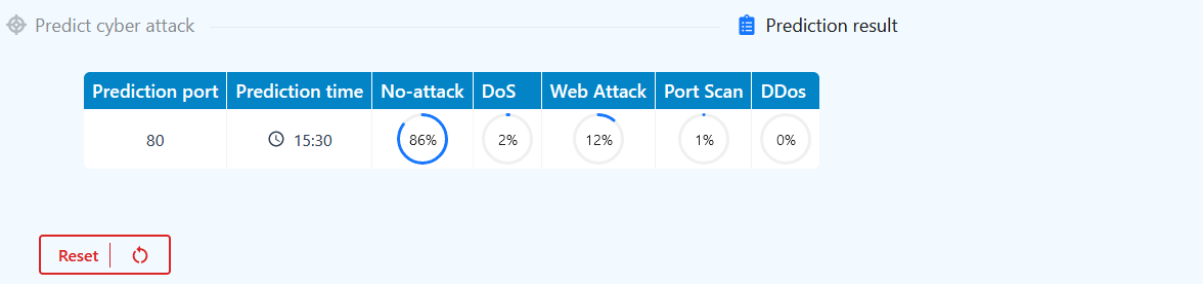


Figure 35: CAF's Forecasting of next minute attacks

4.5.4 Data

As it has been mentioned above, CAF has utilised for training purposes the historical data of CIC-IDS2017 a publicly available dataset. Originally, only two ports (22, 80) have been used for the first training of CAF. In order to expand its forecasting capabilities, the data from the ports 21, 443, 444 of the used dataset have been used (also in order to support the new capability of providing automated forecasting). For the ports with limited data, we assumed the hypothesis that lack of attack information signifies the existence of benign traffic. During the demonstration of the DYNAMO use case, CAF along with SecureAI, will receive network traffic (benign and malicious) from

DYNAMO's end users for further training of its forecasting algorithm and for further demonstration of its capabilities.

4.5.5 Summary and Next Steps

CAF is a tool that can provide forecasting and historical information for DP oriented attacks within a near minute time frame. For the next steps of development, the following aspects will be examined:

- Increase of the type of forecasted attacks
- Provision/suggestion of DPs
- Combined use of information to examine advanced attacks conducted by multiple attack vectors.
- Improvement of the forecasting rate
- Continuation of the search for new datasets

4.6 SecureAI

4.6.1 Tool Overview and concepts

AI has emerged as an indispensable instrument for all companies and organisations to employ in their battle against cybercriminals and other forms of cybercrime. The algorithms that make up artificial intelligence are able to examine vast volumes of data in order to identify and react to potential dangers in real time. In addition, intelligent algorithms are able to recognise patterns and irregularities that would be hard for humans to recognise due to the vast amount of information that is available online. By automating the security procedures with the help of artificial intelligence, organisations are able to proactively protect themselves against attacks. The most recent security threats have brought to light the necessity of implementing comprehensive security measures and procedures. Phishing, distributed denial of service attacks, and ransomware are becoming increasingly common and more difficult to detect and prevent. Therefore, artificial intelligence may be of assistance in a variety of unique ways since it provides an additional layer of protection for organisations in which the security of the systems is non-negotiable.

This is the reason a tool that prioritises security, which is called SecureAI tool, is going to be developed as part of WP4 of DYNAMO (T4.4). SecureAI includes a number of different methods and mechanisms, including Deep Learning, data visualisation, and mitigation suggestions, among others, working as an all-in-one solution for inspection and security that is powered by artificial intelligence. The tool's scope is to identify data abnormalities, threats, and dangers from a wide variety of data sources, such as network traffic, applications logs, system logs, etc. The tool provides anomaly detection, alerts and insights for potential threats, and it also recommends various mitigation suggestions to be taken under consideration for inclusion on DYNAMO's response plans, but also being proposed (in the form of notifications) to the operators for immediate action.

In addition to this, SecureAI offers a customisable dashboard and analytics toolkit that is capable of visualising raw data, providing historical analysis as well as identification of abnormal patterns in the data. This will make it possible for security specialists to do further analysis during attack-based incidents. In addition, the dashboard will provide a clear perspective of the incidents that are currently occurring and the risk-levels of the assets involved in the monitored infrastructure, providing security teams with a tool helping them understand and respond quicker to any threats that may arise.

A live deep packet inspector is incorporated into the tool and analyses packets, network flows, and other sources of data in order to discover aberrant network activity. SecureAI also introduces a neural network-based model tailored for the detection of ransomware within Windows systems through the analysis of log files. This approach utilises neural networks to swiftly and accurately identify anomalous activities within a host, which might be indicative of ransomware attacks.

Users and security professionals are equipped with the information and tools necessary to safeguard their systems and organisations from threats and malicious actions through the usage of SecureAI, which functions as a Cybersecurity Awareness tool within the CTI framework of DYNAMO. Additionally, it contributes to the development and implementation of best practices for handling of cyber-attacks, by providing attack-based mitigation suggestions to the BCM framework (Pianista), for inclusion in the response plans. SecureAI will also provide inputs for possible training materials through its function of attack simulation that can prepare users in identifying and handling relevant cybersecurity issues (through the use of RETA - T3.4).

4.6.2 Architecture

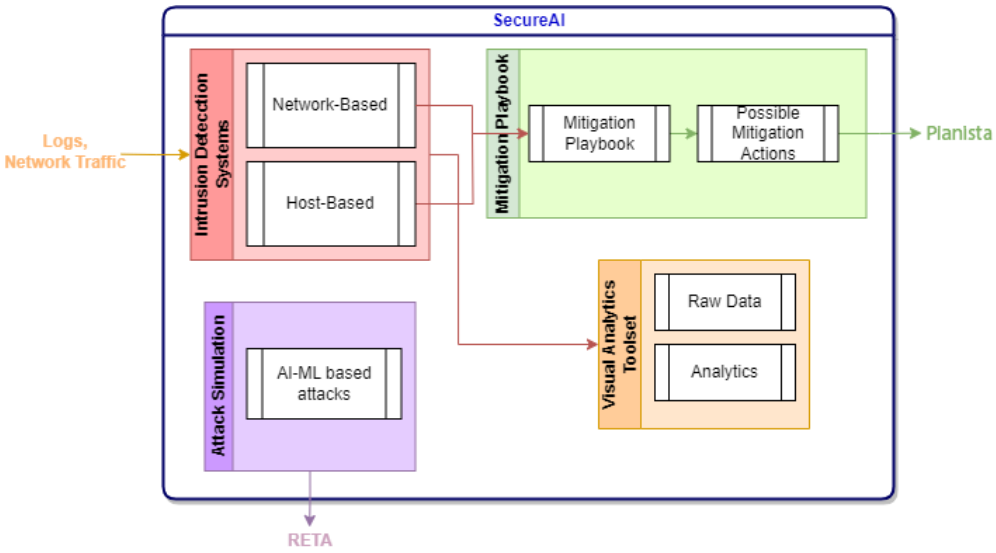


Figure 36: SecureAI Architecture

The core architecture diagram of SecureAI is depicted in Figure 36. The components that will be developed are the following:

- **Intrusion Detection System (IDS)**: For this module, two different IDSs are being developed, one responsible for detecting network anomalies (network-based), and one designed to detect ransomware attacks in Windows systems (host-based). In addition, it is also being examined to add more network IDSs, which employ different AI techniques, to ensure that the network security is more sophisticated and robust.
- **Mitigation Playbook**: This module will provide the user with possible mitigation actions according to the different attack types detected by the IDSs. This information will also act as input to Pianista, DYNAMO's AI-supported goal-oriented response plan generation tool.
- **Attack Simulation module**: attack scenarios will be simulated (by injecting attacks into logs), and then fed to the RETA tool (DYNAMO's training module) which will enable users to be trained on how to use the platform in case an actual attack occurs.
- **Visual Analytics Toolset**: This is the UI of the tool. The dashboard and analytics suite presents a comprehensive toolset for visualising raw data and detecting anomalies, empowering security experts in analysing attacks. It offers real-time insights and alerts into ongoing incidents, facilitating swift response from security teams to emerging threats.

