Grant number: 883286
Project duration: Sep 2020 – Feb 2023
Project Coordinator: Joe Gorman, SINTEF

Horizon 2020: Secure societies
SU-INFRA02-2019
Security for smart and safe cities, including for public spaces
Project Type: Innovation Action

http://www.impetus-project.eu

IMPETUS Project Deliverable: D3.4

Tool Development Final Report

Dissemination Status: Public

Editor: Guillem Garcia, Insikt Intelligence

Authors: Guillem Garcia (Insikt Intelligence), Kéren Saint-Hilaire (IMT), Alexia Comte (UdN), Sandra Cardoso (INS), Ron Ofer (CyberSixgill), Michelangelo Ceci (CINI), Joe Levy (CINEDIT), Bruno Bonomini (CPAD), Paolo Mocellin (UPAD), Thomas de Groot (THA), Sébastien Courtin (UdN), Axelle Cadière (UdN) and Sandrine Bayle (IMT)

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

About IMPETUS

IMPETUS (Intelligent Management of Processes, Ethics and Technology for Urban Safety) is a Horizon 2020 Research and Innovation project that provides city authorities with new means to improve the security of public spaces in smart cities, and so help protect citizens. It delivers an advanced, technology-based solution that helps operational personnel, based on data gathered from multiple sources, to work closely with each other and with state-of-the art tools to detect threats and make well-informed decisions about how to deal with them.

IMPETUS provides a solution that brings together:

- *Technology*: leverage the power of Internet of Things, Artificial Intelligence and Big Data to provide powerful tools that help operational personnel manage physical and cyber security in smart cities.
- *Ethics*: Balance potentially conflicting needs to collect, transform, and share large amounts of data with the imperative of ensuring protection of data privacy and respect for other ethical concerns - all in the context of ensuring benefits to society.
- *Processes*: Define the steps that operational personnel must take, and the assessments they need to make, for effective decision making and coordination - fully aligned with their individual context and the powerful support offered by the technology.

Technological results are complemented by a set of *practitioner's guides* providing guidelines, documentation, and training materials in the areas of operations, ethical/legal issues and cybersecurity.

IMPETUS places great emphasis on taking full and proper account of ethical and legal issues. This is reflected in the way project work is carried out, the nature of the project's results and the restrictions imposed on their use, and the inclusion of external advisors on these issues in project management.

The cities of Oslo (Norway) and Padova (Italy) have been selected as the site of practical trials of the IMPETUS solution during the project lifetime, but the longer-term goal is to achieve adoption much more widely.

The work is carried out by a consortium of 17 partners from 11 different EU Member States and Associated Countries. It brings together 5 research institutions, 7 specialist industrial and SME companies, 3 NGOs and 2 local government authorities (the trial sites). The consortium is complemented by the Community of Safe and Secure Cities (COSSEC) – a group established by the project to provide feedback on the IMPETUS solution as it is being developed and tested.

The project started in September 2020 with a planned duration of 30 months.

For more information

Project web site: https://www.impetus-project.eu/
Project Coordinator: Joe Gorman, SINTEF: joe.gorman@sintef.no
Dissemination Manager: Harald Drager, TIEMS: khdrager@online.no

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Executive Summary

The main objective of this deliverable is to describe all the tools that are being used in IMPETUS and that conform the Cybersecurity Framework, or the Technologies that are being integrated and producing the project's toolkit.

The tools descriptions' goal is to explain how these tools work and which kind of functionalities each of them covers. The description includes:

- The responsible partner
- The tool technological readiness level to understand their maturity in terms of product ready of being implemented and used.
- The data sources to know what the required inputs are of each one of the tools.
- The tool data delivery.
- The tool functionalities, as a brief explanation of the expected outcomes of each one and what are the end-users needs that they cover.

An extensive cooperation with all WP3 participants has been conducted during the writing of this document, when we were mainly focused on understanding and defining the technological and ethical aspects of the technologies for public safety in smart city environments. The definition of the tools is serving a close collaboration with other WPs, that focus on Requirements, Ethics, Technical Development and Platform Integration, Operations and Platform Evaluation, as all of these areas are being developed in an iterative manner. The understanding of how the tools work and what they offer is of key importance for many participants from different backgrounds and perspectives in the project.

During the definition process of the Secure Smart City Tool Development module, we have been able to know first-hand the maturity and readiness of each of the tools that the Cybersecurity Framework will contain. We had the chance to do an assessment of what is possible to obtain by using each of the different tools and how to generate more value to secure smart cities when combining them. Also, a joint work with WP2 to understand how to integrate the tools and how these tools will interact between them has been done, and that uses the information presented in this document as a basis.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Table of Contents

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

List of Tables

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

List of Figures

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

List of Abbreviations

Abbreviation	Explanation
WP	Work Package
TRL	Technology Readiness Level
SaaS	Software as a Service
UI	User Interface
GUI	Graphical User Interface
ATP	Adenosine TriPhosphate
LAN	Local Area Network
SOC	Security Operations Center
SMD	Social Media Detection tool
FD	Firearm Detector tool
BD	Bacteria Detector tool
CTI	Cyber Threat Intelligence tool
PTI	Urban Anomaly Detector tool
WMS	Workload Monitoring System tool
EO	Evacuation Optimiser tool
CTDR	Cyber Threat Detection and Response tool
SIEM	Security Information and Event Management

Table 1: List of Abbreviations

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

1 About this deliverable

1.1 Intended readership/users

The primary audience of the deliverable is the project consortium.

A secondary target audience are the stakeholders of the project: stakeholders of public safety solutions for smart cities, possible end-users, and third-party technology providers.

1.2 Why would I want to read this deliverable?

The main goal of this document is to provide a complete and understandable high-level description of the nine tools that will constitute the IMPETUS integrated toolkit, covering the complete physical and cybersecurity value chain.

The tools are one of the central actives for the development of IMPETUS' Technologies or Cybersecurity Framework, which combined with the Operational and Ethical Framework will form the decision-making solution. A complete and understandable description of them will serve as a common source of information across the project participants and stakeholders. This document aims to serve as a common ground for further discussions, both in Work Package 3 (WP3) and the rest of the project activities.

The description of the tools is key for further and parallel processes in the project, including but not limited to the definition and integration of the platform, the user interface definition, the platform's usability, the requirements (platform, solution, frameworks, ethics) of the project and the project's piloting and validation of the platform.

1.3 Structure

This document consists of a set of sections (Sections 2-11) with the same structure, each of them describing one of the tools in the Cybersecurity Framework for IMPETUS. The structure of the sections was made based on a working document (spreadsheet). Each partner was provided with an identical template to fill the spreadsheet, letting the document authors to fill and format the sections.

Each of these chapters contains:

A high-level, understandable, and clear tool description and, if existing, the tool internal or commercial name.

- The Technology Readiness Level (TRL) by the end of the project determined for the tool, based on the status of each independent tool.
- A description of the data sources that each tool consumes and a brief explanation of how the data produced by the tool is or can be delivered and/or presented to the end-user.
- A structured description of each of the tool's functionalities, that includes the following fields:

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

- Description
- User roles
- Maturity
- Interface (service, methods, data structures)

1.4 Other deliverables that may be of interest

The work done in WP3 and the necessary adaptations of the tools for its integration in the IMPETUS platform, and to be aligned with the scope of the project is fully dependent on the rest of the project's results. However, this is a list of the deliverables that are strongly connected to WP3 and this deliverable:

- D1.2 - Requirements for public safety solutions

  Describes the requirements and needs for the overall solution and the impact of each one of the tools.

- D4.1 - Data analytics and ingestion-based access control initial report

  Describes the data ingestion module to ensure the use of each tool explained here.

- D9.1 - Exploitation strategy and plan

  Explains the exploitation strategy of the whole project, with specific plans for each tool.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

2 Secure Smart City Module Development

2.1 Initial Modular proposal

The initial modular proposal was defined considering the three following work package tasks:

- **Task 3.1 (Detection solutions).** The goal of this task was to combine proactive tools for threat identification, using data sources such as social media, surveillance camera feeds and other similar sensor measurements. Such data sources can be processed with artificial intelligence and statistics techniques (e.g., machine learning and natural language processing), to discover and warn urban safety operators about potential underlying threats affecting their systems.

- **Task 3.2 (Simulation & analysis solutions).** The goal of this second task was to combine the proactive information collected and processed in Task 3.1, with a second batch of tools providing threat analysis and decision-making features, such as automated processing of physical security threats and technical vulnerabilities affecting computational and networked resources, to simulate the consequences of each threat.

- **Task 3.3 (Intervention solutions).** The goal of this third task was to close the loop with reactive tools capable of facing the list of identified threats with the appropriate countermeasures. This includes the combination of different approaches targeting the affected systems, to autonomously evaluate changes in the environment and adapt them prior responding to the threats. For instance, by combining the monitoring process with the selection of response plans, both at physical and cyber layers.

Hence, the initial driving idea was to separate tools into three main modules used to:

- identify threats (T3.1),
- explore the consequences if those identified threats occur (T3.2), and
- prepare an optimized response to neutralize the threats (T3.3).

Following this modularity, nine tools were grouped according to the schema shown in Table 3.

Tool	Task
Social Media Detection (SMD)	Task 3.1: Detection solutions
Firearm Detector (FD)	
Bacteria Detector (BD)	
Breach & Attack Simulation (BAS)	Task 3.2: Simulation & analysis solutions
Cyber Threat Intelligence (CTI)	
Urban Anomaly Detector (UAD)	
Workload Monitoring System (WMS)	Task 3.3: Intervention solutions
Evacuation Optimiser (EO)	
Cyber Threat Detection and Response Tool (CTDR)	

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Table 2. Initial group of tools.

2.2 Decisions made in Deliverable 3.1.

After some individual analysis on the tools, during the WP3 initial discussions, the modular approach (Detection solutions, Simulation & analysis solutions, and Intervention solutions) seemed to be out of date.

While the tools could be superficially categorised by the modules defined, the type of activities, data sources and pace of progress didn't seem to be aligned among them. It made more sense to define the IMPETUS toolkit as a set of nine independent tools.

The platform definition processes that are being carry out in parallel to WP3 in WP1 (functional, non-functional and platform requirements) and WP2 (architecture and technical requirements) are even identifying integration scenarios where some tools that are prone to be sharing data amid themselves were not even included in the same module at the beginning. One clear example for this is the direct collaboration between cybersecurity tools: Cyber Threat Intelligence, Breach & Attack Simulation, and Cyber Threat Detection and Response Tool.

With regards to each preliminary defined module, the planned completion and/or integration of each tool (or some of its functionalities) is not always aligned with the rest of the tools integrated in that module. In terms of internal Work Package and Task organization, this situation made us consider that it was more efficient to track progress for each of the tools independently rather than keeping track of the completion of each of the initial tasks defined (T3.1 to T3.3).

The new WP subdivision is as shown in Table 3.

Current task	Initial task	Tool name (in DoA)	Updated tool/task name
T3.1.1	T3.1: Detection solutions	Social Media Detection (SMD)	Social Media Detection (SMD)
T3.1.2		Weapon and Face Detection (WFD)	Firearm Detector (FD)
T3.1.3		Biochemical Risk Detection (BRD)	Biological Risk Detection (BRD) And lastly to: Bacteria Detector (BD)
T3.1.4	-	-	WP3 Management and strategic planning

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

T3.2.1	T3.2: Simulation & analysis solutions	Breach & Attack Simulation (BAS)	Breach & Attack Simulation (BAS)
T3.2.2		Cyber Threat Intelligence (CTI)	Cyber Threat Intelligence (CTI)
T3.2.3		Physical Threat Intelligence (PTI)	Urban Anomaly Detector (UAD)
T3.3.1	T3.3: Intervention solutions	Human Computer Interaction (HCI)	Workload Monitoring System (WMS)
T3.3.2		Physical Threat Response Optimization (PTRO)	Evacuation Optimiser (EO)
T3.3.3		Cyber Threat Response Optimization (CTRO)	Cyber Threat Detection and Response Tool (CTDR)

Table 3. New WP subdivision.

During the definition process of this toolkit, all WP3 participants also reached the consensus of stablishing the term 'tool' as the correct way to refer to all the different toolkit parts, since terms like 'module', 'component' or 'solution' were leading to confusion during the working sessions. Hence, the agreement was to always use the more appropriate term ('tool').

Moreover, some tool/subtask names were modified to be described more accurately:

- **Firearm Detector tool** was previously called Weapon and Face Detection. Currently, the Firearm Detector Tool constantly anonymizes all the biometric data unless an anomaly is detected (more details in section 4).
- **Bacteria Detector tool** was previously called Biochemical Risk Detection since this tool is a microbial air analyser rather than a biochemical agent's detector.
- **Cyber Threat Detection and Response tool**, previously called Cyber Threat Response Optimization tool has been renamed to reflect better the tool purpose.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

3 TRL Evolution

3.1 Introduction

3.1.1 Definition

TRL assessments was refined using the EARTO extended TRL definition[1]:

The European Association of Research and Technology Organisations (EARTO) has developed an extended definition of the Technology Readiness Level (TRL) by the end of the project framework, which provides a more detailed description of the different stages of technological development. The extended TRL definition is designed to capture the complexity and multidimensional nature of innovation processes in research and technology organizations. Let's go through each level of the extended TRL definition:

Figure 1. Technology Readiness Levels.

- TRL 0 - Basic Principles Observed: This level represents the earliest stage of technology development. It involves the identification of new scientific principles or discoveries that could have technological applications. At this stage, there is no experimental evidence or practical application of the principles.

- TRL 1 - Technology Concept Formulated: In this stage, the basic scientific principles or discoveries are translated into a technological concept. The concept is described in terms of its functionality, purpose, and potential applications. This level focuses on defining the concept's scope and its alignment with broader research and innovation goals.

- TRL 2 - Technology Concept Validated: TRL 2 involves conducting theoretical and experimental validation of the technology concept. This validation may include basic analytical and laboratory studies to demonstrate the feasibility of the concept. The objective is to establish a scientific basis for further development.

[1] https://www.earto.eu/wp-content/uploads/The_TRL_Scale_as_a_R_I_Policy_Tool_-_EARTO_Recommendations_-_Final.pdf

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

- TRL 3 - Proof of Concept Demonstrated: At this stage, the technological concept is validated through experimental demonstrations. These demonstrations aim to prove the practical feasibility of the concept. The focus is on evaluating the key functionalities and identifying potential technical challenges that need to be addressed.

- TRL 4 - Technology Validated in Lab: TRL 4 involves further validation of the technology in a laboratory environment. The technology is subjected to a series of tests and experiments to assess its performance, reliability, and limitations. The objective is to gain a better understanding of the technology's capabilities and refine its design.

- TRL 5 - Technology Validated in Relevant Environment: This level focuses on validating the technology in a relevant environment that closely simulates its intended application. The technology is tested under realistic conditions to evaluate its performance, interactions with the surrounding systems, and potential operational issues.

- TRL 6 - Technology Demonstrated in Relevant Environment: At TRL 6, the technology is demonstrated in a relevant operational environment. This demonstration involves pilot-scale testing or prototype deployment to assess the technology's functionality, performance, and potential impact on the intended application. The objective is to gather data and feedback for further improvements.

- TRL 7 - System Prototype Demonstrated in Operational Environment: This level represents the demonstration of a system-level prototype in an operational environment. The prototype is tested under representative conditions to validate its integration with other systems and to assess its performance, reliability, maintainability, and safety aspects.

- TRL 8 - System Complete and Qualified: TRL 8 involves the completion of the development process, leading to a fully functional system that has undergone qualification. The system is ready for deployment or commercialization, with all necessary regulatory and safety requirements fulfilled. This stage focuses on finalizing the documentation, standardization, and quality assurance processes.

- TRL 9 - Actual System Proven in Operational Environment: The highest level of TRL, TRL 9, represents the actual deployment or implementation of the technology or system in its intended operational environment. The technology is used in real-world applications, and its performance, effectiveness, and impact are evaluated based on operational data and user feedback.

The extended TRL definition provides a more nuanced understanding of the various stages of technological development, from the early conceptualization to the real-world implementation of innovative technologies. It helps research and technology organizations assess the maturity and readiness of their projects and enables better communication and collaboration between stakeholders involved in the innovation process.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

3.1.2 Advantages

We have chosen to adopt the extended version of the Technology Readiness Level (TRL) by the end of the project framework to define the maturity of our tools for two key reasons:

1. Provides specific criteria that simplify assessment: The extended TRL definition offers more detailed and specific criteria for each level, which simplifies the process of assessing tool maturity. With clear guidelines and criteria, it becomes easier to evaluate the progress and readiness of our tools at each stage of development. This specificity helps in setting measurable goals and tracking the advancements made by our tools. By having well-defined criteria, we can effectively communicate the level of maturity and progress to stakeholders, investors, and collaborators. The increased granularity in the extended TRL framework allows for a more accurate evaluation of our tools' capabilities and identifies areas that need further development or improvement.

2. Assessments are now more uniform across tools: The extended TRL framework ensures that assessments of tool maturity are conducted in a more uniform and consistent manner. By providing detailed definitions and criteria for each level, it reduces subjectivity and ambiguity in the assessment process. This uniformity is particularly valuable when comparing multiple tools or evaluating tools developed by different teams or organizations. It allows for a fair and objective comparison, enabling us to make informed decisions about resource allocation, prioritization, and investment. Moreover, a standardized assessment approach facilitates benchmarking against industry standards and best practices, leading to improved alignment with market expectations.

In summary, by adopting the extended TRL framework, we benefit from specific criteria that simplify the assessment process and enable a more uniform and consistent evaluation of tool maturity. This approach provides clarity, transparency, and comparability, allowing us to make informed decisions, prioritize resources effectively, and communicate the progress of our tools to various stakeholders.

3.2 Evolution of the tools TRL statuses

The following table shows the evolution of each tool during the project in terms of the TRL level (at the beginning, during year 2 and at the end of the project.

It also includes a forecast for the following 5 years, regarding the expected evolution in their TRL levels.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

TECHNOLOGICAL RESULTS	EVOLUTION DURING PROJECT			PREVISIONS AFTER THE PROJECT		
	Beginning	Y2	End	1 year	3 years	5 years
Firearm Detector	5 (6)	6	7	8	9	9
Bacteria Detector	5	6	6 (7)(7)	7	TBD	TBD
Urban Anomaly Detector	4 (5)	5	6 (7)(7)	7	8	9
Social Media Detection	6	7	8 (7)	9	9	9
Cyber Threat Intelligence	7 (6)	8	9 (7)(7)	9	9	9
Cyber threat Detection and Response	5	5	6 (7)	7	8	9
Workload Monitoring System	6	6	7 (6)	8	9	9
IMPETUS platform	4	5	6	7	9	9
Evacuation Optimiser	4 (6)	5	6 (7)	7	9	9

Table 4. Evolution of the TRL during and after the project.

The numbers shown in brackets indicate the TRL stated in the DoA (blue colour) and in the PPR (purple colour), where these differ from the assessment here (shown in black).

Some of the tools already started with close to commercial tools (6 or 7):

- Social Media Detection Tool
- Cyber Threat Intelligence
- Workload Monitoring System

The evolution of the Technology Readiness Level (TRL) by the end of the project from 6 at the beginning of the project, to 7 at year 2, and finally reaching 8-9 by the end of the project signifies significant progress and advancement in the development of the technology. They have followed a similar evolution during the project:

a) TRL 6 at the beginning of the project: Starting the project at TRL 6 implies that a system-level prototype has been demonstrated in an operational environment. This indicates that the technology has already undergone testing and validation in relevant conditions, showcasing its functionality and performance. Consequently, the project is in an advanced stage, with a tangible prototype ready for further refinement and optimization.
b) TRL 7 at year 2: Reaching TRL 7 within two years demonstrates substantial progress in the project. At this stage, the technology has been demonstrated in a relevant operational environment, providing valuable insights into its integration, performance, reliability, maintainability, and safety aspects.
c) TRL 8-9 by the end of the project: Reaching TRL 8-9 signifies that the technology is complete, qualified, and ready for deployment or commercialization. The consequences of this high TRL achievement include: Market readiness, scalability and commercialization, real-world impact.

Overall, the evolution of these tools indicates substantial progress, increased confidence, optimization, stakeholder engagement, market readiness, scalability, and real-world impact. This TRL

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

evolution represents the successful development, validation, and deployment of a technology that has the potential to make a significant difference in its target domain.

There is a second group of tools that started at a lower level of maturity (4 or 5):

- Firearm Detector
- Bacteria Detector
- Urban Anomaly Detector
- Cyber Threat Detection and Response
- Evacuation Optimiser

The evolution of the Technology Readiness Level (TRL) by the end of the project from 4-5 at the beginning of the project, to 5-6 at year 2, and finally reaching 6 by the end of the project signifies progressive development and increasing maturity of the technology. They have followed similar evolution during the project:

a) TRL 4-5 at the beginning of the project: Starting the project at TRL 4-5 implies that the technology has undergone initial validation in laboratory environments. It suggests that the technology concept has been demonstrated and basic analytical and experimental studies have been conducted to assess its feasibility.
b) TRL 5-6 at year 2: Reaching TRL 5-6 within two years indicates significant progress and advancement in the project. At this stage, the technology has been further validated in relevant environments, demonstrating its functionality and performance under realistic conditions. The consequences of achieving TRL 5-6 include: Increased confidence, iterative improvement and stakeholder engagement.
c) TRL 6 by the end of the project: Reaching TRL 6 by the end of the project signifies a significant milestone in technology development. It suggests that a system-level prototype has been demonstrated in an operational environment, further validating the technology's capabilities. The consequences of achieving TRL 6 include: Proof of operational feasibility, optimization and fine-tuning, and potential for further development.

In summary, the evolution from TRL 4-5 to 5-6 at year 2, and ultimately reaching TRL 6 by the end of the project, signifies progress, increased confidence, iterative improvement, stakeholder engagement, proof of operational feasibility, optimization, and potential for further development. This TRL evolution suggests that the technology has advanced from the early stages of feasibility validation to a more mature stage, positioning it for potential deployment and commercialization.

All these tools plan to reach TRL level 9 in a maximum of 5 years.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4 Social Media Detection tool (SMD)

4.1 Basic information

The Social Media Detection tool, which commercial name is Insikt Spotlight, is a unique platform, which collects and analyses massive amounts of online public data to help Law Enforcement and Investigative Professionals detect specific written content, powered by Artificial Intelligence methods, Data Mining, Text Mining with Natural Language Processing, Deep Learning, Big Data Analysis, and Social Network Analysis in order to leverage cutting edge algorithms to surface hidden insights and cut through the noise to effectively neutralise and prevent terror, crime and threats affecting cities.

Responsible Partner	**Insikt Intelligence – INS**
Tool internal/commercial name	Insikt Spotlight - https://www.insiktintelligence.com/our-solutions/spotlight-osint/
Technology Readiness Level (TRL) by the end of the project	TRL 8
Data Sources	Social Media Detection will make use of data (text and metadata) from social media -Twitter- and comments sections from local newspapers in Oslo - document.no, reset.no, vg.no and dagbladet.no and Padova – mattinopadova. After the raw data is collected, the data is anonymised or pseudonymised accordingly. If a user/client does not have the permits to recover the raw data, then the system will encrypt the data and will not have a deanonymization key giving him/her no option to deanonymize the data.
Tool data delivery	SaaS and UI.
Description	https://www.impetus-project.eu/index.php/impetus-outputs/the-impetus-solution/14-solutions-eng/40-social-media-detection
Manual	https://www.impetus-project.eu/images/TOOLS_PDFS/Manuals/SMD_User_Manual_47z67g1z3a6q4.pdf

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4.2 Tool Functionalities

4.2.1 Online data acquisition

Description	The data is automatically acquired, given a frequency specified for each project. The sources of the data are: Online Sources from social media: • Twitter • YouTube • TikTok Online Sources from Local Press: • document.no • reset.no • vg.no • dagbladet.no
User roles	Analyst - Investigator who looks for potential threats that are published, organized, promoted or enhanced in online social media and local newspapers.
Maturity	Ready - Scraper of local online newspapers of Padova and Oslo have been developed for the project.
Interface (service, methods, data structures)	By creating a project, the user selects which are the sources that want to include in the investigation. The tool Integrates scrapers adapted to each source. All the acquired data are showed in the dashboard as encrypted or anonymised data and within the different analysis.

Table 5. SMD features.

i. Linguistic feature identification

Description	Seven different methodologies of Natural Language Processing [NLP, i.e., AI applied to text] are applied in 5 languages [English, Italian, Norwegian, French and Arabic] to analyse the content of the public online text from different perspectives: Concepts extraction, Key Ideas extraction, Topic classification, Hate Speech detection, Entities extraction, Hashtag detection and Sentiment analysis.
User roles	Analyst - Investigator who looks for potential threats that are published, organised, promoted or enhanced in online social media and local newspapers.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Maturity	Ready - New list of topics could be customised to improve the analysis if cities are interested. Norwegian NLP has been added for the project.
Interface (service, methods, data structures)	Computational linguistic methodologies: * Post tagging, tokenization and stopword removal for text processing. * Classification of topics based on cosine distances in 300-dimensional word embeddings. * Detection of linguistic patterns to extract the key ideas and concepts of text. * Network analysis applied to describe interactions between users. Dashboard with graphs for user-friendly visualization of the results: bar plots, 2-D graphs, word clouds, tables and network graphs. Word embeddings allow to convert text to numerical vector, which is key for computational linguistic methods. SMD integrates the tokenizers based on XLMRoberta cross-lingual tokenizers and the aligned word embeddings trained in Wikipedia in English, Italian, Norwegian, French and Arabic.

