Funded by the Horizon 2020 Framework Programme of the European Union
PREVISION - Grant Agreement 833115

Deliverable D4.1
Title: Improved Operational and Situational Awareness Applications (Initial Release)

Dissemination Level:	**PU**
Nature of the Deliverable:	**R**
Date:	**30/06/2020**
Distribution:	**WP4**
Editors:	**ITTI**
Reviewers:	**TBD**
Contributors:	**ALL**

Abstract: This document contains the first (initial) description of the improved operational situational awareness applications to be developed as a part of the PREVISION platform. The main goal was to initially describe key functionalities expected in each defined use-case scenario, namely: identification of radicalization and terrorist propaganda, protection of soft targets, fight against illicit trafficking and the analysis of cyber-criminal activities. Moreover, this document also tackles the design of the tools and to identify approaches related to particular applications constituting the operational and situational awareness.

**** Dissemination Level:*** *PU= Public, RE= Restricted to a group specified by the Consortium, PP= Restricted to other program participants (including the Commission services), CO= Confidential, only for members of the Consortium (including the Commission services)*

***** Nature of the Deliverable:*** *P= Prototype, R= Report, S= Specification, T= Tool, O= Other*

Disclaimer

Revision History

Date	Rev.	Description	Partner
19/03/2020	0.1	Initial ToC, Introduction, initial contribution in section 5	ITTI
31/03/2020	0.2	Description of user expectations included in section 2	ITTI
17/04/2020	0.3	Version discussed during the WP4 internal teleconference	ITTI
30/04/2020	0.4	Added subsection to section 5 and 7	ITTI
30/04/2020	0.5	Integrating PARCS contribution	ITTI
07/05/2020	0.6	Improving ITTI's content	ITTI
27/05/2020	0.7	Integrating contribution by IfmPt, UM, and KEMEA	ITTI
03/06/2020	0.8	Comments and contribution by CNRS	CNRS
18/06/2020	0.85	Contribution by UPV integrated	UPV
19/06/2020	0.9	Enhancing UM's content with details	UM
24/06/2020	0.11	Adding two schemas to illustrate 4.2 task by BPTI and CNRS	BPTI, CNRS
01/07/2020	0.12	Contribution by MD SPP integrated	MD SPP
12/07/2020	0.13	Contribution by UPV	UPV
14/07/2020	0.14	Improvements and structure refactoring	ITTI
15/07/2020	0.15	Contribution by SIMAVI integrated	SIMAVI
15/07/2020	0.16	Pre-final version	ITTI
20/07/2020	0.17	Addition by ETRA integrated	ETRA
27/07/2020	0.18	Implementing reviewers' recommendations and comments	UTP

List of Authors

Partner	Author
BPTI	Tomas Krilavičius, Justina Mandravickaitė, Jonas Uus
CNRS	Ngoc Hoang, Josiane Mothe, Olivier Teste, Zia Ulhah
ETRA	Antonio Moreno Borrás
IfmPt	Guenter Okon, Kira Langanki, Michael Schweer
ITTI	Michał Choraś, Marek Pawlicki, Aleksandra Pawlicka, Rafał Kozik, Damian Puchalski
KEMEA	Eleni Darra
MD SPP	Lilia Popovici
PARCS	Axel Kerep
SIMAVI	Nicu Jalba, Gabriela Panzariuc
UM	Misha Glazunov
UPV	Francisco Pérez, Victor Garrido, Alberto García

Table of Contents

Index of figures

Index of tables

Glossary

ACL(s)	Access Control List(s)
AMQ	Advanced Messaging Que
API	Application Programming Interface
AR	Augmented Reality
C2IS	Command and Control Information System
CCTV	Closed-circuit television
CTRL	Conditional Transformer Language
DMARC	Domain-based Message Authentication, Reporting and Conformance
DNN(s)	Deep Neural Network(s)
ETL	Extracted Transformed and Loaded
GIS	Geographic Information System
GPS	Global Positioning System
GUI	Graphical User Interface
HCD	Head-Coupled Displays
HMI	Human–Machine Interface
HTML	Hypertext Mark-up Language
HTTP(S)	Hypertext Transfer Protocol (Secure)
HUD	Head-Up Display
IMAP	Internet Message Access Protocol
ISO	International Organization for Standardization
LDA	Latent Dirichlet Allocation
LEA	Law Enforcement Agency
LTS	Long Term Support
MEMS	Microelectromechanical systems
MITM	Man-In-The-Middle attack
PC	Personal Computer
PIN	Personal Identification Number
POP3	Post Office Protocol v3

RTF	Result Transferability Framework
SMTP	Simple Mail Transfer Protocol
TF-IDF	Term Frequency–Inverse Document Frequency
UC	Use Case
UK	United Kingdom
URL	Uniform Resource Locator
USB	Universal Serial Bus
VM(s)	Virtual Machine(s)
VPN	Virtual Private Network
VR	Virtual Reality
WP	Work Package

Executive Summary

This deliverable is an initial version of the report entitled “Improved Operational and Situational Awareness Applications – Initial Release”. In the document we have identified a list of tools that potentially constitute competition for PREVISION Situational Awareness Applications. In that regards, we have highlighted potential gaps concerning graphical user interfaces, user experience, and methods for data interaction.

We followed up this analysis with the general design concerning operational and situational awareness. In particular, we have identified potential technologies and frameworks that may underpin the core capabilities. One of the key finding here concerns the adaptation of web HMI as main element fostering WP4 applications integration and harnessing various solutions facilitating multi-dimensional data interaction.

Afterwards, details on design, functionalities, background, and position with concerning other PREVISION services for each of the WP4 applications are given. In particular, the following solutions are identified: identification of radicalization and terrorist propaganda, protection of soft targets, fight against illicit trafficking and the analysis of cyber-criminal activities.

1. Introduction

This document is the first version of deliverable entitled "Improved Operational and Situational Awareness Applications – Initial Release", describing key envisaged functionalities and early design and approaches related to particular tools constituting the operational and situational awareness toolset in PREVISION platform. The D4.2 report that is considered as a follow-up of this deliverable is scheduled at M17.

1.1 Motivation

The main goal of this document is to provide a first description, in particular: motivation, context, state-of-the-art and initial design of the tools to be developed in WP4. These tools will aim at enhancing operational and situational awareness of LEAs protecting civilians in different scenarios. Therefore, the document is structured in a similar way as WP4 work is decomposed into tasks.

1.2 Intended audience

The document provides information aiming both technical and LEA partners of the PREVISION. In particular:

- PREVISION technical partners, who will contribute to the PREVISION Platform development, integration, deployment and validation get information on the initial vision of particular components of the PREVISION platform,
- LEA partners involved in the project, get information about the WP4 tools (both from technical as well as functional perspective) that will be deployed within the PREVISION platform in their organisations in the pilot phase.

In general, the majority of the PREVISION partners is involved in one or more WP4 tasks, therefore the intended audience can be considered as the project-wide.

1.3 Relation to other deliverables

This document is strictly related to subsequent WP4 output – D4.2 deliverable (M17) focusing on refinement and further specification of the information provided in D4.1. In addition, different WP4 tools will use findings and output developed in other WPs. In particular, D2.1 (Heterogeneous Data Streams Processing Tools) and D3.1 (Machine Learning and Automation for Crime Prevention and Investigation) serve as an input to the current report. Also, all the finalized outputs of WP1 focused on the end-user perspective, LEAs needs and use cases are taken into account.

1.4 Deliverable structure

As it was mentioned in section 1.1, the structure of the current deliverable is related to the structure of the WP4, following the work distribution into specific tasks:

- Section 3 is related to T4.1 and focuses on operational and situational awareness tools including intelligent HMI solutions, augmented reality, visualization techniques, etc.,

- Section 4 is related to T4.2 in which work concentrates on the identification of radical content spread for terrorist propaganda purposes,
- Section 5 is linked to T4.3 and focuses on techniques enhancing security of crowded places with limited safeguarding capabilities (so-called “soft targets”),
- Section 6 related to T4.4 is focused on the fight against illicit trafficking, and
- Section 7 describes aspects of trend characterization in cybercriminal activities, in particular identification of data sources allowing for detection of specific patterns, large-scale malware analysis and classification, etc.

In addition, section 2 provides general information on key functionalities expected by the end-users expressed at the first stage of the project, while section 8 provides a summary and plan for the D4.2 follow-up report.

2. Situational Awareness Applications

2.1 Exiting solutions

2.1.1 Tools overview

After the consultations with the PREVISION LEAs and practitioners (ELAS and KEMEA for example), the following list of tools for situational awareness has been identified:

- **IBM I2 Analyze (**https://www.ibm.com/products/enterprise-intelligence-analysis **– Commercial):** IBM I2 Analyze is an enterprise intelligence analysis environment used to enable information sharing and intelligence production. It facilitates the analysis of large volumes of data through a secure environment, it enables analysts and operational teams to work collaboratively across boundaries, it has an open and extensible architecture, to complement current processes and procedures and its access is controlled through a highly configurable, fine-grained and pervasive security model.
- **Tovek (Track, Organise and Visualise Elements of Knowledge from any data to provide correct, clear and fast answers** https://www.tovek.cz/en/ **- Commercial):** Tovek Tools is a desktop application for analysts offering them effective search, entity extraction, linking, visualisation and reporting of facts and contexts. It processes the content of various unstructured and structured data, either on a user's computer or connected via Tovek Servers.
- **QGIS (A Free and Open Source Geographic Information System** https://qgis.org/en/site/ **- Free):** QGIS is an open-source cross-platform desktop geographic information system (GIS) application that supports viewing, editing and analysis of geospatial data. It allows users to analyze and edit spatial information, in addition to composing and exporting graphical maps. QGIS supports vector data, shapefiles, coverages, personal geodatabases, dxf, MapInfo, PostGIS, Web services, including Web Map Service and Web Feature Service and other formats.
- **Web-IQ Voyager ((crawler clear & dark web) - Voyager Web-IQ's turnkey web crawling and intelligence platform -** https://web-iq.com/products/voyager **- Commercial):** Voyager captures online data sources and builds actionable intelligence by analyzing the captured data and by allowing teams to find new evidence, get valuable insights and discover trends. Some of its features include the capture of online sources, the building of intelligence as well as getting valuable insights and trends discovering.
- **Big Data Analysis tools:** There is an extensive list of tools used to analyze datasets to draw results. These tools can be both commercial and/or free.
- **Open sources analytics tools:**
    - **Tweetdeck:** it is a social media dashboard application for management of Twitter accounts and
    - **Hootsuite:** it is a social media management platform which interface takes the form of a dashboard and supports social network integrations for Twitter, Facebook, Instagram, LinkedIn and YouTube.

- **Detection tools**: network-based tools such as Nmap and Nessus used to scan and ping for vulnerabilities using a graphical user interface.

2.1.2 Identified Gaps and Opportunities

Based on our knowledge and desktop research, GUI and user experience are not primary and key factors in designing/development of current/past tools for law enforcement and crisis management. Current tools and their operators face such limitations as:

- Many tools are equipped with various GUI, characterized by extremely different layouts and way of working with them,
- Lack of intuitiveness and user-friendliness impacting the effectiveness of the operator work,
- Relatively less number of web-based, cloud-based and microservices-based solutions than standalone, desktop tools.

Some of D1.1 requirements are related to the user experience with regards to the PREVISION platform. The following expectations have been highlighted by the end-users.

- Tools constituting PREVISION platform must be integrated in a way ensuring that a work of the LEA operators will be as intuitive and user-friendly as possible, facilitating the different security management operations in incidents that could be considered critical, and in their respective phases of execution.
- Dynamic dashboard and friendly-looking interface adapted and scaled to different end-stations (desktop PC, tablets, smartphones, etc.).
- Filtering functionalities (e.g. filtering collected data based on: relevance criteria, positive match of the content, geo-fenced by region, e.g. country/city/area, languages, etc.).
- Web-based architecture rather than desktop platform.

2.2 User expectations regarding key functionalities

2.2.1 Expectation regarding Operational and Situational Awareness Tools

According to the operational and situational awareness tools, the aim is to create a platform that encompasses multiple tools for the visualization and processing of large volumes of data. To achieve this objective, advanced visualization techniques are applied, such as immersion, graphical representation and/or augmented reality, in order to improve the perception and understanding of security agents and thus facilitate decision-making. In other words, the aim is to carry out a direct interaction between users and data in order to facilitate the general understanding of the situation according to the possible cases that may arise.

To this aim, a variety of tools will be developed to help improve the processing of PREVISION data. an HMI will be created that will allow the users to interact with the entire platform since it will be the user interface that will include the tools developed by and for PREVISION.

To generate a solution as scalable and adaptable as possible, a container-based solution is proposed (it could be by means of Docker or VMs), in this way the tools can be developed in the desired language

and can be easily integrated into the platform. PREVISION tools will be able to be executed, visualized and analysed through screens designed specifically for each one of them, that is why the HMI will present a lateral bar that will allow navigating in a fast and simple way through the multiple options.

Communication between the tools can be generated in the following ways:

On the one hand, all communication will be done through an asynchronous communication bus that will allow each tool to be published and subscribed to the topics that are of interest to it in order to be able to take action when relevant information is published. On the other hand, it will be possible to expose the tools through REST services.

Specifically, the purpose of task T4.1 is to provide different applications that offer the user the possibility of having global knowledge of the situation in order to perform one task or another.

The main tools that will be used for the visualization and the interaction between the user and the platform will be:

a. **HMI (Human Machine Interface):** Global platform that will house most of the tools developed in order to be able to generate an interface capable of facilitating the work of LEAS by being able to visualize alerts, interact with data and manage situations.
b. **Leap Motion:** Interactive technology that brings to the project the capability to manipulate and navigate through the data using gestures instead of using the mouse.
c. **AR Mobile Application:** Mobile application capable of visualizing and representing data in augmented reality.

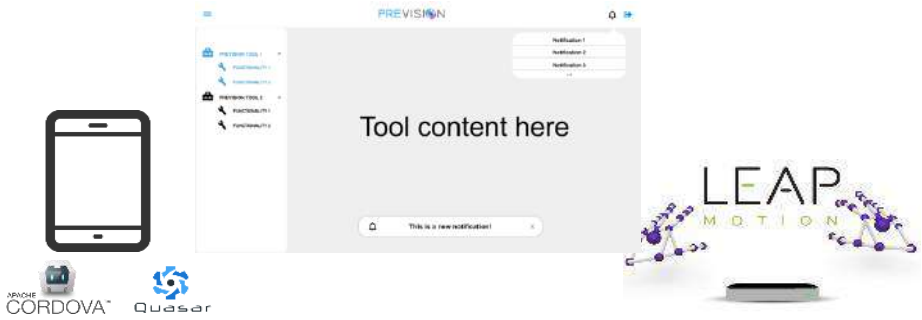

Figure 1: Situational awareness tools

2.2.2 Expectation regarding Soft Target protection

Soft Targets are those places that are in an incapacity to counteract attacks, initiated by terrorists, aggressive individuals or extremists. The terrorism phenomenon gains territory in Europe day by day and thus, stadiums and other public spaces where first persons in the state are present, on different occasions, make MD SPP responsible for protecting them from serious terrorist assaults.

The stadiums also represent places where people congregate in impressive number and more likely to become a target of such incidents, being less secured. This challenge should be addressed before it

occurs or immediately after it takes place. Before the assault the risks and threats analysis of possible dangers should be undertaken. Forecasting is somehow possible in places where the probability of occurrence is bigger.

The major rule for securing soft targets Is the efficiency of security technologies in order to protect people. In this case surveillance systems offer an overview of the environment, rooms and people entering this soft target and their behaviour. If the number of cameras is limited then the recording device is switched on for further evidence. Big advantages are cameras with analytical functions, including face detecting, abnormal behaviour etc.

ID scanners may also prove efficient verification of people entering the target place. Raising security awareness of the personnel, cooperation between all the forces deployed, exchange of information and valuable details together may cumulate the success of the operation.

Media being very attractive to attackers also may create an image of a secured zone before the event, thus discouraging their intentions. Also, early detection is possible by collecting information from media and social networks where one may depict clues that can lead to a possible terrorist attack, organized in this manner.

Suspicious behaviour can be detected after an active search of suspicious signs in the area that come in contrary to the usual course of events. Thus, recordings and statistics of such events and similar events that can lead to an attack should be considered.

Finally, cooperation with other services involved in the same type of security missions and with other soft targets represent the prerogatives for the success in preventing, detecting, deterring, promptly answering and in diminishing the consequences of different forms of terrorist attacks.

The document D1.1, delivered at the M3 was the first report devoted to specify project use-cases and to describe end-user and system requirements to be covered by the PREVISION platform. The detailed specification of the Soft Target protection scenario (UC1) and the set of defined functional and operational requirements expressed by the PREVISION end-users allowed us to cluster user needs concerning Soft Target protection components into several expectations:

a. The **treatment of video, images and audio** coming from **different sources and formats**, with the possibility of **processing and applying algorithms** that facilitate the operators in **the identification** of people, situations, events and suspicious patterns, in real-time.
b. Analysis of **social networks and Internet communications** in which possible threats are identified and detected early.
c. **Geo-referencing of the threats and assets** to be protected by the LEAs, as an aid to the planning and management of public security and protection.
d. **Tools for the treatment and analysis of data**, both structured and unstructured, relevant to the prevention, mitigation and investigation of criminal acts (including tools against information hiding in data streams).

e. **Integration of tools into a work platform** for the LEA operators that is as intuitive and user-friendly as possible, facilitating the different security management operations in incidents that could be considered critical, and in their respective phases of execution.