Moreover, the SecureAI platform utilises a structured database system to efficiently store diverse data types such as logs, user information, raw data, and predictive outputs, facilitating user interactions with historical events. A data gathering module will serve as a bridge between SecureAI and the "Cyber-threat intelligence gathering and extraction" tool from T4.1, ensuring a flow of

information for comprehensive threat analysis. At its core, SecureAI incorporates AI mechanisms, enabling sophisticated anomaly detection within gathered data while dynamically adapting (utilizing reinforcement learning) to changing data patterns, ensuring continual optimisation and enhanced resilience against potential data drift scenarios.

Finally, it should be noted that the two IDS modules comprising SecureAI's detection engine (host-based and network-based) were trained using synthetic datasets – generated based on descriptions coming from DYNAMO's use cases and agreed upon with the end-users, while for the the network-based detection Lycos-IDS2017 open dataset [29] was also utilised[30]. The data to be used for testing of the modules will be collected and shared from DYNAMO's end-users.

4.6.2.1 Communication Interfaces and I/O

The corresponding inputs/outputs of SecureAI are the following:

Inputs: Heterogeneous data collected and forwarded from testbeds (network traffic from monitoring tools installed in the test environment and host-based logs collected from extraction modules running on individual devices within the test environment).
Outputs: Abnormality detection, predictive analytics on the input data, alerting and mitigation suggestions, and attack simulations through the dashboard.

The communication interfaces needed for SecureAI are REST APIs and HTTP requests (between the tool's sub-components) while Kafka integration is planned for the data exchange with the several collaborating DYNAMO's components. In terms of the specific data formats, the tool expects inputs in JSON format and will use the same format for the output, while both will be STIX compliant.
In more detail, the process of obtaining the outputs related to attack detection and the corresponding mitigation actions recommended by the tool is as follows:

- **Data Collection**: SecureAI collects two types of logs: network logs from monitoring tools installed in the test environment and received through the Kafka broker and/or manually importing files through DYNAMO's GUI, as well as host-based logs collected from extraction modules running on individual devices within the test environment. These inputs are aggregated and analysed to provide comprehensive insights into potential security threats.
- **Analysis and Detection**: Through advanced AI algorithms and deep learning mechanisms, the collected data is processed by the tool, in order to identify patterns, detect anomalies and recognize potential threats. This analysis involves reviewing both network-based and host-based logs received during the data collection phase. The detected abnormality is being pushed to DYNAMO's Kafka broker within the relevant kafka topic.
- **Mitigation suggestions**: Upon detecting a potential threat, SecureAI classifies the type of attack based on predefined attack patterns and historical data and proposes mitigation suggestions, derived from a predefined mitigation playbook, which aims at providing insights on suggested course of actions to prevent or minimise the impact of security threats on the examined infrastructure [30], [31], [32]. The suggested mitigations are made accessible to DYNAMO's response plan tool (Pianista) through a REST endpoint, while they are also forwarded to the operator as direct insight through DYNAMO's portal.
- **Dashboard and Analytics**: The analysed data, alerting about potential identified threats and relevant mitigation suggestions, is visualised and presented through the tool's dashboard. This interface provides clear visualisations of raw data, potential irregularities, ongoing incidents, and predictive analytics. While the tool's core visualisations will be available directly through DYNAMO's platform (data will be forwarded via REST APIs), SSO integration with the platform is envisioned in order for SecureAI's dashboard to be accessible to the user.
- **Output Format**: The analysis results, including attack detection results and mitigation suggestions, are compiled into a structured JSON format, using appropriate STIX bundles,

[30] Available at: https://maupiti-git.univ-lemans.fr/lycos/lycos-ids2017

making it compatible with collaborating DYNAMO's components. The output, containing details about the attack detection and mitigation suggestions, is being forwarded to DYNAMO's portal and the collaborating components through the "abnormality" topic of the deployed Kafka broker. Additionally, by accessing DYNAMO's interface, users can specify their query parameters and request relevant data for a specific time-period. SecureAI then generates and presents the data, enabling users to gain insights into potential security threats during the specified period.

Overall, SecureAI's output involves collecting and processing data, performing analysis, visualising the results, and providing the information through a dashboard interface, offering users actionable insights to strengthen their cybersecurity stance.

4.6.3 Mockups / Screenshots

The UI mockups, pivotal for designing and implementing the user interface, are an integral aspect of the tool's development; however, due to the complexity and dependencies of the component and the need for refinements to adapt to the final platform's needs, are currently under development and they will be reported in detail in D4.2 "Final version of the cyber-threat intelligence gathering, extraction, sharing components and AI-based solutions".

4.6.4 Summary and Next Steps

In summary, SecureAI is an AI-driven inspection and security solution that has the capability to recognise risks, threats, and irregularities in data derived from a wide range of data sources. In addition to providing users with the knowledge and tools they require to protect themselves from online threats and malicious software, it also assists organisations in the process of establishing and maintaining secure networks. SecureAI also facilitates the generation of training material through its user interface by infusing data emulating selected attacks, thus triggering the system on demand (for training purposes). This is done with the intention of better equipping users with the knowledge and skills necessary to prepare and respond against cyberattacks. SecureAI is being developed under Task 4.4 (M12-M30). The table below displays a brief timeline for the development of the task.

Time Period	Milestone Description
M12-M16	State of the Art analysis/ Research on available ML technologies for Intrusion Detection Systems and possible mitigation suggestions.
M14-M18	Design, Architecture and High-Level Requirements of the system – Provision of early version
M18-M20	Development and dissemination of a standardised format for gathering data directly from those who use the product or service
M18-M24	Delivery of integration ready prototype
M24-M28	First Integration, Testing and Refinements
M28-30	Documentation, Reporting and final Integration

Table 3: Timeline for SecureAI's development under Task 4.4

Finally it is worth mentioning that the detection and forecasting framework of DYNAMO (developed under T4.4) includes two components – SecureAI and 4.5. Both components are focusing on network data for Intrusion Detection and forecasting for the following reasons:

- Timely Threat Detection: An AI-driven analysis of network traffic can identify suspicious patterns and anomalies which constitute potential indicators of an intrusion. This allows for immediate response to potential threats that can decrease the impact of vulnerabilities.
- Accuracy: Applied on historic data, machine learning algorithms can improve over time. This continuous training process of the tools enhances the accuracy of threat detection and of the forecasting and reduces the number of false positives that appear on a network traffic file.
- Resource Optimisation: By automating the detection process and by deploying the appropriate algorithms on network files, AI reduces the workload required to achieve appropriate results by tools. For example, CAF manages to achieve forecasting results with the use of a selected number of features which are easily extracted from a network traffic file. This action significantly reduces the analysis required by the tool.

4.7 E-EWS

4.7.1 Tool Overview and concepts

The ECHO Early Warning System (E-EWS) aims at delivering a security operation support tool enabling personnel to coordinate and share cyber-incident information in near real-time. Operators will thus be able to assess impacts, identify and evaluate mitigation actions and prioritise response measures.

Although multiple incident handling tools are available on the market today, the E-EWS expands on common functionality by providing a sharing capability allowing incident management between disparate operational units across organisational boundaries, including the coordination of management workflows. With the E-EWS platform, organisations can retain their fully independent management of their own incident and related cybersecurity data management through the E-EWS, or through their own integrated incident handling tools, while also supporting collaborative information sharing of said incidents.