Table 6. SMD text analysis features.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

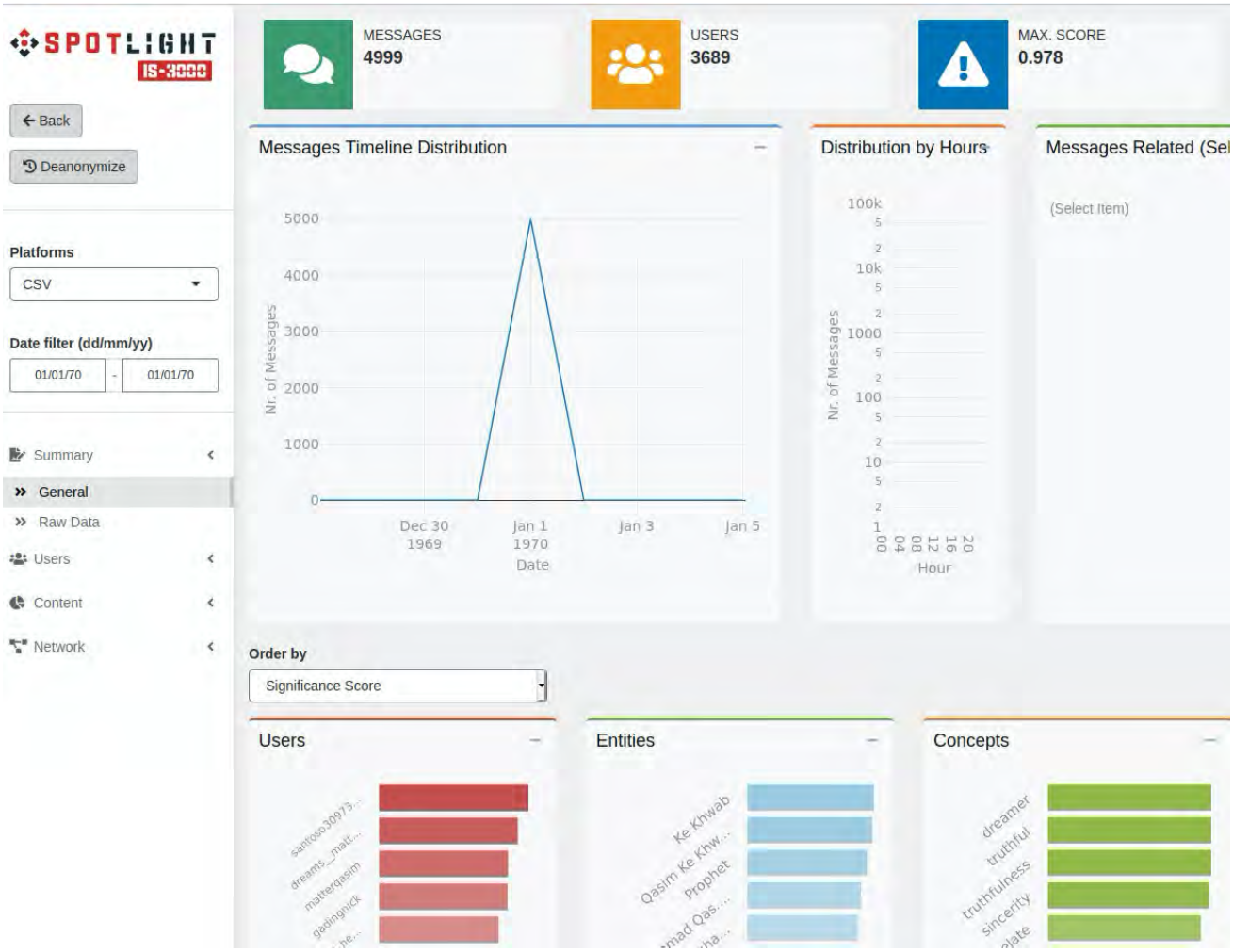


Figure 2. Social Media Detection (SMD) main dashboard.

ii. Hate Speech Detection

Description	The texts extracted from the social media and the local newspapers are analysed to evaluate the degree of hate speech they have.
User roles	Analyst - Investigator who looks for hate speech in online social media and local newspapers.
Maturity	Ready - These models can be improved along the project with new data sets of the cities.
Interface (service, methods, data structures)	Supervised machine language models trained from XLMRoberta cross-lingual pre-trained model. Transfer learning methods applied to multilingual text classifiers of hate speech content.

Table 7. SMD hate speech features.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

iii. Real-Time Environment

Description	The functionalities of Insikt Spotlight have been developed to run in a low processing time in all the workflow.
User roles	Analyst - Investigator who looks for potential threats that are published, organised, promoted, or enhanced in online social media and local newspapers.
Maturity	Ready - Low processing time in all the workflow
Interface (service, methods, data structures)	Low processing time in all the workflow. Near-real time.

Table 8. SMD real-time analysis.

iv. Deanonymization

Description	Functionality that allows to recover pseudonymised personal data by authorised parties. This is only available for the users that have this option active and the raw data has been pseudonymised. In the case of the anonymised data, this feature is not available.
User roles	Authorised analyst - Authorised parties for accessing to personal data.
Maturity	In progress - We are working in the implementation of this functionality and its integration in the dashboard.
Interface (service, methods, data structures)	The dashboard shows pseudonymised data. It means that all usernames and nicknames used in SM and local newspapers have been previously pseudonymised and these codes are the information shown. These codes can be converted to the original personal data by using the deanonymization functionality. The deanonymization will be only possible for those authorised user roles who will have a key that will allow the de-encryption. The system only decodes ONE data at a time, but it will be shown throughout the whole dashboard. So every instance of that data in the dashboard will be shown in its raw form while the project is open. If the user leaves or closes the project, the data will reappear pseudonymised. The data is NOT deanonymized in the database, it is only visible temporarily in the user's local machine. This information is not stored. In order to do the deanonymization, the system will ask for the user's password to verify that he/she is the valid user. After validating, the system will ask for the specific data that the user wants to decrypt and a motive for doing so. This is kept as a log for future reference to monitor all the deanonymization that takes place in the system.

Table 9. Anonymization process.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

5 Firearm Detector tool (FD)

5.1 Basic information

SAMSON is an AI (artificial intelligence) that uses already installed surveillance cameras to detect small magazine fed weapons as well as assault rifles. The instant a weapon enters the camera field of view, SAMSON shares a real-time alert. Here is a detailed motion graphics workflow of the proposed solution: https://workspace.cimediacloud.com/r/GfA1BqXGH26U

The anonymized Firearm Detector Tool constantly obfuscates/anonymizes people including their biometric data such as clothing, gender, face. When an anomaly is detected such as a small magazine fed handgun or assault rifle, then all biometric data is revealed and shared in real-time with the SOC's (Security Operation Control) dispatcher.

Responsible Partner	**Cinedit www.cinedit.com**
Tool internal / commercial name	SAMSON
Technology Readiness Level (TRL) by the end of the project	TRL 7
Data Sources	Data Sources: 4K (8 Mega Pixels) and 5 Mega Pixels surveillance cameras with IR leds (night mode on monochrome) with a shutter speed of at least 120th/sec.
Tool data delivery	UI, edge device and AI retraining.
Description	https://www.impetus-project.eu/index.php/impetus-outputs/the-impetus-solution/14-solutions-eng/32-firearm-detector-eng
Manual	https://www.impetus-project.eu/images/TOOLS_PDFS/Manuals/FD_User_Manual_9w423yj42v2sp.pdf

5.2 Tool Functionalities

5.2.1 Firearm Detector

Description	Using already installed surveillance cameras, SAMSON detects small magazine fed weapons. The instant a weapon enters the camera field of view, an alert is instantly shared with the relevant teams.
User roles	Dispatcher at SOC

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Maturity	SAMSON is mature for indoor environments using 2MP, 5MP and 8MP (Mega Pixels) surveillance cameras. It can successfully detect weapons as long as the pixel density across the weapon is 25*25 pixels.
Interface (service, methods, data structures)	Alerts are displayed in our UI dashboard (see Figure 3 and Figure 4) and are pushed to a local directory.

Table 10. FD features.

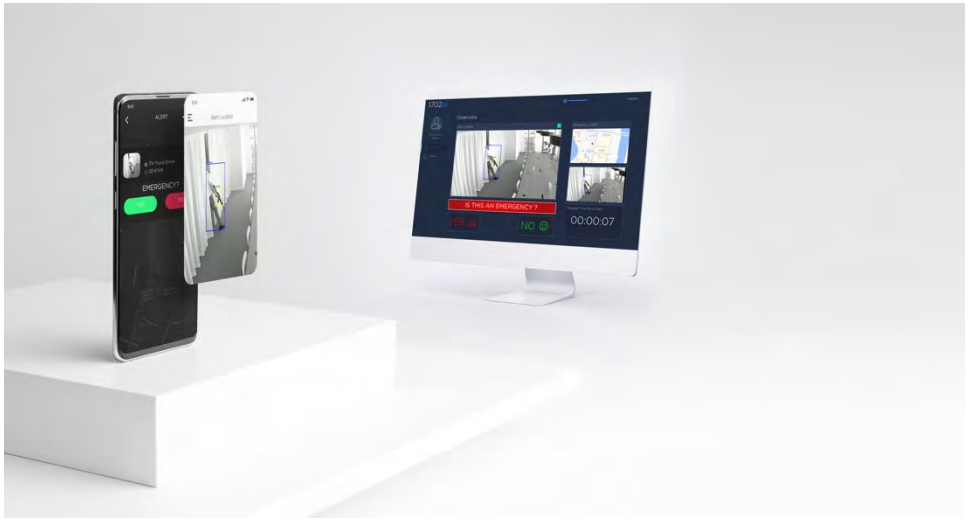


Figure 3. FD UI.

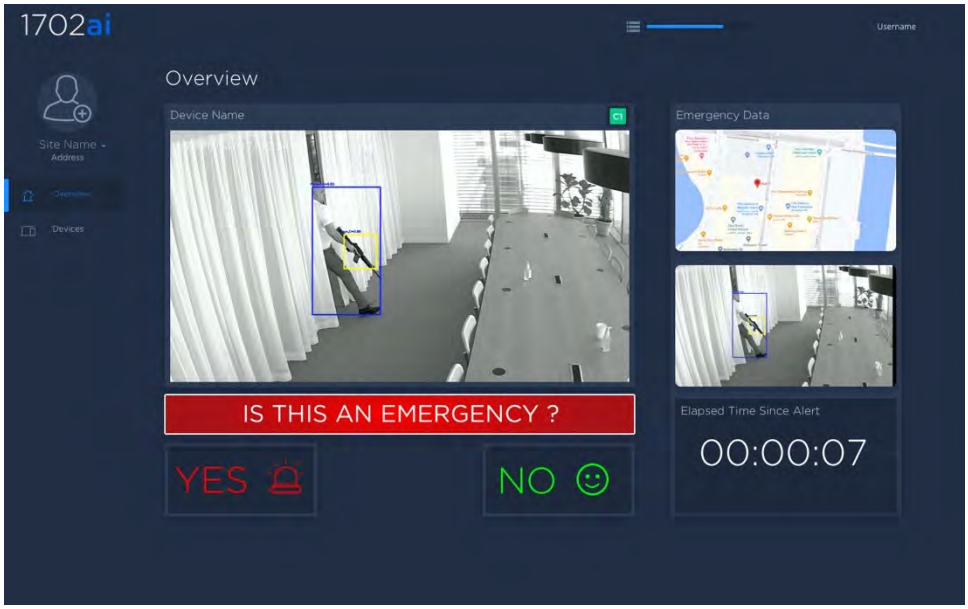


Figure 4. Detection of firearms.

Figure 5. Detection of firearms with the FD tool.

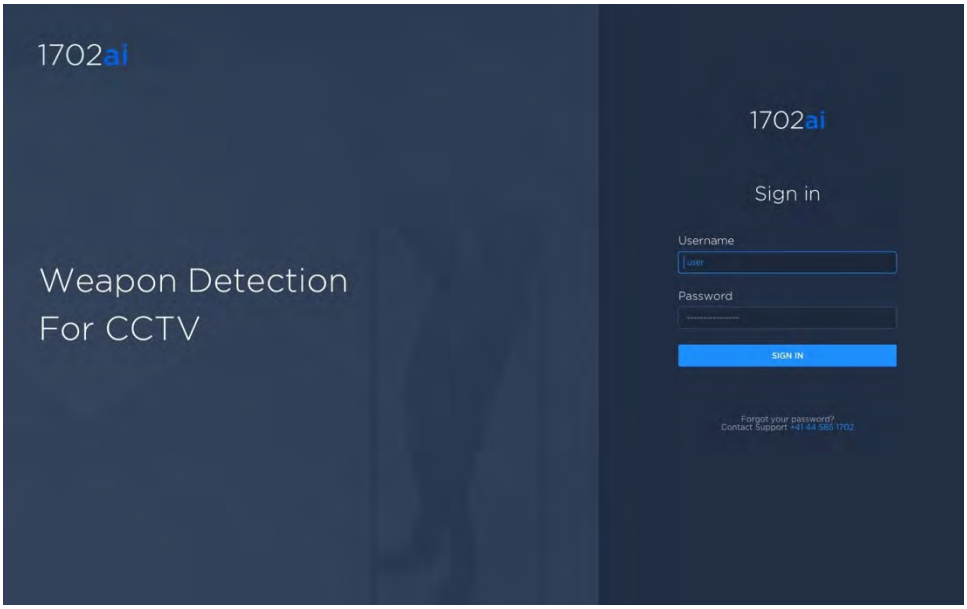


Figure 6. Login to the FD tool.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

6 Bacteria Detector tool (BD)

6.1 Basic information

The BD is an air analyser aimed at detecting microorganism's concentration in public area (train station, subway, hospital, festival, theatre). The device will transmit these data to a monitoring station.

The tool is made up of two parts that are respectively air-biocollector and an ATP analyser. The biocollector collects and catches the airborne particles in water volume, this is the sample. The ATP analyser part is designed by GLBiocontrol and called GLOW'N'CARE. The method used is ATP-metry to quantify the microorganisms in the air. Indeed, ATP (Adenosine TriPhosphate) is easy to use, this organic compound provides energy to drive many processes in living cells and is found in all known forms of life as human cell, bacteria cell and fungi cell. It can therefore be assumed in case of biological threat, the concentration in ATP will be higher due to the bacteria concentration in the air.

The ongoing step is to *physically* and *digitally* connect both tools for creating the BD. Indeed, the air caught in the volume is transferred in the ATP-analyser and the data is sent to the platform. The BD prototype is aimed at reaching TRL 6 by being deployed in real environment. This procedure allows the BD to be property calibrated. Data aggregation is necessary in both partner cities to define what the relevant threshold is in case of a bio-terror attack. Additionally, the next goal is to deploy the BD so it can be remotely controlled using the partner cities LAN.

Responsible Partner	**UdN and IMT**
Tool internal / commercial name	Microbial air analyser
Technology Readiness Level (TRL) by the end of the project	TRL 6
Data Sources	The data collected is to define the concentration of microorganisms suspended in the air. Given our environmental data, we do not use anonymization or any pseudonymization.
Tool data delivery	The data will be delivered using the Apache Kafka as a framework.
Description	https://www.impetus-project.eu/index.php/impetus-outputs/the-impetus-solution/14-solutions-eng/36-bacteria-detector
Manual	https://www.impetus-project.eu/images/TOOLS_PDFS/Manuals/BD_User_Manual_5s5d4f411v74j.pdf

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

6.2 Tool Functionalities

6.2.1 Biocollector

Description	The air-biocollector collects and catches air microorganisms in a small volume of water. The air-biocollector is compounded in an inlet for an automatic filling and an outlet to transfer the sample to the ATP analyser.
User roles	Technician or operator.
Maturity	TRL 6
Interface (service, methods, data structures)	A program has been developed to control a remote sequence. Plus, a local and html interface has been designed to drive the BD remotely.

Table 11. Bicollector features.

6.2.2 ATP- analyser

Description	The ATP-analyser is a commercial device. It was modified and adapted to receive the air sample from the air-biocollector. The ATP-analyser is compounded in an inlet for receiving the air sample and a cell for measurements to define the concentration of microorganism present in the air.
User roles	Technician or operator
Maturity	TRL 7
Interface (service, methods, data structures)	A specific program allows to communicate and to automate all the measurements.

Table 12. ATP analyzer features.

6.2.3 Data sending

Description	The sent data includes four types of information: ATP levels, the internal standard values, ATP concentrations as well as the ATP concentration per unit of air.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

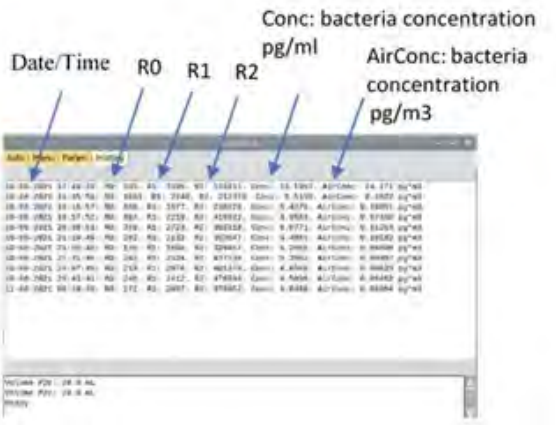
User roles	Technician or operator
Maturity	TRL 6
Interface (service, methods, data structures)	Below is a sample of shared data output:  R0: Background signal noise. R1: Measure of ATP in RLU (Relative Light Unit) R2: The measure in RLU after addition of internal standard (1000 pg of ATP). Conc: Quantity in pg of ATP/ml AirConc: Concentration of ATP per unit of air (m^3) The analysis results are stored in a .csv file in the BD and sent through secured FTP to the IMPETUS platform. The data history can be downloaded from the BD dashboard as a log file.

Table 13. Data sending features.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

7 Cyber Threat Intelligence Tool (CTI)

7.1 Basic information

Cybersixgill continuously collects and exposes the earliest possible indications of risk, moments after they surface on the clear, deep, and dark web.

Our proprietary algorithms extract data from a wide range of sources, including content from limited-access deep and dark web forums, underground markets, invite-only messaging groups, code repositories, paste sites and clear web platforms, as well as an unparalleled archive of indexed, searchable historical data from as early as the 1990s.

This data is processed, correlated, and enriched with machine learning techniques to create profiles and patterns of malicious threat actors and their peer networks, and deliver critical insight into the nature, source, and context of each threat.

Our extensive body of threat intelligence data can be consumed through various solution offerings and integrations, each addressing critical customer pain points and use cases. These solutions are scalable, searchable, and seamlessly integrated into existing security stacks, arming teams with critical insights to proactively block threats before they materialize into attacks.

Responsible Partner	**Sixgill – SG**
Tool internal / commercial name	Cybersixgill - https://www.cybersixgill.com
Technology Readiness Level (TRL) by the end of the project	TRL 9
Data Sources	Cyber threat intelligence tool provides market-leading coverage of underground sources and platforms. The tool collection spreads from the Clear, Deep and Dark web, and beyond Clear web: Paste sites, Reddit, 8chan, NVD, Twitter, GitHub Deep and dark web: Open and closed (invite-only) forums, Markets, Credit card markets, Paste sites, IRC channels, Dread, Zeronet. Beyond: Open and closed (invite-only) Telegram groups and channels, Open and closed (invite-only) QQ groups, Rogue apps, Phishing domains.
Tool data delivery	SaaS and UI.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Description	https://www.impetus-project.eu/index.php/impetus-outputs/the-impetus-solution/14-solutions-eng/49-cyber-threat-detection-and-response
Manual	https://www.impetus-project.eu/images/TOOLS_PDFS/Manuals/CTI_User_Manual_s548rg1j7p43a.pdf

7.2 Tool Functionalities

7.2.1 Actionable Alerts – main feature for the IT operators

Description	Pre-configured and automatically updated alerts and insights according to vertical and use case. Automatic mapping of the organization's assets, for triggering imminent and emerging alerts.
User roles	IT operators/specialist
Maturity	TRL-9
Interface (service, methods, data structures)	Cybersixgill Portal is a cloud-based platform (SaaS). Allowing the user to log in from any location. Its only need is an internet connection. The alerts will be displayed in the IMPETUS dashboard, allowing the user to quickly see the status of it and easily access the portal and resolve them.

Table 14. CTI Actionable alerts.

7.2.2 Investigation module

Description	Deep dive into any escalation in real-time and understands the context. Research threat actor's profile and activity history. Review and analyze across languages, sites, timeframes, types of products, topics, entities, and more. Allow the user to track and manage an ongoing investigation by attaching pieces of information under a specific case, as well as sharing this information and progress with other colleagues
User roles	IT operators/specialist
Maturity	TRL-9

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Interface (service, methods, data structures)	Cybersixgill Portal is a cloud-based platform (SaaS). Allowing the user to log in from any location. Its only need is an internet connection.

Table 15. CTI Investigation module.

7.2.3 CVE module

Description	Provides users total context regarding trending CVEs in the deep and dark web, and predicts the immediate risks of vulnerabilities with a higher probability of being exploited.
User roles	IT operators/specialist
Maturity	TRL-9
Interface (service, methods, data structures)	Cybersixgill Portal is a cloud-based platform (SaaS). Allowing the user to log in from any location. Its only need is an internet connection.

Table 16. CTI CVE Module.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

8 Urban Anomaly Detector tool (UAD)

8.1 Basic information

Algorithms for the construction of anomaly detection models and event classification models.

Responsible Partner	**CINI**
Tool internal / commercial name	The tool includes two Machine Learning algorithms: Spark-GHSOM [1] and DENCAST [2] which can perform Anomaly Detection and Event Detection. The algorithms are now being integrated into the UAD tool.
Technology Readiness Level (TRL) by the end of the project	TRL 6
Data Sources	Structured (possibly labelled) data from sensors. The data are automatically generated by the sensors, and, in the general scenario of IMPETUS, they should not contain any personal data. If any personal data is provided, these data will be removed or anonymized by means of rolling hashing functions. The tool can process geo-referenced time series for any measure that can be collected by available sensors (e.g., temperature, PM10, pedestrians flow, traffic).
Tool data delivery	1st version delivery: Sept 2021. 2nd version delivery: July 2022
Description	https://www.impetus-project.eu/index.php/impetus-outputs/the-impetus-solution/14-solutions-eng/39-urban-anomaly-detector
Manual	https://www.impetus-project.eu/images/TOOLS_PDFS/Manuals/UAD_User_Manual_8m72t57m7xie7.pdf

8.2 Tool Functionalities

8.2.1 Identify if the current sensor data is anomalous or normal

Description	Functionality that allows the system to notify when an anomalous phenomenon occurs. An anomaly is identified when a time series does not follow the expected behaviour, according to historical data and according to

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

	the behaviour of data in the (spatial) neighbour. The algorithm Spark-GHSOM is adapted for this functionality.
User roles	Analyst - Who can check what caused the anomaly to emerge by analysing the importance of the variables under analysis.
Maturity	Ready – Integrated with the other tools and integrated in the Big Data Analytics platform in WP4 (INPUT) and the IMPETUS platform (OUTPUT). The method is capable to process a dataset by catching possible and interpretable anomalies through a ranking of the variables under analysis.
Interface (service, methods, data structures)	The IMPETUS platform receives the alert, which represents the anomaly, from the tool. The dashboard can ask the tool to notify which sensor and/or variable of analysis are relevant for the anomaly and ask some aggregate data to show the trend of such variables.

Table 17. UAD features.

8.2.2 Identify the class of threats of an unclassified sensor data

Description	Functionality that allows to notify different users when a specific threat is identified from the sensor data. The tool can process geo-referenced time series for any measure that can be collected by available sensors (e.g., temperature, PM10, pedestrians flow, traffic). The algorithm DENCAST is adapted for this functionality.
User roles	Analyst - Who can check what type of threat the current data is referring to among the set of predefined threats. If data are provided in streaming, the alerts can be generated in quasi-real time and the analyst can act immediately.
Maturity	Ready - The event classifier method is implemented and tested. Integrated in the Big Data Analytics platform in WP4 (INPUT) and the IMPETUS platform (OUTPUT).
Interface (service, methods, data structures)	The IMPETUS dashboard receives the class, which represents the identified threat, from the tool. The dashboard can ask the tool to notify which sensor and/or variable of analysis are relevant for the anomaly and ask some aggregate data to show the trend of such variables.

Table 18. UAD type of threats.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

8.2.3 Interaction with anomaly detectors and event classifiers via REST APIs

Description	The machine learning model states could be changed to explicitly switch from one state to another depending on the specific needs
User roles	Data Analyst - Investigator who looks at the current state of the anomaly detector or event classifier and decides to switch from a particular state to another for a particular reason. For example, the user can manually force the learning process (training phase) to update the models on the basis of new data arrived. If the user does not manually force the start of the learning process, it is automatically started according to some periodicity.
Maturity	Ready
Interface (service, methods, data structures)	The methods take data from data sources via REST APIs and generate output in JSON format, which is sent to the IMPETUS platform via Kafka queues.