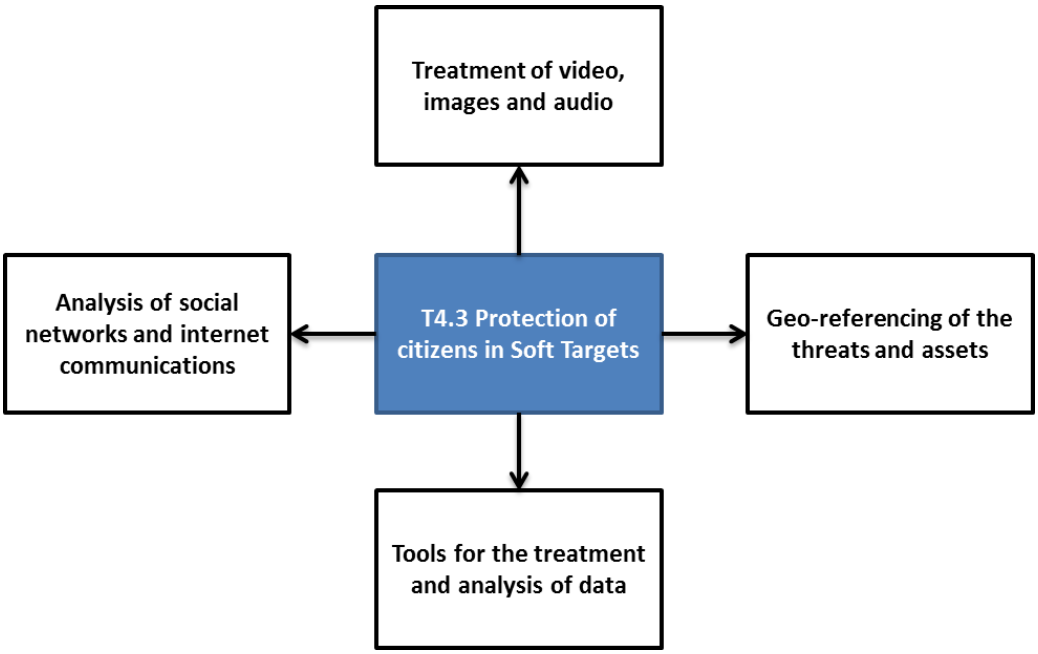

Figure 2: User expectations regarding the protection of citizens in Soft Targets (source D1.1)

2.2.3 Expectation regarding Fight against Illicit Trafficking

As described in WP1, the Law Enforcement Agencies face new rising activities of trafficking of Cultural objects coming from conflict zones all over the world and massive metal-detecting plundering in European countries.

Lacking the specific technical solutions and the necessary resources on that, the LEAs need modern and user-friendly tools to apprehend the complex field of Art and Antiquities Market and to identify unknown and illicit artefacts. Moreover, they need to be able to confront forged or fraudulent documentation of artefact for sales, that are, in fact, items coming from Blood Archaeology or of Art Theft.

Specifically, the purpose of task T4.4 is to provide an application that best meets the needs expressed in WP1 and Use Case 5. Such functions reflect the LEAs needs, and a web scanning tool provide information useful in the fight against illicit trafficking.

The main axes of the solution are:

a) **Centralized homogeneous and normalized Database.**
Design of a unique central database with standardized items descriptions and objects typologies, to be used by intelligent browser that can ease and speed the investigation and identification processes.

This centralized database is also useful for LEA since they often do not have a powerful engine and efficient Internet access.

- **A unique device that reduces time and brings expertise to field officer.**

b) **Smart Browser.**

Design a unique smart browser that, with image and text processing, can directly interrogates centralized data base and, if possible, multiple and heterogenous sources, web or local and renders a coherent and unique response.

The tool will be able to tell if an artefact is amongst stolen object (theft) or in an existing positive database, thus creating an alert report to the positive database owner.

(in collaboration with WP3).

It will also, with a typology approach, can identify object origin and types. This allows the LEAs to have a better idea of the nature of the artefacts.

It also allows the LEAs to confront vague, wrong or fraudulent description of objects in sales

The Tool will match found type with ICOM Red List typologies of endangered objects.

- **A unique accessible platform that reduces investigations time and alert on theft objects.**
- **Indicates if endangered types of objects**

c) **Automated Predictive Web scanning.**

The LEAs will be able to program a selected web scanning automated task for the machine.

The machine will then search and identify art and archaeological artefacts on the web.

It will confront object descriptions with typology matched in the central database.

It will point incoherent and/or missing description/ origin with actual typologies.

It can also scan and track a missing object that is recorded in the stolen object database.

The various results of these tasks will produce **Automated Alert & Predictive Reports** from detected stolen objects, possible fake, incoherent typology VS origin and description on merchant websites and auction houses.

The machine will issue a list of alert reports with links to objects and websites to be processed in the second investigation by the LEAs.

The objects detected will then be integrated as new entries in the UC5 database

- **Automated processes that save time, energy and manpower**
- **Goes from reactive to a proactive strategy with a predictive solution**

2.2.4 Expectation regarding Trend Characterisation in Cybercriminal Activities

Based on D1.1 document and Use Case Scenario 4 that is related to Cyber-enabled crime, the main focus is the trend characterisation in cybercriminal activities. The key challenges include the detection of specific patterns and the characterisation of advanced cyber threats. Task 4.4 will mainly compose a visual analytics solution aiming at increasing situational awareness in the area of cyber-related activities characterization.

The main expectations regarding the trend characterization in cybercriminal activities include the following:

- **Proactive analysis of trends and social media activity** with the possibility to identify suspicious events, patterns and people who act or interact illegally with others in order to buy stolen credit cards in the dark/ surface web. This will be achieved following the processing and application of different algorithms in order to predict and deter illegal actions as the prementioned ones.
- **Proactive analysis of open sources for suspicious activity**: similar to the previous one, the proactive analysis of open sources will lead the LEAs to identify suspicious and illegal actions in the Dark marketplace as regards to stolen credit cards
- **Social behaviour analysis and advanced social media analytics**: the analysis of social networks and the social behaviour of suspected people will lead to the early identification and detection of cyber threats.
- **Malware traffic analysis:** During the investigation, LEAs have the ability through different tools, to first identify a few suspicious sites and then to determine the activities of each suspicious individual. In malware traffic analysis, it is vital to collect data such as running processes, open ports, and memory in order to understand the way the malware interacts, identify the type of information being targeted and finally identify the suspects.
- **Alerts for suspicious sites for frauds and data thefts:** Various tools can produce automated alerts and predictive reports from detected stolen cards.
- **Tools for fraud detection, prevention, and analysis of data** for the investigation of criminal acts.

Moreover, additional challenges steaming from organisational matters and certainly projecting onto to functional capabilities of these components are the following:

a. Improvement of **communication between different organizations**.
b. Better **time management**.
c. Improvement of image and video quality, better **detection and processing of biometrics and behaviour patterns,** more specifically **malicious and abnormal behaviours**.
d. Breakdown of the cyber-activities and communication performed on **social networks and the Internet** (Clearnet and Darkweb), identification and detection of cyber-crime threats, including analysis of threats based on information hiding methods.

3. Operational and situational awareness tools

3.1 General context of Operational and situational awareness tools

The current threat landscape is rapidly evolving, challenging the authorities, investigators and LEAs. Cyber capabilities, knowledge, means and novel visualization techniques are key components that can lead to an effective defence.

Nowadays criminals are considered to have more advanced capabilities, knowledge and powerful tools to attack. On the other hand, defenders, e.g. cyber security professionals, law authorities or investigators need to be equipped with similar tools to improve their situational awareness, capabilities and knowledge and address the ever-evolving threat landscape.

Developing powerful tools is essential to enhance defenders' situational awareness and to promote effective response strategy across the different stages: preparation, detection, analysis, containment, eradication, recovery and post-incident activity.

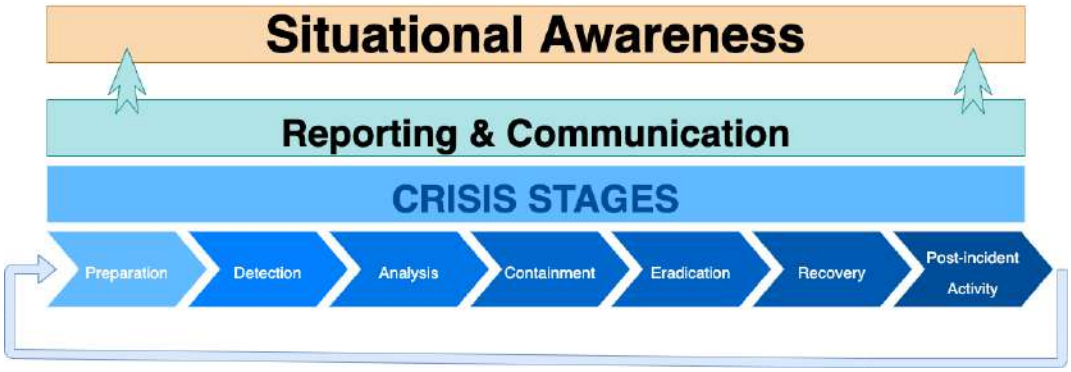

Figure 3: Event management plan

One of the key parts of PREVISION project is to provide in some way the capability to deal with complex operations in a single space. To achieve it, this chapter reviews the different information sources and the structure of the data ingested by PREVISION platform and how this information could be organized and displayed in different ways in order to provide to the investigators a better case understanding and increase their knowledge and situational awareness.

As proposed solutions, PREVISION will provide different visualization tools depending on the requirements of each case, and the main WEB HMI where the LEAS will interact with the information stored in PREVISION platform and the different services provided by the deployed tools, reducing the execution times in daily tasks.

3.2 Multi-dimensional data interaction

3.2.1 Telecom/cyber data

In order to provide multi-dimensional interaction with the telecom and cyber data, we plan to use Kibana frontend application that sits on top of the Elastic Stack. At this moment we plan to visualise various volumetric statistic concerning the collected data as well as final and intermediate results produced by malware/botnet detection micro-services. The list includes (but is not limited to only that items):

- Volumetric statistics regarding the most active hosts, most frequent internet protocols, and most popular destination services (see Figure 4).
- Sankey diagrams, which displays active hosts. Host pairs are joined together by coloured bars representing flows. The client hosts on the left edge of the diagram. Similarly, the servers are located on the right. Bar width is proportional to the amount of traffic exchanged. The wider the bar, the higher the traffic exchanged between the corresponding pair of hosts (see Figure 5).
- Geo-localization of destination services (see Figure 6)
- Various statistics concerning detected threats and anomalies.

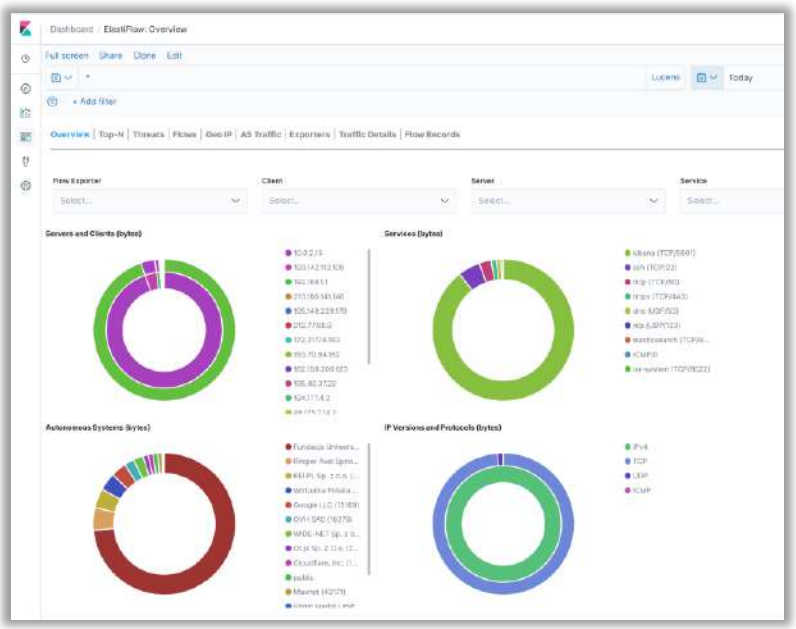

Figure 4: Volumetric statistics for collected traffic (frequent destination services, most active host, frequently used protocols, etc.)

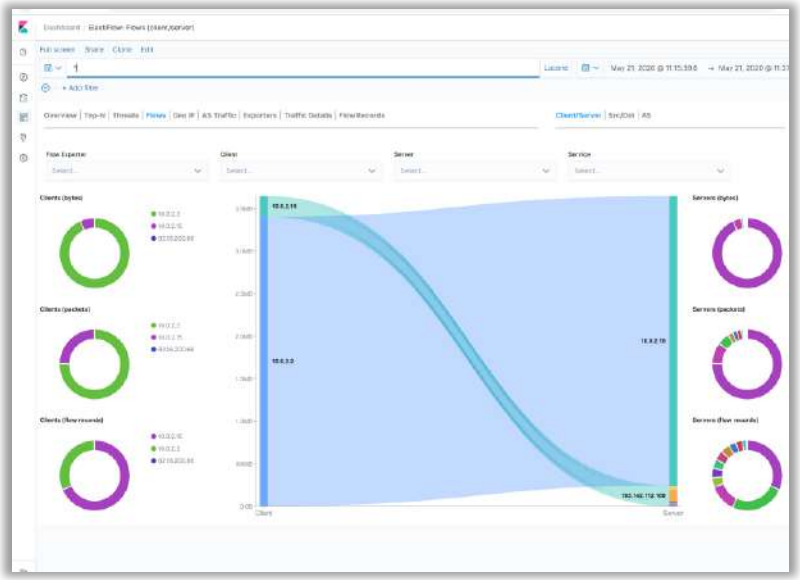

Figure 5: Sankey Diagram of "top talkers"

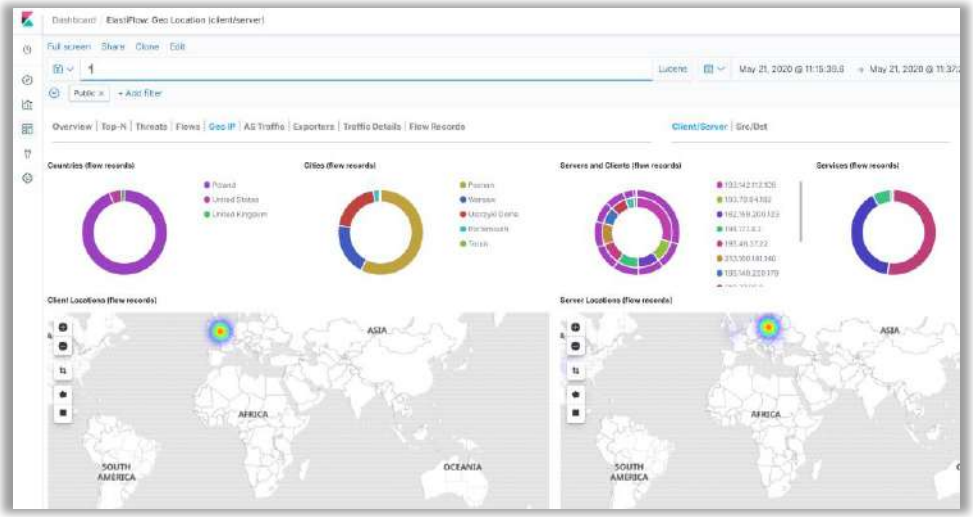

Figure 6: Destination services visualized along with the geo-localization of the IP addresses

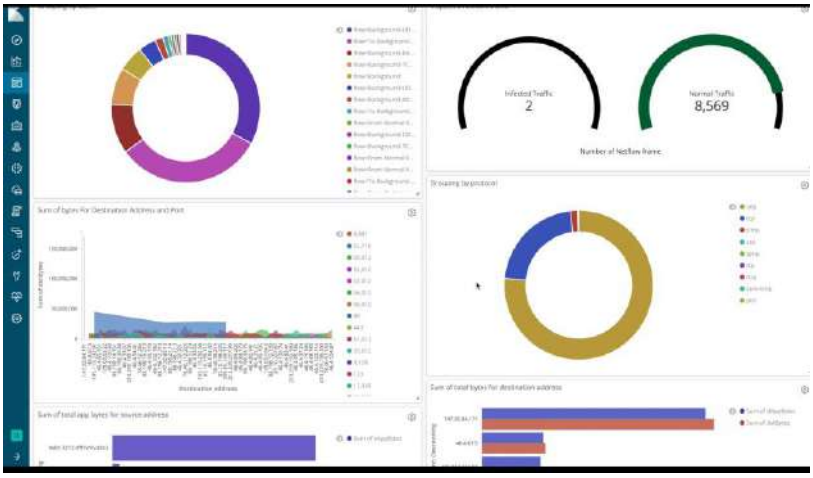

Figure 7: Statistics regarding anomalies and treats that have been detected by analyzing services.

3.2.2 Soft target protection data

From the application point of view, user-system interaction is one of the important aspects that need to be properly addressed. The soft target protection application is about bringing together information and data provided by various services and tools developed in PREVISION project. Here we take the opportunity to highlight the most relevant aspects of interaction with multi-dimensional data concerning the soft target protection.

Application-wide interaction

These may include all kind of steps the users may take before calling specific functionality (video, image, social media). We may imagine that all capabilities will leverage a map-like view that the user may interact with. Therefore, the interaction may include such activities as zooming, panning, layers switching, etc. The relevant information (assets, threats, warnings, etc.) may be displayed on the map as pins. When the user clicks on such a pin, more details may be displayed.