The E-EWS under Dynamo platform is integrated with the Data Anonymisation and Fine-Grained Access tool to secure and enrich the sharing process with anonymisation of information and encryption/decryption of the cyber-ticket ensuring a high confidentiality of the information.

The following table presents the main capabilities of E-EWS.

Functionality	Description
Cyberticket Management	In response to cyber incidents a tracking tool is required to guide the coordination within the organisation's cyber operations team. E-EWS allows the definition of a cyberticket to be customised to adhere to the organisation's needs. This includes defining the cyberticket fields as well as applying a workflow that reflects the organisations internal cyber operations process. Next to the organisational cyberticket processes the E-EWS also allows collaboration on ongoing incidents in near real-time across the organisational boundary. The sharing module supports selective sharing of cyberticket data through a secure channel.
Warning Management	In addition to the reference data, we have a model for structured exchanged of CTI based on a sharing in communities, the Warning. The difference from cybertickets is that the warnings record a structured model made for sharing either early detections of anomalies or for sharing of confirmed threats. In this it makes the information exchange more like reference data, but with the exception that it is built on a fixed set of attributes, can be constructed directly in the tool, and that it can be shared in one or multiple constituents, where the reference data automatically is available globally and normally comes from OSINT.

Signal Management	In a network with shared resources, it might be interesting to share real-time metrics about these resources to dependent collaborators. For instance, if we have one point in the network, we measure the current transmission rate, and collaborators are dependent on this link we can distribute this through signals or if we want to track the last known location of a certain entity and distribute to partners. The signal is made so that it can model a resource and distribute information that is time-dependant. It is based on two sets of attributes, the signal attributes, and the signal value attribute, the signal attributes make it possible to record a context about the signal (information that is not time-dependant) and where the values will hold the current and historic values.
Notification subscriptions	E-EWS has the capability to send real time notification for a sizeable number of events. The E-EWS user can take a subscription on the type of events (like Ticket created, Ticket updated, Warning created …) that he wants to receive, as well as the channel on to which the notification should be delivered (UI notification directly in the E-EWS GUI, email notification to the user's configured email address).
Comment	Within the context of a cyberticket, a wiki page or a warning, comments can be exchanged to discuss ongoing actions and coordinate. Attachments can be added to the comment. The comment threading functionality supports replying to comments as well.
Information sharing	General information and knowledge can be recorded and shared with constituents. Either within the level of the organisation itself or shared between users that are part of different organisations. The information includes structured information about threats and vulnerabilities (e.g., STIX, CVE) and unstructured information (e.g., Wiki pages, knowledgebase).
Reference data sharing	Vulnerability and threat data is used in the cyber domain to report on cyber related events and information that can help to provide context to cyber information. The library is shared between the organisations within the E-EWS and acts as a collaborative database of information. Reference data is periodically retrieved from a CVE database that is configured during the installation of the E-EWS which will ensure the availability of the latest vulnerability information for all the constituents in the E-EWS.
Plugin Management	E-EWS provide a plugin framework. The plugins will have the capability to extract data from the E-EWS repository in a secure manner, but also to add information to the system. For this, the public API of the E-EWS can be used. Each plugin will have to be registered into the E-EWS before it can be used. Plugins can be developed independently provided they adhere to the specified interface and protocols. This integration with existing 3rd party tooling can increase the possible data sources of the E-EWS.
Software Development Kit (SDK)	E-EWS provides an SDK to allow external systems to interact with its internal data. It exposes a REST- interface to allow various mode of operation i.e., create, read update, and delete actions (also known as CRUD). These operations allow a client to connect, access and manipulate data inside the incident management. SDK interface uses Basic authentication along with HTTPS protocol, which is a simple but well-established protocol for authentication.

Tenants Management	The E-EWS is a multi-tenant application which supports the operation of disparate organisational units within the same server. The management of organisations is complete separated and allows each organisation to customise the E-EWS for their particular needs.

Table 4: E-EWS Main Capabilities

4.7.2 Architecture

4.7.2.1 Hich-Level Architectural View

The high-level architecture of the E-EWS is presented in Figure 37. It contains the E-EWS Server and its sub-components. The components represent the major functionalities to be provided by the E-EWS server.

The Web User Interface acts as the main point for user interaction. It allows the general operation of the key features of the E-EWS as well as administrative tasks. The web UI will hide unauthorised options from the user depending on the assigned role. The Web UI is hosted by the E-EWS Server.

The third-party tool plugin framework allows external tools to be securely connected to the E-EWS public API. The plugins can be used to retrieve or change information in the E-EWS.

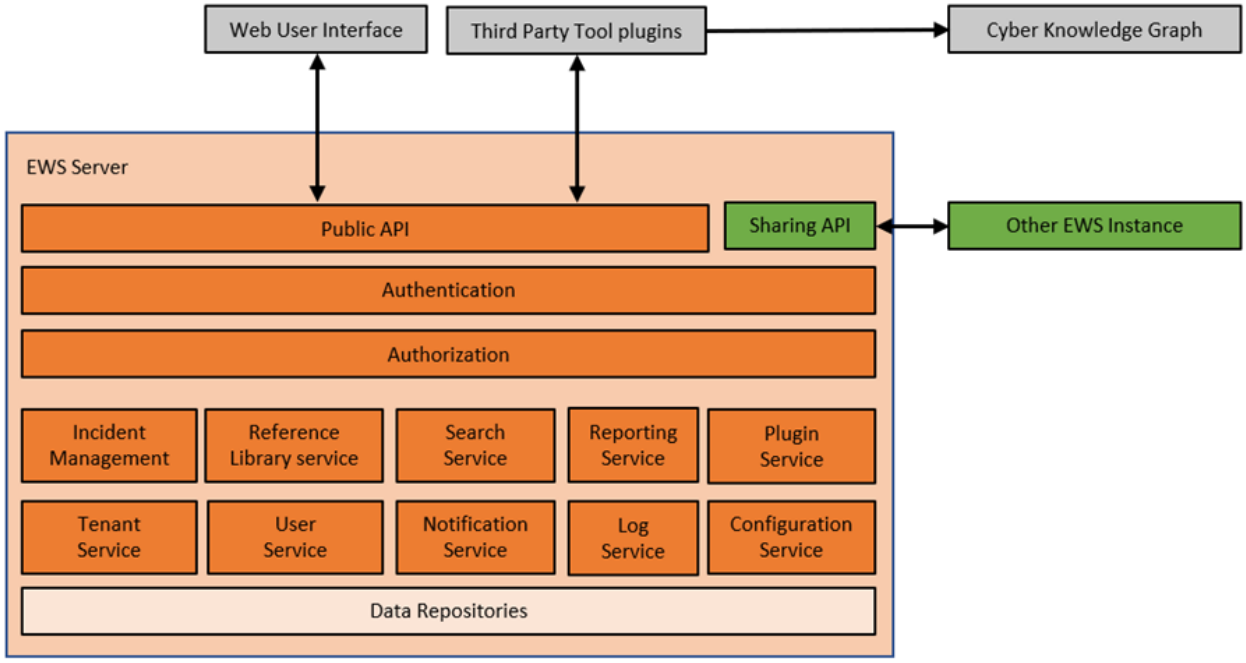


Figure 37: E-EWS High-level system architecture

4.7.2.2 Distributed deployment

The E-EWS server is a multi-tenant system where the data of multiple organisations can reside in one server installation. The organisations are grouped into constituents, representing a common domain of interest. Within a constituent, the organisations can securely share data. Whereby the constituents can be composed of multiple organisations also including organisations that reside on another E-EWS server. This distributed case will allow secure information sharing across the organisational or even national boundary. The constituents are managed by a governance body, organisations will be added upon request, and with the agreement of the already existing constituent partners.