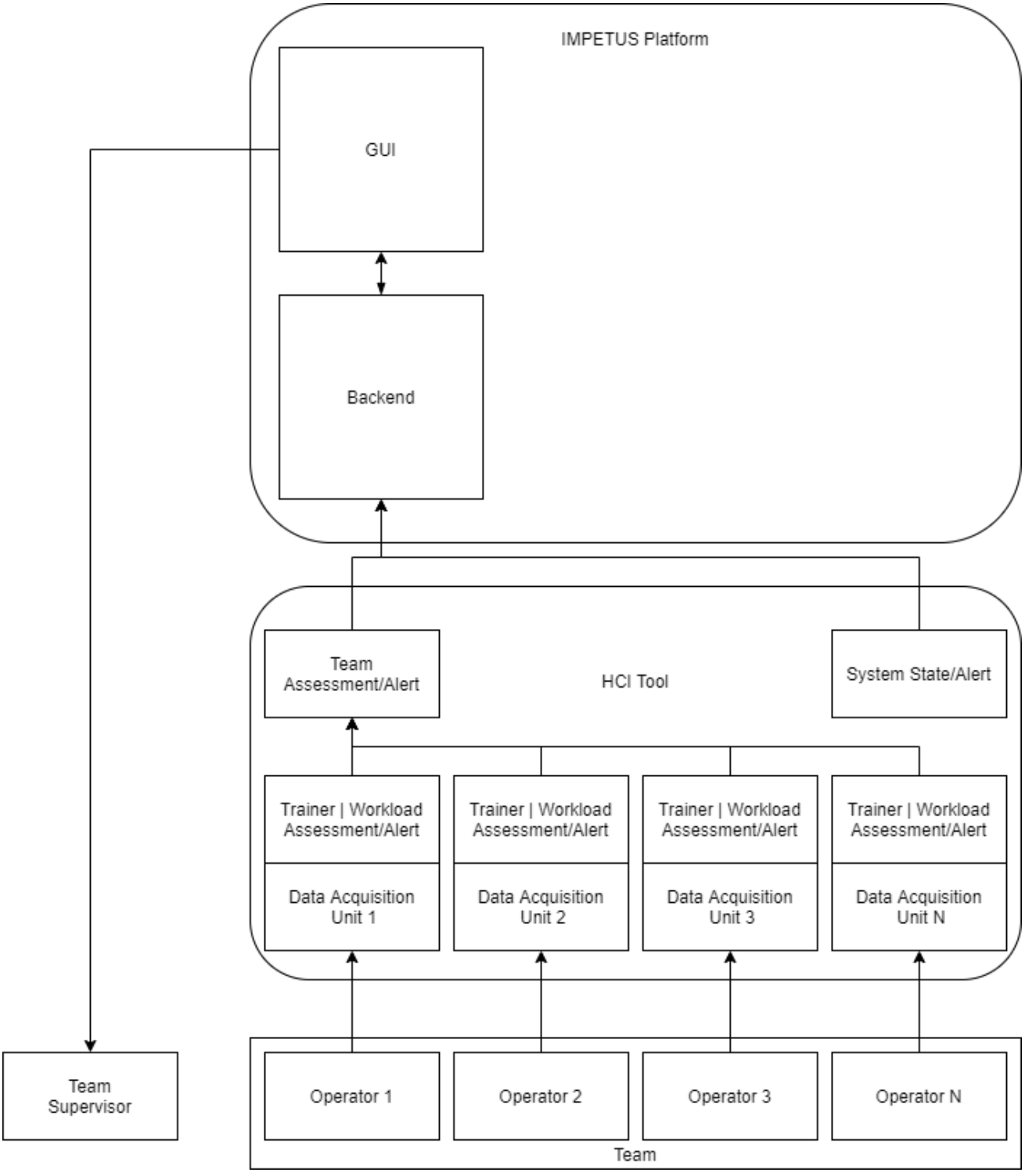
Table 19. UAD REST API.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

9 Workload Monitoring System tool (WMS)

9.1 Basic information

Workload Monitoring System (WMS) tool assesses human operator mental workload based on neuro-physiological measurements.



The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Figure 7. Integration of WMS in the IMPETUS platform.

Responsible Partner	**Thales**
Tool internal / commercial name	WMS / Workload Monitoring System
Technology Readiness Level (TRL) by the end of the project	TRL 7
Data Sources	The WMS Tool is dependent on the operator's neuro-physiological signals that are captured through the tools' sensor set. The sensors will be chosen depending on the user preference and will likely encompass an EEG (ElectroEncephaloGram) sensor, to capture brain activity, and a PPG (Photoplethysmogram) sensor to measure heart rate. Both sensors are built in a single device (a headband). The training data will be collected once, on a personal basis, to create a personalized model for each user. A set of features will be extracted from that data. This feature set data will be transferred to Thales premises where it will be used to train the workload assessment machine learning models. The assessment model will be deployed on a secured USB drive, which will be in the possession of the individual who the model belongs to.
Tool data delivery	The WMS Tool sends real-time Assessment and Alert Data on the operator and team's workload state. The tool can also send Data Quality Alerts and information about the WMS tool system state. The WMS Tool service will be deployed as a module within the IMPETUS Platform Product on premisses. The sensor set and corresponding data acquisition units will be deployed on site.
Description	https://www.impetus-project.eu/index.php/impetus-outputs/the-impetus-solution/14-solutions-eng/43-workload-monitoring-system
Manual	https://www.impetus-project.eu/images/TOOLS_PDFS/Manuals/WMS_User_Manual_29w38t94vz0va.pdf

9.2 Tool Functionalities

9.2.1 Custom Sensor Set

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Description	A custom sensor set is designed based on the task and user requirements, as described in chapter 9.4.
User roles	SOC Operator, WMS tool Technician.
Maturity	Off the shelf sensors, signal acquisition software operational.
Interface (service, methods, data structures)	physical contact, one set per desk.

Table 20. WMS functionalities.

9.2.2 Realtime Data Acquisition and Quality Check

Description	The Data Acquisition is used to acquire sensor data from an operator and check for data quality.
User roles	SOC Operator, WMS tool Technician.
Maturity	In progress, first version operational.
Interface (service, methods, data structures)	debug UI, alerts to IMPETUS PLATFORM.

Table 21. WMS data acquisition and quality check.

9.2.3 Data Feature Extraction for personalized ML model

Description	Realtime feature extraction for ML model.
User roles	WMS tool Technician.
Maturity	In progress, first version operational.
Interface (service, methods, data structures)	debug UI, LSL (Lab Streaming Layer).

Table 22. WMS features extraction.

9.2.4 Personal Model Trainer

Description	Train the personal workload assessment Machine Learning model. Done one time per operator, using a custom calibration task.
User roles	WMS tool Technician.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Maturity	In progress, first version operational.
Interface (service, methods, data structures)	debug UI, model put on secured USB stick.

Table 23. WMS personal model trainer.

9.2.5 Assessment

Description	Realtime assessment of operator sensor data.
User roles	SOC Operator, SOC Team Supervisor.
Maturity	In progress. Framework has been developed. Optimizing on feature selection.
Interface (service, methods, data structures)	personal unit with model loaded (secured USB stick).

Table 24. WMS assessment.

9.2.6 Alert System

Description	Operator alerts on basis of the assessments.
User roles	WMS tool Technician.
Maturity	In progress. Module must be adapted to the IMPETUS platform (under design).
Interface (service, methods, data structures)	Tool internal message communication bus (MQTT), alerts sent to IMPETUS PLATFORM (Kafka).

Table 25. WMS Alert System.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

10 Evacuation Optimiser tool (EO)

10.1 Basic information

Responsible Partner	**CPAD**
Tool internal / commercial name	Thunderhead Enginerring – Pathfinder (https://www.thunderheadeng.com/pathfinder), third party software for egress simulation
Technology Readiness Level (TRL) by the end of the project	TRL 6
Data Sources	People counters, when available, which defines the total number of people within an area.
Tool data delivery	Set of guidelines
Description	https://www.impetus-project.eu/index.php/impetus-outputs/the-impetus-solution/14-solutions-eng/44-evacuation-optimiser
Manual	https://www.impetus-project.eu/images/TOOLS_PDFS/Manuals/EO_User_Manual_38k0nn21vz0g8.pdf

10.2 Tool Functionalities

Please look at EO user manual for details.

10.2.1 Egress simulations

Description	Egress simulations are performed via Pathfinder software to identify area guidelines for the users
User roles	EO tool Technician
Maturity	TRL 9 – Pathfinder is a commercial product
Interface (service, methods, data structures)	Pathfinder exports some raw data to be interpreted – not to be used directly by SOC operator.

Table 26. EO simulations.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

10.2.2 Egress guidelines

Description	Guidelines can be used by SOC operator to optimize action.
User roles	SOC operator.
Maturity	TRL 6
Interface (service, methods, data structures)	Guidelines are available via web interface.

Table 27. EO guidelines.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

11 Cyber Threat Detection and Response tool (CTDR)

11.1 Basic information

The Cyber Threat Detection and Response tool is based in Prelude OSS, the freeware version of Prelude SIEM[3], which is a Security Information and Event Management (SIEM) tool, for the generation and reporting of cybersecurity alerts. Under the scope of the IMPETUS project, we will refer to either Prelude OSS or Prelude SIEM as Prelude, for simplicity reasons.

It is composed of monitoring software that collects and processes events created by other tools (e.g., events stored in the log files of an antivirus, network firewalls or intrusion detection systems). Additionally, alerts can also be displayed and explored by security analysists via graphical user interface (GUI) dashboards.

In the IMPETUS project, Prelude is extended with the use of the ELK stack[5], which is the abbreviation for three open-source projects, namely Elasticsearch, Logstash and Kibana. The goal of the Cyber Threat Detection and Response tool is to add further processing and reporting capabilities to Prelude, as well as to provide new dashboards to cybersecurity analysts.

The alerts from the SIEM are mapped with an attack graph generator output and a vulnerability ontology to prioritise actions that should be taken to mitigate the risk. Attack graph generation is possible based on output of Nessus scan. It was previously planned to enrich and process inputs from other cybersecurity tools (e.g., XM Cyber's BAS tool). The following figure represents the framework before XM Cyber departure.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

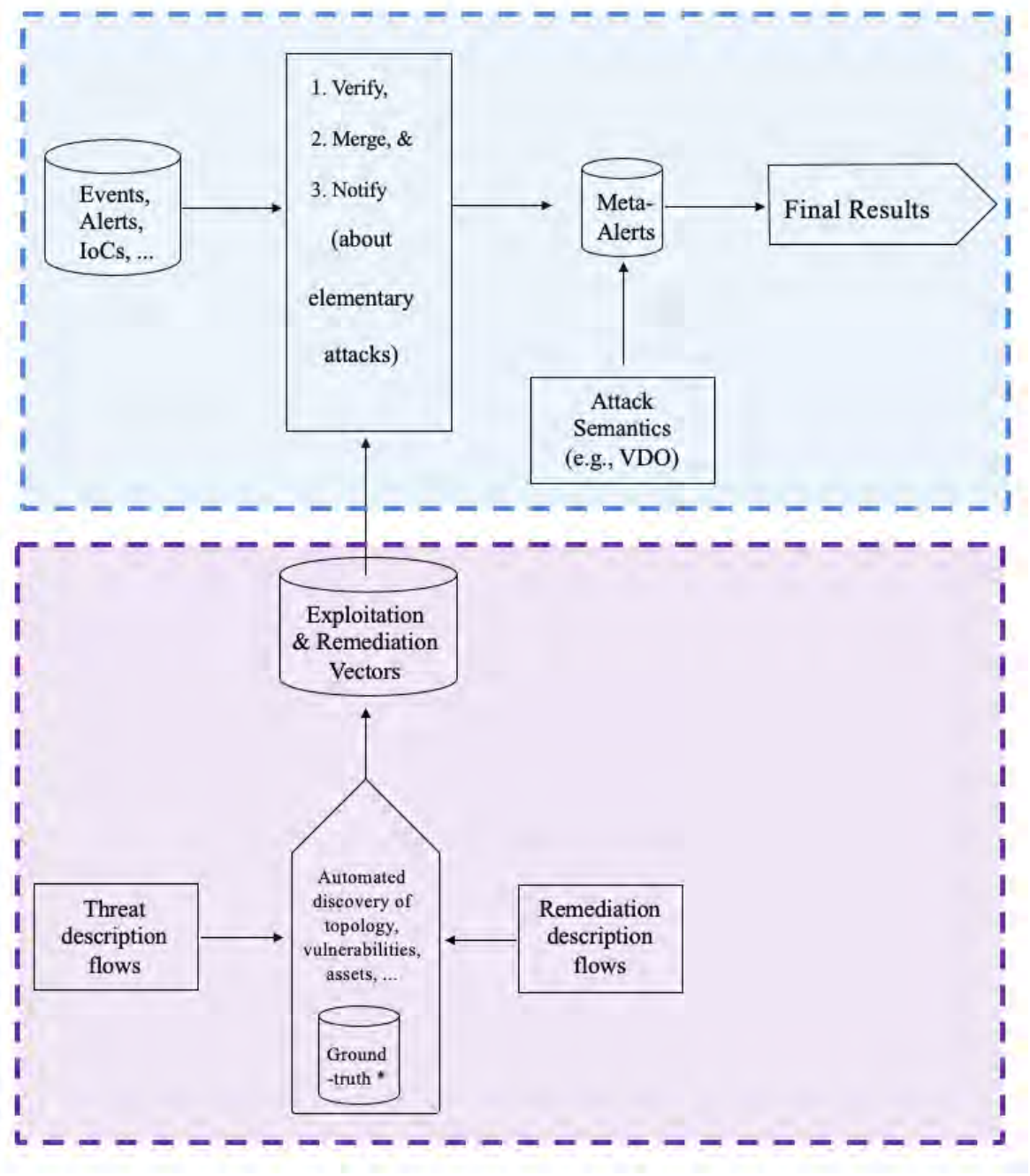


Figure 8. Prelude-ELK flow chart showing the inputs, processes, and outputs.

The figure represents the framework after XM Cyber departure. Functionality and results provided by Prelude-ELK are highlighted in blue. Functionality and results by vulnerability scanners like Nessus are highlighted in orange. Missing functionality and results, previously covered by the tool of XM Cyber, are highlighted in red.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

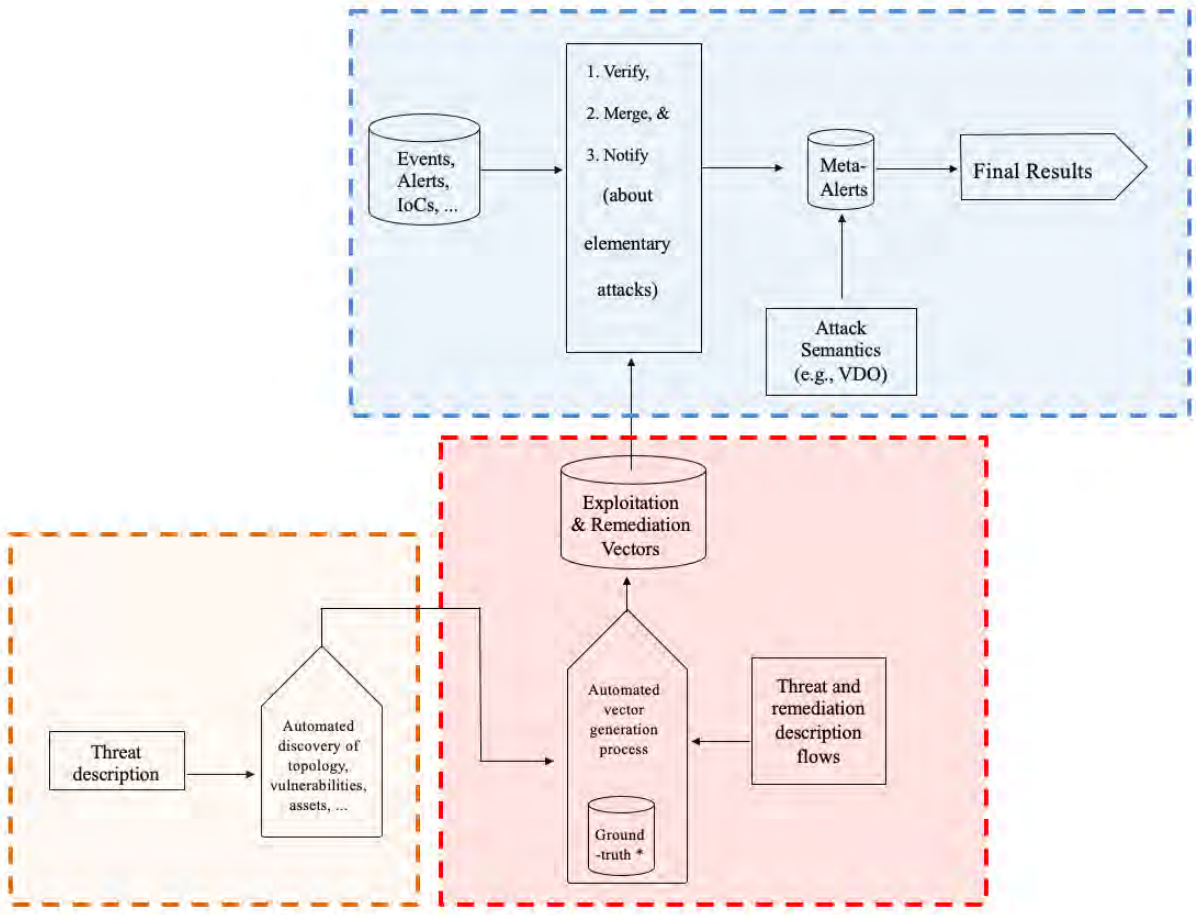


Figure 9. Prelude-ELK flow chart after XM Cyber departure showing the inputs, processes and outputs.

Elasticsearch allows indexing and processing unstructured data. It also provides a distributed web interface to access the resulting information. Logstash is the parsing engine associated with Elasticsearch for collecting, analysing, and storing logs. It can integrate many sources simultaneously. Finally, Kibana is a data visualization platform that provides visualization functionalities on indexed content in Elasticsearch. Users can create dashboards with charts and maps of large volumes of data.

Nessus Essentials is a free a security scanner. It is used to discover vulnerabilities. The output of a Nessus scan can be exported as an XML file that then is exported in the attack graph generator to generate a pro-active attack graph. The mapping between a pro-active attack graph, the alerts and the ontology allow updating the attack graph to have an assessment that consider the system changes and the position of an adversary in the system.

Responsible Partner	**IMT**
Tool internal / commercial name	Prelude SIEM (https://www.prelude-siem.com/) – This tool is developed by a third party. Prelude OSS (https://www.prelude-siem.org/) – Also developed by a third party.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

	ELK Stack (https://www.elastic.co/what-is/elk-stack) – Developed by third parties. MulVAL (https://github.com/risksense/mulval) - Developed by third parties. Vulnerability Database Ontology (https://github.com/usnistgov/vulntology/blob/master/CONTRIBUTING.md) - Implemented by third parties. Attack Graph Generator – Developed by IMT.
Technology Readiness Level (TRL) by the end of the project	TRL 6
Data Sources	Unstructured recording of events (e.g., log files from an operating system or network device) using the syslog format (cf. RFC 5424[6]). Nessus scan output.
Tool data delivery	Web interface
Description	https://www.impetus-project.eu/index.php/impetus-outputs/the-impetus-solution/14-solutions-eng/46-cyber-threat-intelligence
Manual	https://www.impetus-project.eu/images/TOOLS_PDFS/Manuals/CTDR_User_Manual_a5y8t7o0q7234.pdf

11.2 Tool Functionalities

11.2.1 Receive logs

Description	Prelude-ELK is installed as a service on a Docker container, configured to receive syslog files from the components of the monitored system, using events messages on an IP network.
User roles	Security analysts via dashboards
Maturity	TRL 6 - Dockerised version of Prelude-ELK, composed of containers and configurations.
Interface (service, methods, data structures)	Use of RFC 5424 (syslog) interface for the collection of logs.

Table 28. CTDR logs.

11.2.2 Generate alerts

Description	Based on automated rules to generate cybersecurity IDMEF alerts.
User roles	Security analysts via dashboards.
Maturity	TRL 6 - Dockerised version of Prelude-ELK, composed of containers and configurations.
Interface (service, methods, data structures)	Prelude-ELK already installed on the containers. An alert database, in which the alerts of Prelude-ELK will be stored, is also installed, and configured as a container.

Table 29. CTDR Alerts.

11.2.3 Correlate alerts

Description	Based on automated rules to correlate and generates cybersecurity alerts. For example, when a user tries to get remote access several time on a machine, Prelude-ELK will correlate previous alerts into a new 'Brute Force' alert.
User roles	Security analysts via dashboards
Maturity	TRL 6 - Dockerised version of Prelude-ELK, composed of containers and configurations.
Interface (service, methods, data structures)	Prelude-correlator tool and required scripts, already installed on a container in the Dockerised version of Prelude-ELK.

Table 30. CTDR alerts correlations.

11.2.4 Visualize alerts

Description	Visualization of events and processed alerts, either in raw, IDMEF or additional formats.
User roles	Security analysts via dashboards.
Maturity	TRL 6 - Dockerised version of Prelude-ELK, composed of containers and configurations.
Interface (service, methods, data structures)	Standard and extended Prelude-ELK dashboards, already installed on containers in the Dockerised version of Prelude-ELK.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Table 31. CTDR alerts visualization.

11.2.5 Timeline visualization

Description	Visualization of events and processed alerts, either in raw, IDMEF or additional formats, and its distribution across time (e.g., using chart bar diagrams).
User roles	Security analysts via dashboards.
Maturity	TRL 6 - Dockerised version of Prelude-ELK, composed of containers and configurations.
Interface (service, methods, data structures)	Standard and extended Prelude-ELK dashboards, already installed on containers in the Dockerised version of Prelude-ELK.

Table 32. CTDR timeline visualization.

10.6.6 Attack Graph Generation

Description	A graphical representation of the attacker paths to compromise the system.
User roles	Security analysists via dashboards.
Maturity	TRL 6 - Dockerised version of Prelude-ELK, composed of VMs, containers and configurations.
Interface (service, methods, data structures)	Attack graph generator, already installed on container in the Dockerised version of Prelude-ELK.

Table 33. CTDR attack graph generation.

11.2.6 Attack-Defense Graph

Description	A graphical representation of the attacker paths and the countermeasures that can be applied on the vulnerabilities to mitigate them.
User roles	Security analysts via dashboards.
Maturity	TRL 6 - Dockerised version of Prelude-ELK, composed of containers and configurations.
Interface (service, methods, data structures)	Standard and extended Prelude-ELK dashboards, already installed on containers in the Dockerised version of Prelude-ELK.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Table 34. CTDR attack defense graph.

10.6.8 Alerts to IMPETUS

Description	When a vulnerability is exploited, alerts are sent to the IMPETUS platform.
User roles	Security analysts via dashboards.
Maturity	TRL 6
Interface (service, methods, data structures)	IMPETUS Platform

Table 35. CTDR alerts to Impetus.

10.6.9 Attack Graph Enrichment based on attack simulation

Description	The attack graph is updated when a vulnerability exploited can cause the exploitation of another vulnerability or when new paths conducing to the goal are discovered.
User roles	Security analysts via dashboards
Maturity	TRL 6 – Simulation on virtual network allow demonstrating the attack graph enrichment.
Interface (service, methods, data structures)	Standard and extended Prelude-ELK dashboards, already installed on containers in the Dockerised version of Prelude-ELK.

Table 36. CTDR attack graph enrichment.

10.6.10 Attack-Defense Graph Enrichment based on attack simulation

Description	The attack-defense graph is updated when a vulnerability exploited can cause the exploitation of another vulnerability or when new paths conducing to the goal are discovered.
User roles	Security analysts via dashboards.
Maturity	TRL 6 – Simulation on virtual network allow demonstrating the attack-defense graph enrichment.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Interface (service, methods, data structures)	Standard and extended Prelude-ELK dashboards, already installed on containers in the Dockerised version of Prelude-ELK.

Table 37. CTDR attack-defense graph enrichment.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

12 Conclusion

At this point, a general observation regarding the tools for the IMPETUS project is that we have reached the expected maturity of the toolkit overall. If we compare the Technology Readiness Level (TRL) of the starting point for each of the tools in the project's definition, we will see that many of the tools have achieved the expected level of maturity by the end of the project.

Regarding the EO tool, it has been modified after detecting some conflicts with the scope of the specifications included at the beginning of the project. To align with the user requirements (in WP1) and to optimize the EO tool impact, this tool has required additional development and adaptation. To reflect the situation, we consider it more appropriate to categorize EO tool maturity as TRL 5.

The UAD TRL level included in the previous deliverable (3.1) was considering the maturity of the algorithms used in it (Spark-GHSOM8 and DENCAST9) but not the tool overall. We have corrected the situation in this document by indicating the actual TRL of the tool (TRL 7).

Another relevant perspective regarding diversity, in terms of maturity and divergent projection of each of the tools, is to highlight that not only the maturity level is planned to be different for all the tools at the beginning and the end of the project. The partners have also different objectives regarding their tools. Some of the tools are mature enough to be integrated into the IMPETUS platform, with no other development expected. However, other tools required modification of components -such as data connectors or operational capabilities- that bring the opportunity to the responsible partners to obtain a much more mature product.