Batch and Near-real-time Video and Image Analysis

These functionalities are being developed under WP2. As specified in D5.1[1] the module can provide such capabilities as abnormal behaviour detection, face recognition and identification or crisis event detection. From the user perspective, it will be important to know how source data (e.g. video streams) for such tools will be provided. Also collecting the results may turn out to be a tedious task to do, as some methods would require some time to finalise the analysis and/or computations.

Web and social media crawling

These tools and functionalities will also be provided by WP2 tools (within T2.1 Crawling tools and T2.4 Darknet, Web and Social Networks Data Analysis). The main concern regarding the interaction is related to results presentation. Some parts of the data could be geo-referenced and easily presented on the

[1] D5.1 Initial PREVISION Architecture – PREVISION W5 Deliverable

map. On the other hand, such things as “detection of communities”[1] shall be presented in a different way.

The efficiency of the interaction through user interfaces with the whole system is firstly determined by the easiness of utilization. Another significant aspect is that the interface will offer intuitive and specific response.

Ensuring the interoperability between the treatment of video, images and audio, analysis of social networks and Internet communications, Geo-referencing of the threats and assets, etc. tools offered by PREVISION platform will end up in valuable connections, results capable to help the LEAs in depicting better management coordination plans of actions in organizing events and by gaining valuable time in preventing, deterring, responding and mitigating the attacks in soft targets.

Thus, following the analysis of an event and after determining a possible risk that is likely to occur represents the prerogatives for searching connections with other tools and the whole information they might offer. As a result, we will obtain all data available referring to the event, person or behaviour. This event will determine the mapping and clear geo location depicted from surveillance systems and video recordings, if they offer an appropriate view of the entire zone. Mapping each event will enable better time management in responding and even preventing different dangerous attacks.

3.2.3 Traffic, Financial and Telecom data

This module aims to inform whoever has to use the data, e.g. LEAs, operator, etc, of what information is in the system so that they can be more efficient later on when using the data to perform any type of operation with it that is useful for the purpose of PREVISION. The visualization of the information comes from the reading of the datasets and/or streaming that has entered the system and has been stored in the database in task 3.2 or read directly from some source and uploaded to the Elastic Stack. Elastic Stack has been chosen to store and display the data as it offers a fast access indexed database and a visualization of the data through Kibana on panels and dashboards, and also most partners have chosen this visualization technology. There will be an index for each topic, financial, telecommunications and traffic, and you will be able to see the fields that are considered most important and that can associate some records with others in order to visually find relationships that can later be exploited by the system. The conceptual architecture of this module is presented in Figure 8, while an example of a financial index is shown in Figure 9.

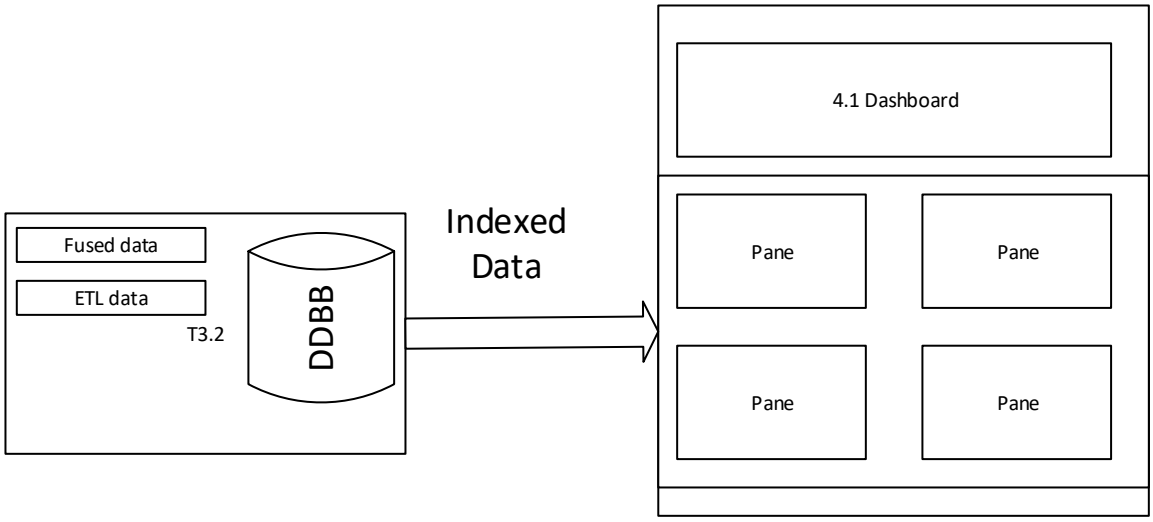

Figure 8: Structure of the system

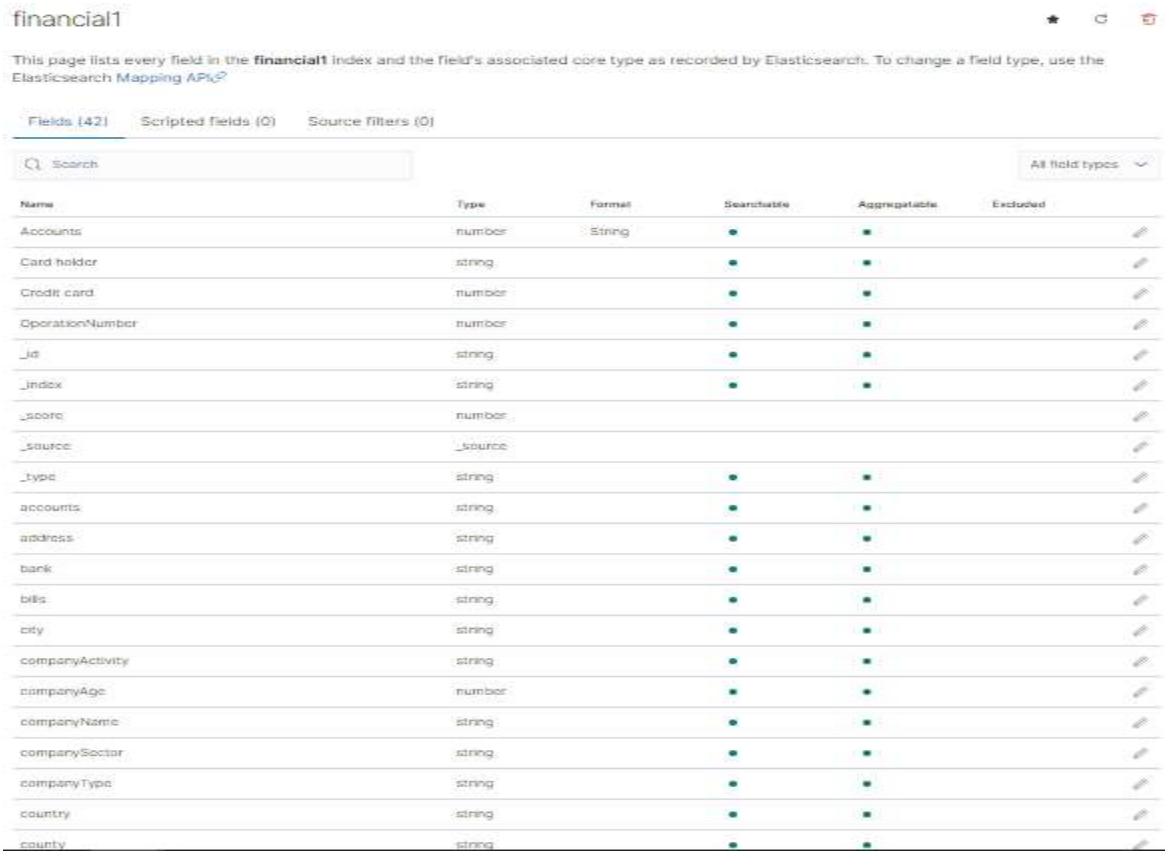

Figure 9: An example of a financial index

The necessary panels in a dashboard showing the fields considered most relevant for the users of the system, we show a small example of data and how they could be shown in Financial, Telecom and Traffic data (Figure 10).

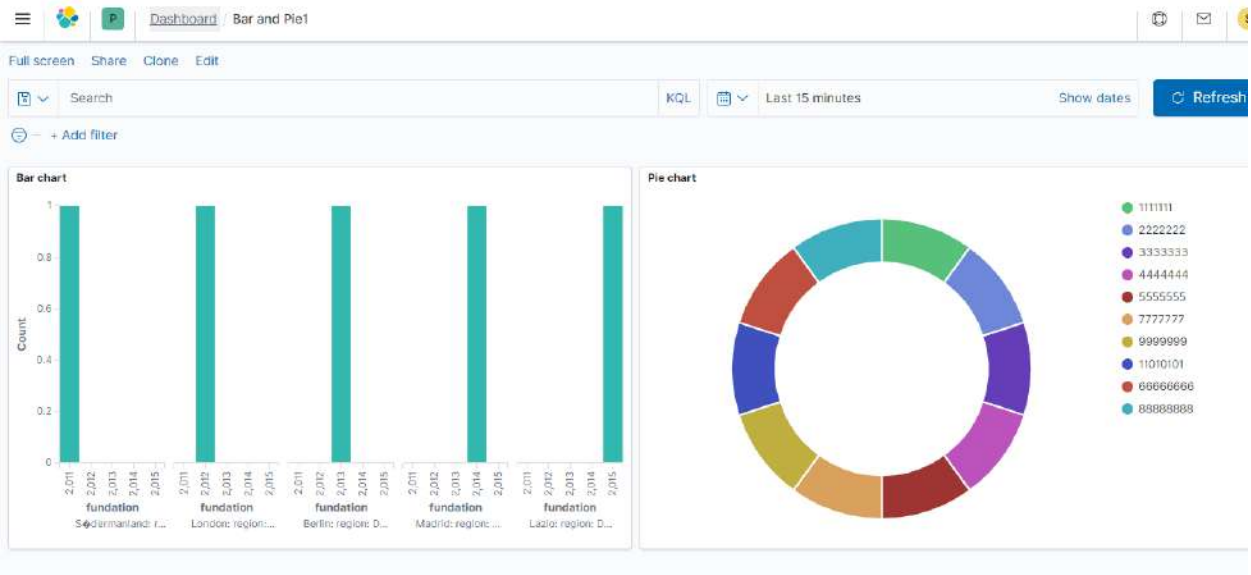

Figure 10: Example of the FTT dashboard

3.3 Visualization of large amount of data from heterogeneous sources

This section is a review of possible visualization tools provided by the technical partners and how the tools will deal with PREVISION services and the heterogeneous data stored in PREVISION databases. The tools will be analysed to understand if it is possible to integrate them in a single visualization WEB interface, or if they require additional services or hardware for proper usage.

3.3.1 New graphs representation

The fast evolution of Big Data technology in different branches such as economic, health, security has made incredible growth of data type and size. As a result, it is required to enhance the analysis, storage searching among others. Because the data analysts are more comforted with visual representations, provide interactive data visualization became obvious. The benefits of visualization tools are remarkable even more if we take into account decision-making, data sharing, reports generation and other features that remark the growing need for these tools.

Making a small research in the market, there are several open-source and private solutions that are able to display a huge amount of data in different formats using many different techniques and representation ways. In fact, we can split them into two categories. Isolated tools that can be used out of the unified PREVISION Web HMI, and integrable components usable in web dashboards.

Isolated Dashboards:

- **Kibana** is an open-source software that sits on Elastic Stack and provides great visualization and search capabilities for data indexed in Elasticsearch. Kibana has integrated the Logstash data

processing tool to complement its processing and data entry functions. As Kibana is a JavaScript application, it can be used on any platform and in various formats, without affecting the visualization and functionality of the system.

On the other hand, Kibana also offers visualizations through geographical maps that allow a flexible and cost-effective filtering, thus becoming a versatile and dynamic software for users.

It is of special interest to use the Kibana tool because:

- Search, view and visualize indexed data in ElasticSearch and analyse it by generating graphs to perform various processes:
  - Logging and analysis
  - Metrics and monitoring
  - Geospatial data analysis
  - Security analysis
  - Business Analytics
- Monitor and secure Elastic instances on the interface
- Centralize access for integrated solutions
- It has a large community of users
- Addresses many use cases

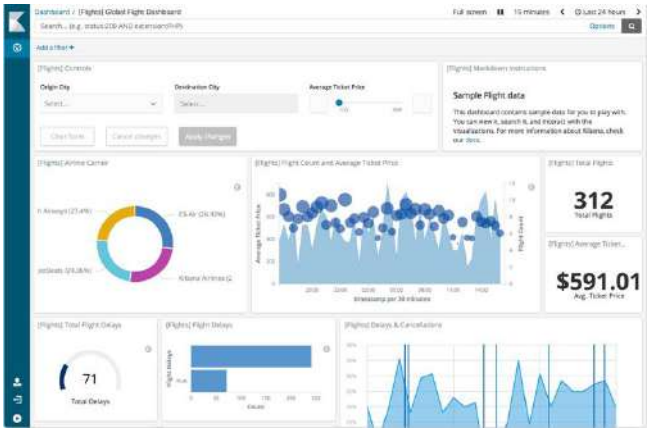

Figure 11: Example of Kibana dashboard

- **Grafana** is an open-source tool developed, specifically with Apache 2.0 license. Grafana is written in Go Language and Node.js LTS, and with a strong Application Programming Interface (API). It is a time-series data visualization tool.

  Grafana's dashboard allows you to consult, visualize, alert and understand your metrics no matter where they are stored. Create, explore and share dashboards with your team and foster a data-driven culture.

  Grafana has many advantages when interacting with your tool:

  - Elegant graphics for visualization
  - Dynamic and reusable panels
  - It is highly extensible
  - Allows data exchange between different dashboards
  - Allows you to receive alerts through notifications

Figure 12: Example of Grafana dashboard

- Tableau is a private solution for interactive data visualization. It is mainly used in business intelligence because is easy to learn and can be easily implemented in whatever type of project. The solution also includes R and Python modules to improve with custom visualizations. The cost of Tableau is reduced compared to other regular tools, such as OBE by Oracle, Business object by SAP and others. The solution provides several representation types such as:
  - Tree Map: used to show hierarchical data using size and colour.
  - Word Cloud: The words in the cloud are sized according to some measurement of their occurrence frequency.
  - Gantt Charts: used to show the overlap of tasks over a progression of time.
  - Box Plots: used to compare different sets of data and their variations.
  - Histograms: used to track the occurrence of a specific variable in a large data set.
  - Bubble Charts: data is shown in a cluster of circles.
  - Scatter Plot: used in statistics.

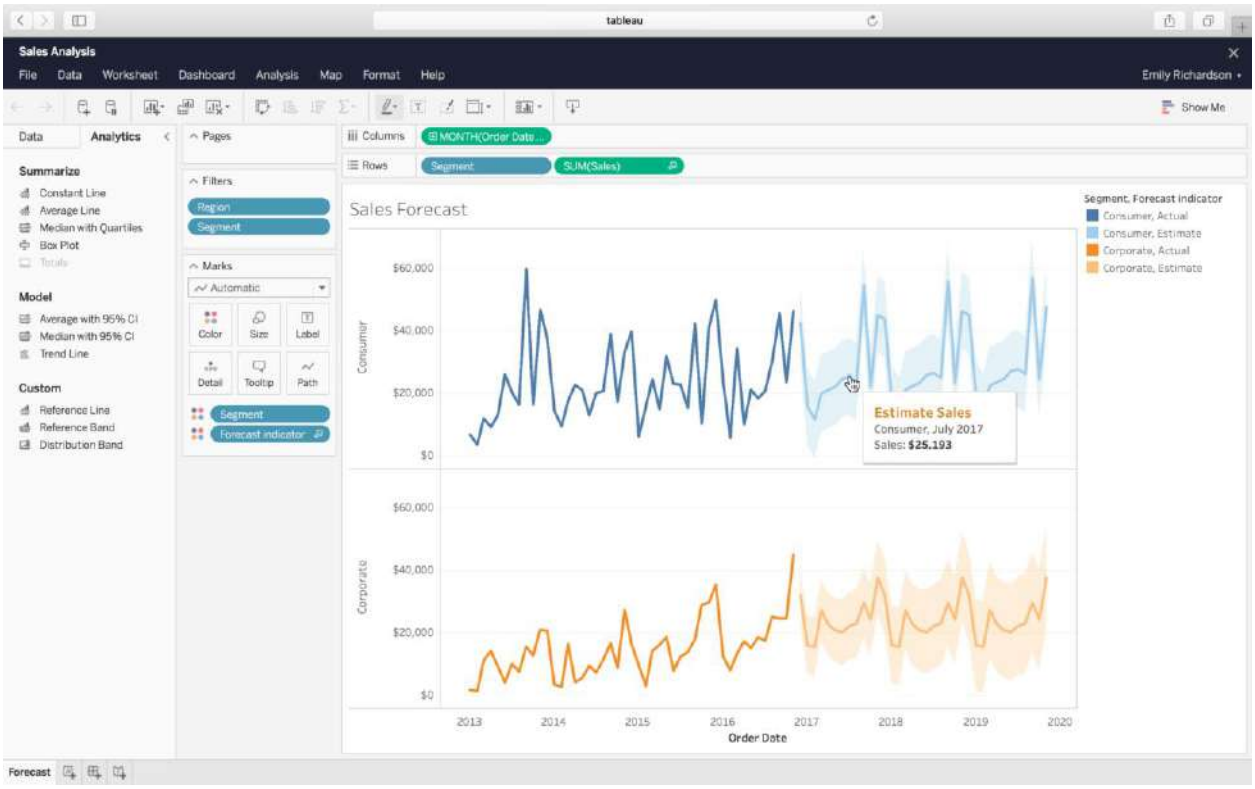

Figure 13: Example of Tableau dashboard

Integrable Components:

- **D3.js** is an open-source JavaScript library for manipulating documents based on data. It provides data visualization in different ways. In addition, the library provides data loading, data binding and analytic transformation. D3 provides a customized mapping rule and according to the needs of the users the graphs could be modified (type, colour, size …).
  It is an important remark that D3 library is compatible with all actual browsers in different operating systems and it is also compatible with mobile phones.
  Some of the most outstanding representations of the D3.js library, are:
  - Bubble Chart: encodes data in the circumference area, by size and colour.
  - Circle Packing: a grouping diagram that emphasizes the hierarchies between sets.
  - Dendrogram: hierarchical diagram of nodes and vertices in the form of a tree.
  - Force-Directed Graph: a node and vertex diagram using a positioning algorithm that minimizes vertex crossing.
  - Tree Map: map generated by recursive subdivision where the value of each node is given by the area of its rectangle.
  - Hierarchical Edge Bundling: a network of nodes that highlights inbound (dependent) and outbound (dependent) links.
  - Word Cloud: encodes by size the occurrence of the most frequent words in a text.