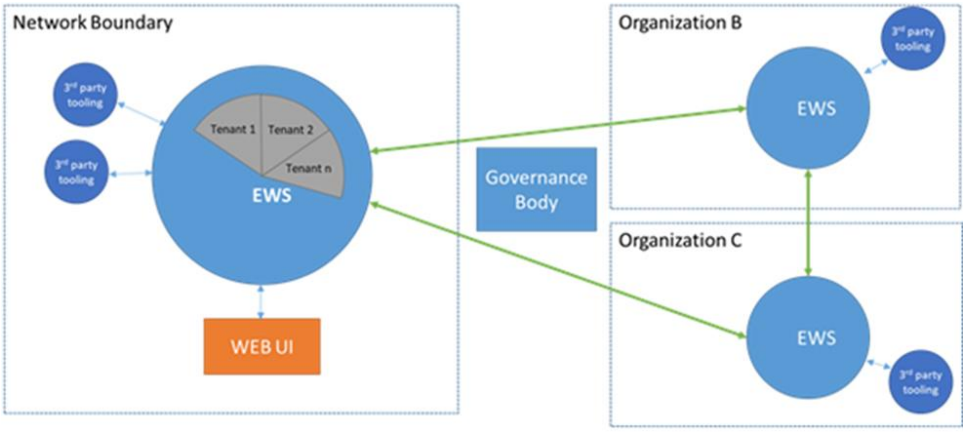


Figure 38: E-EWS Distributed deployment

Information shared within the E-EWS is not accessed from a remote site, as cloud infrastructure. Instead, the shared information is replicated to each disparate data repository. This allows the information to be accessible from the E-EWS even if a node would be unreachable. The information is only available to the nodes with which it is shared.

4.7.3 Mockups / Screenshots

Graphical User Interface

Users can access the E-EWS Client though the Single Sign-On provided by Keycloak under the Dynamo platform. From the Dynamo portal web page, the user can choose in the menu the EWS voice and it will be redirected to the E-EWS client which will ask the organisation to login and proceed with authorisation.

After a correct login, the main E-EWS interface is presented to the user. As a general design rule, the top bar shows the user info, while the left menu bar shows actions and features available. The main area is where said features can be managed.

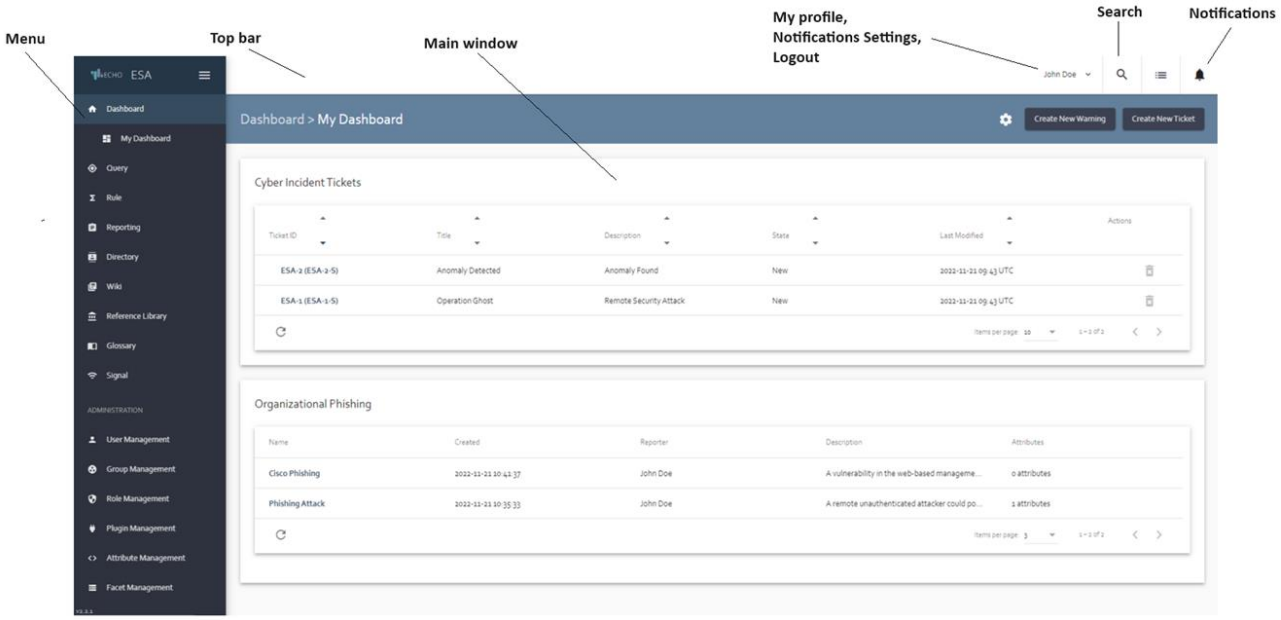


Figure 39: E-EWS Application overview

4.7.4 Summary and Next Steps

E-EWS is designed as a security operations support tool, allowing personnel to coordinate and share cyber-incident information in real-time. This enables operators to assess impacts, identify mitigation actions, and prioritise response measures. The platform stands out by providing a sharing capability, facilitating incident management between different operational units across organisational boundaries, including coordination of workflows. Organisations can maintain independent incident management through E-EWS or their integrated tools while supporting collaborative information sharing.

For effective response to cyber incidents, E-EWS includes a customisable cyberticket tracking tool. Organisations can define cyberticket fields and workflows according to their needs, aligning with internal cyber operations processes. The platform allows real-time collaboration on ongoing incidents, supporting selective sharing of cyberticket data through secure channels.

E-EWS serves as a repository for general information and knowledge, allowing sharing within an organisation or between users in different organisations. This encompasses structured information about threats and vulnerabilities (e.g., STIX, CVE) as well as unstructured information (e.g., Wiki pages, knowledge base). The platform's library acts as a collaborative database of information, periodically updating vulnerability data from a configured CVE database, ensuring constituents have the latest information on vulnerabilities. This collaborative approach enhances situational awareness and response effectiveness in the cyber domain.

E-EWS provides the capability to add external and internal references to a cyberticket. An external reference can be an URL pointing to a resource within the organisation network, a resource on the public Internet. Internal references relate the cyberticket to information present in E-EWS. This can be:

- cyberticket
- wiki page entry
- reference library entry
- signal
- warning

Including these references in a cyberticket enhances the amount of information available and improves the overall understanding of the incident. For the DYNAMO project, the following improvements, focused on the reference library feature, will be implemented:

- Import CTIs in STIX format from Dynamo message broker
- Integrate E-EWS with Data Anonymisation tool
- Integrate E-EWS with Fine Grained Access tool
- Integrate E-EWS with Dynamo platform Keycloak SSO

4.8 Fine Grained Access

4.8.1 Tool Overview and concepts

The fine-grained access control solution based on attribute-based encryption (ABE) provides a secure and flexible approach to data sharing between components that provide data providers and components that consume.

The ABE technology enables fine-grained access control by allowing encryption and decryption of data based on attributes. Access policies can be defined and enforced at any level of a data structure, providing a granular control mechanism. This means that access control can be applied to individual data items, specific directories, or even entire data sets, depending on the requirements of the system.

Furthermore, the proposed ABE system achieves an important level of robustness in terms of both availability and security. This is accomplished by distributing the management of access control among multiple attribute authorities (AAs). These authorities collaborate to enforce the access policy by collectively managing the generation of cryptographic keys necessary for accessing the shared data.

Finally, the proposed system ensures efficient access revocation. When it becomes necessary to revoke access for specific entities, the access management authorities collectively modify the access policies and re-encrypt the affected data items. This ensures that the revoked entities can no longer decrypt and access the shared data.

4.8.2 Architecture

The fine-grained access control solution involves five entities: A certificate authority (CA), multiple AAs, data providers, data consumers, and a cloud server.