To summarize, during the development of WP3, we established TRL Scale 7 as the standard for determining the maturity level of each of the tools and all the tools have achieved this level or higher, except EO.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Members of the IMPETUS consortium

SINTEF	SINTEF, Strindvegen 4, Trondheim, Norway, https://www.sintef.no	Joe Gorman joe.gorman@sintef.no
Institut Mines-Télécom	Institut Mines Telecom, 19 place Marguerite Perey, 91120 Palaiseau, France, https://www.imt.fr	Joaquin Garcia-Alfaro joaquin.garcia_alfaro@telecom-sudparis.eu
UNÎMES	Université de Nimes, Rue du Docteur Georges Salan CS 13019 30021 Nîmes Cedex 1, France, https://www.unimes.fr	Axelle Cadiere axelle.cadiere@unimes.fr
cini consorzio interuniversitario nazionale per l'informatica	Consorzio Interuniversitario Nazionale per l'Informatica, Via Ariosto, 25, 00185 – Roma, Italy, https://www.consorzio-cini.it	Donato Malerba donato.malerba@uniba.it
Università degli Studi di Padova	University of Padova, Via 8 Febbraio, 2 - 35122 Padova, Italy, https://www.unipd.it	Giuseppe Maschio giuseppe.maschio@unipd.it
Entrepreneurship Development Centre for BIOTECHNOLOGY and MEDICINE	Biotehnoloogia ja Meditsiini Ettevõtluse Arendamise Sihtasutus, Tiigi 61b, 50410 Tartu, Estonia, https://biopark.ee	Sven Parkel sven@biopark.ee
SIMAVI Software Imagination & Vision	SIMAVI, Complex Victoria Park, Corp C4, Etaj 2, Şos. Bucureşti – Ploieşti, nr. 73 – 81, Sector 1, Bucureşti, Romania, https://www.simavi.ro	Gabriel Nicola Gabriel.Nicola@simavi.ro Monica Florea Monica.Florea@simavi.ro
THALES	Thales Nederland BV, Zuidelijke Havenweg 40, 7554 RR Hengelo, Netherlands, https://www.thalesgroup.com/en/countries/europe/netherlands	Johan de Heer johan.deheer@nl.thalesgroup.com
CINEDIT INTELLIGENT VIDEO ANALYTICS	Cinedit VA GmbH, Poststrasse 21, 8634 Hombrechtikon, Switzerland, https://www.cinedit.com	Joachim Levy j@cinedit.com

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

INSIKT INTELLIGENCE	Insikt Intelligence, Calle Huelva 106, 9-4, 08020 Barcelona, Spain, https://www.insiktintelligence.com	Dana Tantu dana@insiktintelligence.com
SIXGILL	Sixgill, Derech Menachem Begin 132 Azrieli Tower, Triangle Building, 42nd Floor, Tel Aviv, 6701101, Israel, https://www.cybersixgill.com	Benjamin Preminger benjamin@cybersixgill.com Ron Shamir ron@cybersixgill.com
Comune di Padova	City of Padova, via del Municipio, 1 - 35122 Padova Italy, https://www.padovanet.it	Enrico Fiorentin fiorentine@comune.padova.it Stefano Baraldi Baraldis@comune.padova.it
Oslo	City of Oslo, Grensen 13, 0159 Oslo, Norway, https://www.oslo.kommune.no	Osman Ibrahim osman.ibrahim@ber.oslo.kommune.no
Institut za sigurnosne politike ISP	Institute for Security Policies, Kruge 9, 10000 Zagreb, Croatia, http://insigpol.hr	Krunoslav Katic krunoslav.katic@insigpol.hr
TIEMS	International Emergency Management Society, Rue Des Deux Eglises 39, 1000 Brussels, Belgium, https://www.tiems.info	K. Harald Drager khdrager@online.no
UniSMART Fondazione Università di Padova	Unismart – Fondazione Università degli Studi di Padova, Via VIII febbraio, 2 - 35122 Padova, Italy, https://www.unismart.it	Alberto Da Re alberto.dare@unismart.it

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

13 Annex 1: Results Descriptions

Firearm Detector

Continuously monitors surveillance camera feeds and automatically creates an alert if a firearm is detected in a public space

WHAT PROBLEM DOES THE TOOL HELP SOLVE?

Dangerous scenarios and extreme events involving use of weapons, sadly, do occur in our cities. The purpose of this tool is to use surveillance cameras to detect firearms in real-time and improve the physical security of open spaces.

Without this tool:

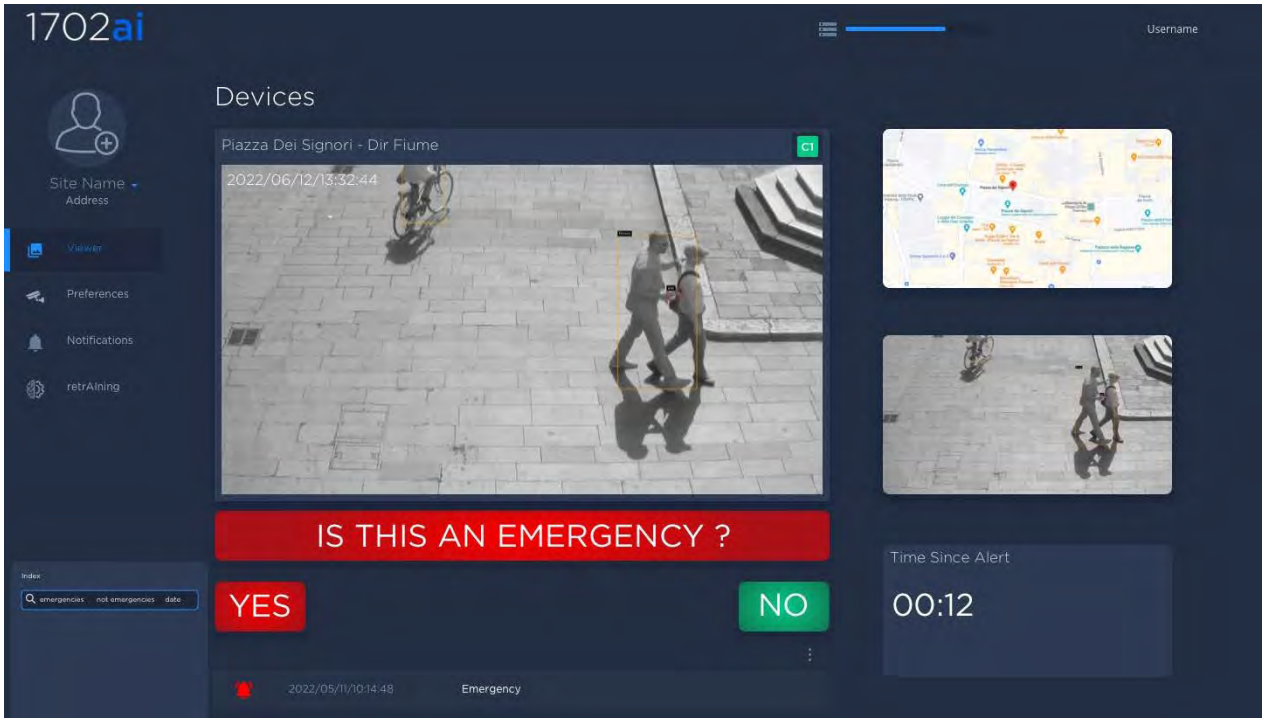
- Law enforcement is hindered due to the lack of detailed situational awareness (delays and uncertainties in reporting, lack of information about exact location)
- Response times can be lengthy – and in situations where every second count, this can lead to loss of life

With this tool:

- Immediate supply of images and location data enables super-fast response times
- The risk of loss of life is significantly reduced
- SOC operations are significantly improved

HOW IS IT DEPLOYED IN IMPETUS?

- **Who are the users**: Dispatcher at Security Operations Centres and first responders.
- **What are the critical situations for deployment**: The tool is continuously deployed to monitor and look out for weapons (without any operator intervention). If a weapon is detected, an alert is presented to the security operator who can decide how to respond.



HOW DOES IT WORK?

The instant a weapon enters the surveillance video camera's field of view, an alert is shared with the Security Operations Center. Each alert provides immediate situational awareness. The tool is GDPR, NATO and DHS (Department of Homeland Security) compliant.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883286.

Bacteria Detector

Continuously monitors air samples to detect abnormally high concentrations of airborne bacteria

WHAT PROBLEM DOES THE TOOL HELP SOLVE?

The Bacteria Detector continuously monitors bacterial concentration in the air to help protect citizens from biological hazards. It communicates with the IMPETUS platform to raise alerts with the authorities.
Without the tool:

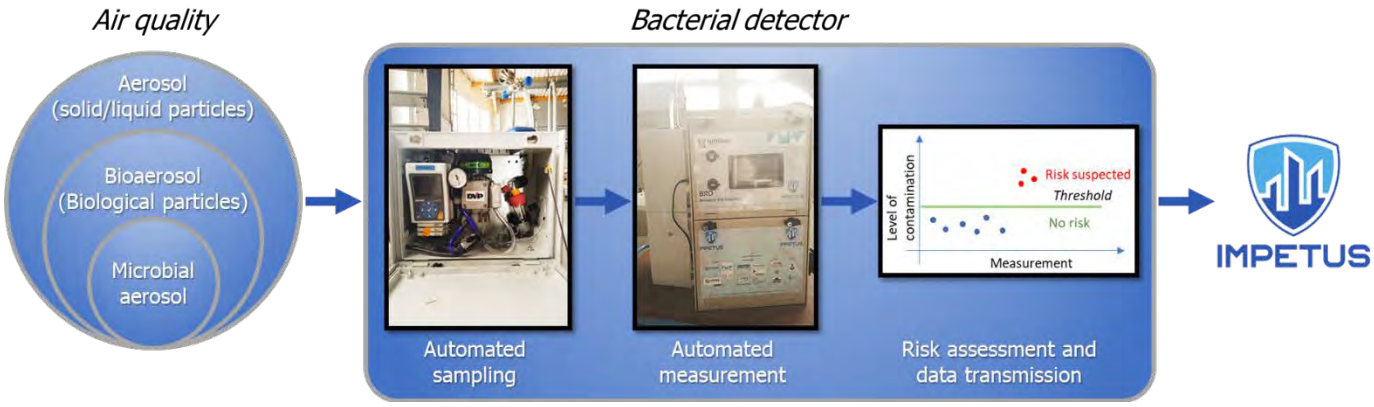
- One person can infect 1–10 other people, depending on the pathogen
- Physicians need to take samples from patients to find a suitable treatment, which prolongs treatment
- Hospital staff are not protected, and an epidemic can be declared the day after the disease appears

With the tool:

- Only those present at the point of infection are contaminated
- Samples are taken in the room and from patients (with a result in <4 hours)
- Physicians readily adapt their procedure and treatment plan, thus saving time
- Hospital staff are protected, and the risk of spreading is limited

HOW IS IT DEPLOYED IN IMPETUS?

- **Who are the users**: Trained technicians operate the equipment. Security centre operators and stakeholders in hospitals, government officials, senior level management, etc. receive early notification of possible contamination threats and infectious bacterial outbreaks through online monitoring.
- **What are the critical situations for deployment**: Continuous: the main purpose of the tools is to provide constant situational awareness and raise alerts when needed.



HOW DOES IT WORK?

This tool combines an air biocollector (developed by IMT Alès / University of Nîmes) and a bacterial concentration measurement device*. Firstly, air is sampled using an impinger and any bacteria trapped on the device are resuspended in water. Secondly, the water is analysed to measure bacteria in the air. Finally, the data is sent to the IMPETUS platform and an alert is triggered if the measurement exceeds a defined threshold.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883286.

Urban anomaly detector

Continuously monitors data gathered from multiple city sensors and detects cases deviating from the norm - indicating possible cause for concern

WHAT PROBLEM DOES THE TOOL HELP SOLVE?

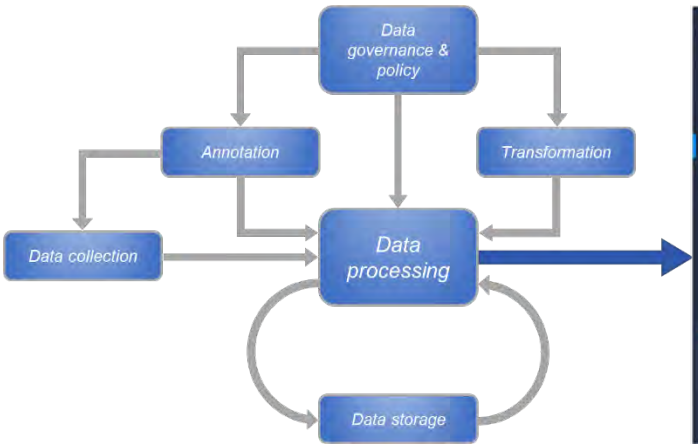
Smart cities continuously gather data from multiple sensors throughout the city. While variations in the data can be a sign of possible problems, the volumes of data are typically so huge that it is not feasible to monitor manually, or easily detect anomalies.

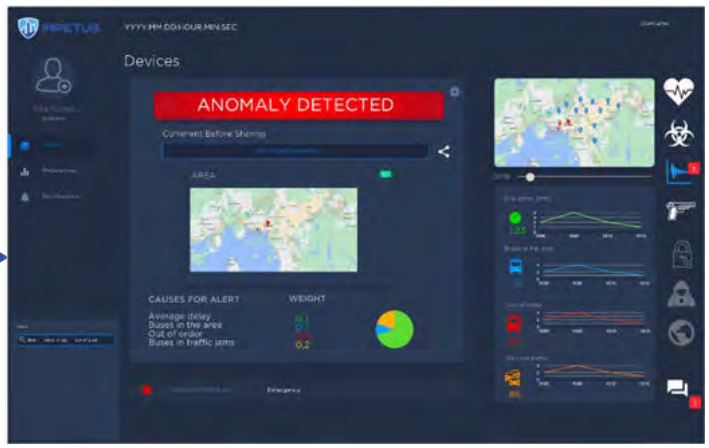
The tool uses AI (Artificial Intelligence) techniques to gather data from multiple sources over long time periods to recognise patterns and recognise what is "normal" at different times and places. It then uses that knowledge to detect anomalies when they occur, even if they have not been observed before. The tool can categorize anomalies and let a human operator evaluate whether they represent a real danger.

- Without the tool: abnormal events or situations can go unnoticed because humans are unable to process the amount of data needed to identify a threat when it occurs, which can lead to chaos and possibly disaster.
- With the tool: any unusual developments are quickly and automatically identified, and steps can then be taken to assess the situation and, maybe, mitigate a disaster.

HOW IS IT DEPLOYED IN IMPETUS?

- **Who are the users**: Security, transport and operational personnel monitoring impending physical threats, traffic flow and/or security infringements before and after any abnormal event; other stakeholders such as city managers, government officials, senior level official, etc.
- **What are the critical situations for deployment**: Continuous. The tool aims to provide constant situational awareness – anomalies can arise at any time.





HOW DOES IT WORK?

Large quantities of data are constantly collected from several sources, e.g., CCTV, sensors, municipal properties (details will vary from city to city). These data are processed using policy awareness, analytics and visualisation. If anomalies are detected, a visualisation – showing what is "unusual" – is sent as an alert to the IMPETUS platform, for the attention of emergency operators.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883286.

Social Media Detection

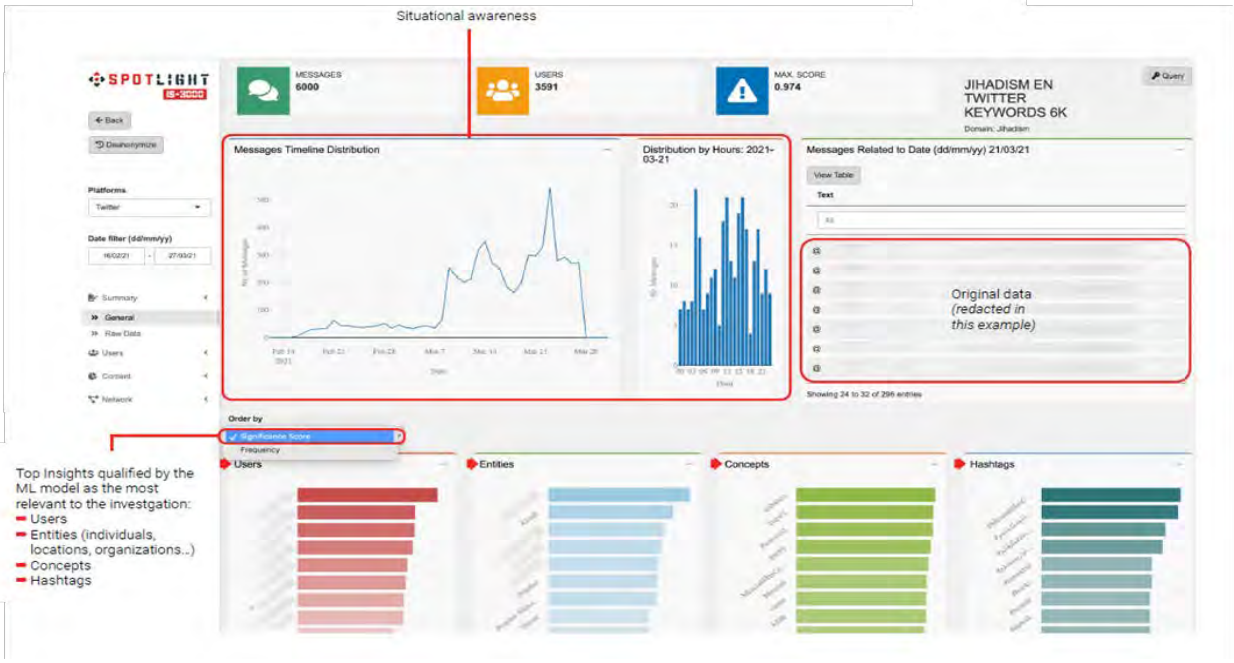
Scans large volumes of text on social media and other public online sites, looking for topics/keywords that might indicate potential trouble or threats

WHAT PROBLEM DOES THE TOOL HELP SOLVE?

The huge amounts of data on social media++ can contain vital information that is relevant for people responsible for public safety – but this information very likely goes unnoticed because it is not humanly possible for people to monitor and analyse the huge volumes. Warnings of possible issues go unnoticed. The purpose of the tool is to increase efficiency and capacity when searching for accurate and relevant insights in the ocean of data published on the open web. As the software expedites data analyses, the user can run multiple search projects, thus expanding and/or fine-tuning their search to obtain more relevant outputs.

HOW IS IT DEPLOYED IN IMPETUS?

- **Who are the users**: Intelligence analysts, tasked to give security centre operators early notice of possible dangerous situations/threats or monitoring the aftermath online, which can be of interest to other stakeholders such as government officials, senior level management, etc.
- **What are the critical situations for deployment**: A 3-step process:
    1. Create a project of interest
    2. Acquire and analyse data
    3. Use the dashboard to send alerts when anomalies are detected



HOW DOES IT WORK?

The analyst first creates a project of the topic of their interest using search criteria, e.g. keywords. The tool retrieves massive volumes of data from social media platforms, websites, forums, etc. based on the search criteria. The tool analyses the data, removing unrelated content, and presents the most relevant insights/information for each project. The user receives a notification through the IMPETUS platform that the results have generated. The analyst can then filter and fine-tune the search criteria and results to get more specific and more relevant information. This tool will aid the end user in identifying any hidden threats, or notify the user if unrest is brewing.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883286.

Evacuation Optimiser

Provides instant advice to emergency staff on how to effectively manage an evacuation, based on simulations of different evacuation scenarios

WHAT PROBLEM DOES THE TOOL HELP SOLVE?

The main purpose of the tool is to pre-optimise and support the management of controlled crowd movement in public spaces in complex events, to prevent any injury and/or loss of life, e.g. in an emergency evacuation.
Without the tool:

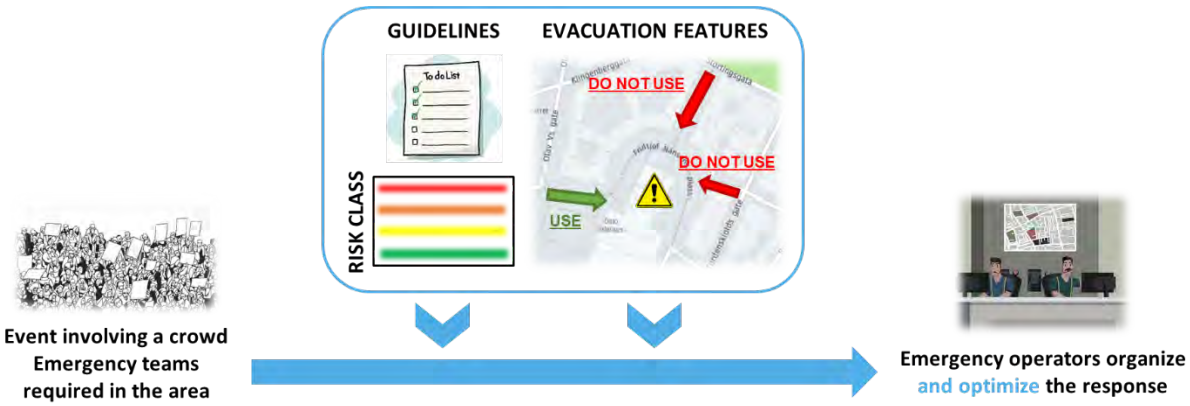
- The adequacy of number and size of exit routes is unidentified
- Specific gateways for emergency services are not known
- The total evacuation time and risk associated with evacuation remain unknown

With the tool:

- The number and direction of exit routes for the size of the crowd is evaluated
- Gateways for emergency services are identified
- An accurate calculation of total evacuation time and risk is presented to emergency operators via the IMPETUS platform
- Successful evacuation procedures

HOW IS IT DEPLOYED IN IMPETUS?

- **Who are the users**: First responders and security centre operators who are tasked with early notification of possible dangerous situations/threats, or online, real-time monitoring of the event/emergency; other stakeholders such as government officials, senior level management, city managers, etc.
- **What are the critical situations for deployment**: The tool facilitates coordination between different agencies, staff in control rooms and staff on location, and members of the public in preparation of and during a critical event. It can help dispatch required resources as efficiently as possible. The tool also facilitates planning of and execution of evacuations by mapping the quickest, most direct route for crowd control and movement.



HOW DOES IT WORK?

- **Preparation for an emergency**: Using data from people-counting sensors, the tool pre-simulates evacuation scenarios from a public space under different circumstances and provides general operative guidelines for managing the exit of a crowd in the different scenarios.
- **During an emergency**: Based on data from earlier simulations, the size of the crowd, the number of entry/exit point and the capacity of the evacuation routes, the tool estimates the time needed to evacuate the crowd, and estimates the risk involved. Guidelines on optimal entry and evacuation routes are presented to emergency personnel and security operators via the IMPETUS platform.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883286.

Cyber Threat Intelligence

Detects, classifies and helps mitigate cyberspace threats to an organisation's IT assets

WHAT PROBLEM DOES THE TOOL HELP SOLVE?

The purpose of the tool is to continuously expose the earliest indication of cyber risks to an organization's network from deep and dark web fora and markets, as well as private messaging groups.
Without the tool, analysts will have to cope with a lot of manual work, regarding:

- Collecting domain, IP and third-party data
- Indexing, tagging and metadata analysis of collected data
- Extracting relevant data and restructuring and packaging for data storage in a database maintained by the tool provider (Cybersixgill)

With the tool, you are able to:

- Receive and use a queue of asset-based alerts
- Conduct offline and discreet investigation of ongoing threats and events in cyberspace
- Receive contextual information of – and mitigate – the threats to the organization (who, where, what)

HOW IS IT DEPLOYED IN IMPETUS?

- **Who are the users**: IT specialists tasked with giving Security Operations Center operators and other stakeholders (government officials, senior level management, etc.) early notice of possible threats posed to the organization's assets.
- **What are the critical situations for deployment**: Regular: scans would typically be performed daily. The tool provides comprehensive insights into the nature and source of cyber threats, and as these can emerge rapidly it essential to keep up to date.

HOW DOES IT WORK?

There are 3 main steps:

1. **Data collection** – Finding all relevant sources, sign-in closed access forums and groups, and inquire the data (by crawling).
2. **Data processing and analysis** – The tool runs several processes on every newly collected item: indexing, enrichment, tagging, entity extraction, metadata, restructuring and saving the data into a database.
3. **Data lake query** – Automated and manual processes are running on our extensive database of cyber incidents and threat actors' activity.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883286.

Cyber threat Detection and Response

Detects cyber vulnerabilities in IT Systems: raises alerts and suggests countermeasures if they arise

WHAT PROBLEM DOES THE TOOL HELP SOLVE?

Information systems typically have so many vulnerabilities that it is not feasible to continuously monitor or manually manage all of them. Moreover, there are complex dependencies between vulnerabilities. For example: some vulnerabilities only become critical when some other vulnerability has been exploited (i.e., there has been a successful attack). This tool:

- Identifies exploited threats and potentially exploited vulnerabilities
- Prioritises actions to tackle the exploited threats and any exploitable vulnerabilities based on criticality of the situation

Without the tool:

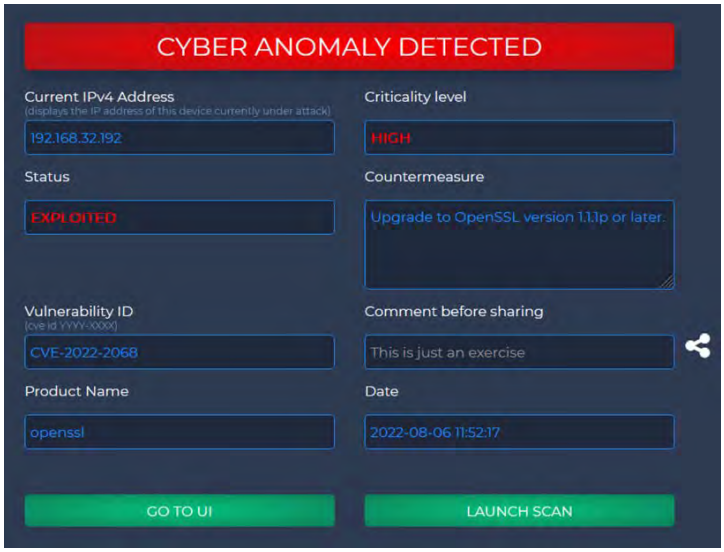
- Users' manual analyses of the system identify only a fraction of the vulnerabilities inherent within the system
- Users are not aware of how inter-linked vulnerabilities could expose the system
- Users are not aware when a vulnerability has been exploited

With the tool:

- Users can scan complex systems to identify all vulnerabilities and their relationships
- Users can monitor systems in real-time and receive an alert on the IMPETUS platform when a vulnerability has been exploited
- Countermeasures can be prioritized based on the criticality of the threat

HOW IS IT DEPLOYED IN IMPETUS?