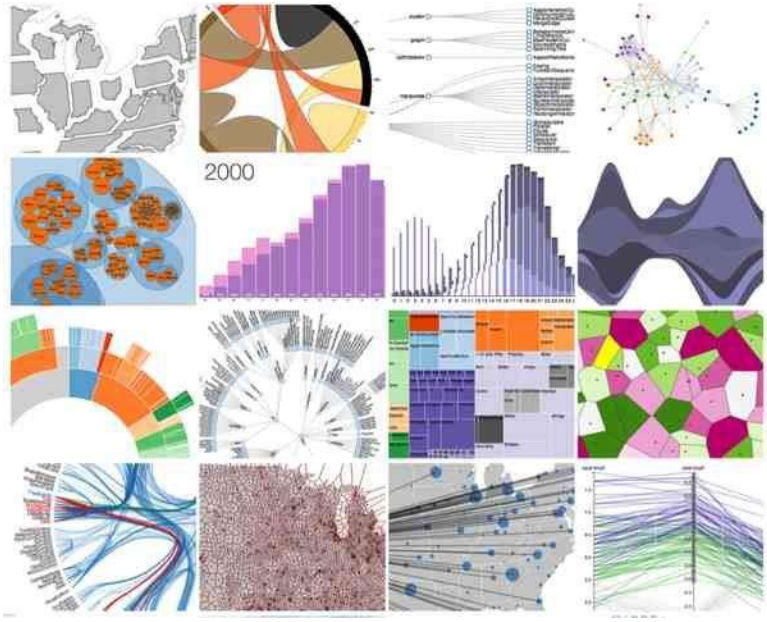

Figure 14: D3.js charts examples

- **Chart.js** is also an open-source JavaScript charting library for designers and developers that produces offline dynamic charts. The library as D3.js is compatible with all the modern browsers using HTML Canvas elements.
  Some of the most outstanding representations of the Chart.js library are:
  - Bar-funnel: bar funnel chart type.
  - Boxplot: boxplot and violin plot chart type.
  - Error-bars: diverse error bar variants of standard chart types.
  - Financial: financial chart types such as a candlestick.
  - Geo: geographic map chart types such as choropleth and bubble map.
  - Graph: graph chart types such as a force-directed graph.
  - Matrix: matrix chart type.
  - Pcp: parallel coordinates plot chart type.
  - Sankey: Sankey diagram chart type.
  - Smith: smith chart type.
  - Treemap: treemap chart type.
  - Ven: Venn and Euler chart type.

Figure 15: Chart.js charts examples

3.3.2 Immersive technologies

Today, immersive visualization techniques such as AR and virtual reality (VR) have become very important for the representation of information in the field of security.

The projection of three-dimensional visualizations often has problems with overlapping information that could be avoided by distributing the connections between nodes in the three dimensions. In particular, VR provides a natural environment for the display of 3D graphics that enhances the user's sense of feeling by allowing them to navigate and interact more intuitively between complex data structures.

The complete and diverse range of VR displays on the market, also called head-mounted displays (HMD) and head-coupled displays (HCD), distinguishes two main types of devices (whose main features and products are summarized in Table 1): Tethered virtual reality devices (physically connected to the PC) and mobile virtual reality devices (directly integrated with the smartphone).

Table 1. VR main features.

	VR tethered	VR mobile
Advantages	• Very accuracy motion tracking (external sensors (6-DoF)). • Powerful graphics • Room-scale tracking • Fully immersive experience	• Greater comfort and freedom of movement. • Additional hardware not required. • Interaction via Bluetooth or touch
Disadvantages	• Lower comfort. • Requires high performances PCs. • High prices	• Less accurate motion tracking. • Limited quality image.
Outstanding products	• Oculus Rift • HTC Vive	• Samsung Gear VR • Google Daydream • Google Cardboard

The most innovative proposals in this area now focus on VR-independent devices as an alternative to mobile displays but without the need for the mobile device itself (Oculus Go, Lenovo Mirage Solo, etc.). On the other hand, there are also mixed reality solutions that integrate both VR and AR capabilities in the same device (Samsung Odyssey, Microsoft Windows Mixed Reality, etc.).

Immersive visualization using VR can be applied, among many others, to the different existing representation techniques. In fact, GIS solutions such as Cesium are already compatible with VR devices such as Oculus Rift thanks to the development of plug-ins such as Cesium-VR/Cesium-Leap. Furthermore, Oculus has found a key strategic partner in Unity, a platform for the development of scenes and interactive 3D content. Through them it is possible to design immersive environments for the representation and navigation between complex graphics or three-dimensional data structures.

3.3.3 Augmented reality enhancing the perception and cognition

Augmented Reality (AR) allows superimposing computer-generated image into a physical real-world environment. It is able to enhance natural environments or situations and offer perceptually enriched experiences through elements that provide extra information of it is viewing in real-time. With the help of advanced AR technologies (e.g. adding computer vision, incorporating AR cameras into smartphone applications and object recognition) the information about the surrounding real world of the user becomes interactive and digitally manipulated.

There are several technologies used in augmented reality rendering, including optical projection systems, monitors, handheld devices, and display systems, which are worn on the human body. Most used are the following.

- **Eyeglasses**: Is a wearable gadget that creates AR content inside the area of the client's perspective. Users can see their physical surrounding similarly as on account of conventional

glasses. AR smart glasses like Google Glass superimpose additional content to whatever users see.

- **HUD (Head-up display):** is a transparent display that shows visual content within the scene of a user viewpoint. These solutions were originally developed for pilot training to help them learn how to fight in the air. Their innovation stack includes a projector, which transmits an image onto a display, combiner for capturing projected light, and video processor to produce visual information.
- **Handheld**: A Handheld display employs a small display that fits in a user's hand. All hand-held AR solutions thus far choose video see-through. Initially hand-held AR employed fiducial markers, and later GPS units and MEMS sensors such as digital compasses and six tiers of freedom accelerometer–gyroscope. Handheld display AR guarantees to be the first commercial fulfillment for AR technologies. The two main benefits of hand-held AR are the transportable nature of handheld devices and the ever-present nature of digital camera phones.

Thanks to the ease of use and portability offered by augmented reality on mobile devices, it was decided to make a prototype by developing a mobile application that integrates augmented reality as a technology for enhancing the perception and capabilities of end-users.

The key functionality of this app is to represent assets with different types of visualization. One through a 2D map and another through AR, which allows to activate the smartphone camera and visualize the position of the assets that appear or disappear as we move. This allows a quick, easy and intuitive way to recognize and reach the different georeferenced assets that are stored in the database.

For the development of a prototype, AR JavaScript library will be used for the AR services. This library will be integrated into the frontend to be developed with JavaScript framework Vue.js and Quasar for the style. Last, for hybrid (runnable in both Android and IOS operating systems) applications deployment it will be used Apache Cordova as can be seen in the following figure.

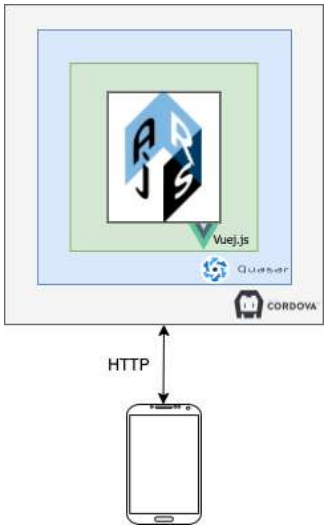

Figure 16: AR Application Scheme

3.3.4 Multi-dimensional Web GIS

Data processing and analysis is aimed at obtaining information from large complex data sets. Visualization of information provides general knowledge and contributes to increasing human understanding through visual exploration and efficient interaction with the information represented.

There is a need to understand the potential of the 3D visualization tools to address the challenges of communication, mutual understanding and mediation in situation-based planning. However, it should be noted that it is understood as a complementary tool for the analysis and elaboration of possible solutions.

The Geographic Information System (GIS), is initially focused on the representation and visualization of georeferenced data and spatial information in real-time which is still today the essential component of almost all C2IS systems.

These systems are in great demand, so the market offers an increasingly wide range of GIS tools that can be adapted to all types of applications, for example:

- commercial GIS solutions (ArcGIS, Luciad, etc.)
- open source (QGIS, gvSig)
- map servers (GeoServer, MapServer, etc.)
- proprietary (Google Maps, Bing Maps)
- free (OpenStreetMap) web mapping services
- frameworks for geospatial web applications (Cesium, OpenLayers, CartoDB, among others.)

The following is a brief description of the GIS proposal that best meets the requirements of the project.

Cesium

Cesium is a powerful open-source JavaScript library. It is used to create and design 2D/3D maps in web environments. What makes Cesium.js stand out is its interoperability with various map sources such as Google Maps, Microsoft Bing Maps or OpenStreetMap. As well as the multiplicity of supported data types such as terrain information, images, vector data or 3D models. As far as 3D models are concerned, Cesium has recently implemented 3D-Tiles, a very efficient open-source 3D data format. There is also a large number of add-ons that can be integrated into Cesium and give it extra functionality.

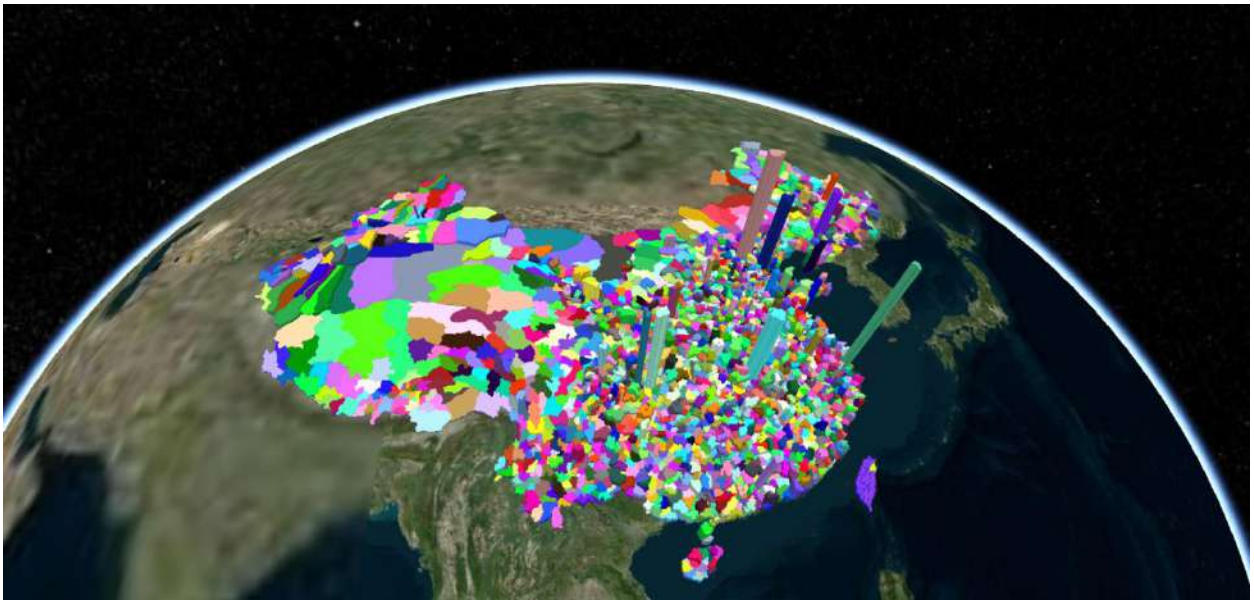

Figure 17: Cesium 3D visualization

Table 2. GIS features comparative.

GIS	Open Source	3D Visualization	3D -Multi-Layer	Web Integration
Luciad	NO	YES	YES	YES (RIA)
Google Maps	YES	NO	NO	YES
OpenStreetMap	YES	NO	NO	YES
Cesium	YES	YES	YES	YES

3.3.5 Multipurpose haptic devices

As the years progress, VR technology is advancing and providing its users with increasingly real-world experiences, although there is still a lack of devices that allow user interaction with VR objects. Gradually, different types of devices, both, for the game industry and general-purpose, are appearing on the market, although many of them are still under development.

For a better understanding of haptic technologies, they are described as technologies that provide tactile feedback to the user. These devices involve a two-way process between the user and a computer-generated device, giving the user the feeling that they are touching something that is not physically there. Therefore, haptics can be used as a simple simulation mode. For example, it could be simulated that the user can interact with the alerts and scroll through the map to visualize the information about them. One of the most interesting utilities is its remote control of the devices and provide a more realistic interaction with the images.

Haptic devices not only provide a way for the user to receive information from a computer/simulation, but the user provides information to the computer in response to a felt sensation, be it pressure, vibration or temperature change. Haptic interfaces can be divided into two main categories: tactile and force feedback.

Markets for technological innovation are growing rapidly, as demand for professional purposes is increasing. VR and haptic technologies have been considered particularly useful for very specific environments such as medical or police training. With these technologies it is possible to simulate risk situations in a controlled environment and to be able to analyse in advance the possible actions to be taken.

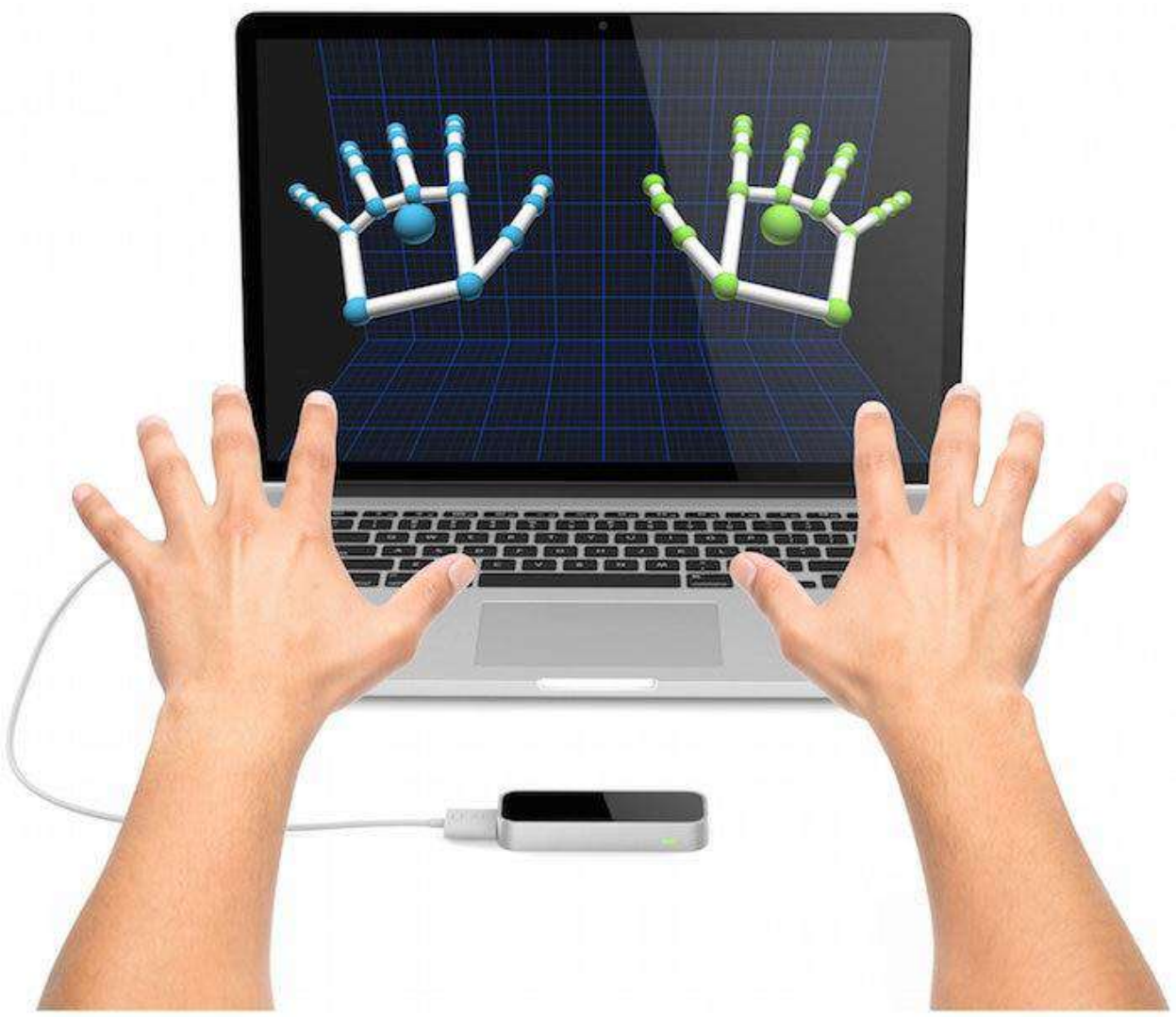

Figure 18: Leap Motion hands detection.