- CA is a blockchain-based PKI management system that is charged of setting up the system parameters such as the cryptographic parameters to be used, the set of attributes and their respective public keys. CA is responsible of registering AAs as well as data consumers. It is also responsible of choosing the robustness level that should be satisfied, i.e., the number of AAs that should collaborate to issue a decryption key.
- Attributes authorities are responsible of issuing decryption keys to data consumers. They collaborate together with the CA to set up the master public key of the system.
- Cloud Service Provide is an entity that provides data storage capabilities.
- Data providers are the entities aiming to share their data. It encrypts the data to be shared using a chosen access structure formulated over a set of attributes that defines who can access the shared data.
- Data consumers are the entities that will access and use the shared data. They are labelled by a set of attributes by the AAs.

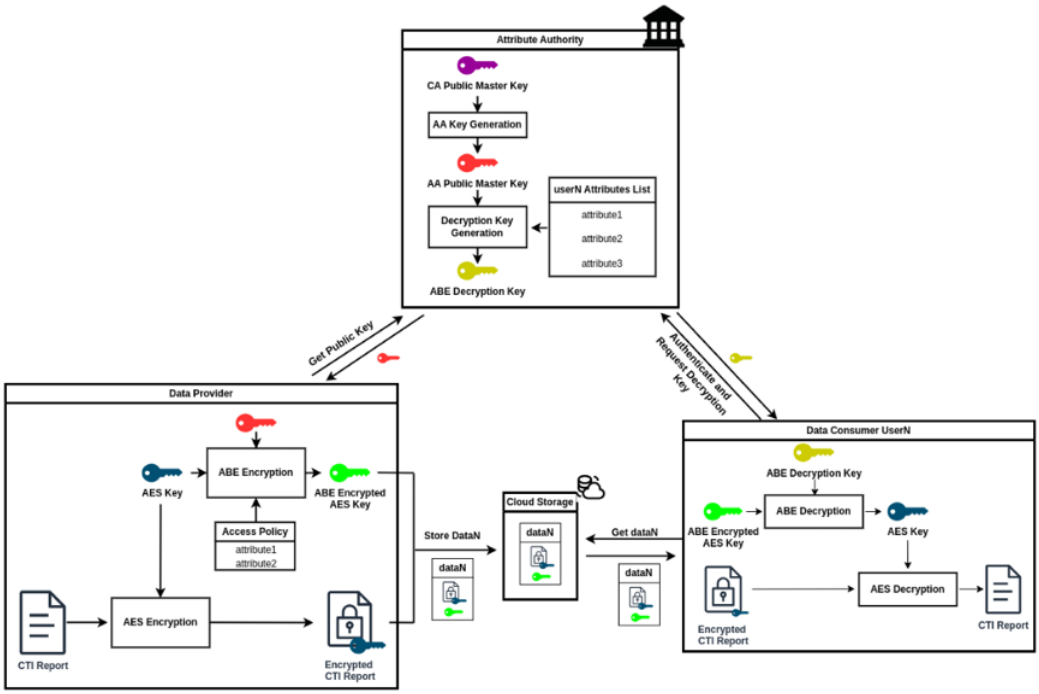


Figure 40: High-level Attribute-Based Encryption Framework

As shown in Figure 40, each CTI report will be firstly encrypted using a secure symmetric key algorithm, such as AES. The Data Provider then needs:

- An access policy that should be enforced.
- The AA's Public Master Key, which is generated by the AA with the CA Public Master Key (The interactions with the CA are abstracted in this example).
- The symmetric encryption key is then encrypted using an Attribute-Based Encryption algorithm.
- The Data Provider then sends the encrypted data bundle containing the Encrypted CTI Report, and the Encrypted AES Key to a trusted cloud server for storage.
- All communications are encrypted and signed using certificates issued by the CA.
- The data consumer, after being registered by the CA, will ask the AA to get a decryption key. The Data Consumer will authenticate himself with a signed query containing their ID and their certificate. The AA can then verify the signature of the user. If the user is legitimate, the AA will generate a decryption key using the set of attributes accorded to the user by the Authority.
- The Data Consumer can then download the encrypted data, this query is also signed and encrypted. The Data Consumer can also identify the Cloud Storage using its certificate. The user can then decrypt the symmetric key to decrypt the encrypted data.

For the implementation of the tool, each ABE Entity provides a communication interface through a REST API.

4.8.3 Mockups / Screenshots

Figure 41 illustrates how the tools configuration wizard is configuring multiple roles in different occasions.

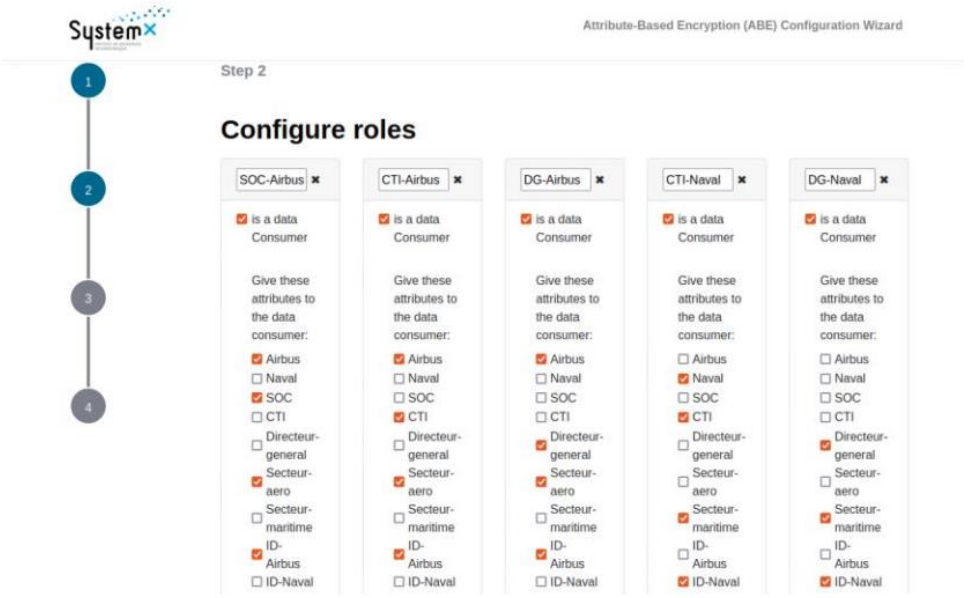


Figure 41: Example of the Fine-Grained Access Configuration Wizard with multiple roles

4.8.4 Summary and Next Steps

The fine-grained access control solution based on ABE provides a secure and flexible approach to data sharing between data providers and data consumers. It allows the implementation of a secure and granular sharing mechanism. The current version of the ABE tool is based on Public Key Infrastructure (KPI), and we plan to enrich it to allow its integration with the E-EWS by implementing a unified authentication solution such as Keycloak in order to provide secure encryption for DYNAMO's information sharing solution.

4.9 Data Anonymisation Tool

4.9.1 Tool Overview and concepts

Anonymisation is a process by which personal data is made non-personal. It refers to a set of techniques whose aim is to remove any possibility of being able to identify the individual to whom the personal data belongs. By convention, anonymisation is considered irreversible. Once the data has been changed, it is impossible to restore it to its original state. Thus, the difficulty lies in finding the right compromise: you cannot eliminate too much information, or the data will no longer be of interest.

In a database or dataset including personnel information, data could be classified as:

- Identifiers: an attribute that allows unambiguous re-identification of the individual to whom the record refers (e.g., social security number, passport number, etc.). To avoid direct re-identification of an individual, "identifiers" attributes must be deleted or encrypted.
- Quasi-identifiers: unlike an identifier, a quasi-identifier alone cannot be used to re-identify a record. However, its combination with other quasi-identifier attributes can enable unambiguous re-identification of certain individuals. Quasi-identifiers cannot be deleted, unlike identifiers, as they are often necessary to perform useful analysis of the data.
- Confidential attributes: attributes containing sensitive information about the people involved in the data collection process (e.g., salary, health status, sexual orientation, etc.). The main aim of microdata protection techniques is to prevent intruders from learning or inferring confidential information about a particular person.
- Non-confidential attributes: any attribute that does not belong to the above categories.