- **Who are the users**: (A) IT specialists responsible for protecting IT infrastructure against possible cyber-attacks (through analysis, monitoring and mitigation); (B) System operators and Security Centre operators who need notification of imminent threats/problematic situations.
- **What are the critical situations for deployment**: Regular: scans and analyses would be performed periodically. The tool is designed to provide up to date situational awareness.



HOW DOES IT WORK?

The tool monitors network traffic data and correlates it with vulnerabilities discovered from a network scan. When an anomaly threatening a vulnerability on the system is detected, remedial actions are prioritised based on the severity of the threat. A cyber-security alert is generated, which is sent to the IMPETUS platform. Users can then take the prescribed action to mitigate the threat. For example, when a user tries to remotely access a machine several times, the tool will generate an alert to the IMPETUS platform suggesting the necessary countermeasures.

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883286.

Workload Monitoring System

Measures mental workload and stress of emergency operators using a brain-computer interface, raises alerts if anomalies arise

WHAT PROBLEM DOES THE TOOL HELP SOLVE?

A SOC (Security Operations Centre) can be a highly stressful working environment, and staff may react slowly or even make mistakes if stress goes unnoticed. The opposite situation – too little to do – can lead to boredom and inattentiveness.
This tool minimizes potential human error and improves human–machine teaming performance by monitoring the physical, emotional and mental workload status of operators while they perform their duties. It provides an early notification of an individual and/or a team's workload capability and ability to cope with stressors during emergencies.
Without the tool:

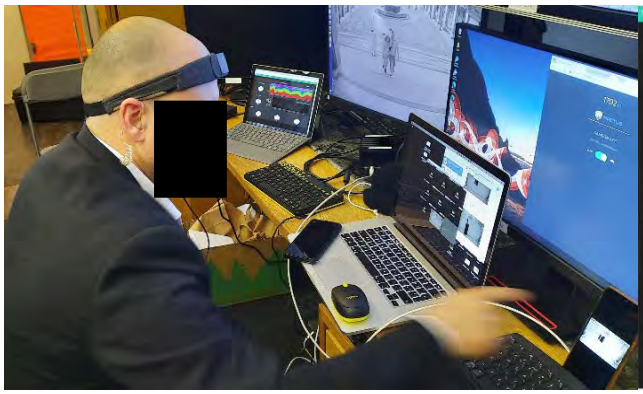
- Workload perception is implicit, subjective and sporadic

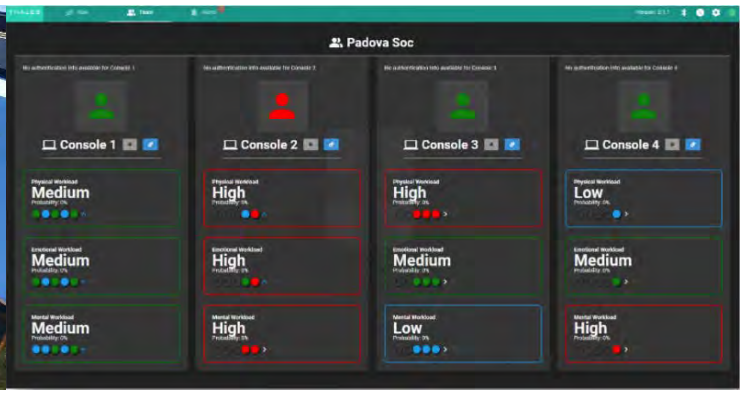
With the tool:

- Workload assessment is explicit, objective and continuous

HOW IS IT DEPLOYED IN IMPETUS?

- **Who are the users**: SOC operators and supervisors, IT specialists, behavioural scientists, stress analysts.
- **What are the critical situations for deployment**: The tool and its sensors are unobtrusive and can be deployed continuously while operators are working, including during emergencies.





HOW DOES IT WORK?

Each operator wears an unobtrusive wearable headband which detects bio-signals (pulse, brain waves) and transmits these to the tool. Operator workload is predicted based on personalized, pre-trained (machine learnt) models. The tool can be used at individual and team levels. The supervisor is alerted when an anomaly is detected.
The graphical user-interface provides the supervisor with an overview of:

- Workload status of each team member, including trends over-time
- Alerts related to:
  - sensor data availability (e.g. in case of sensor failure)
  - workload (too high/too low) for any of the operators

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883286.

The IMPETUS Platform

Integrates multiple tools in a unified interface

WHAT PROBLEM DOES THE TOOL HELP SOLVE?

People involved in security operations often need to deal with multiple tools at the same time. At a given moment they may be interacting directly with just one specific tool – but they need to be made immediately aware of critical situations that other tools may have detected. If tools interact with users via separate interfaces, it can be very difficult for staff to work effectively, especially in stressful situations. Also: different users may have different perceptions of the overall situation depending on which tools they happen to be using.

The IMPETUS platform provides a way to combine multiple tools in a unified interface, so that users who need to interact with multiple tools can do so in one place. It shows the status of all the tools (example: an urgent alert has been raised) and allows an operator to interact with a specific tool to get more information. It supports common situational awareness as different operators have the same overall view. It also offers possibilities to produce customised interfaces fine-tuned to the needs of different users (depending on their role, some users might be primarily interested in different subsets of the tools available).

The platform already supports integration with the tools developed in the IMPETUS project, but it is designed in an open way so that other tools (ones already in use by an organisation, or new ones they might acquire in future) can also be integrated.

HOW IS IT DEPLOYED IN IMPETUS?

- **Who are the users**: Emergency and security centre operators and their supervisors; IT analysts and technicians; other staff responsible for monitoring and dealing with urban security.
- **What are the critical situations for deployment:** Continuous: Security is a 24/7 operation.

HOW DOES IT WORK?

The platform provides a central dashboard integrating tools to allow monitoring of potential threat events as they arise. There is a main dashboard showing the overall status of all tools, and tool-specific dashboards to allow more detailed interaction with specific tools.
Alerts are shown with different levels of priority, and indications of whether they have been acknowledged and/or resolved. Where feasible, data is presented graphically for easy visualization, and a map is provided to show where the event occurred. Comments can be associated with alerts and shared with other users.
The platform was implemented using the Snap4City platform: http://www.snap4city.org/

This project receives funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883286.

14 Annex 2: User Manuals

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

http://www.impetus-project.eu

BD
Bacteria Detection

Authors: Alexia Comte (UdN), Mathieu Tur (UdN), Sébastien Courtin (UdN), Axelle Cadière (UdN), Sandrine Bayle (IMT)

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Table of contents

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

1 Table of figures

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

2 List of tables

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

3 References

ATP-metry. (n.d.). Retrieved from http://www.atp-metry.com/

GLBiocontrol. (n.d.). Retrieved from https://www.gl-biocontrol.com/

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

1.General Information

1.1 Goal

Role of Bacteria Detection (**BD**) tool is to measure continuously bacteria concentration in the ambient air to prevent citizens from biological risks. It is connected to IMPETUS platform for centralizing data with local authorities.
This manual intends to guide the operator through BD user interface to process automated measurements and handle maintenance.

1.2 Principle

BD tool principle is to combine an air biocollector (developed by IMT Alès / University of Nîmes) and a bacteria concentration measurement device (Glow'N'Care (**GNC**) from GLBiocontrol Company [1][1]). First one acquires a sample of ambient air and traps the bacteria into water. Then second one analyses this water to retrieve ambient air bacteria concentration. The data is finally sent to IMPETUS platform and an alert is trigged if the measurement exceeds a defined threshold (Figure 1):

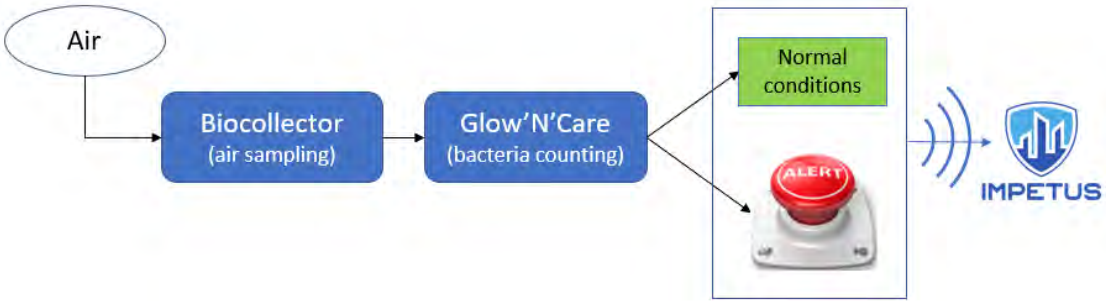


Figure 1: BD Principle

1.2.1 Air Biocollector

The air biocollector was designed for IMPETUS project to collect and concentrate air microorganisms in a water volume, it relies on a glassware called impinger (Figure 2):

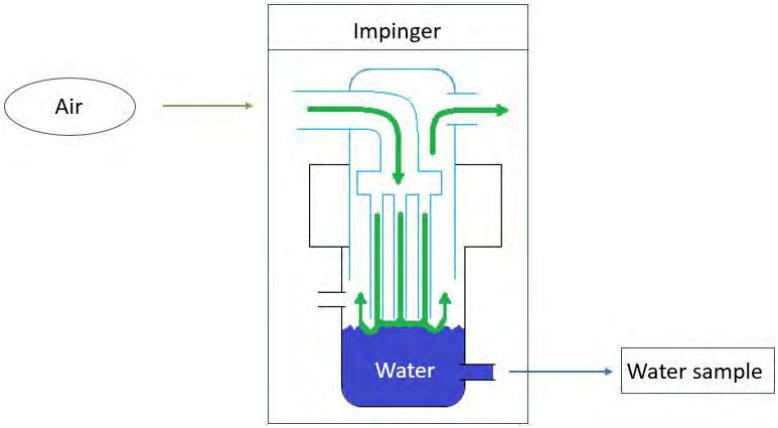


Figure 2: Biocollector principle

[1] GLBiocontrol company: http://www.gl-biocontrol.com/

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

1.2.2 Glow'N'Care device

This bacteria concentration measurement device is initially designed for monitoring concentration of microorganisms (biomass) in water, it relies on ATP-metry technology [2][2]: measurement of ATP concentration by bioluminescence (Figure 3):

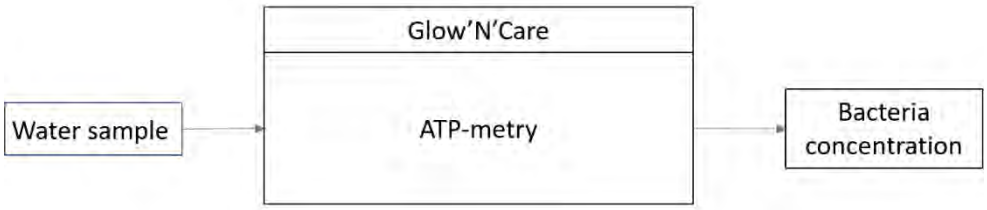


Figure 3: Glow'N'Care (GNC) principle

[2] ATP-metry technology: http://www.atp-metry.com/

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

2. Description

Biocollector and Glow'N'Care device are both included into two electrical enclosures, as it can be seen in Figure 4:

Figure 4: BD picture

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

2.1 Biocollector impinger box

Biocollector role is to collect air microorganisms in a volume of water. This is done by using a pump to collect an air sample from the BD inlet and impact it on a water surface contained in the impinger. In the meanwhile, peristaltic pumps oversee cleaning, bleaching, rinsing, and filling the impinger to ensure getting water free of residual bacteria (Figure 5):

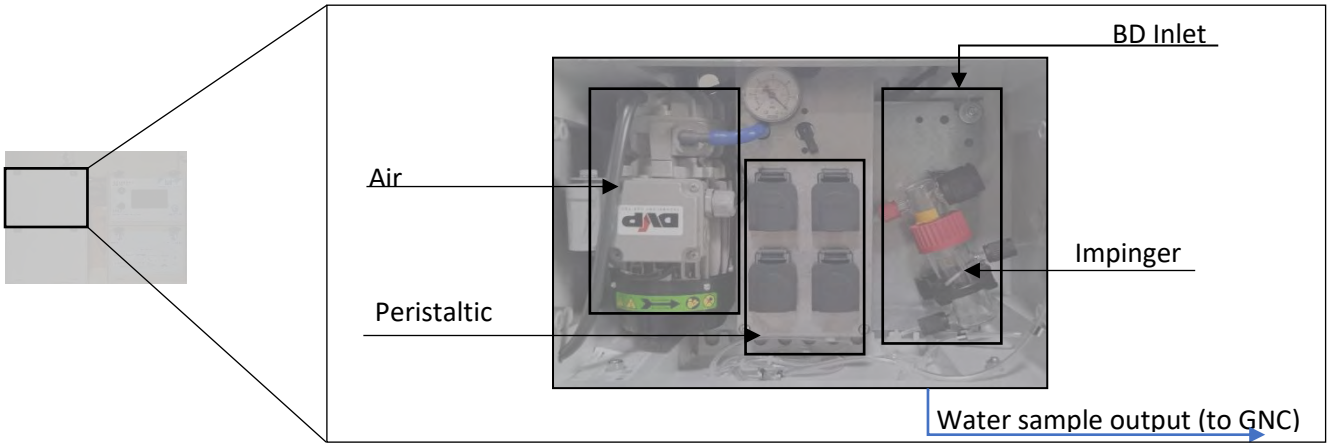


Figure 5: Biocollector Impinger box

Once impinger water has collected bacteria from the air sample, water sample is ready to be sent to GNC device.

2.2 Biocollector supplies box

This box contains four tanks used for handling liquids involved in biocollection: water, bleach, cleaning solution and sewer, tanks positions are shown in Figure 6:

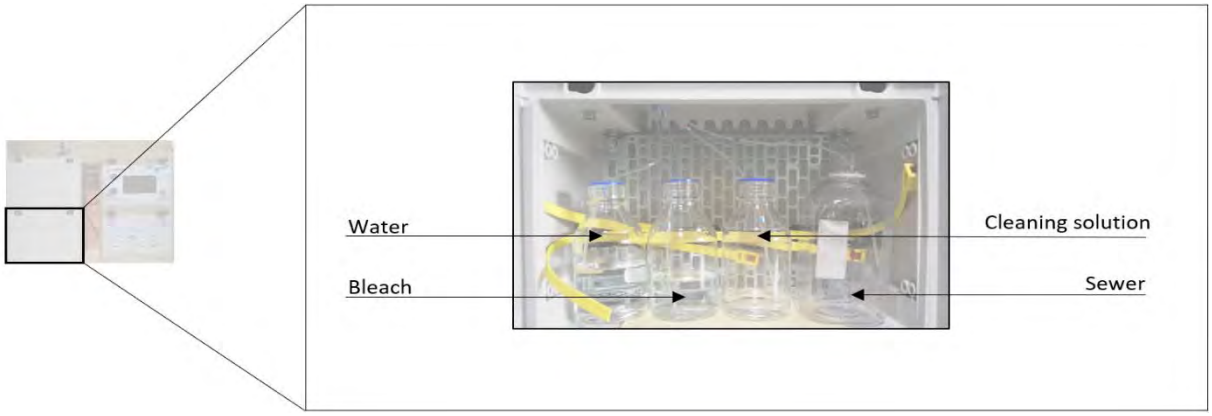


Figure 6: Biocollector supplies box

These tanks need to be checked before starting every automated measurement sequence.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

2.3 Glow'N'Care front panel box

This box is responsible for managing and monitoring automated measurement sequences through a user interface presented on a touch screen computer where software is implemented (Figure 7):

Figure 7: Glow'N'Care front panel

Inside the box: electrical and electronical circuits controlling sensors and actuators of the whole BD device. The operator will be involved in checking temperature and stirrer settings (Figure 8) before starting measurements:

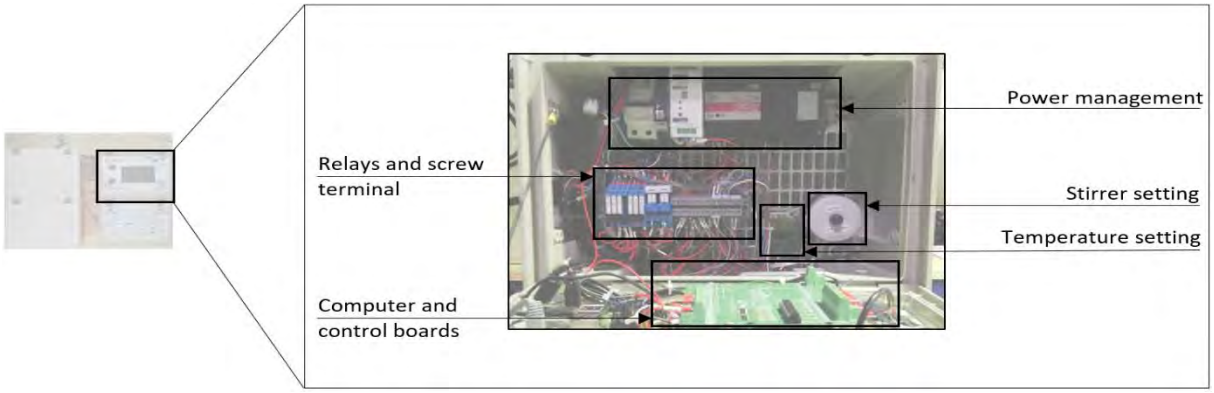


Figure 8: Glow'N'Care front panel under the hood

This box is not intended to be opened while using BD, risk of electric hazard!

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

2.4 Glow'N'Care machinery

GNC machinery receives the water sample from biocollector and measures the bacteria concentration using ATP-metry [2], which involves pumps, circuitry, cold chemical reagents, and a photon counting (measurement) cell as shown in Figure 9:

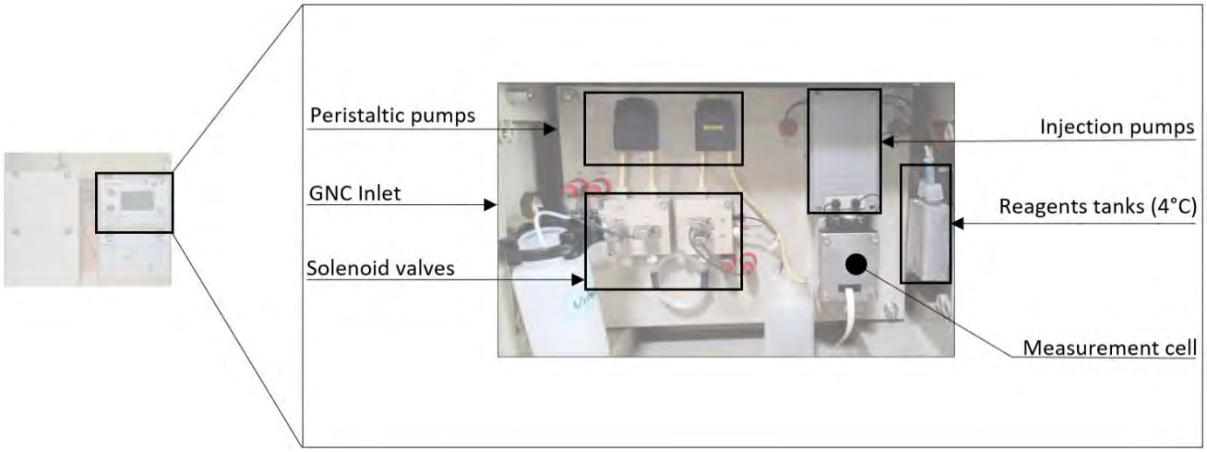


Figure 9: Glow'N'Care machinery

Peristaltic pumps manage water sample acquisition with the help of solenoid valves and fills/empties measurement cell while injection pumps act on adding reagents inside it.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

3. User interface

On Glow'N'Care front panel box there is a touch-screen computer where is displayed BD tool user interface, for controlling and monitoring automated measurement sequences (Figure 10):

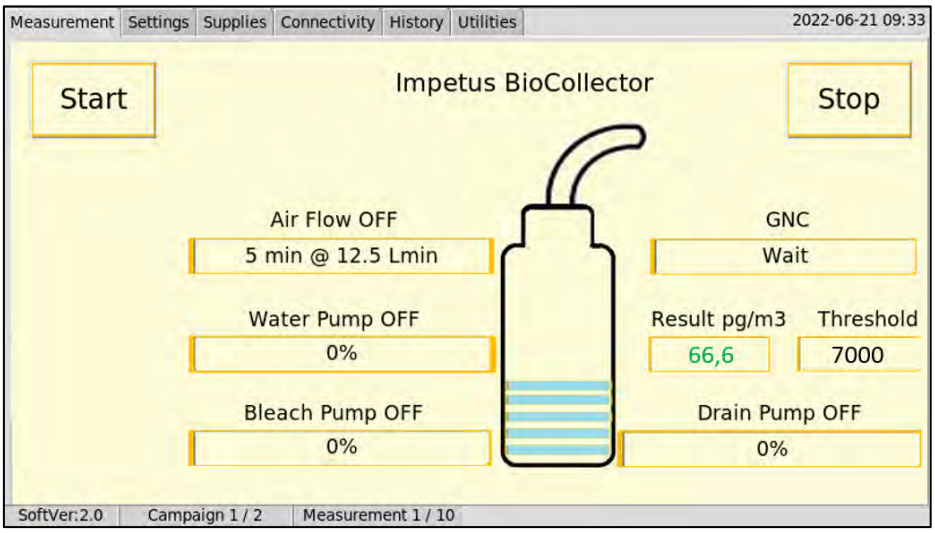


Figure 10: BD user interface

Several options are available through six tabs: "Measurement", "Settings", "Supplies", "Connectivity", "History" and "Utilities".

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

3.1 Measurement tab

Main tab of the BD user interface, from where the operator can start an automated measurement sequence and follow progress of its several steps. A dynamic display allows to monitor main actuators statuses continuously, as described in Figure 11 and Table 1:

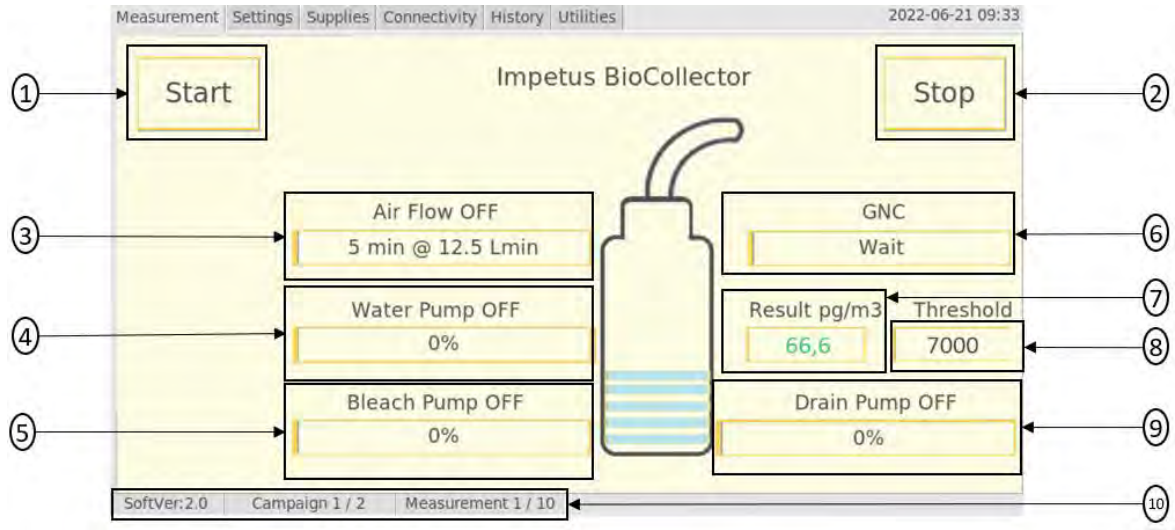


Figure 11: Measurement tab

Ref.	*Type*	*Name*	*Action / Description*
1	*Button*	*Start*	Start an automated BD measurement sequence
2	*Button*	*Stop*	Stop an automated BD measurement sequence
3	*Indicator*	*Air Flow*	Air flow status, settings, and progress
4	*Indicator*	*Water Pump*	Water pump status and progress
5	*Indicator*	*Bleach/Cleaning Pump*	Bleach/Cleaning pump status and progress
6	*Indicator*	*GnC*	Glow'N'Care device current step
7	*Indicator*	*Result pg.m^{-3}*	Last bacteria concentration measured
8	*Indicator*	*Threshold*	Acceptable bacteria concentration level setting
9	*Indicator*	*Drain Pump*	Drain pump status and progress
10	*Indicator*	*Status bar*	Software version and sequence status

Table 1: Measurement tab user interactions

This tab display will update continuously during the automated measurement process, monitoring steps executed by the BD. A detailed description is available in paragraph 5.2 Sequence chronology.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

3.2 Settings tab

This tab allows the operator to set automated measurement sequence parameters, detailed in Figure 12 and Table 2. It is also useful while setting up BD tool.