In the case of PREVISION, a LEAP Motion sensor has been chosen. It is a small device that provides a first approach to integration with many interesting features.

LEAP Motion allows the user to track and interact with hands and gestures in three dimensions with virtual and augmented reality. In this way the LEAP Motion controller allows the user to interact with the virtual elements of a simulation directly with their hands. The motion controller is a small and inexpensive USB-shaped device that can be placed in any space.

The Leap Motion controller is based on an infrared-based stereo camera. It works by illuminating a space close to the camera's infrared light, so it can detect the user's hands. The reported standard deviation of the position is less than one millimeter.

The importance of the LEAP motion controller and software is the compatibility of this simple device with VR/AR to provide motion tracking, including Oculus Rift. Other versions of the software are now capable of detecting pinch and grab gestures, providing smoother interaction with visual elements and the ability to interact and manipulate them.

3.4 Web HMI for multiple tools integration

For end-users to be able to interact with the platform and its different tools, the design and implementation of an interactive panel is a key part.

On the one hand, the design of the access interface to the PREVISION platform has been opted for the development of a WEB application since this allows the interaction of multiple users on the platform, without the need to install any type of application on the operating system.

On the other hand, the facilities provided by the development of a WEB HMI that allows to access all the services provided by the platform, practically from any type of device, whether it is a computer, smartphone or tablet, in addition to reducing to the maximum the hardware requirements for its execution, makes this solution the optimal one for the designed platform.

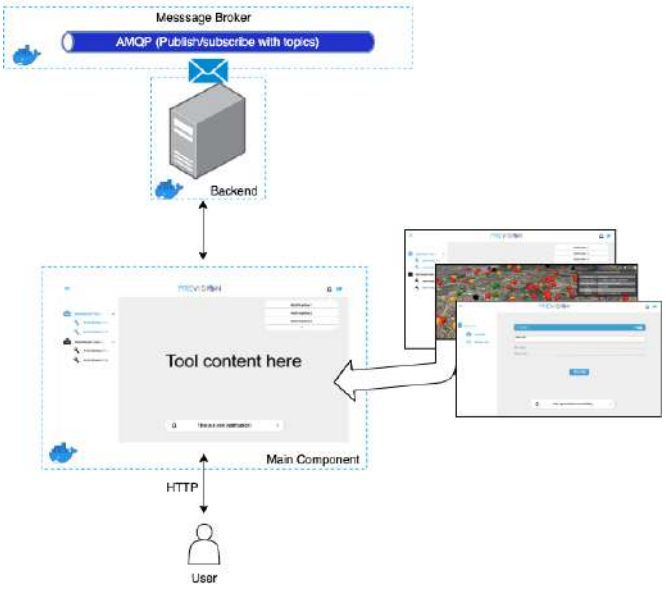

Figure 19: Web HMI Architecture

The visualization interface follows the scheme that can be seen in Figure 19. On the one hand, there is a Backend module that is subscribed to the Rabbit Broker to receive and send any type of message to the platform, and on the other hand there is a Frontend module that displays the WEB application with the different screens and services.

It is decided to Dockerize the different elements that compose the architecture. In this case there are three Docker containers: frontend, backend and message broker.

The aim of deploying a Docker architecture is to facilitate the deployment and modularity of the platform. In this way, it is possible to add or modify elements, as well as different PREVISION Tools, without altering the system.

3.4.1 Backend

For the Backend, a development in NodeJs that uses the Javascript language for the implementation of its multiple functions has been chosen. This multi-platform execution environment has been chosen, due to a large number of modules and services available, as well as the flexibility it provides when making server-side WEB applications.

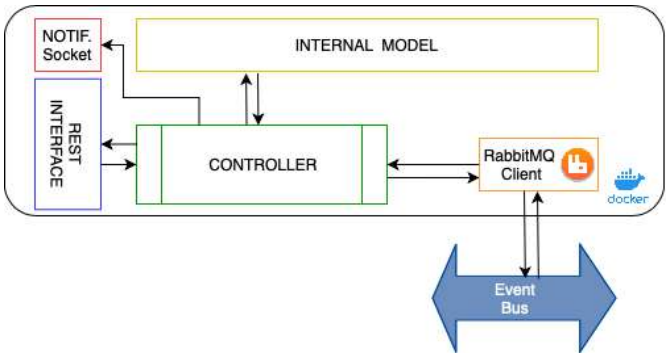

Figure 20: Backend Scheme

In Figure 20 we can see the different sub-components that have been designed. It should be noted that we have two interfaces for communication with the frontend module.

The first one is a REST interface provided through the Express application server that runs inside our module in NodeJs. In it, the different services that will be consumed in request/response form by the Frontend, are exposed. In this way any synchronous query made by the user will be sent to the platform. In case the answer to this request is asynchronous, either because of the duration of the request, or because the process is waiting for its execution in a queue, the user will receive the answer once the process is finished through the Notification Socket.

For the communication with the platform, a Rabbit client has been chosen that subscribes the different events broadcast by the analysis and management modules.

3.4.2 Frontend

Frontend module is also encapsulated within a docker container to allow fast deployment in new environments. For its development the VueJS Javascript framework is used which allows to develop interfaces with a very attractive and professional result. Moreover, its modular development capacity is perfect for designing the screens individually, allowing to integrate new modules when new functionalities are incorporated into the platform.

In order to make requests and receive responses from the backend, AXIOS library is used which allows us to establish request/response type connections and Socket.io to receive asynchronous notifications from the platform.

3.4.3 Mockups

Once the frontend/backend architecture has been designed it is important to provide end-users with an interface that is attractive and easy to use. For this reason, a series of mockups has been designed that allow the user to get an idea of what the main interface would look like.

The main idea is to provide an operator with the intuitive dashboard, built around the central window displaying content provided by the actually triggered service/tool. The remaining components of the display will be limited to the main navigation area (with menus and sub-menus, including tool settings)

and the upper bar, including customizable notifications box and standard login/logout button. The first mock-up of the PREVISION platform GUI is presented in Figure 21.

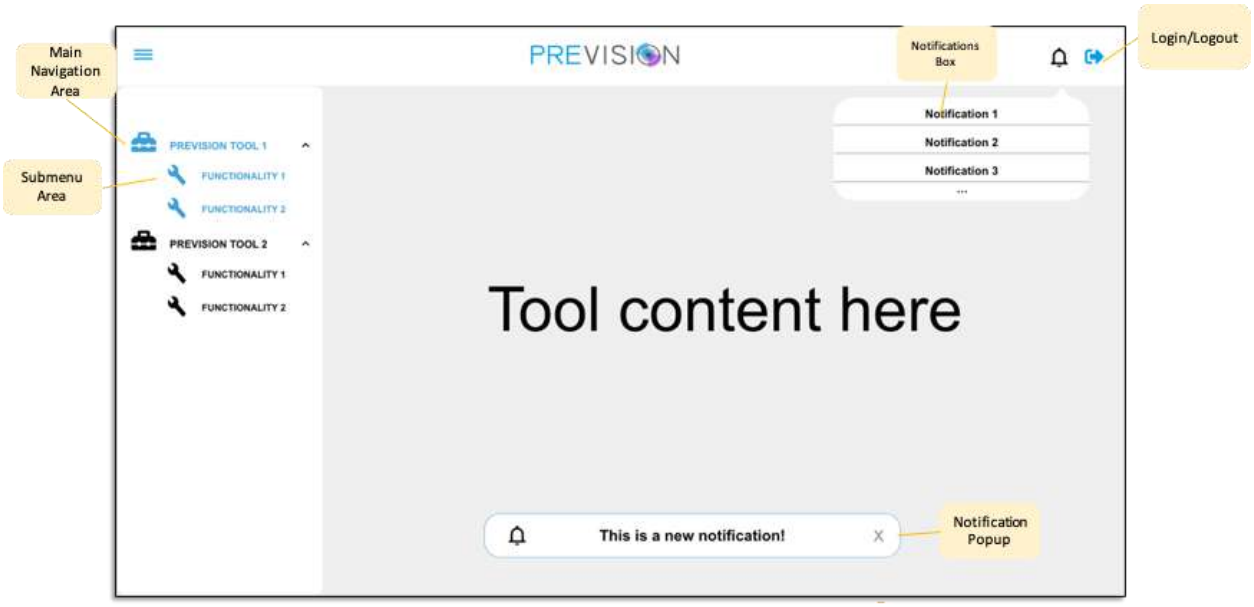

Figure 21: PREVISION GUI – Main Dashboard Layout

The main areas that can be seen in the layout are:

- Main navigation area: for this a tab-based navigation was chosen.
- Sub menu area: represented by a list of selectable tool functions.
- Component area: this area is assigned to the tool functions selected by the user – it can be further partitioned: an optional function-specific navigation area on the top and the main area, for the actual user interaction, below.
- Notifications box: specific notifications can be shown to the user here.

Other general functions are located at the top of the HMI:

- Login/Logout: allows the login/logout of a user.
- Notifications Panel: allows the notifying of logged users of certain system events.

Following the model of the main frame, an example component is designed. In this case it is a component that allows to upload files to the platform.

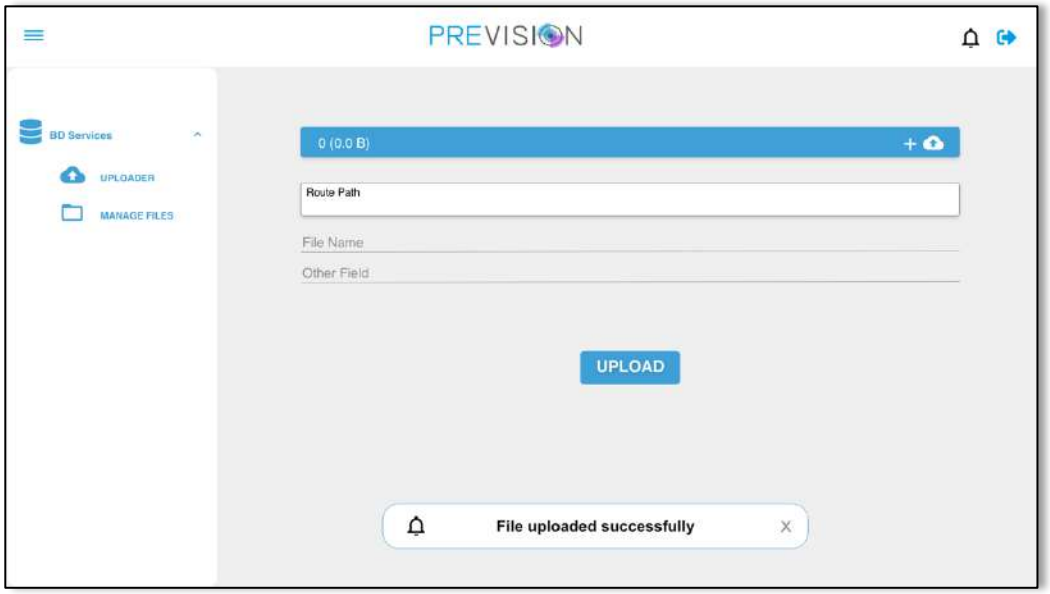

Figure 22: PREVISION GUI – Component Example

On the other hand, the access interface to the platform has been designed using a form through which users should authenticate themselves with their username and password.

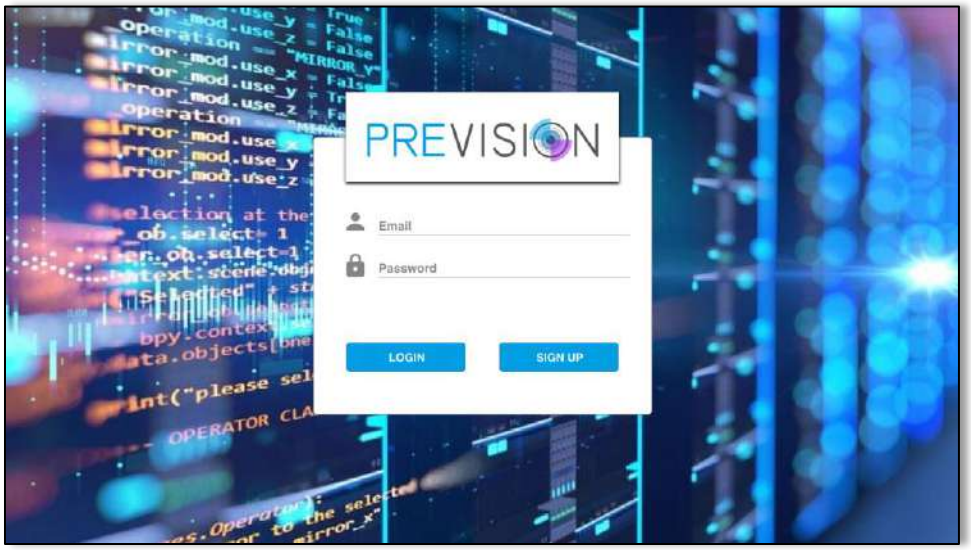

Figure 23: PREVISION GUI – Login Mockup

These mockups will be distributed to all partners to collect feedback in order to layout can be modified and adapted to take into account their comments. Once the final mock-ups are designed, each partner will be responsible for designing its interfaces to interact with its components.

4. Identification of radicalization and terrorist propaganda

4.1 General context

The Internet and social networks are increasingly becoming media of extremist propaganda. Extremists of all colours spread their ideologies and world views on homepages, in forums or chats, which are often contrary to the basic liberal democratic values of the European Union. It is not uncommon that violence is used against those of different faiths, those who think differently and members of social minorities. Especially in times of crisis such as the Corona pandemic, conspiracy theorists and radicals are increasingly popular.

There are many forms of violent radicalizations [1] such as ultra-right-side (associated with fascist, racialist/racist, ultra-nationalist motives), politico-religious (associated with a political reading of religion and the defense of a religious identity, whatever the religion is), ultra-left side (articulated around anti-capitalism and the transformation of a political system perceived as a generator of social inequalities), unique cause ones (e.g., environmental, animal rights, anti-abortionists, homophobic, anti-feminist, etc.).

There are also **different media** where radical people express themselves: social media, web 2.0, specific newspapers, chat boxes, forums in **different languages or dialects** (each media has also its specific form of expression e.g., tweets vs. newspapers).

The Internet and social media are so attractive to extremist groups for various reasons:

a. They represent an ideal opportunity for self-expression and communication.
b. They can reach millions of addressees around the world in a very short time.
c. They offer an ideal opportunity for networking like-minded people.
d. Social control, not only by the security authorities, but also by the social environment is made considerably more difficult.

The following objectives are pursued in the task:

1. Biographical research will be used to identify risk factors that can initiate or accelerate radicalization processes This involves an individual risk analysis or risk management.
2. A set of instruments is to be developed in order to support security authorities in identifying extremist propaganda on the Internet and classifying it in terms of its degree of danger. This concerns both extremist content on freely accessible Internet pages and content in closed chats.

Extremist Internet propaganda

The linguistic approach can play a central role in the assessment of the degree of radicalization and the potential danger of extremist groups.

Considering the question of how radicalization processes can progress in the context of computer-mediated communication and how people in forums, chat rooms and social networks are susceptible to extremist and radical content, different explanatory models and theories can be used. These include the Social Identity Model of Deindividuation Effects (SIDE-Model) and the Social Identity Approach. In this research area, the focus is particularly on "processes of group formation and group dynamics that can be triggered by certain characteristics of the Internet (such as the possibilities for anonymity)" [2].

The SIDE-Model assumes that when an individual's identity is predominant, the perceived group homogeneity (of the group relevant to the individual) decreases and the individual orientates himself or herself primarily towards his or her individual norms and values. In contrast, when an individual's social identity (feeling of belonging to a group) is predominant, the perceived group homogeneity increases and the individual then orientates him- or herself primarily to the group norms and values. Circumstances of computer-mediated communication such as anonymity and identifiability play an important role in this context. Anonymity reinforces the described processes and a low level of identifiability leads to a person's orientation towards his or her norms and values. With increasing identifiability, however, the orientation of the individual towards group norms increases [3]. In short, it assumes that "processes of social identity and identification with groups can lead to group-conform behaviour" [4].

The theory of social identity assumes that individuals strive for positive social identity and acquire it through membership of one or more groups and the emotional significance of this membership [5]. Group members gain or lose prestige through a comparison with other (relevant) groups, which serves to strengthen their own social identity.

The author proceeds from the following theses:

- Extremist propaganda on the Internet or in social media has an influence on individual radicalization processes.
- Hate speech or racist expressions and stigmatization of ethnic, religious and social minorities by extremist actors on the Internet encourage concrete acts of violence.
- If fantasies of violence are expressed, the risk of concrete violent action increases. Investigation approach!
- If communication in social media is stopped, the danger of concrete violent action increases. Investigation approach!

The degree of radicalization of Internet-based propaganda can be measured by various indicators. In addition to the degree of radicalization, the degree of dissemination and the target groups are also of analytical interest. Young people in particular are receptive to Internet-based propaganda.

Exemplary indicators for the degree of radicalization:

- The banning of a site by state authorities is an indicator of a high degree of radicalization.