DAT consists of a unified web interface and a simple, flexible and configurable anonymisation engine. It will enable configuration of the dataset and the chosen anonymisation models. The configuration will be conducted step by step, to guide the non-expert user through the anonymisation process.

Anonymisation models are defined in an abstract way. This means they can be configured independently of the techniques used. These configurations can be saved and reused by other variants of the same privacy model.

4.9.2 Architecture

In this section, we describe the architecture of the DAT. We first present an overview of the functional/logical view of the DAT (Figure 42). Then, we briefly describe the communication interfaces.

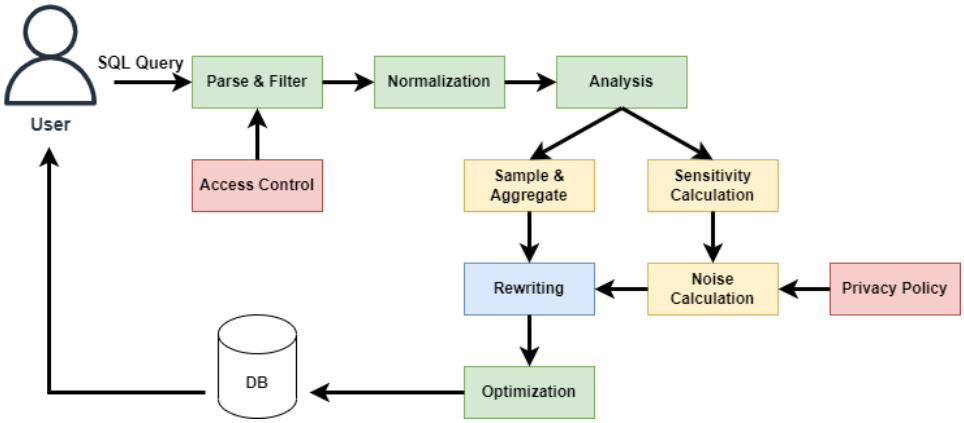


Figure 42: DAT Queries processing

DAT's abstraction enables it to interconnect easily with existing APIs. To add an interconnection, an interface has been implemented and the external API has been called. The current implementation already interconnects with the ARX API.

The ARX adapter takes care of adapting the APT configuration to that of ARX, and vice versa, adapting the ARX results to that of APT. Everything is transparent to the end user.

For k-anonymity & ℓ-diversity model implementations, we plan to use our own implementations or reuse algorithms already implemented by other tools (e.g. ARX, UTD or CAT).

4.9.3 Mockups / Screenshots

Figures 43 and 44 provide an overview of DAT's interface along with anonymised entries in a database, anonymised by DAT.

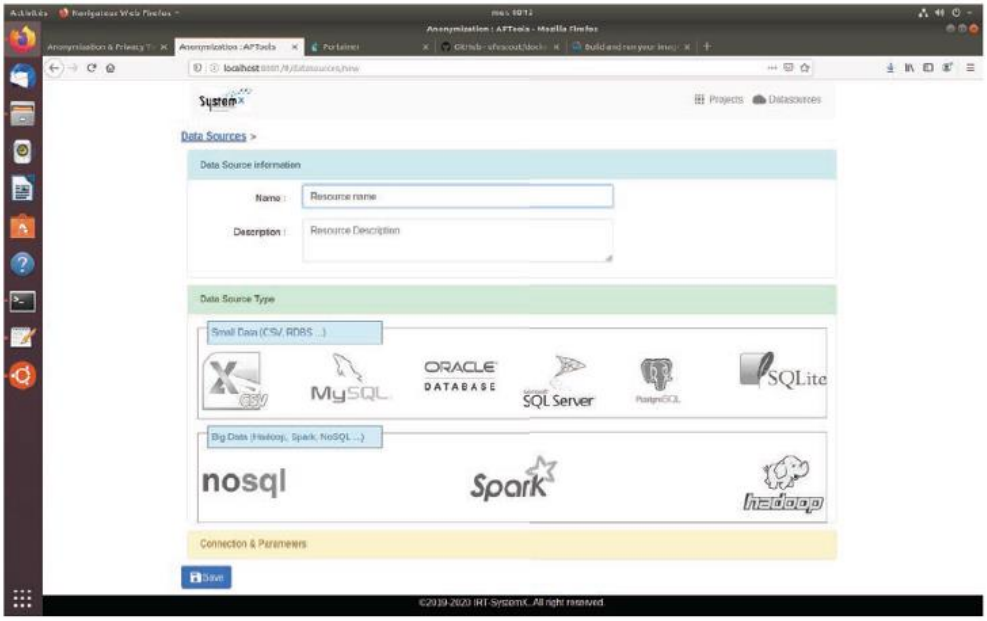


Figure 43: DAT Interface Overview

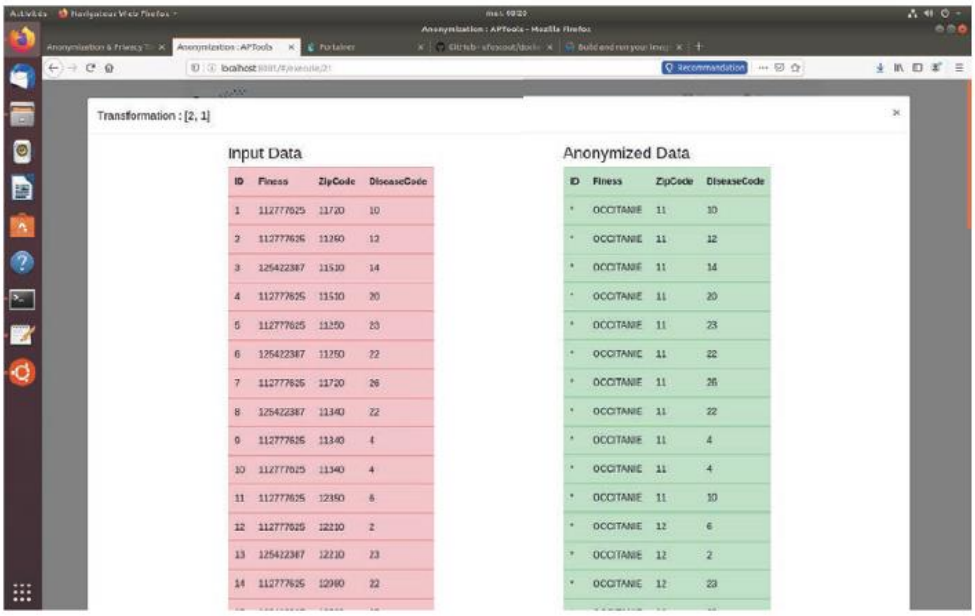


Figure 44: Example of anonymised results in a database

4.9.4 Summary and Next Steps

DAT consists of a unified web interface and a simple, flexible and configurable anonymisation engine. It enables configuration of the dataset and the chosen anonymisation models. The current version uses SQL database and CVS file. We plan to extend its functionalities to allow anonymisation of files and in particular CTI reports.

In detail, as depicted in 43, the DAT extracts and anonymises sensitive data in the CTI report received from the EWS before it is shared with other users. The extraction step applies NLP techniques to automatically extract a predefined set of entity. In a second step, the tool applies to each extracted entity, an appropriate anonymisation technique according to a defined policy.