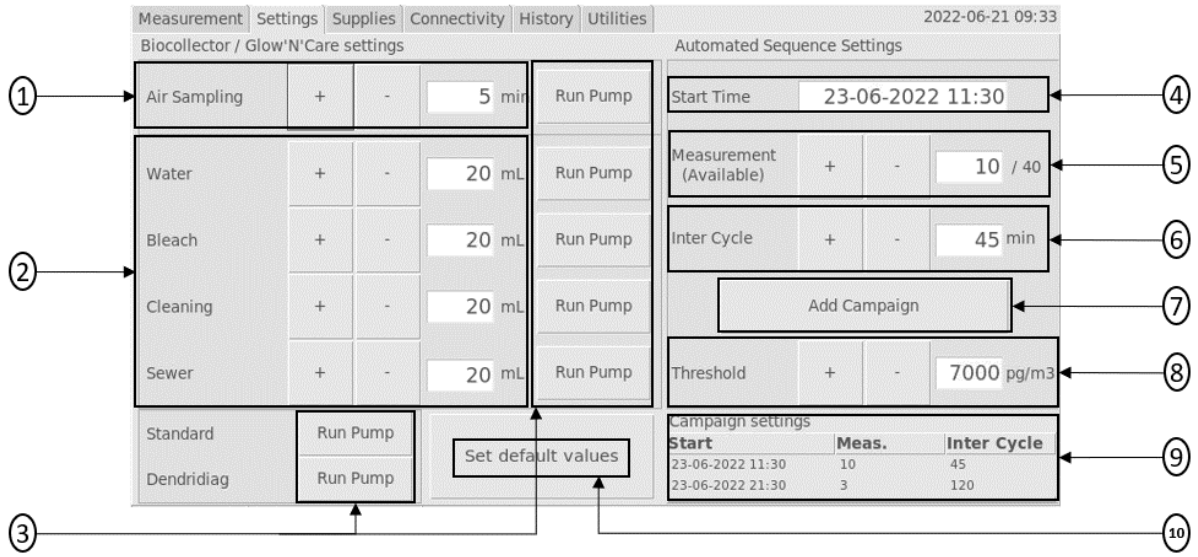


Figure 12: Settings tab

Ref.	*Type*	*Name*	*Action / Description*
1	*Control*	*Air Sampling*	Ambient air sampling time
2	*Controls*	*Biocollector Volumes*	Respective volume involved in biocollector
3	*Buttons*	*Run pump*	Launch respective pump
4	*Control*	*Start Time*	Desired start time
5	*Control Indicator*	*Measurement (Available)*	Number of automated measurement(s) in the sequence / Measurements available
6	*Control*	*Inter Cycle*	Sleep time between two measurements
7	*Button*	*Add Campaign*	Allows to add another sequence rhythm*
8	*Control*	*Alert Threshold*	Acceptable bacteria concentration level setting
9	*Indicator*	*Campaign settings*	Display of sequence settings
10	*Button*	*Set default values*	Record all numeric controls of the tab as default values, even after BD shutdown

Table 2: Settings tab user interactions

* BD allows to use a second rhythm, e.g., for night/day sequences (see paragraph 4.2.5 Setting measurement parameters).

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

3.3 Supplies tab

This tab allows the operator to monitor and handle BD supplies, for checking if BD is going to run out of one of them and updating their levels when refilling (Figure 13 and Table 3):

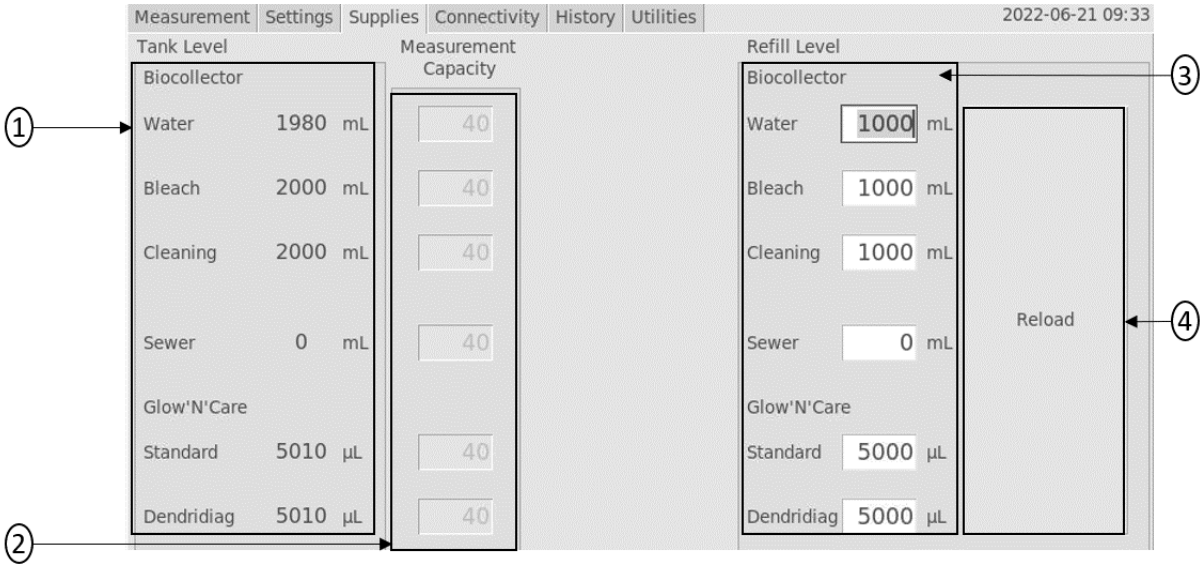


Figure 13: Supplies tab

Ref.	*Type*	*Name*	*Action / Description*
1	*Indicator*	*Tanks level*	Available volume in respective tank
2	*Indicator*	*Measurements Capacity*	Number of measurements feasible with respective available volumes
3	*Control*	*Refill levels*	Fill these fields with corresponding volumes when reloading supplies
4	*Button*	*Reload*	Click this button to refresh display with refilled volumes

Table 3: Supplies tab user interactions

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

3.4 Connectivity tab

As BD tool is a connected device (thanks to its internal LTE router), this tab is intended to set and test communication parameters with IMPETUS platform, but it can also be used to set and test sending results file to an ftp server or an e-mail address at each end of measurement sequence (Figure 14 and Table 4):

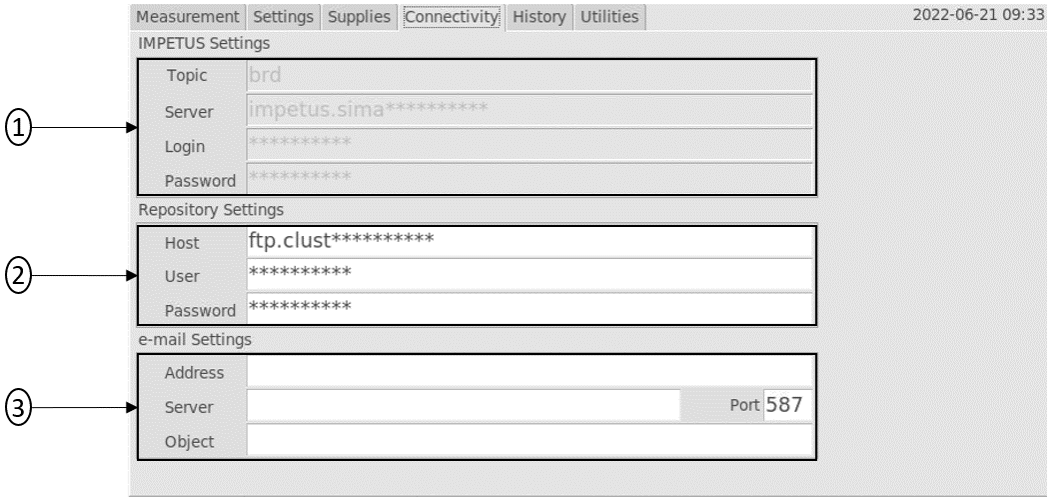


Figure 14: Connectivity tab

Ref.	*Type*	*Name*	*Action / Description*
1	*Control*	*IMPETUS settings*	Credentials given by IMPETUS to communicate with the platform
2	*Control*	*Repository settings*	If desired, enter here host address and credentials of your choice to upload a results file (.csv) on an ftp server of your choice at each end of automated sequence
3	*Control*	*E-Mail settings*	If desired, enter here e-mail settings for receiving a results file (.csv) at each end of automated sequence

Table 4: Connectivity tab user interactions

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

3.5 History tab

A tab showing every measurement result since BD tool has been turned on (Figure 15):

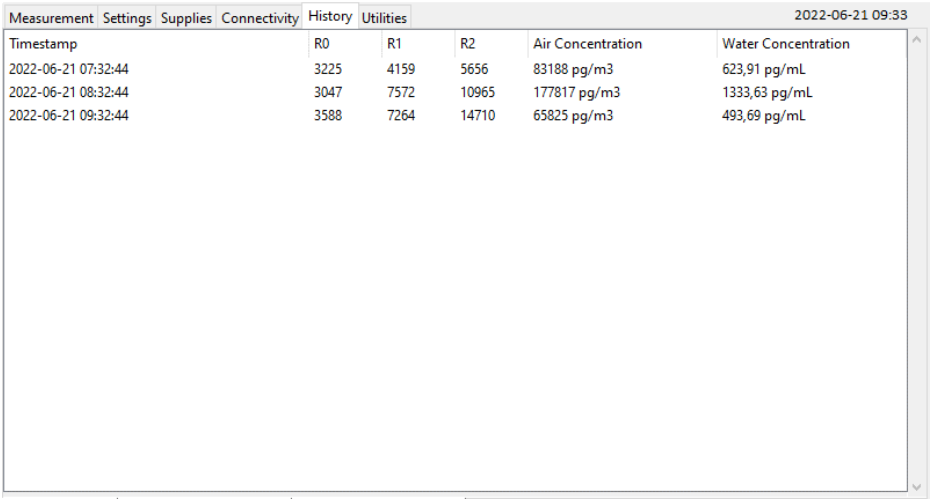


Figure 15: History tab

Values logged by the system are timestamp, intermediate ATP-metry [2] values (R_0, R_1 and R_2), bacteria concentration in water and in air sample.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

3.6 Utilities tab

This tab is intended to be mainly used for maintenance purpose, as it monitors continuously BD elements statuses such as computer processor temperature or memory usage. It also allows to exit application window or going through detailed monitoring and actions, as described in Figure 16 and Table 5:

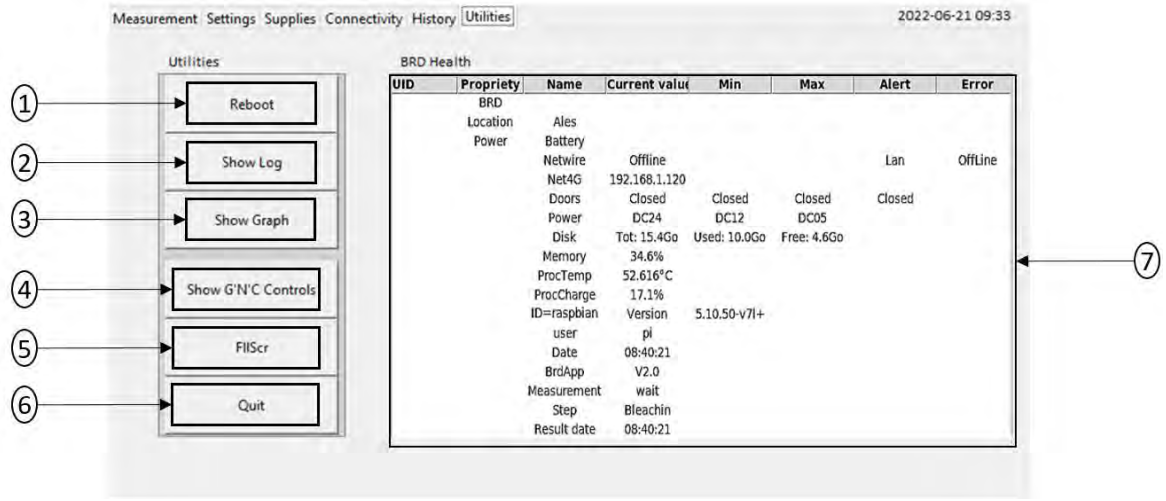


Figure 16: Utilities tab

Ref.	*Type*	*Name*	*Action / Description*
1	*Button*	*Reboot*	Reboot BD computer
2	*Button*	*Show Log*	Display a pop-up window showing every low-level action done by the BD*
3	*Button*	*Show Graph*	Display a pop-up window monitoring continuously photo counting device value*
4	*Button*	*Show GnC controls*	Display a pop-up window allowing to control independently GNC actuators *
5	*Button*	*Fullscreen*	Toggle user interface full screen / window mode
6	*Button*	*Quit*	Quit BD application*

Table 5: Utilities tab user interactions

* Pop-ups contents are detailed in **¡Error! No se encuentra el origen de la referencia.**

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4. Getting started

4.1 BD Setup

To setup BD properly, first identify biocollector and Glow'N'Care device (

Figure 17):

Figure 17: Identifying biocollector and Glow'N'Care device

Put the two elements side by side and identify connections of biocollector (Figure 18 a) and GNC device (Figure 18 b):

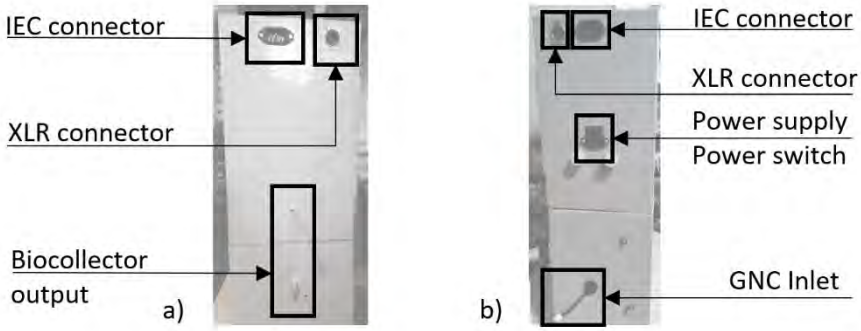


Figure 18: Biocollector (a) and GNC (b) connections

Connect IEC and XLR cable between the two elements, then screw the fitting from biocollector output and GNC inlet (Figure 19 a and b):

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

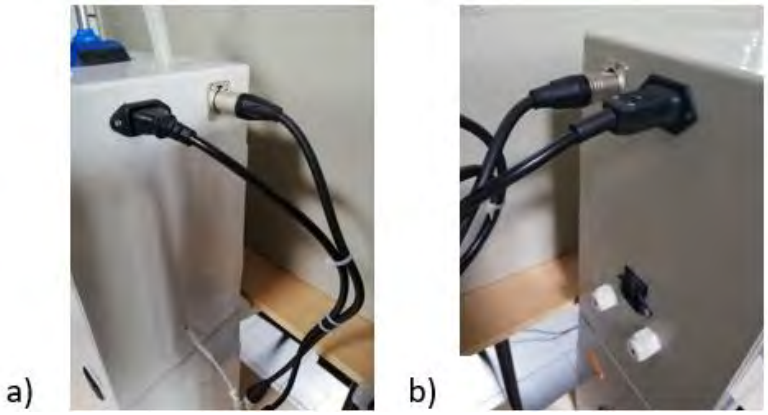


Figure 19: Biocollector (a) and GNC (b) connected

The BD can now be plugged in and turned on (switch near the power supply input). The CNG front panel box boots up and displays the user interface (Figure 20)

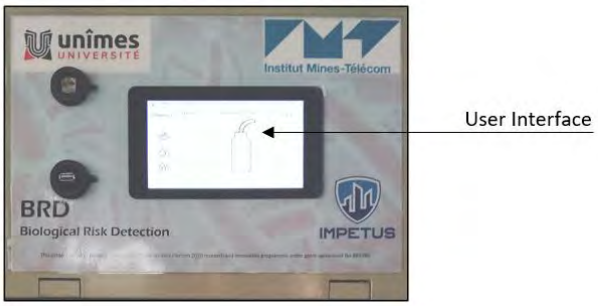


Figure 20: User interface display

Open carefully GNC front panel box and check temperature and stirrer settings (Figure 21):

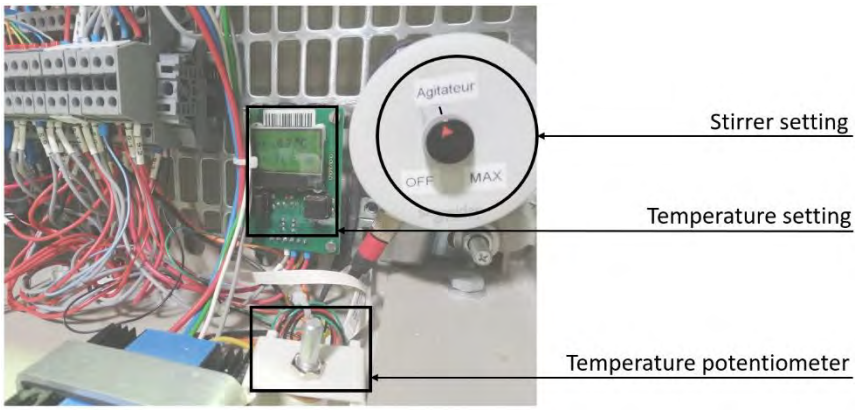


Figure 21: Stirrer and temperature settings

Adjust manually stirrer potentiometer position to be as close as possible to Figure 21 (adjust to black line), then set temperature setting between 4°C and 8°C with the help of the potentiometer.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4.2 Getting started

Once that BD tool is setup, this sub-chapter describes the routine to prepare every automated measurement sequence. The operator, by following each iteration, will be able to check the system, set parameters and start the sequence.

4.2.1 Biocollector supplies checking

Biocollector involves several liquids which will be overseen by four peristaltic pumps. To avoid damage in case of misuse, the operator must ensure that, before starting, supplies are ready to be used properly, by opening biocollector supplies box and filling/emptying tanks as described in Figure 22 and below:

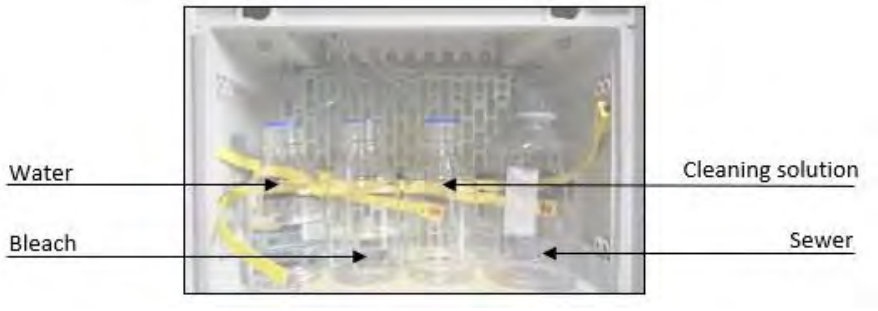


Figure 22: Supplies level checking

Water, bleach, and cleaning solution tanks must be filled to their maximum capacity while sewer tank needs to be empty.

All volumes introduced need to be now enquired in BD software, by going to the user interface in "Supplies" tab and entering values in all corresponding fields. Click then on "Reload" button, the tab will refresh with new levels, this will let BD know how many measurements are achievable.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4.2.2 Biocollector pumps priming

Biocollector liquids management is overseen by four peristaltics pumps, handling respectively water, bleach, cleaning solution and sewer (drain). Before starting, operator needs to ensure that pumps are correctly primed by following next steps:

Still in the biocollector impinger box, identify corresponding pumps (Figure 23) and prepare a recipient to receive liquid overflows.

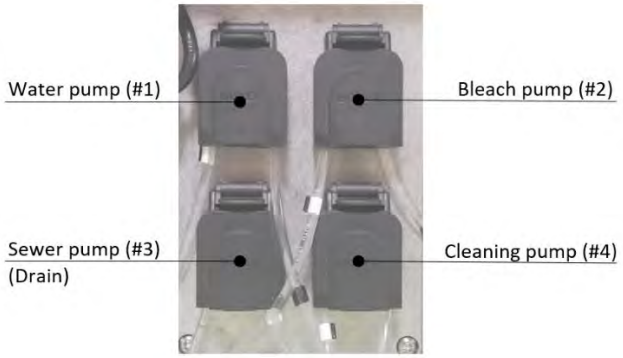
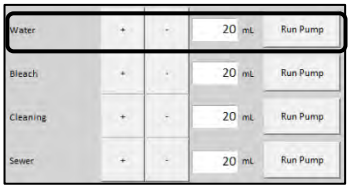


Figure 23: Biocollector pumps

First, unplug water pipe between water pump (#1) and impinger and point it in the recipient, then go to user interface "Settings" tab and click on "Run Pump" button corresponding to water. Wait until no more air bubbles appear in the circuit, then replug the pipe, water circuit is now in charge.



Then, do the same for bleach and cleaning solution pumps, note that the sewer (drain) pump doesn't need priming.

4.2.3 Biocollector checking

To ensure ambient air sample acquisition in good conditions, the operator needs to ensure air-flow is well tuned and impinger well setup. It has to be done by opening the biocollector impinger box (

Figure 24) and following next instructions:

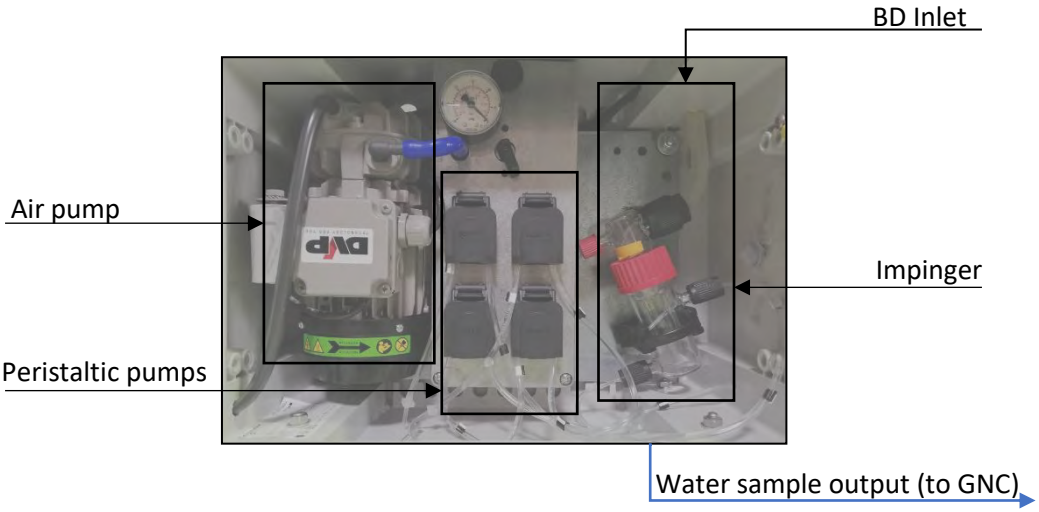


Figure 24: Biocollector checking

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Check impinger position

The impinger is made of two parts, a receiver and a cap which can manually slides up and down (Figure 25):

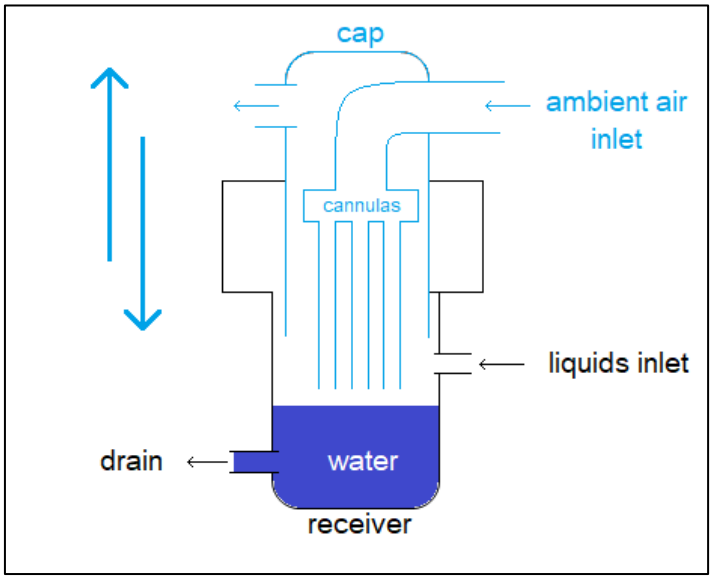
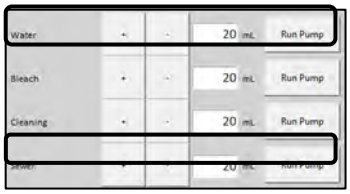


Figure 25: Impinger diagram

For an optimal use of the BD, the cannulas have to be close to the water (5-10mm) in order for ambient air to impact its surface and exchange a maximum of bacteria. Before starting an automated measurement sequence, operator needs to check impinger cap position by doing the following:

Go to user interface “Settings” tab:



1. If the impinger is not empty, click on “Run Pump” corresponding to sewer.
2. Once impinger empty, click on “Run Pump” correponding to water and wait a few seconds.
3. Adjust manually lowest cannula level sliding the cap up or down in order to get as close as possible to the following setup (Figure 26):

Figure 26: Optimal Impinger cap position

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Check air pump flow

The BD measurement process requires an incoming air flow of 30L.min^{-1}, before starting, operator needs to check it directly at BD inlet, while setting air pump until matching requirements:

Go to user interface “Settings” tab :

1. Point an air-flow meter at the BD inlet ,then click on “Run Pump” corresponding to air sampling.
2. Set air pump flow regulation valve (Figure 27) until air-flow meter indicates 30L.min^{-1}, then click again on “Run Pump” button.