- If a site calls for violence against people and/or objects, this is also an indicator of a high degree of radicalization.
- If a propaganda site explicitly calls on people to join extremist/terrorist groups, this is an indicator of a high degree of radicalization.
- If a propaganda site calls for sympathy with persons or groups who have been involved in politically motivated (violent) acts in the past.
- The more conspiratorial information is shared on the net (darknet), the higher the degree of radicalization of the acting actors.

4.2 Envisioned PREVISION tools and services to be adapted

There are several sub-targets to answer this challenge. One of them is to **build linguistic resources** (per language, per radicalization type) that will be useful for radicalization detections in the case they do not exist. Indeed, a very few resources are available on this topic and for that purpose, while they exist, for example, in the case of depression detection. The next one is to **detect the radicalization and its level**.

4.2.1 Building linguistic resources (per language, per radicalization type) that will be useful for radicalization detections

The requirements as defined by CNRS-IRIT are as follows:

- Can be adapted to various languages with little effort or rely on tools that already exist in many languages (e.g., word frequency analysis, word stemming, lemmatization, named entity recognition, etc.);
- Is useful to treat new types of radicalization (e.g., while dictionaries and other linguistic resources exist for sentiment analysis, aggressive texts, there are not necessary for any new type of radicalization);
- Need a human interaction: LEAs grade the extracted terms (non-radical, extremely radical, radical, both radical and non-radical, not known) ;
- Is a mean for LEA to build re-usable resources;
- Is a mean for other tools to be more efficient (e.g., ML models to detect radical content);
- Could serve for cross-lingual analysis and/or parallel lexicons;

The results of Task 4.2 and with the contribution of BTPI and CNRS, will be as follow (See also Figure 24):

- Workflow description (PREVISION components used and order)
- Implementation of the workflow
- Description on how to handle a new language and/or a new radicalization type (to be used by the LEAs themselves)
- Linguistic resources for some languages, some radicalization types

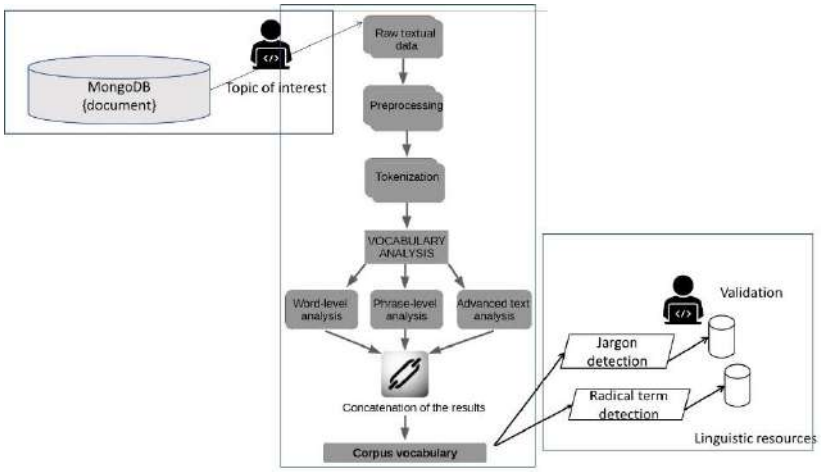

Figure 24: Building linguistic resources

With regard to the development, testing and evaluation, CNRS and BTPI will start with:

- Corpus of British radical environmentalist texts (1992-2003)[2], corpus articles of radical environmental UK groups, scanned text from a magazine, full text, open access. Need to be completed with "non-radical" related data of the same writing category (magazine in British English).
- MOROCO – Moldavian and Romanian dialectal corpus, formal texts; language detection. The dataset contains Moldavian and Romanian samples of text collected from the news domain. The samples belongs to six topics such as culture, finance, politics, science, sports, and technology. The dataset includes the training, validation, and testing subsets of samples, which are useful to perform the task such as discriminating the Moldavian and Romanian dialects.
- Dictionaries: various dictionaries or lexicons of radical terms for some types of radicalization exists (LEA can provide some as well e.g., Germany) that could serve as baseline or ground-truth when evaluating the developed methods.

The experimental design is as follows:

- Use a corpus associated with a form of radicalization, select terms and their possible polarity, evaluate using a ground-truth
- Same but with LEA feedback

4.2.2 Detect the radicalization and its level

The objective is to provide alarms and/or visual tools to help detecting possible radical content from a textual data set. The expected results are as follows (See also Figure 25):

[2] https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/PNK7AB

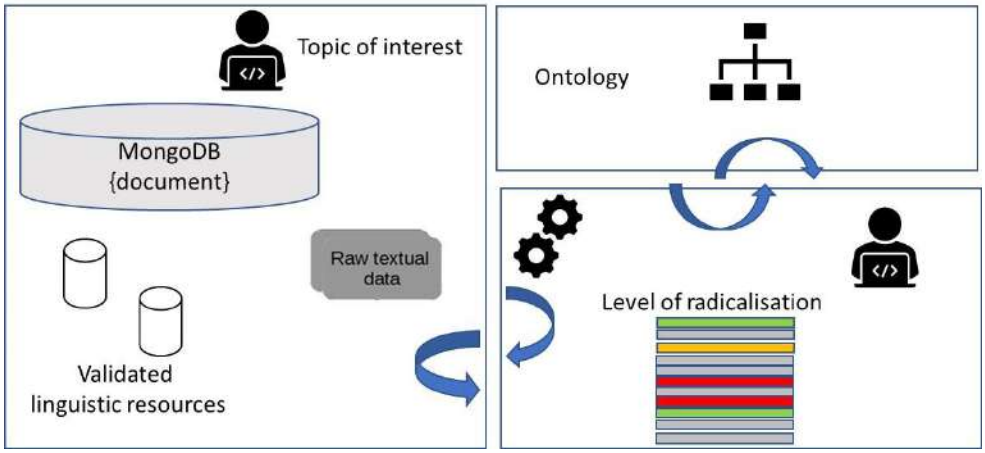

Figure 25: Detect radicalization and its level

- Workflow description (PREVISION components used and order)
- Implementation of the workflow
- The description on how to handle a new corpus for LEAs

The evaluation will be on environmental radicalization in English to start with (as a POC)

Other Data:

In order to achieve the goals outlined above, appropriate data will be required. The provision of data has proved difficult in the past. There were several reasons for this:

- The PREVISION project does not involve security authorities that explicitly deal with state security offences.
- In general, security authorities show little willingness to make police data from the field of extremism/terrorism available for research purposes.
- In the case of propaganda sites, there is a tendency for extremist groups to increasingly switch from public sites to closed forums. This is not insignificantly related to the fact that the pressure from the security authorities to prosecute has increased significantly in recent years, and the pressure on providers to delete extremist content promptly has increased.

In order to solve the problem of data samples, the following procedure was followed:

IfmPt has created a database with currently 58 biographies. These are the biographies of real people who have been exposed to politically motivated crimes in the past. The real names have been replaced by aliases. All contents concerning the individual persons were compiled from officially accessible sources. The database is continuously being expanded to include additional persons.

In addition, "normal biographies" are created to enable the system to distinguish between persons with an extremist biography and a normal biography. So far, 15 "normal biographies" have been created for control purposes.

IfmPt provides examples of extremist propaganda sites in English, which can be used as a starting point for further identification of radical propaganda on the net. These are exclusively pages that are publicly accessible.

Extremist content from so-called "secure messengers" (Telegram), social networks (Facebook) or blogging services (Twitter) cannot be made available because the data protection requirements are not met.

The model to be developed for classifying extremist propaganda should not be limited to freely accessible online material, but should also be able to be applied to content from closed forums and circles. It would even be desirable for the latter content to be made available for the development of the software.

In order to further expand the data basis, IfmPt is in contact with experts from security authorities and scientific institutions.

Technical requirements:

Web crawler: IfmPt creates a list of extremist propaganda sites as a starting point. With the help of these pages, the web crawler identifies further pages and stores them, taking the taxonomy into account. IfmPt creates an algorithm that can be used to classify the degree of radicalization of a page. The taxonomy is probably provided by the TENSOR project, the crawling tool probably by SPH.

When further developing taxonomies, it must be taken into account subcultural characteristics that are not unusual for extremist milieus.

Text analysis can also be extended by image analysis to identify extremist symbols. Symbols used can also provide clues to the degree of radicalization of an individual or an actor.

With the help of artificial intelligence, machine learning and deep learning, the algorithms can be further developed and optimized.

As a first step, it is planned to analyse extremist propaganda in English. Propaganda sites usually address the target groups in the national language, especially with regard to right-wing and left-wing propaganda. In a further step, the taxonomy and the models would then have to be adapted to country-specific characteristics. However, this is not part of the current assignment.

5. Protection of citizens in soft targets

5.1 General context of soft targets protection

The term „soft targets" typically describes vulnerable places, group of civilians or individuals that can be targeted by terrorists or criminals to maximize the psychological effect of their efforts. Vulnerability in this context indicates lack of strong safeguarding mechanisms and procedures applied on daily basis. In opposition, so-called "hard targets" are defended by security officers and other security mechanisms. Examples include military installations, government premises, airports and some critical infrastructures such as power stations. On the contrary, soft targets are characterized by the possibility of gathering large numbers of unarmed civilians in places to which access is not restricted or defended. Examples may be touristic places such as national monuments, hospitals, schools, public transportation, cultural sites [6].

As indicated in [6], development of some security frameworks covering a vast spectrum of soft target types and standardization efforts in this matter are challenging due to the different threats associated with different sites, specific characteristics of some soft targets, different design principles affecting possibility to defend them, etc. However, some publicly available guidelines, recommendations and checklist are existing regarding the security of soft targets. Many of them are focused on a specific area, e.g. transportation or education and published by national governance authorities, therefore the majority of them are scoped and apply to the national context.

One of the more comprehensive available guidelines is the document entitled "Basics of soft targets protection guidelines", developed and published by the Czech Soft Targets Protection Institute [7]. The document covers several aspects of the soft target's protection. Firstly, the common definition of soft targets is adopted: soft targets are places with high concentration of people and low degree of security against assault. The second aspect is categorization and prioritization (based on the severity) types of possible attacks: from bombing attacks (including suicide bombing, bombs delivered by mail and with the use of vehicles), chemical attacks (arson), gun/shooting attacks to assaults with use of knife and vehicle running into the target. The authors define also timeline for the incidents, differentiating three phases:

- before incident focusing on prevention and deterrence,
- during incident including detection and immediate response, and
- after incident, i.e. mitigation of impact.

The separate chapter and analysis is devoted to the security components in the context of soft targets protection. In general, three basic categories include security personnel (human component), electronic devices and mechanical devices. As the primary mean from electronic devices category camera surveillance systems and security alarm systems are listed. The other ones are Surveillance and Alarm Receiving Center (remote surveillance and control center), public emergency system (for broadcasting messages), x-ray scanners, different types of detectors (metal, explosives), entry and attendance control systems.

According to the guidelines, the essential part of soft targets protection is risk analysis and in particular security diagnostics of the potential soft targets. Several factors are considered as the important criteria for evaluation of desirability for a potential attacker and feasibility of security measures. They are: public accessibility of the given site, presence and availability of security personnel, mechanisms and police forces, concentration of unarmed civilians, presence of the media and symbolic value. Capability for self-protection of potential soft targets depends on three general factors: organizational structure, resources and funds deployable for security reasons and ability to risks identification [7].

5.2 Envisioned PREVISION tools and services to be adapted

Soft targets are significantly more difficult to monitor for security threats than places with controlled access. Therefore, we anticipate various PREVISION tools to be adapted in order to facilitate the mechanism for increasing situational awareness of all kind of actors dealing with soft target protection.

5.2.1 Characterisation of tools and services

As it has been emphasised by LEAs in the Soft Target Protection Use-case description, data feeds such as videos, social media, and geo-referencing of threats are the elements that need particular attention form the application perspective point of view. This application is intended to bring together information from WP2 (mainly data in video format) and WP3 (semantic processing and anomaly detection). In that light, the tools and services to be adapted in the soft target protection application could be the following:

- **Batch and Near-real-time Video and Image Analysis**, in order to access to the available CCTV (both streaming in real-time and snapshots) including capabilities for identification and re-identification of person
- **Crawling tools and Darknet** in order to run sentiment analysis on social media feeds
- **Web and Social Networks Data Analysis** in order to collect relevant background data that is related to the protected target
- **Smart Fusion and Incomplete Data Handling** in order to deal with structural heterogeneity and run search queries on multiple datasets simultaneously.

5.2.2 Tools orchestration

Some of the above-mentioned interaction will require some sort of tools and **services orchestration** in order to be available as the functionality of the PREVISION platform. In other words, the user requests coming from the graphical interface need to be translated to adequate query/call targeting the backend tools and services. It will require appropriate request transformation and services and tools orchestration (e.g. output obtained from one tool need to be passed to another, etc.).

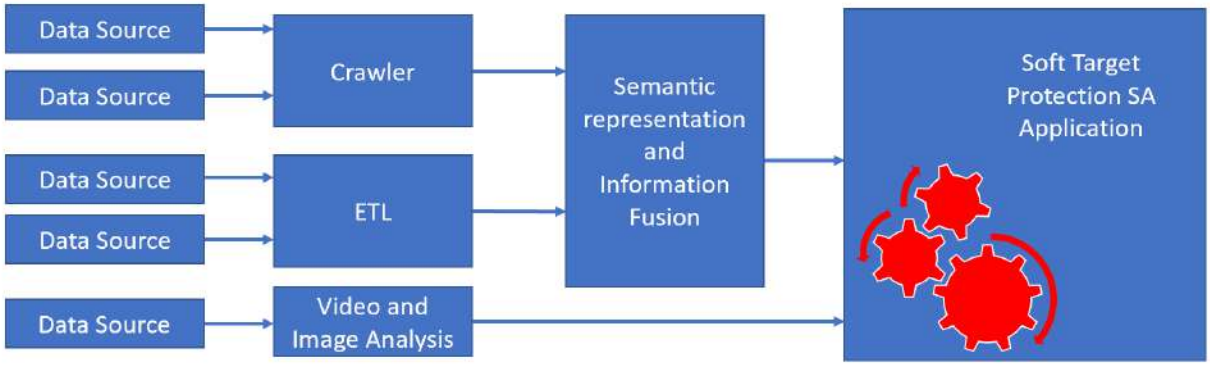

Figure 26: Information flow for Soft Target Protection SA Application.

We anticipate the following information flow (as depicted in Figure 26):

1. Indicate data sources to be ingested.
2. Configure probes, analysers, services, pre-processors, etc. to ingest the data from the indicated sources. Each tool (or group of tools) has its configuration, thus a specific procedure need to be developed in that regard.
3. Data obtained by different sources might be combined using various techniques (e.g. by specific entity, date, sematic meaning, etc).
4. At the same time data coming from other data sources can be Extracted Transformed and Loaded (ETL) into a database to serve additional data feed for information fusion.
5. At the end of the entire pipeline there is an application that makes the data and analysis results available for the user.

In order to implement the tools orchestration a good approach is adopting the Enterprise Integration Patterns (EAI) [21]. By including a component that implements these patterns in the architecture of PREVISION Platform on top of backend tools we can provide for the presentation layer functionalities such as:

- Message construction;
- Message filter;
- Message routing;
- Message transformation;

A more comprehensive list of features that can be obtained using EAI can be found in [22] .

Besides orchestration and transformation, this layer will ensure a weak coupling between the presentation layer and the backend applications.

5.2.3 Visualisation of results

At the same time, we may anticipate that we will adopt **various visualisation components** facilitating GIS-like capabilities, identification and tracking of LEAs officers and staff operating in the field. This method of visualisation is likely to ease the interaction with the system, as it allows the developer to pack a significant amount of information in a single screen.

6. Fight against illicit trafficking

6.1 Task & workgroups within T4.4 and connected WPs.

The workgroups can be organized within the different connected functionalities as depicted in Figure 27**.**

Figure 27: T4.4. Diagram of Tasks and Work Groups

The preliminary task was to gather heterogenous Databases to fill the data for the tests cases and to provide typologies of objects to test the solution. the preliminary task is to gather heterogenous databases in order to launch the first data tests and to provide the corresponding archaeological objects typologies

(Actors All)

The data providing has been permitted by the partnership agreement.

(Actors CNRS, PARCS, MCA)

6.1.1 Task 1: Central Homogeneous and Normalized Database Design

The primary task is to build and design a Central homogeneous and normalized DATABASE in WP2 and T4.4, a key feature for T4.4 but also for WP3.

1) The **structure** will be designed under Postgres, according to the nomenclature, the hierarchy of data and the attributes and relationships defined by the team's experts.
   (Actors PARCS TEAM)
2) The **Integration and standardization of the heterogeneous** dataset previously collected, will be processed by setting the structure of the ontology and the object description form.
   (Actors PARCS and CNRS)
3) The **ontology** will be developed under CIDOC-CRM ontology, the properties and classes will be monitored By PARCS and CNRS. See properties and classes list in the annex of T4.4 documentation.
   As discussed, the Ontology used in Art and Archaeology is the CIDOC CRM, developed by the ICOM association (25000 museums). In order to have the tool used by most Heritage professionals we have decided to use this ontology under the ISO norm 211127. The bridge between CIDOC-CRM and MAGNETO Ontologies will be developed aside, if necessary.
   (Actors PARCS CNRS)
4) The **Functionalities** will be provided by the whole UC5 members and will drive the design of the API.
   The **API** will be developed by PARCS.
5) The **Ergonomy** of the API will be processed under recommendations and past experiences of all members.
   The Ergonomy aims to develop the best user-friendly environment and intuitive handling for the interface of the user.
   (Actors all)
6) The **Typologies** selection and the data used are chosen after strict criteria.
   - Typologies, period and objects must be well studied and qualified by academics
   - The objects must have minimum quality documentation (exploitable pictures, texts)
   - If possible, the objects must have clear origin and provenance.
   - The selected typologies must have minimum of iteration within the corpus: a minimum of couple of hundreds of entries of objects of same typologies is required for test and big data analysis.