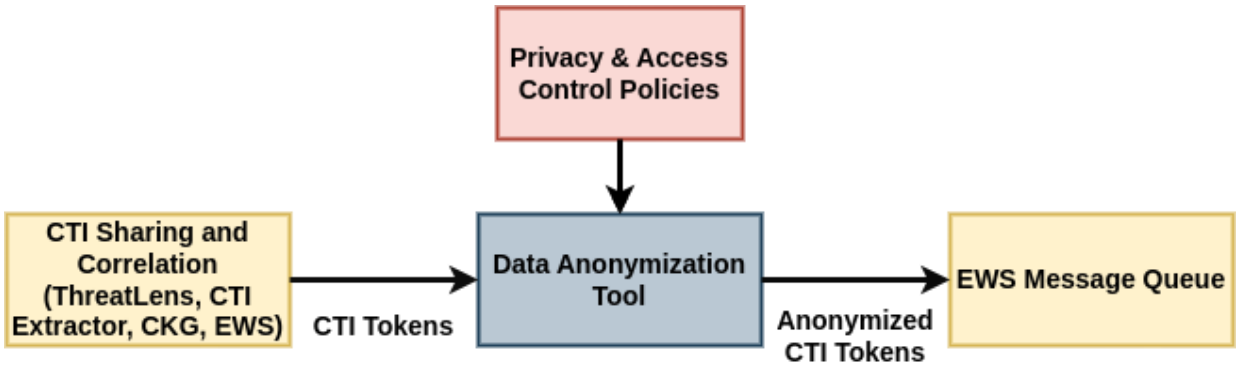


Figure 45: CTI Anonymisation Flow

Its new version, known as PARIS (Privacy-Aware Reports and CTI Sharing) is an advanced tool designed to automatically detect and anonymize sensitive data in CTI reports. It utilises a NER model specifically adapted to capture both general entities and additional CTI-specific entities, enhancing its ability to identify critical information.

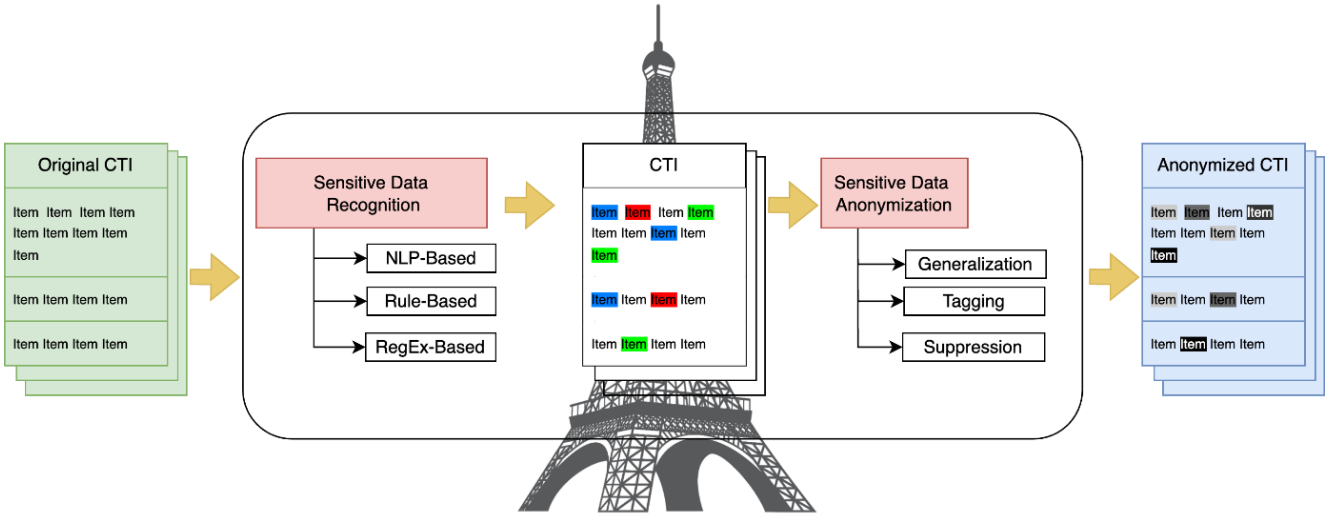


Figure 46 : PARIS Architecture

The process begins with the tool ingesting a CTI report as input, from which it extracts sensitive data using the NER model. we use two datasets created from CTI reports, APTNER [33] and DNRTI [34] to train the model. The APTNER dataset contains 21 predefined entity categories, which impacts the training of any model on this dataset while the the DNRTI dataset contains only 13 predefined entity categories.

Exhaustive lists of known ransomware and APTs are added to the pre-trained spaCy NER model. This model can then recognize built-in entities and newly added entities. A rule-based approach is

used to extract other sensitive data such as IP address, email address, domain address, and vulnerability ID.

This model identifies and categorizes various sensitive entities present in the report. Based on the sensitivity of these entities, PARIS applies appropriate anonymisation techniques, such as pseudonymisation or masking, to protect the data without compromising the report's usefulness. Once the sensitive information is anonymised, PARIS generates an anonymised version of the CTI report, ready for secure sharing. The tool’s architecture (as depicted in the accompanying figure 46) efficiently handles the end-to-end process of identifying and anonymising sensitive data, ensuring a balance between security and information sharing.

Chapter 5 Summary and Conclusion

This document reports the development status of the CTI prototypes. It highlights the work performed by the members of WP4 team in the development phase of the tasks. It also serves as proof that the CTI prototypes are aligned with the current state and needs in CTI and information-sharing. The tools are also addressing the challenges in collecting, analysing and utilising threat information. These actions are supported and strengthened with the use of AI-based tools, as explained in Chapter 2. The consideration of the AI act, along with NIS2 and GDPR vital to the CTI Framework due to the nature of the information that is being used or produced by its tools.

Collaboration with the BCM framework of the DYNAMO platform is an important objective for DYNAMO and the CTI in particular. The utilisation of the received threat intelligence and the protection of the infrastructure against threats and attacks is materialised through the use of the BCM Framework. Prior to this, the collaboration with the BCM Framework, will allow a more specified search of CTI and a more tailored identification of threats by the CTI Framework.

The main outcomes are:

- An initial description of the CTI prototypes
- An overview of the User Interface or Mockups of the prototypes
- Detailed development status of the prototypes
- An overview of the prototypes' position within DYNAMO's architecture
- A description for the future development steps

This report will be used until the completion of the development stage within the project as a reference for the next version of the prototypes and any activities.

Chapter 6 Bibliography

[1] ENISA, ENISA THREAT LANDSCAPE 2023, 2023.

[2] N. Sun, M. Ding, J. Jiang, W. Xu, X. Mo and Y. Tai, “Cyber Threat Intelligence Mining for Proactive Cybersecurity Defense: A Survey and New Perspectives,” *IEEE Communications Surveys and Tutorials,* vol. 25, no. 3, pp. 1748-1774, 2023.

[3] C. Yucel, I. Chalkias, D. Mallis, E. Karagiannnis, D. Cetinkaya and V. Katos, “On the assessment of completeness and timeliness of actionable cyber threat intelligence artefacts,” in *Multimedia Communications, Services and Security*, Krakow, Poland, 2020.

[4] T. D. Wagner, K. Mahbub, E. Palomar and A. Abdallah, “Cyber threat intelligence sharing: Survey and research directions,” *Cyber threat intelligence sharing: Survey and research directions,* vol. 87, no. C, 2019.

[5] M. Arazzi, D. R. Arikkat, S. Nicolazzo, A. Nocera and M. Conti, “NLP-Based Techniques for Cyber Threat Intelligence,” arXiv, 2023.

[6] N. Goel and N. Sethi, “CYBER THREAT INTELLIGENCE: A SURVEY ON PROGRESSIVE TECHNIQUES AND CHALLENGES,” in *International Conference on Big Data, IoT, Cyber Security and Information Technology (ICBDICSIT)*, Pune, India, 2022.