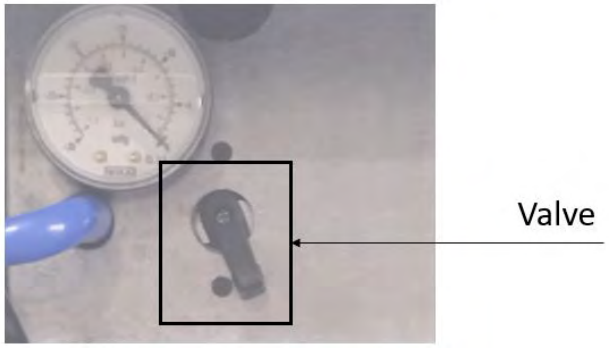


Figure 27: Air pump flow regulation valve

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4.2.4 Glow'N'Care reagents management

To process ATP-metry [2] in order to reveal bacteria presence in the water sample from biocollector, GNC involves two chemical reagents that operator should have in hands: "Standard" solution and "Dendridiag" solution, supplied by the GLBiocontrol company [1].

As for supplies liquids seen before, operator needs to check also that chemical reagents pumps are primed before starting, by opening Glow'N'Care machinery box (Figure 28) and following next steps:

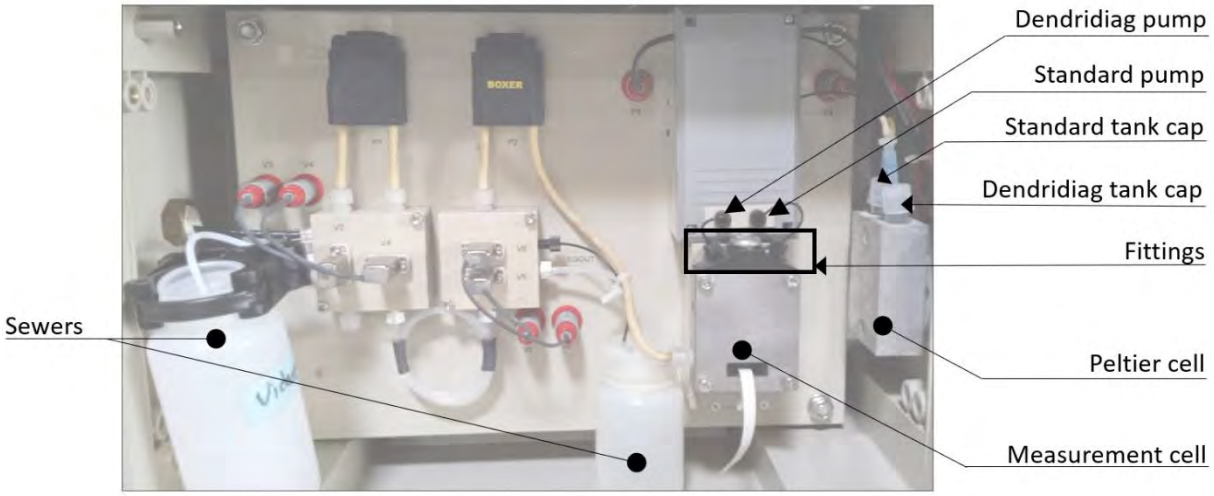


Figure 28: Glow'N'Care reagents management

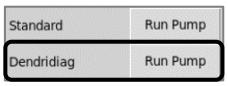
Uncap the 2 sewer tanks and empty them if needed, this will prevent from liquids overflow.

In the Peltier cell, put the "Standard" solution in the tank (the back one) and put the needle from Standard pump into it.
Unscrew the upper right black fitting (follow standard pump output) of the measurement cell and prepare a recipient at output in order to receive the overflow.

Go to user interface "Settings" tab and click on "Run Pump" corresponding to standard. Wait until the reagent drops out of the fitting and screw it back on.

Still in the Peltier cell, put the "Dendridiag" solution into the container (the front one) and put the needle from pump 4 (Dendridiag) into it.
Unscrew the upper left black fitting (follow dendridiag pump output) of the measurement cell and prepare a recipient at output in order to receive the overflow.

Go to user interface "Settings" tab and click on "Run Pump" corresponding to Dendridiag. Wait until the reagent drops out of the fitting and screw it back on.



The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4.2.5 Setting measurement parameters

Now that BD is ready to run properly, operator can set parameters of the automated measurement sequence by going to user interface in "Settings" tab (Figure 29), available options are described in Table 6: Automated measurement sequence settings:

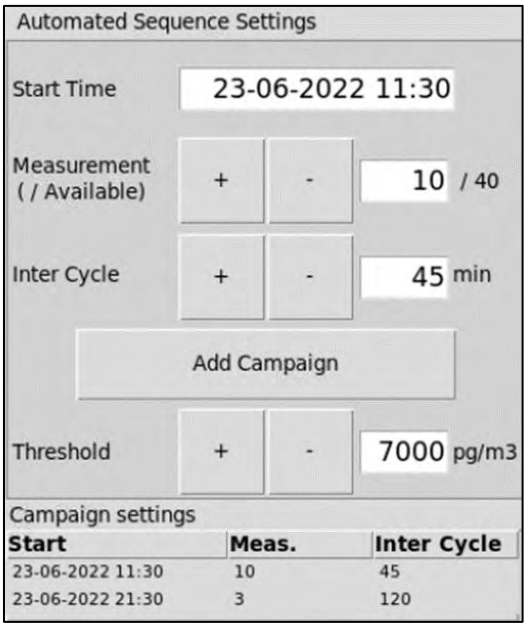


Figure 29: Automated measurement sequence settings

Type	*Name*	*Action / Description*
Control	*Start Time*	Desired start time
Control	Measurement *(/ Available)*	Number of automated measurement(s) in the sequence / Measurements available
Indicator	Available	Number of available measurements
Control	Inter Cycle	Sleep time between two measurements
Button	*Add Campaign*	Allows to add another sequence rhythm*
Control	*Threshold Alert*	Acceptable bacteria concentration level setting
Indicator	*Campaign settings*	Display of sequence settings

Table 6: Automated measurement sequence settings

* Use this option to configure another sequence after the first one (planning will be displayed in "Campaign settings"). Figure 29 shows an example where BD will start at 11:30 and run 10 times with a pause of 45 minutes between measurements, then restart at 21:30 and run 3 times with a pause of 120 minutes.

Operator is now able to check if BD tool can start by going to user interface "Measurement" tab and check if "Start" button is enabled. A greyed "Start" button means that BD is not ready, please check chapter 6: Troubleshooting to understand what is causing unexpected behavior.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

5. Automated measurement sequence monitoring

This chapter describes the behaviour of the User Interface during BD automated measurement sequence.

5.1 Starting a sequence

The BD automated measurement sequence starts when the operator clicks on "Start" button in the "Measurement" tab in the user interface as shown in Figure 30:

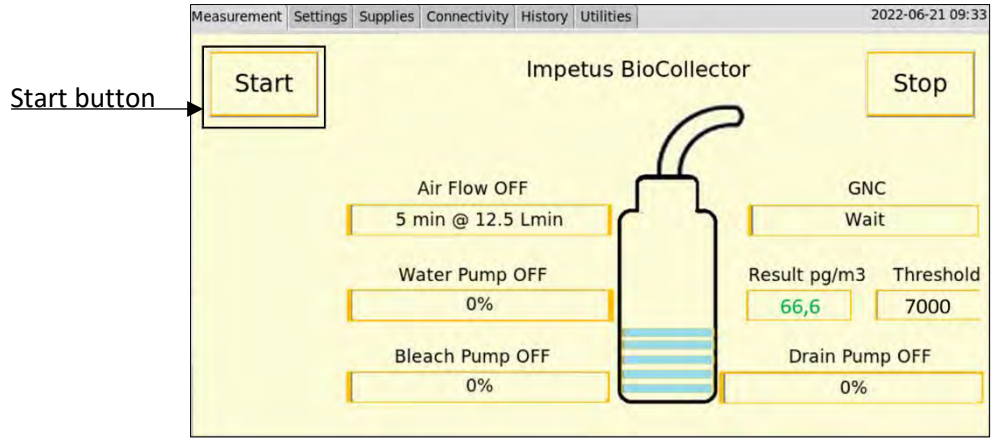


Figure 30: Starting an automated measurement sequence

5.2 Sequence chronology

As described in chapter 1.2, BD tool combines a biocollector and GNC device, both are overseen simultaneously by the software. Once sequence is launched, operator will be able to follow every step through indicators "Biocollector step" and "GNC step", user interface also monitors main BD actuators status as described in Figure 31:

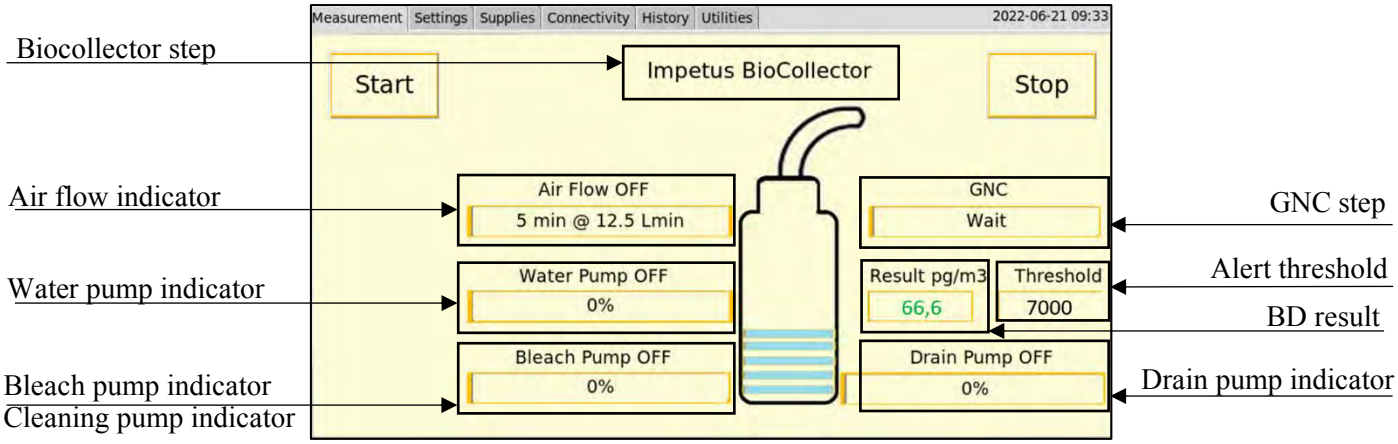


Figure 31: Sequence display

Progress of the automated measurement sequence is described in the next subchapters and Figure 49.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

5.2.1

Step 1: Biocollector and GNC Purge

Impinger residual content is removed (Figure 32), idem for GNC measurement cell (Figure 33):

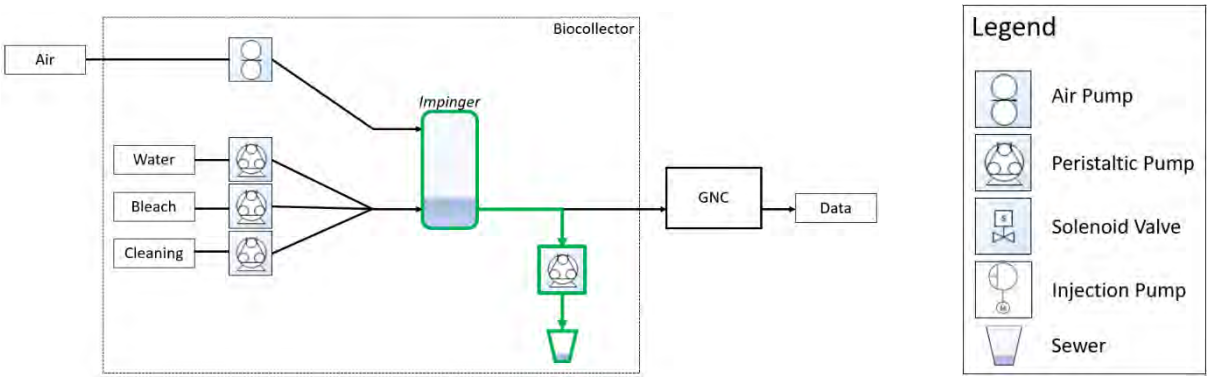


Figure 32: Biocollector step 1: Purge

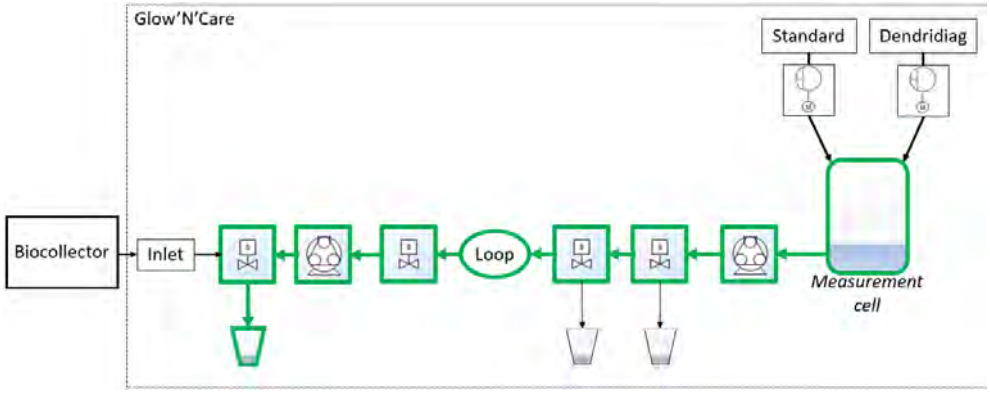


Figure 33: GNC step 1: Purge

5.2.2 Biocollector

step 2 and 3: Cleaning / Purge

Impinger is filled up with cleaning solution (Figure 34), then purged (Figure 35)

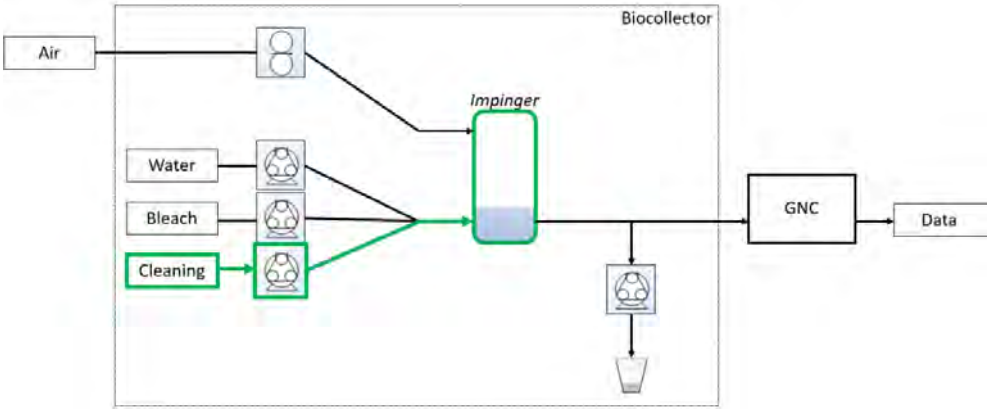


Figure 34: Biocollector step 2: Cleaning

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

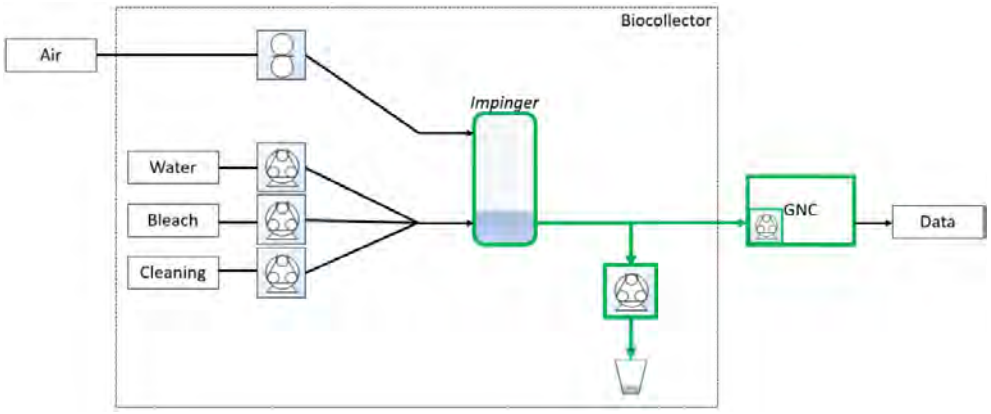


Figure 35: Biocollector step 3: Purge

During the purge, 8 mL volume of cleaning solution is acquired by GNC for cleaning itself.

5.2.3 GNC

step 2 and 3: Cleaning / Purge

Cleaning solution from biocollector is sent through the whole GNC circuit (Figure 36), then purged (Figure 37):

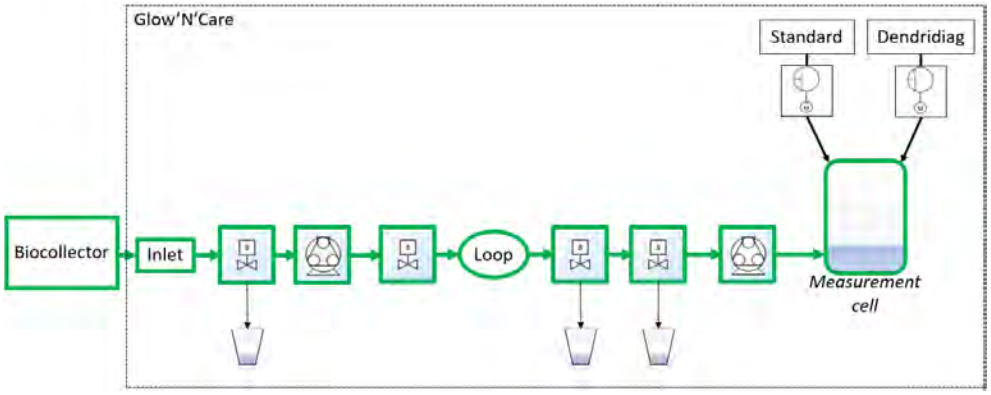


Figure 36: GNC step 2: Cleaning

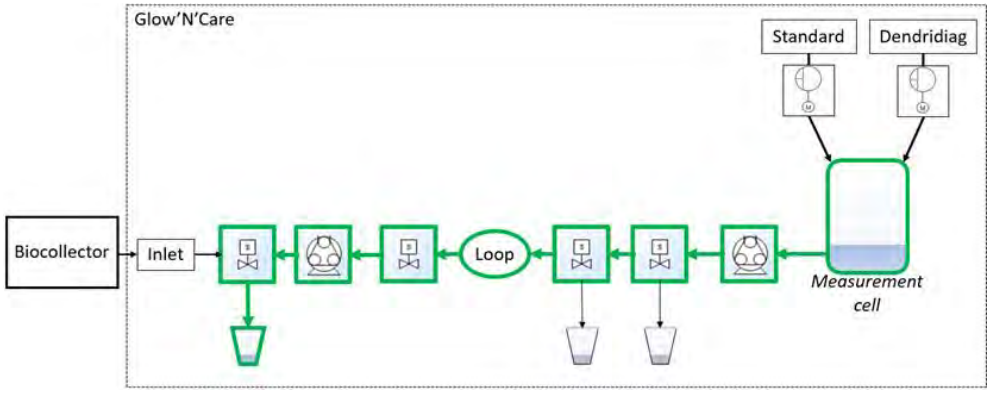


Figure 37: GNC step 3: Purge

5.2.4 Biocollector

step 4 and 5: Rinsing / Purge

Impinger is filled up with water (Figure 38), then purged while 8 mL volume of water is acquired by GNC for rinsing itself (Figure 35).

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

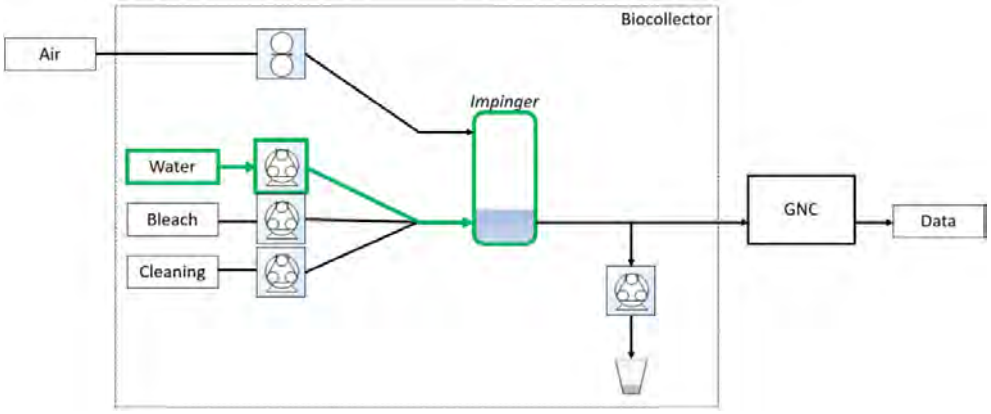


Figure 38: Biocollector step 4: Rinsing

5.2.5 GNC

step 4 and 5: Rinsing / Purge

Water from biocollector is sent through the whole GNC circuit (Figure 36) and then purged (Figure 37), as in paragraph 5.2.3 GNC

step 2 and 3: Cleaning / Purge

5.2.6 Biocollector step 6 and 7: Bleaching / Purge

Impinger is filled up with bleach (Figure 39), then purged while 8 mL volume of bleach is acquired by GNC for bleaching itself (Figure 35)

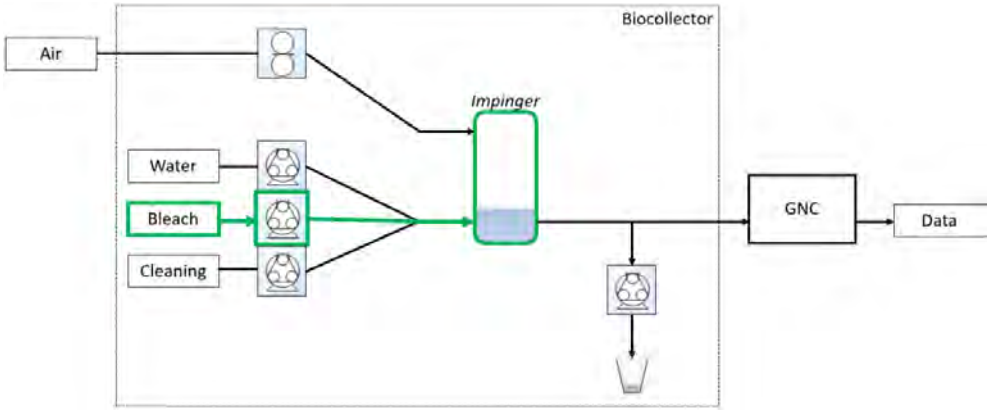


Figure 39: Biocollector step 6: Bleaching

5.2.7 GNC

step 6 and 7: Bleaching / Purge

Bleach from biocollector is sent through the whole GNC circuit (Figure 36) and then purged (Figure 37), as in paragraph 5.2.3 GNC

step 2 and 3: Cleaning / Purge

5.2.8 Biocollector

step 8 and 9: Rinsing / Purge

Impinger is filled up with water (Figure 38), then purged while 8 mL volume of water is acquired by GNC for rinsing itself (Figure 35) as in paragraph 5.2.4 Biocollector

step 4 and 5: Rinsing / Purge

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

5.2.9 GNC

step 8 and 9: Rinsing / Purge

Bleach from biocollector is sent through the whole GNC circuit (Figure 36) and then purged (Figure 37), as in paragraph 5.2.3 GNC

step 2 and 3: Cleaning / Purge

5.2.10 Biocollector

step 10 and 11: Biocollector filling / Air sampling

Impinger is filled up with water (Figure 40) to be ready for air sampling (Figure 41)

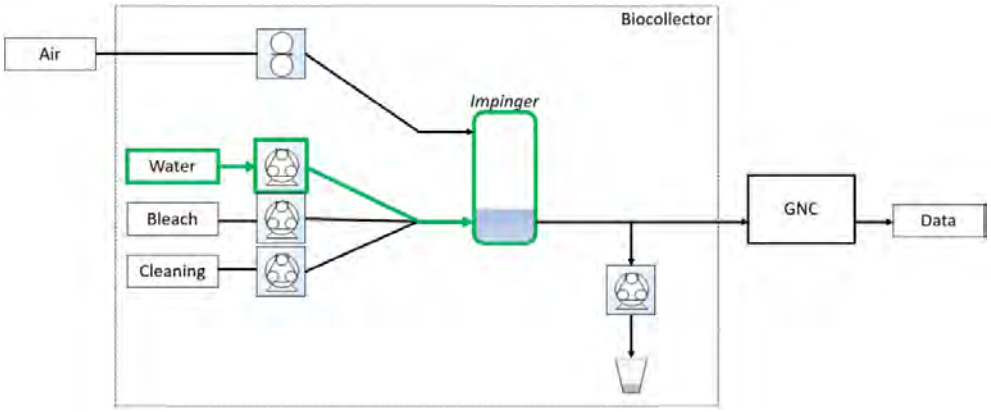


Figure 40: Biocollector step 10: Biocollector filling

This step acquires the ambient air sample from BD inlet to collect bacteria into impinger water (Figure 41):