   (Actors ENSP, MCA, PARCS, CNRS)

7) The **Terminology** of the database represent the body of terms used to describe the object of study. It proposes categories and different technical descriptive words and jargons for an artefact. To allow this tool to be universally used by professional of different fields (scientific, LEA, Art professional) the T4.4 team will propose a set of linked various vocabularies that describes an object but with different level of expertise and approaches.
   The terminology will increase in precision as the user go deeper into the description and documentation of the object.

The “tunnel of description”, to help LEA to narrow down their descriptive process, will be developed in the functionality’s workgroup.
This to allow the solution to be used by experts as well as non-expert.
(Actors MCA, CNRS and ENSP)

6.1.2 Task 2: Queries & Responses

It is necessary to define the acronym, the synonym, the professional and jargon translation (primary French to English) and potential Key Word that are used to search for an object in the central database or with the smart browser (**see task 5**). The vocabulary of questions will be broad in order to hit the maximum chances of success, despite the presumed non-expertise level of the LEA.

The Team must define the questions, queries and events that will occur in the LEA investigations, this in order to create realistic search and test cases.

Along with the Queries, we also have to define the expected and theoretical responses. This will also design the verification protocol.

This is an important task of the Search engine. This will be used in search through **the Smart browser** developed in T 4.4 and WP3.

(Actors ALL)

6.1.3 Task 3: Image & Text Processing / Typology Matching.

1) The WP2 module will help the design of T4.4. In coordination with CERTH in WP2, SPH will contribute with their technologies and those of other members to develop an Image Processing technology tool for the use of PREVISION. We will coordinate our technicians with their in this task to adapt their technology to our specifications.
2) We will define the algorithm to process the image and text and the Object typology within the specific process. (PARCS and CNRS).
3) We will analyze potential technique (ex: bound boxes) and protocols to capture details within series, to help to identify the object by matching them to academic model and close typology. This task will be executed by PARCS and CNRS
   (Actors WP2. CNRS, PARCS, CERTH)

6.1.4 Task 4: Web Scanning and Crawling Tool

In coordination with SPH in WP2, we will contribute with our technologies and those of other members to develop an efficient web scanning tool for the use of PREVISION. We will coordinate our technicians with their in this task to adapt their technology to our specifications.

1) We will work with CNRS to define the algorithms to track and recognize objects on the web.

2) The team will set the criteria of the searches, the level of details of the scan, the nature and complexity of the reports, as the criteria for alerts.

(Actors WP2. CNRS, PARCS, SPH)

6.1.5 Task 5: Smart Browser.

As described in WP3, the team will design a Smart browser to search through various and heterogeneous sources (online or local), usually the digital database of Heritage group and institutions available on the web. The tool will also address the social networks, the merchant websites and, if possible, within the duration of the tender: the DarkWeb.

The tool will search identify, match possible illicit items hidden behind vague, missing or fraudulent data and description.

(Actors WP3, CNRS, IRRIT, PARCS)

6.1.6 Task 6: General Processing.

The important task of organizing the different services under automated functionalities will be carried by PARCS.

The attached diagram (Figure 28) shows the general kinematics and the sequence of the different functions.

The general architecture is organized into “services” which interact via standardized interfaces.

(Actor PARCS)

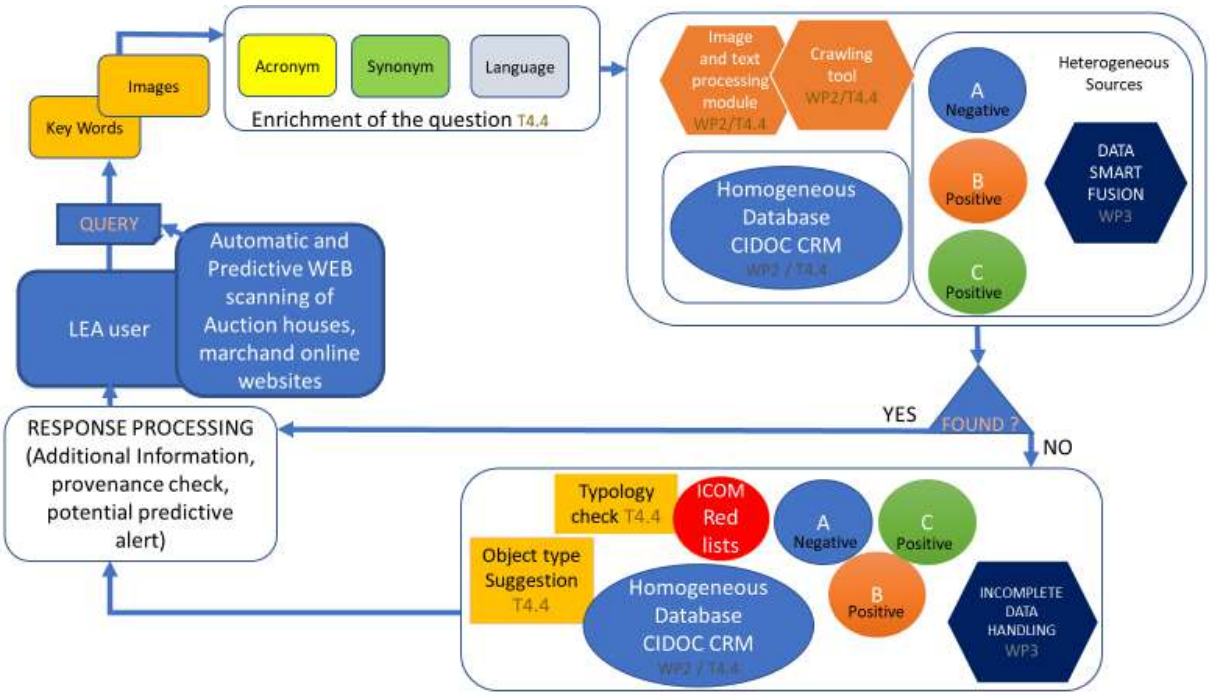

Figure 28: T4.4 General Processing

6.1.7 Task 7: Test Cases & Check Protocol. Evaluation

The test cases will be defined extensively to cover most of the expected and occurring events of these specific investigations.

The test case must address the maximum of issues and responses possible to be valid for the evaluation.

To confront the results of the queries we need to test them with well-defined and well-studied **referent database**.

We will run the specific pertinence tests and events check on few selected typologies within the databases.

So, we have, to create case analysis, to constitute the test environment, to do testing, operational certification and integration.

(Actors PARCS, ENSP, MCA, CNRS)

Evaluation will be led and processed by PARCS on specific platform. At the end, servers would be Hosted in SF data center and or in PREVISION platform.

6.1.8 Task 8: Integration to MAGNETO and other platforms

In the perspective of interoperability within PREVISION and the universal usage of our solution we plan to design bridges or interface to integrate our Ontology and services to Magneto's system.

The aim is to also anticipate its integration to other type of platform.

(Actors WP5. PARCS / UPV)

6.1.9 Task 9: Coordination with WP 6 & 7.

In order to obtain a certified application and reach durability and sustainability of the solution, the T4.4 Team will be in constant relation with WP6 and WP7 leaders to integrate and anticipate their questions, demands and observations on Budget, Business plan, dissemination plan, etc.

(Actor PARCS)

7. Trend characterization in cybercriminal activities

7.1 General context

Cybercrime affectsa wide range of cases and has quite complex taxonomy. Here in the PREVISION project, we are interested in two main types of it. It has been depicted in Figure 29.

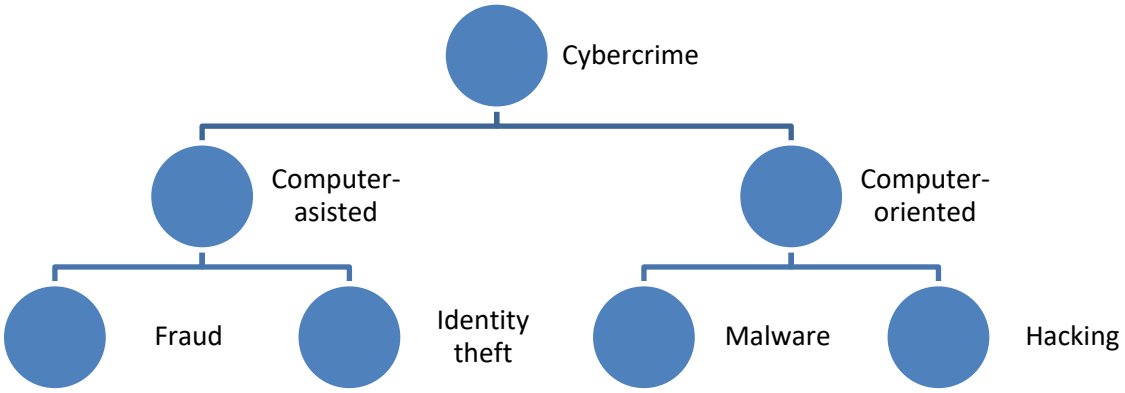

Figure 29: Types of cybercrime

It must be noted that, in this task, we are heavily focusing on **computer-oriented cybercrime activities**. We believe that the left strand of activities presented on the figure above can be addressed by other PREVISION tools for textual data processing and analysis. Nonetheless, in this section an overall characterisation of both is provided.

7.1 Computer-assisted cybercrime

In this section, the various types of computer-assisted fraud that are to be mitigated in PREVISION is described. The act of getting unlawful access to data utilizing special software or through human interactions or user actions is also being mentioned in this report. This definition of fraud encompasses various social engineering techniques and computer-assisted attacks. A particular type of fraud, namely, identity theft attack, is going to be addressed in more detail due to its high importance and potentially severe impacts of this type of attack.

7.1.1 Fraud

We concentrate on two major types of fraud:

- Social engineering: psychological manipulation aimed to trick victims into giving away sensitive information or making security mistakes. The most prominent examples of this type of fraud are:
    - Phishing: specially forged emails to trick users into sharing their confidential data or into paying for some fake service.
    - Phone scams: fake phone calls from banks, insurance companies, etc., in which the user is asked to provide their account number, PIN, or other sensitive information.

    - Fake Microsoft and other software companies' support: for instance, support scams that can provide criminals with access to the LEAs' databases.
- Rogue software: examples include fake antivirus – a malicious application that tries to pass for a legitimate antivirus, pretending to have found an infection in the system, or a phony system/connection performance booster. The objective is to force the user into downloading a malicious software masquerading as a security patch/update or as a utility that enhances the overall system performance.

7.1.2 Social engineering

The term "social engineering" is used for a wide range of malicious activities accomplished through human interactions. Social engineering attacks use psychological manipulation to trick users into breaking security practices or giving away sensitive information. This type of fraud is usually based on usage of the available technological stack without almost any custom software development (e.g., employing phishing emails, out-of-box malware, trojans, phone scams, legal remote access software such as TeamViewer, etc.). Sometimes a minimal development may be implied in cases where an attacker tries to imitate the design and layout of a legitimate webpage or email to trick the user; however, the main emphasis is put on psychologically tricking people into sharing their sensitive data.

Types of social engineering attacks include:

- Baiting: the victim is promised a reward, e.g., a lottery prize.
- Malware: the victim is tricked into believing that their device is infected with malware, and if they pay, the malware will be removed.
- Scareware: victims receive threats that their system is infected with viruses and suggest downloading fake antivirus software.

More sophisticated social engineering attacks may happen in several steps. The attackers first investigate the intended victim to gather all the necessary information. Then the attackers use this information to gain the victim's trust and manipulate him/her into making security mistakes, e.g., revealing sensitive information or providing access to essential resources. Social engineering attacks are especially dangerous because they rely on human error, rather than vulnerabilities in software and operating systems. Such mistakes are hard to identify and prevent. A significant amount of social engineering attacks are phishing emails, phone scams, and fake software support scams.

7.1.2.1 *Phishing*

Phishing means specially forged email messages that try to trick a user into clicking a link or open an attachment that usually aims at either stealing personal information of the user or at installing malicious software on the user's machine. This malicious software will subsequently trace the user's keyboard typing, take screenshots and forward them to the attacker, etc.

7.1.2.2 *Phone scams*

Phone scams are a prevalent form of cyber-attack and are probably the most widespread form of social engineering attack. Phone scams can be live or automated. Scammers usually pose as representatives of government agencies, technical, or financial companies or even try to impersonate victim's relatives.

They attempt to threaten people into disclosing sensitive information or bait them by promising a reward (e.g., winning a lottery prize).

7.1.2.3 *Fake software support*

Tech support scammers trick the users into believing that there is a severe problem with their computer. They may use a phone, email, online ads, pop-up windows on websites and listings in search results while posing as legitimate tech support (most often, Microsoft tech support). Using phone calls or online chats, the scammers trick the victim into allowing remote access to their computer.

7.1.3 Rogue software

This type of fraud involves more sophisticated software engineering approaches in comparison with social engineering. Namely, a separate software package is developed to pretend to perform some of the advertised functionality together with the more advanced skills of using JavaScript to promote or automatically download the rogue software to the user device.

7.1.3.1 *Fake antivirus and system performance boosters*

When surfing the Internet, people may encounter a pop-up message advertising a new advanced antivirus software or informing the user that their computer is infected or merely is underperforming due to the inadequately used resources. Such messages encourage the user to install a new antivirus software/performance booster, click the pop-up to update the existing software or to clean the system from the supposed virus. Once installed, they infect the system with malware and make it vulnerable for further attacks.

7.1.4 Identity theft

As has been described above, the types of fraud may have different targets as their final goals. In PREVISION, we concentrate on the most dangerous one, which is identity theft.

Identity theft occurs when an attacker steals a person's personal information that is subsequently used without prior permission. The personal information may include full name, billing address, email address, driver's license number, passport number, Social Security Number, bank account, any login credentials, credit card sensitive information, etc. This information may be further used to commit fraud or maybe just sold via the dark web to other cybersecurity criminals.

The identity theft may be done in a variety of ways; however, in PREVISION, we will mitigate the following most commonly used approaches for identity theft:

1. Social engineering employing phishing, phone scams, and fake software support.
2. Rogue software.
3. Insecure web connections.
4. Breaches in the organizational security policy.

Social engineering and rogue software have been described in detail in the previous section.

Insecure web connections allow the man-in-the-middle (MITM) attacks that provide all the network traffic exchanged between the victim machine and the remote hosts to the eavesdropper. For instance,

it may happen when a user connects to a public WiFi hotspot to send or get emails without the usage of the encrypted communication channels.

Breaches in the security policy usually occur due to the lack of proper authentication mechanisms and inadequate backup procedures.

The suggested defence against identity theft attacks takes into consideration all the approaches mentioned above and, as a result, it consists of several layers.

The first layer of defence includes the following items:

- An automatic intelligent solution that allows us to identify the phishing emails and immediately block them.
- Automatic blocking of the phone calls, based on the call-blocking and fraud detection applications (e.g., Hiya).
- Setting-up an anti-spoofing authentication protocol such as Domain-based Message Authentication, Reporting, and Conformance (DMARC), makes it impossible for the attacker to use the user's organization domain in email spoofing attacks.
- Control the information that is shared via the websites and social networks in order not to expose the sensitive data that can be subsequently maliciously exploited by the attacker.

The idea of this first layer is, on the one hand, to automatically block as many phishing emails and phone scams as possible and, on the other hand, to prevent the attacker from the easy access to the information that may significantly simplify the process of forging a phishing email that may bypass the automatic filters (e.g., by means domain spoofing). An attacker can provide additional credibility by using the publicly available sensitive data shared on websites and social networks during the scam phone call.

The second layer of defence deals with the appropriate securing of the connections:

- The VPN connection should always be used, and all machines should be automatically connected to the VPN-host during machine start-up, and no internet connection should be available until the VPN channel is up and running.
- HTTPS connections should be established whenever it is possible. This can be achieved by installing individual browser plugins on every user's machine that allows them to automatically check if the connection is performed via a secure channel.
- All emails within the organizational domain should be automatically sent only in an encrypted manner. All the emails outcoming from the organization should also be encrypted as much as possible.

The idea of the second layer is to mitigate the potential MITM identity theft attacks. This can be achieved by available software tools and security suites such as Cisco VPN routers and clients, available web browser extensions (e.g., "HTTPS Everywhere"), and encryption solutions for emails (e.g., encryption functionality of the MS Outlook software).

The third layer of the suggested defences consists of the following items:

- The operating systems should be configured in a proper way for automatic updates, as well as the appropriate antivirus and antimalware security suites that should be deployed on every system.
- All the applications on mobile and desktop devices should be timely updated together with antiviral databases.
- Use of anti-banner services to make Internet surfing more secure.
- Use of two-factor authentication.
- Regular offline backups of the critical organization databases and software.

The third layer aims at keeping the defence of the organization and users' machines up to date. It also includes the means for a possible rollback of the whole organizational systems in case of a successful attack, which is possible thanks to the daily regular offline backups that cannot be accessed by the attacker since they are made and stored entirely offline. Moreover, two-factor authentication makes it almost impossible for the attacker to get authenticated in the system even the MITM attack has been successfully implemented.

The fourth layer relates to the increase of the awareness of the users about the potential threats related to the identity theft attacks.

Traditionally, organizations are focused on technical aspects of cybersecurity, but dealing with such fraud often requires taking a human-centric approach to cybersecurity awareness. Most often, it is the user who is the weakest link in the security chain. Automatic computer defences are meaningless if people ignore safety standards.