[7] B. Stojkovski, G. Lenzini, V. Koenig and S. Rivas, “What’s in a Cyber Threat Intelligence sharing platform?: A mixed-methods user experience investigation of MISP,” in *ACSAC '21: Proceedings of the 37th Annual Computer Security Applications Conference*, New York, USA, 2021.

[8] P. Ranjit, S. Borah, A. K. Bhoi, F. I. Muhammad , M. Pramanik, Y. Kumar and R. H. Jhaveri, “A consolidated decision tree-based intrusion detection system for binary and multiclass imbalanced datasets.,” *Mathematics,* vol. 9, no. 7, p. 751, 2021.

[9] M. A. Umer, K. N. Junejo, M. T. Jilani and A. P. Mathur, “Machine learning for intrusion detection in industrial control systems: Applications, challenges, and recommendations.,” in *International Journal of Critical Infrastructure Protection,*, Berkeley, USA, 2022.

[10] P. Saraswat, “Supervised Machine Learning Algorithm: A Review of Classification Techniques,” in *International Conference on Intelligent Emerging Methods of Artificial Intelligence & Cloud Computing*, 2021.

[11] M. P. Malek, S. Naderi and H. G. Garakani, “A review on internet traffic classification based on artificial intelligence techniques,” *International Journal of Information and Communication Technology Research,* vol. 14, no. 2, pp. 1-13, 2022.

[12] S. J. Moore, “Deep learning for network intrusion: A hierarchical approach to reduce false alarms,” *Intelligent Systems with Applications,,* vol. 18, pp. 200-215, 2023.

[13] L. Santos, R. Goncalves, C. Rabadao and J. L. B. R. Martins, ""A flow-based intrusion detection framework for internet of things networks," *Cluster Computing,* pp. 1-21, 2023.

[14] U. Imtiaz and M. Qusay H. , "A Two-Level Hybrid Model for Anomalous Activity Detection in IoT Networks," *Electronics,* vol. 9, no. 3, p. 530, 2020.

[15] R. Vijayanand and D. Devaraj, "A Novel Feature Selection Method Using Whale Optimization Algorithm and Genetic Operators for Intrusion Detection System in Wireless Mesh Network," *IEEE Access,* vol. 8, pp. 56847-56854, 2020.

[16] C. Zhang, X. Yue, R. Wang, N. Li and Y. Ding, "Study on Traffic Sign Recognition by Optimized Lenet-5 Algorithm," *International Journal of Pattern Recognition and Artificial Intelligence,* vol. 34, p. 2055003, 2020.

[17] R. Ahmad, I. Alsmadi, W. Alhamdani and L. Tawalbeh, "Towards building data analytics benchmarks for IoT intrusion detection," *Cluster Computing,* vol. 25, pp. 2125-2141, 2022.

[18] I. Ullah and Q. H. Mahmoud, "A two-level flow-based anomalous activity detection system for IoT networks," *Electronics,* vol. 9, no. 3, p. 530, 2020.

[19] C. Zheng, M. Zang, X. Hong, B. Riyad, S. Vargaftik, Y. Ben-Itzhak and N. Zilberman, "Automating In-Network Machine Learning," arXiv, 2022.

[20] E. Chatzoglou, G. Kambourakis, C. Kolias and C. Smiliotopoulos, "Pick Quality Over Quantity: Expert Feature Selection and Data Preprocessing for 802.11 Intrusion Detection Systems," *IEEE Access,* vol. 10, pp. 64761-64784, 2022.

[21] F. Kaiser, T. Budig, E. Goebel, T. Fischer, J. Muff, M. Wiens and F. Schultmann, "Attack Forecast and Prediction," in *C&ESAR 21 - Computers & Electronics Security Applications Rendez-vous*, Rennes, France, 2021.

[22] A. Swaminathan, B. Ramakrishnan, M. Kanishka and R. Surendran, "Prediction of Cyber-attacks and Criminality Using Machine Learning Algorithms," in *2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies, 3ICT 2022*, Bahrain, 2022.

[23] A. Casolaro, V. Capone , G. Iannuzzo and F. Camastra, "Deep Learning for Time Series Forecasting: Advances and Open Problems," *Information 2023 New Deep Learning Approach for Time Series Forecasting,* vol. 14, no. 11, 2023.

[24] . K. Kowsari , K. J. Meimandi, M. Heidarysafa, S. Mendu , L. Barnes and D. Brown, "Text Classification Algorithms: A Survey," *Machine Learning on Scientific Data and Information),* vol. 10, no. 9, 2019.

[25] S.-W. Kim and J.-M. Gil, "Research paper classification systems based on TF-IDF and LDA schemes," *Human-centric Computing and Information Sciences,* vol. 9, 2019.

[26] G. Aivatoglou, M. Anastasiadis, G. Spanos, A. Voulgaridis, K. Votis, D. Tzovaras and L. Angelis, "A RAkEL-based methodology to estimate software vulnerability characteristics & score - an application to EU project ECHO," *1180: Cybersecurity, Intelligent Multimedia Systems for Threat Detection and Data Protection,* p. 9459–9479, 2021.

[27] A. Giorgos, M. Anastasiadis, G. Spanos, A. Voulgaridis, K. Votis and D. Tzovaras, “A tree-based machine learning methodology to automatically classify software vulnerabilities,” in *IEEE International Conference on Cyber Security and Resilience (CSR)*, 2021.

[28] M. Anastasiadis, G. Aivatoglou, G. Spanos, A. Voulgaridis and K. Votis, “Combining text analysis techniques with unsupervised machine learning methodologies for improved software vulnerability management,” in *2022 IEEE International Conference on Cyber Security and Resilience (CSR)*, Rhodes, Greece, 2022.

[29] A. Rosay, F. Carlier, E. Cheval and P. Leroux, “From CIC-IDS2017 to LYCOS-IDS2017: A corrected dataset for better performance,” in *WI-IAT '21: IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, New York, 2022.

[30] T. McIntosh, A. S. M. Kayes, Y. P. P. Chen, A. Ng and P. Watters, “Ransomware mitigation in the modern era: A comprehensive review, research challenges, and future directions.,” *ACM Comput. Surv. (CSUR),* vol. 54, no. 9, pp. 1-36, 2021.

[31] B. Alhijawi, S. Almajali, H. Elgala, H. B. Salameh and M. Ayyash, “A survey on DoS/DDoS mitigation techniques in SDNs: Classification, comparison, solutions, testing tools and datasets.,” *Comput. Electr. Eng,* vol. 99, p. 107706, 2022.

[32] W. B. Shahid, B. Aslam, H. Abbas and S. Khalid, “An enhanced deep learning based framework for web attacks detection, mitigation and attacker profiling.,” *J. Netw. Comput. Appl.,* vol. 198, p. 103270, 2022.

[33] X. H. S. X. Z. W. X. J. Z. C. S. J. J. Wang, “Aptner: A specific dataset for ner missions in cyber threat intelligence field,” in *2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design (CSCWD) pp. 1233–1238 (2022)*, 2022.

[34] X. L. X. A. S. L. N. J. Z. X. Z. X. Z. X. M. Z. X. Wang, “Dnrti: A large-scale dataset for named entity recognition in threat intelligence,” in *2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom). pp. 1842–1848. IEEE (2020)*, 2020.

[35] C. C. Aggarwal and Z. C. Xiang, “A survey of text classification algorithms.,” in *Mining text data*, Springer, 2012, pp. 163-222.

[36] L. Man, T. C. Lim, S. Jian and L. Yue, “Supervised and traditional term weighting methods for automatic text categorization.,” *IEEE transactions on pattern analysis and machine intelligence,* pp. 721-735, 2008.

[37] J. Read, B. Pfahringer, G. Holmes and E. Frank, “Classifier chains for multi-label classification,” *Machine learning,* vol. 85, no. 3, pp. 333-359, 2011.