To deal with such threats, the following policies must be implemented:

- Raise people's awareness by security training programs. When people understand how easy it is to be tricked or scammed by a social engineering attack, they are more likely to be suspicious of emails, phone calls, or other cyber-attack approaches. So periodic security training with final testing should be introduced.
- Avoid pirated software and torrenting.
- Strict data separation so that people don't have access to data or systems they don't require for their work. Sensitive data should be configured using special ACLs (Access Control Lists) for different levels of employees.

Effectiveness of the suggested policies is mostly based on prevention of potential security risks. Such measures as automatic blocking of phishing emails and phone scams, data encryption, encrypted channels for data transfer, different levels of access control, employee training, etc. To prevent data loss, backup policies are to be introduced. Operating system automatic updates together with dedicated network and software security packages help in detecting cyber threats or malware.

7.1.5 Envisioned PREVISION tools and services to be adapted

To counteract potential identity theft attempts, an intelligent system is going to be developed for PREVISION. Since most of the identity theft attacks are made via phishing emails, the system will concentrate on the email infrastructure defences.

It will be comprised of three major part:

1. phishing detection
2. automatic response generation
3. reporting about current phishing attempts via a chatbot/email

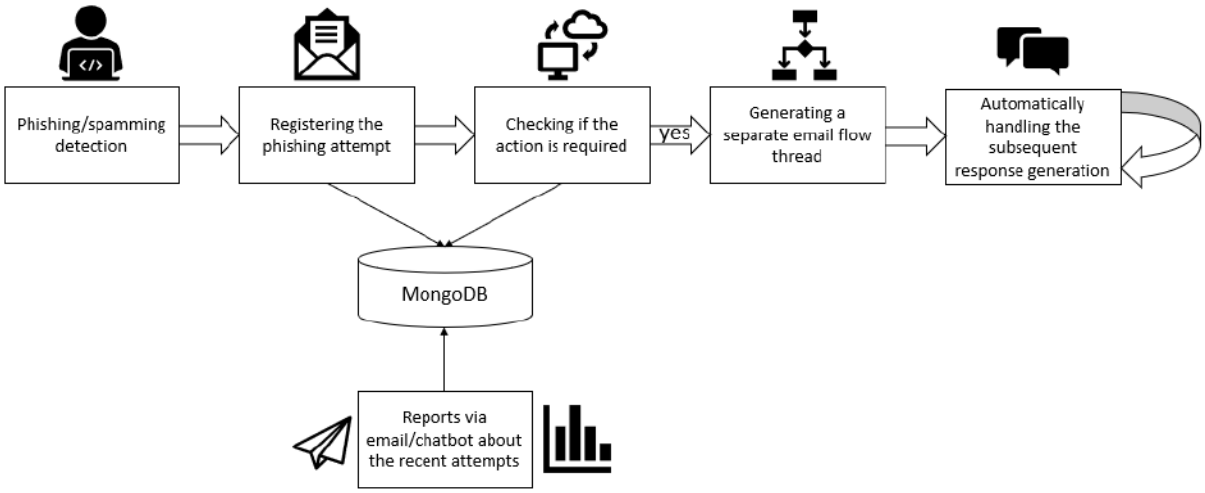

Figure 30: Overview of the information flow within the system, central database storage and the reporting of the incidents

Phishing detection is a complicated task, especially in the case of spear-phishing attacks. There have been numerous attempts of dealing with this problem [12]. Most of them concentrated on the identification of the relevant features of such emails which is usually time-consuming and inefficient in terms of classification accuracy. In our system we will switch from manual feature engineering to the more recent deep learning approach delegating the feature engineering task to the deep neural network (DNN). In particular, we will concentrate on two approaches:

- Universal language modeling utilizing transformer-based DNNs [13]
- Deep generative modeling based on Bayesian inference [18]

Both approaches are claimed to demonstrate promising results in the literature. It should be noted that we won't discard manual features at all since the system will rely on the online available phishing URLs databases, namely, PhishTank [15], OpenPhish [14] and Google Safe Browsing [16]. The match of the URL in the blacklist with the one in the email is a clear indication of the phishing attack and will be immediately classified as one. However, more difficult cases will be handled using the DNNs.

Another issue with classification into phishing versus non-phishing emails relates to the lack of the representative balanced datasets. To counteract this issue, we will use available fraudulent email datasets [19][20] for phishing emails together with the negative dataset of normal non-fraudulent emails of the balanced size.

The second step is an automatic response generation. This step aims at taking the time of the human phishers/spammers for reading the response email automatically generated by the system and potentially writing the reply. It will allow spending phishers' resources on targeting the artificial intelligent agents instead of the real persons exploiting their time and workload. It consists of the following stages:

- Topic modeling for emails and classification of the email in question within the particular topic
- Keywords extraction that the most relevant for the given email
- Conditional-response generation based on the previously classified topic and keywords
- Post-processing of the automatically generated result
- Sending the final response to the phisher

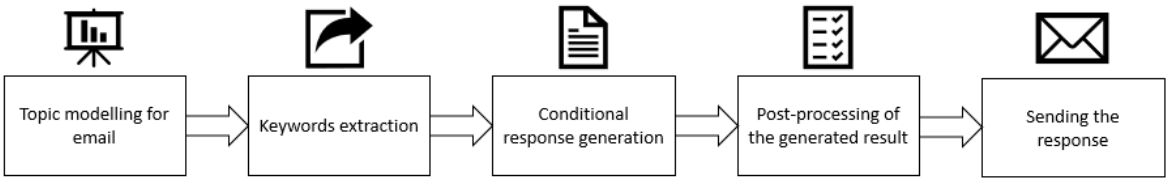

Figure 31: General concept of adapting PREVISION technology stack for (computer-oriented) cyber trends characterization

Topic modeling will be developed based on a generative probabilistic model exploiting the Latent Dirichlet Allocation (LDA) [8]. LDA represents a hierarchical Bayesian model where every email is going to be modeled as a mixture of the topics discovered within the whole email collection. Such an approach would allow us to update dynamically the topic categories as the systems evolve which should increase the precision of the clustering depending on the changes of the strategies used by the phishers/spammers.

Keywords extraction will be implemented employing two techniques:

- Calculating a ratio of term frequency divided by inverse document frequency (TF-IDF) [9]
- Extracting named entities from the email through modern transformer-based universal language models [10][11]

This step will allow us to get a short email summary putting the significant weights on the most relevant and the terms specific for the email.

The most important stage of conditional response generation will be based on the deep universal transformer-based language model CTRL [17]. This model allows a controlled response generation, i.e.,

in addition to the standard prompt used in any generative model numerous content-specific control codes can be applied to customize the generation output. This customization will be based on both previous stages, namely, topic modeling and keywords extraction.

Finally, it will be possible to configure the notification about the current phishing attempt utilizing communication with a Telegram chatbot. This would allow us to deal with the situational awareness of the users under the phishing attack.

Figure 32: An example of chatbot dialog with the user about the detected incident handling

It should be noted that one of the steps of the system's operational flow concerns the integration of the system into the email infrastructure used in the organization under protection. This step depends on a particular email server and software in use together with the exploited email protocols (e.g., SMTP, POP3, IMAP, etc.). The system aims at integration with the specific email infrastructure in question and it will be separately specified when more details will be available about this part of the project.

7.2 Computer-oriented cybercrime

Network and information security are now one of the most pressing problems of security, as they affect the economy, citizens, and whole societies directly. It is universally observed that the number of successful attacks on information, civilians, even seemingly secure financial systems and most importantly critical infrastructures is still growing. The ongoing growth in the complexity of malicious software has rendered the long-established solutions for cyberattack detection inadequate. Specifically, at any time novel malware emerges, the conventional security systems prove inept until the signatures are brought up to date. Moreover, the immense growth of Internet users generated a plethora of adversaries who abuse the Internet's framework. Currently, the number of successful attacks on various IT systems is still growing. Some attacks are performed by malicious users acting alone, some are carefully arranged invasions performed by groups of compromised machines.

As use case scenario, we consider the problem of malware and botnet detection by means of analysing the data in form of network flows. The problem of malware and botnets is related to the situation where massive numbers of computers have been infected, through an array of methods, like e-mail attachments, drive-by downloads etc. The infected machines from a kind of network controlled by a bot-master, who issues a fire signal to cause malicious activities. The problem of botnets is highly relevant, as these can be responsible for DoS attacks, spam, sharing or stealing data, fraudulent clicks and many other.

The general technology stack comprised of PREVSION tool and services has been depicted in Figure 33. The idea for detecting the cyber-oriented threats is to analyse the traffic coming from the IT network by means of various dedicated micro-services. The data can be ingested into the system using various probes and network sniffers. There are various options such as frobe[3], nProbe[4] or CICFlowMeter[5] that are capable of emitting flow statistics out of the raw traffic data. Another option is to harvest the date from NetFlow-enabled[6] network devices via the LogStash[7].

[3] http://manpages.ubuntu.com/manpages/bionic/man8/fprobe.8.html
[4] https://www.ntop.org/products/netflow/nprobe/
[5] http://netflowmeter.ca/
[6] https://support.solarwinds.com/SuccessCenter/s/article/Devices-that-support-NetFlow
[7] https://www.elastic.co/logstash

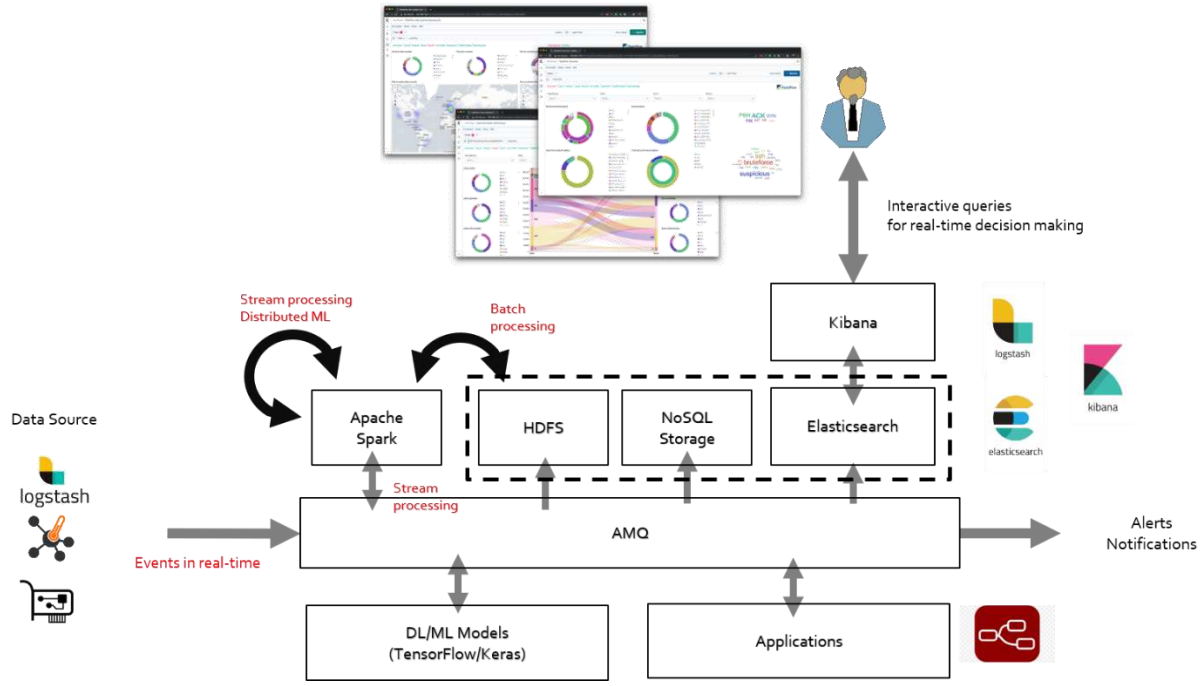

Figure 33: General concept of adapting PREVISION technology stack for (computer-oriented) cyber trends characterization

An important element sitting in the middle of the architecture diagram is the AMQ (Advanced Messaging Que) system (e.g. RabbitMQ or Apache Kafka). It allows retaining incoming data in the distributed queueing system for further processing. This approach also eliminates tide coupling between processing services, as they are not explicitly contacted together and thus can be easily interchanged or replicated without impact on the entire system. In that regard, such tools as Elasticsearch and Kibana[8] can be connected to the system and build advanced dashboards and visualisations.

[8] https://www.elastic.co/

8. Summary and conclusions

The main goal of this document was to provide the first description, in particular: motivation, context, state-of-the-art and initial design of the tools to be developed in WP4. These tools will aim at enhancing operational and situational awareness of LEAs protecting civilians in different scenarios. To ease the document development and its future refinement, the report has been structured as WP4 work is decomposed into tasks.

Firstly, after consultations with the PREVISION end-users (LEAs) and practitioners, existing solutions currently in use have been identified and current gaps and opportunities analysed.

Subsequent chapters of the report are focused on particular applications related to WP4 tasks. General context of the operational and situational awareness tools has been described together with the aspects addressing the key expectations of our end-users, such as multi-dimensional data interaction, visualisation of big data characterized by heterogeneity and integration of multiple tools with the use of a web HMI. Similarly, applications to be developed for identification of radicalization and propaganda, protection of citizens in soft targets and for cybercrime scenarios have been described, including identification of envisioned tools and services that can be adapted to address the end-users' expectations.

For the tool related to fight against illicit trafficking scenario, a detailed workplan divided into sub-tasks and corresponding tool functionalities have been defined.

The refined release of the current deliverable is scheduled at M17 and will be reported as D4.2.

9. References

[1] Types De Radicalisation (in French), available online: https://info-radical.org/fr/types-de-radicalisation/

[2] Translated by the author from German; Boehnke, K., Ö. Odag & A. Leiser, 2015: Neue Medien und politischer Extremismus im Jugendalter: Die Bedeutung von Internet und Social Media für jugendliche Hinwendungs- und Radikalisierungsprozesse. Stand der Forschung und zentrale Erkenntnisse themenrelevanter Forschungsdisziplinen aus ausgewählten Länder, at: https://www.dji.de/fileadmin/user_upload/bibs2018/Boehnke_Odag_Leiser_2015_Neue_Medien_Extremismus.pdf [last access: 27.05.2020], p. 48.

[3] Hogrefe 2020, at: https://portal.hogrefe.com/dorsch/side-modell/ [last access: 27.05.2020].

[4] Translated by the author from German, Boehnke, K., Ö. Odag & A. Leiser, 2015: Neue Medien und politischer Extremismus im Jugendalter: Die Bedeutung von Internet und Social Media für jugendliche Hinwendungs- und Radikalisierungsprozesse. Stand der Forschung und zentrale Erkenntnisse themenrelevanter Forschungsdisziplinen aus ausgewählten Länder, at: https://www.dji.de/fileadmin/user_upload/bibs2018/Boehnke_Odag_Leiser_2015_Neue_Medien_Extremismus.pdf [last access: 27.05.2020], p. 48.

[5] Skrobanek, J., 2004: Regionale Identifikation, negative Stereotypisierung und Eigengruppenbevorzugung. Wiesbaden: Springer, p. 73.

[6] Karlos, V., Larcher, M., & Solomos, G. (2018). Review on Soft target/Public space protection guidance.

[7] Kalvach, Z. (2016). Basics of Soft Targets Protection—Guidelines. *Soft Targets Protection Institute*, 45.

[8] Blei, D. M., Ng, A. Y. & Jordan, M. I. (2003). Latent Dirichlet allocation. *J. Mach. Learn. Res.*, 3, 993--1022. doi: http://dx.doi.org/10.1162/jmlr.2003.3.4-5.993

[9] Jones, K.S. (1972). A statistical interpretation of term specificity and its application in retrieval, *Journal of Documentation*, *volume 28*.

[10] Devlin, Jacob, et al. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv:1810.04805 [cs], arXiv.org, http://arxiv.org/abs/1810.04805.

[11] Burtsev M. et al. (2018). DeepPavlov: Open-Source Library for Dialogue Systems. *Proceedings of ACL*, *System Demonstrations*.

[12] Das, Avisha, et al. (2019). SOK: A Comprehensive Reexamination of Phishing Research from the Security Perspective. arXiv:1911.00953 [cs], arXiv.org, http://arxiv.org/abs/1911.00953.

[13] Vaswani, Ashish, et al. (2017). Attention Is All You Need. arXiv:1706.03762 [cs], arXiv.org, http://arxiv.org/abs/1706.03762.

[14] OpenPhish. (2020). OpenPhish - Phishing Intelligence. https://openphish.com/.

[15] PhishTank. (2020). PhishTank | Join the fight against phishing. https://www.phishtank.com/.

[16] Google Sage Browsing. (2020). APIs to access the Google Safe Browsing lists of unsafe web resources. https://developers.google.com/safe-browsing/.

[17] Keskar, Nitish Shirish, et al. (2019) CTRL: A Conditional Transformer Language Model for Controllable Generation. arXiv:1909.05858 [cs], arXiv.org, http://arxiv.org/abs/1909.05858.

[18] Chien, J.-T. (2019). Deep Bayesian Natural Language Processing. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics: Tutorial Abstracts*, Association for Computational Linguistics, 25--30.

[19] Radev, D. (2008), CLAIR collection of fraud email, ACL Data and Code Repository, ADCR2008T001, http://aclweb.org/aclwiki

[20] PhishCorpus dataset. (2020). https://monkey.org/~jose/phishing/

[21] Enterprise Integration Patterns - https://www.enterpriseintegrationpatterns.com/

[22] https://www.enterpriseintegrationpatterns.com/patterns/messaging/