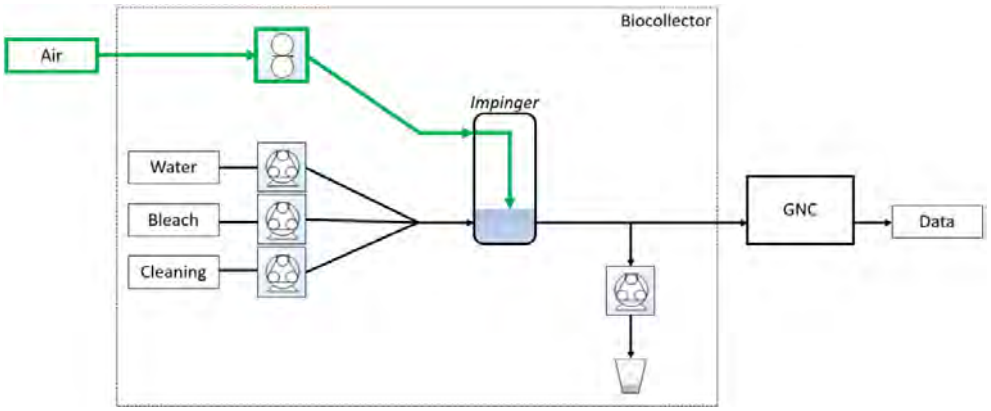


Figure 41: Biocollector step 11: Air sampling

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

5.2.11 Biocollector

step 12: Measure

Impinger water is pumped by GNC (Figure 42) for bacteria counting:

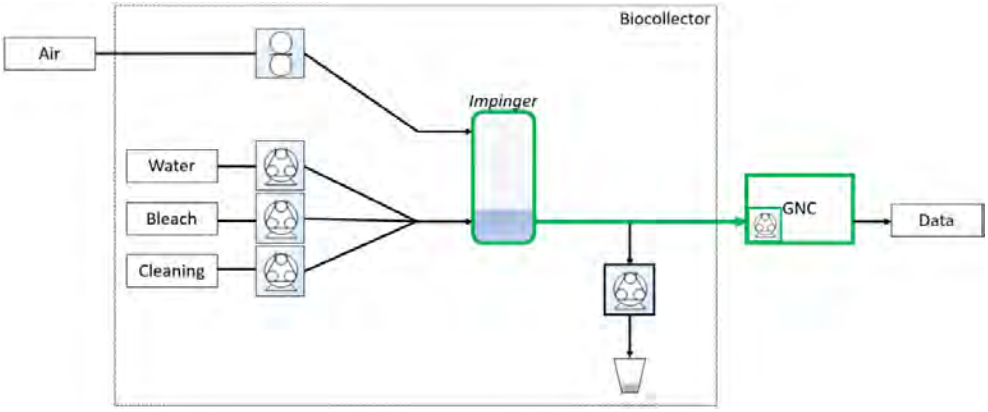


Figure 42: Biocollector step 12: Measure

5.2.12 GNC

step 10: Loop filling

Biocollector water sample is stored in a first part of GNC circuit (loop) as shown in Figure 43:

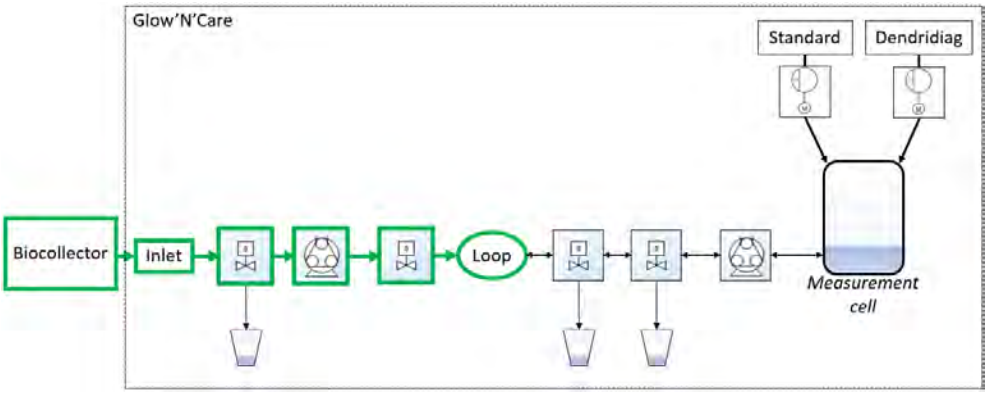


Figure 43: GNC step 10: Loop filling

5.2.13 GNC

step 11: Cell filling

Measurement cell is filled up with 1mL of loop content (Figure 44):

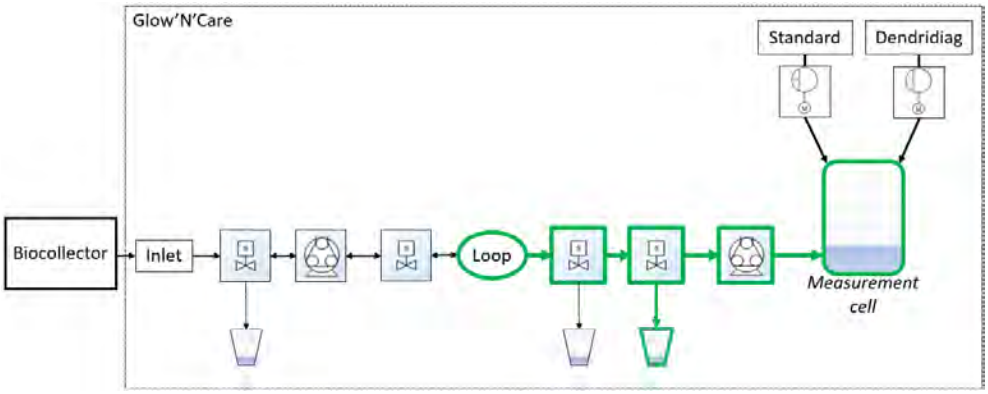


Figure 44: GNC step 11: Cell filling

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

5.2.14 GNC

step 12: R0 Measure

Content of measurement cell is stirred; emitted light is measured by the photomultiplier to obtain R_0 (background light) value (Figure 45):

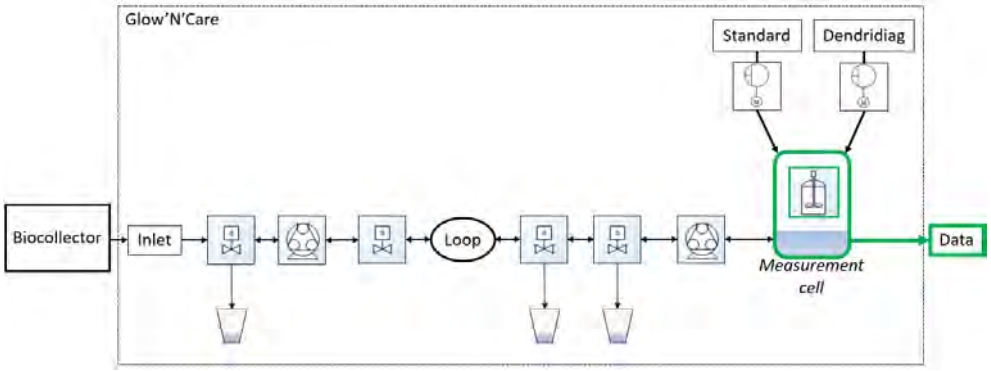


Figure 45: GNC step 12: R_0 Measure

5.2.15 GNC

step 13: Dendridiag

Dendridiag reactant is introduced into measurement cell and stirred with water sample (Figure 46):

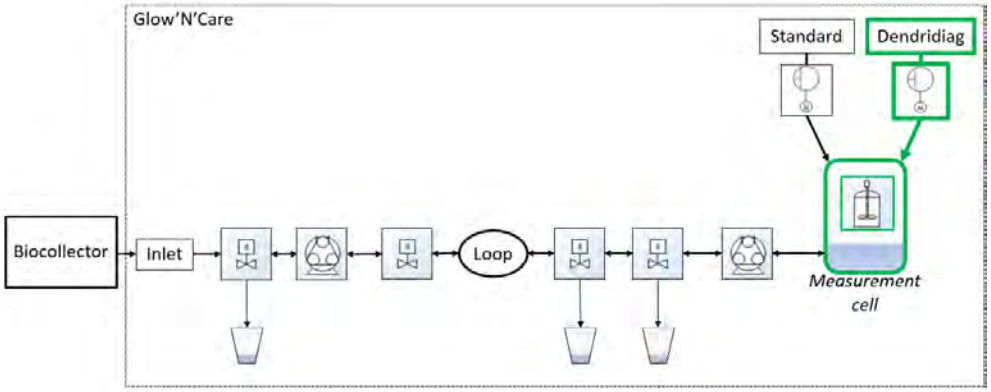


Figure 46: GNC step 13: Dendridiag

5.2.16 GNC

step 14: R1 Measure

Same step as paragraph *5.2.14 GNC*

step 12: R0 Measure, to obtain R_1 value.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

5.2.17 GNC

step 15: Standard

Standard reactant is introduced into measurement cell and stirred with mixed water sample and Dendridiag (Figure 47):

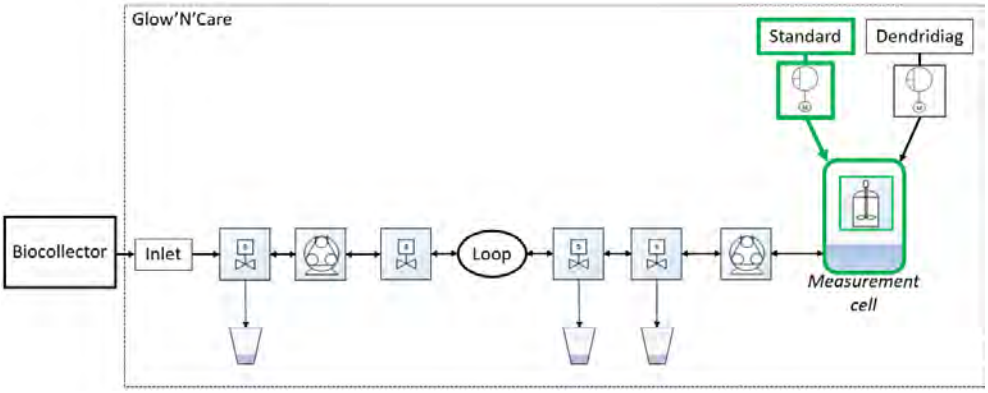


Figure 47:GNC step 15: Standard

5.2.18 GNC

step 16: R2 Measure

Same step as paragraph 5.2.14 GNC

step 12: R0 Measure to obtain R_2 value.

After this step ambient air bacteria concentration is known, computed from bacteria concentration in the water sample, obtained from photon counting device measurements (R_0, R_1 and R_2).

If the concentration exceeds the threshold specified in user interface "Settings" tab, BD will show a "red alert" and send it to IMPETUS platform immediately.

Otherwise, if the bacteria concentration is normal a 'green alert' is sent then, BD will automatically check the status of its actuators and supplies before starting a new measurement process according to parameters set.

In every case, BD results, alert level and measurements parameters are logged into BD and sent to IMPETUS platform (with parameters set in user interface "Connectivity" tab). If configured, data will also be sent to an e-mail address or ftp server.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

5.2.19 GNC

step 17: Purge cell

Measurement cell content is removed (Figure 48):

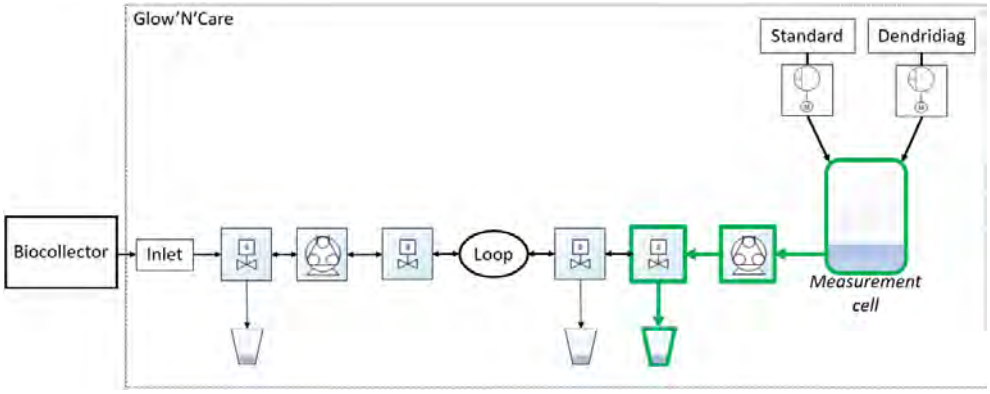


Figure 48: GNC step 17: Purge cell

5.2.20 Biocollector

step 13: Purge

Impinger water is removed (Figure 32) as in paragraph *5.2.1*

Step 1: Biocollector and GNC Purge.

5.2.21 Biocollector

step 14 and 15: Bleaching / Purge

Impinger is filled up with bleach (Figure 39), then purged while 8 mL volume of bleach is acquired by GNC for bleaching itself (Figure 35), as in paragraph 5.2.6 Biocollector step 6 and 7: Bleaching / Purge.

5.2.22 GNC

step 18 and 19: Bleaching / Purge

Bleach from biocollector is sent through the whole GNC circuit (Figure 36) and then purged (Figure 37), as in paragraph 5.2.7 GNC

step 6 and 7: Bleaching / Purge.

5.2.23 Biocollector

step 16 and 17: Rinsing / Purge

Impinger is filled up with water (Figure 38), then purged while 8 mL volume of water is acquired by GNC for rinsing itself (Figure 35) as in paragraph 5.2.4 Biocollector

step 4 and 5: Rinsing / Purge.

5.2.24 GNC

step 20 and 21: Rinsing / Purge

Water from biocollector is sent through the whole GNC circuit (Figure 36) and then purged (Figure 37), as in paragraph 5.2.5 GNC

step 4 and 5: Rinsing / Purge.

After this step, if the sequence is not finished, BD will start the new measurement at paragraph5.2.4 Biocollector

step 4 and 5: Rinsing / Purge. If the sequence is finished, BD will end sequence properly following the next steps.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

5.2.25 Biocollector

step 18 and 19: Cleaning / Purge

Impinger is filled up with cleaning solution (Figure 34), then purged (Figure 35) as in paragraph 5.2.2 Biocollector

step 2 and 3: Cleaning / Purge.

5.2.26 GNC

step 22 and 23: Cleaning / Purge

Cleaning solution from biocollector is sent through the whole GNC circuit (Figure 36), then purged (Figure 37) as in paragraph 5.2.3 GNC

step 2 and 3: Cleaning / Purge.

5.2.27 Biocollector

step 20 and 21: Rinsing / Purge

Impinger is filled up with water (Figure 38), then purged while 8 mL volume of water is acquired by GNC for rinsing itself (Figure 35) as in paragraph 5.2.4 Biocollector

step 4 and 5: Rinsing / Purge.

5.2.28 GNC

step 22 and 23: Rinsing / Purge

Water from biocollector is sent through the whole GNC circuit (Figure 36) and then purged (Figure 37), as in paragraph 5.2.5 GNC

step 4 and 5: Rinsing / Purge.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

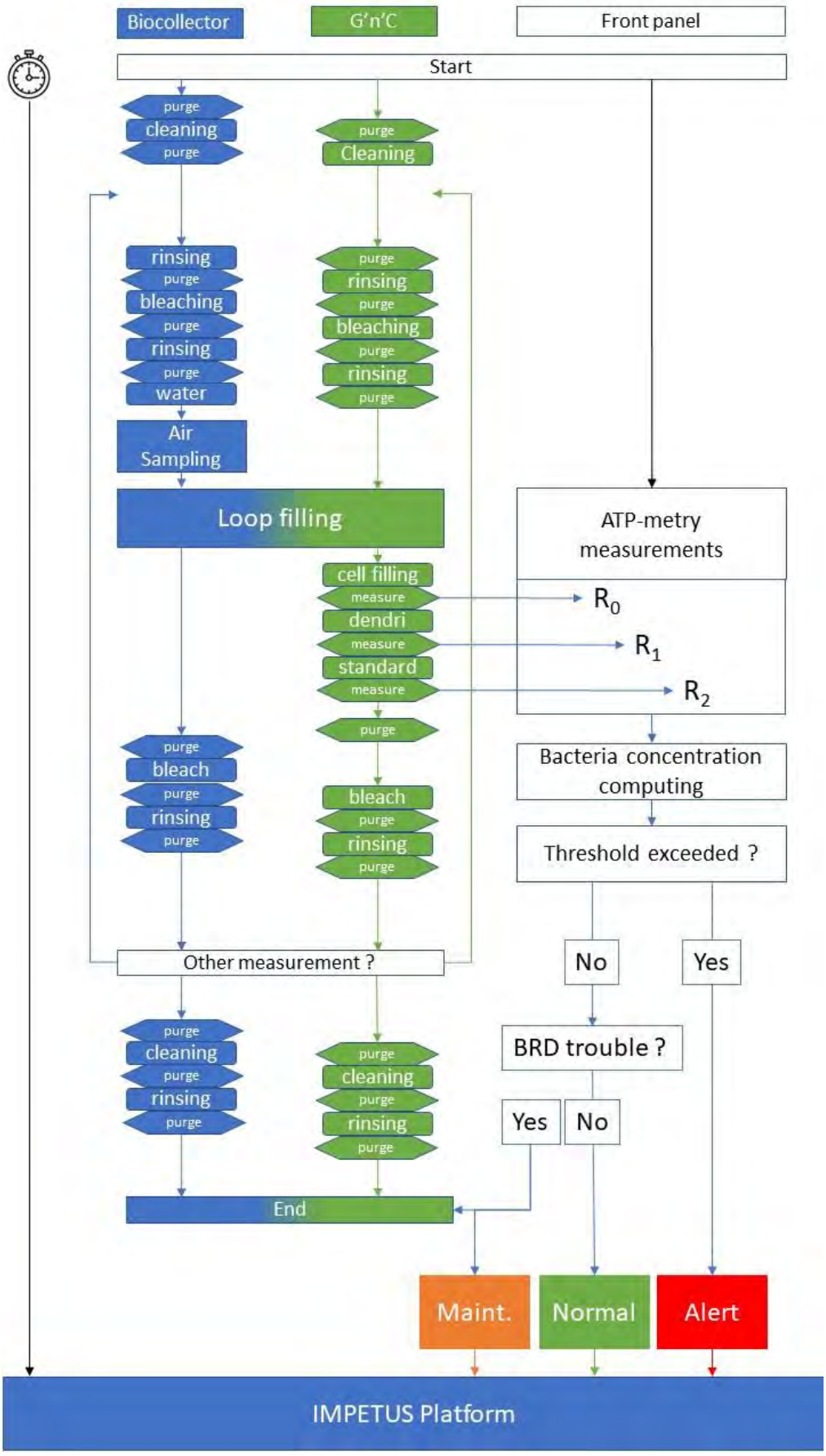


Figure 49: BD Measurement process diagram

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

6. Troubleshooting

This chapter describes unexpected behaviours which can be encountered while using BD tool, and how to solve them.

Maintenance alert	**Description and troubleshooting**
Supplies Tanks level	At least one supply tank shows a low level - Check tanks level in the UI tab “Pumps” - Open box #4 (Supplies) and check tanks levels as in paragraph 4.a - Click on “Reloading” button
Reagents level	At least one reagent tank shows a low level - Open box #3 (GnC) and check reagents tanks levels
Nucleo	Connection lost between BD embedded computer and photon counting device, ask for technical administrator help
Door opened	Saturation of the photon counting device - Check that box #2 (GNC) is properly closed
Power supply	Indicates that BD is running on its internal UPS (Uninterruptible Power Supply) - Check power network

If the operator is not able to resolve the issue using this diagram, he/she must ask help from a technical administrator.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Members of the IMPETUS consortium

SINTEF	SINTEF, Strindvegen 4, Trondheim, Norway, https://www.sintef.no	Joe Gorman joe.gorman@sintef.no
Institut Mines-Télécom	Institut Mines Telecom, 19 place Marguerite Perey, 91120 Palaiseau, France, https://www.imt.fr	Joaquin Garcia-Alfaro joaquin.garcia_alfaro@telecom-sudparis.eu
UNÎMES	Université de Nimes, Rue du Docteur Georges Salan CS 13019 30021 Nîmes Cedex 1, France, https://www.unimes.fr	Axelle Cadiere axelle.cadiere@unimes.fr
cini consorzio interuniversitario nazionale per l'informatica	Consorzio Interuniversitario Nazionale per l'Informatica, Via Ariosto, 25, 00185 – Roma, Italy, https://www.consorzio-cini.it	Donato Malerba donato.malerba@uniba.it
Università degli Studi di Padova	University of Padova, Via 8 Febbraio, 2 - 35122 Padova, Italy, https://www.unipd.it	Giuseppe Maschio giuseppe.maschio@unipd.it
Entrepreneurship Development Centre for BIOTECHNOLOGY and MEDICINE	Biotehnoloogia ja Meditsiini Ettevõtluse Arendamise Sihtasutus, Tiigi 61b, 50410 Tartu, Estonia, https://biopark.ee	Sven Parkel sven@biopark.ee
SIMAVI Software Imagination & Vision	SIMAVI, Complex Victoria Park, Corp C4, Etaj 2, Șos. București – Ploiești, nr. 73 – 81, Sector 1, București, Romania, https://www.simavi.ro	Gabriel Nicola Gabriel.Nicola@simavi.ro Monica Florea Monica.Florea@simavi.ro
THALES	Thales Nederland BV, Zuidelijke Havenweg 40, 7554 RR Hengelo, Netherlands, https://www.thalesgroup.com/en/countries/europe/netherlands	Johan de Heer johan.deheer@nl.thalesgroup.com
CINEDIT INTELLIGENT VIDEO ANALYTICS	Cinedit VA GmbH, Poststrasse 21, 8634 Hombrechtikon, Switzerland, https://www.cinedit.com	Joachim Levy j@cinedit.com

	Insikt Intelligence, Calle Huelva 106, 9-4, 08020 Barcelona, Spain, https://www.insiktintelligence.com	Dana Tantu dana@insiktintelligence.com
	Sixgill, Derech Menachem Begin 132 Azrieli Tower, Triangle Building, 42nd Floor, Tel Aviv, 6701101, Israel, https://www.cybersixgill.com	Benjamin Preminger benjamin@cybersixgill.com Ron Shamir ron@cybersixgill.com
	City of Padova, via del Municipio, 1 - 35122 Padova Italy, https://www.padovanet.it	Enrico Fiorentin fiorentine@comune.padova.it Stefano Baraldi Baraldis@comune.padova.it
	City of Oslo, Grensen 13, 0159 Oslo, Norway, https://www.oslo.kommune.no	Osman Ibrahim osman.ibrahim@ber.oslo.kommune.no
	Institute for Security Policies, Kruge 9, 10000 Zagreb, Croatia, http://insigpol.hr	Krunoslav Katic krunoslav.katic@insigpol.hr
	International Emergency Management Society, Rue Des Deux Eglises 39, 1000 Brussels, Belgium, https://www.tiems.info	K. Harald Drager khdrager@online.no
	Unismart – Fondazione Università degli Studi di Padova, Via VIII febbraio, 2 - 35122 Padova, Italy, https://www.unismart.it	Alberto Da Re alberto.dare@unismart.it

http://www.impetus-project.eu

CTDR – Cyber Threat Detection and Response

Authors: Kéren Saint-Hilaire, Joaquin Garcia-Alfaro (IMT)

Table of Contents

1 General Information

1.1 Overview

The objective of this manual is to show the main capabilities of the Prelude-ELK in the following two use cases:

- Monitoring an organization network to detect attacks and incidents to alert the cybersecurity experts of Padova and Oslo cities.
- Mapping of information from monitored network with attack graph generated and ontology

Data sources in these use cases are:

Use case	Data sources
Padova	Simulated network system
	Network system
Oslo	Simulated network system
	Network system

1.2 What is a SIEM?

A Security Information and Event Management (SIEM) system is composed of monitoring software for the analysis and management of events created by cybersecurity tools (e.g., log events created from antivirus tools, network firewalls and intrusion detection systems). The processed events are stored and managed as cybersecurity alerts.

1.3 What is Prelude?

Prelude is a SIEM that collects and centralizes the security information of an organization to offer a central point of control. It provides analysis and correlation of cybersecurity logs and triggers alerts about cyberattack attempts in real-time. Under the scope of the IMPETUS project, we will use the open-source (freeware) version of the Prelude SIEM, Prelude OSS, available at https://www.prelude-siem.org/ (GPLv2 version of https://www.prelude-siem.com/). Hereinafter, we will refer to Prelude OSS as Prelude, for simplicity reasons.

1.4 What is ELK?

Under the scope of the IMPETUS project, we will use a version of Prelude extended with ELK, which is an abbreviation for three open source projects, namely Elasticsearch, Logstash and Kibana. Hereinafter, we will refer to Prelude-ELK to Prelude+ELK, for simplicity reasons.

Elasticsearch allows indexing and processing unstructured data. It provides a distributed web interface to access the resulting information. Logstash is the parsing engine associated with Elasticsearch for collecting, analyzing, and storing logs. It can integrate many sources simultaneously. Finally, Kibana is a data visualization platform that provides visualization functionalities on indexed content in Elasticsearch. Users can create dashboards with charts and maps of large volumes of data.

1.5 Collection and analysis of data

The addition of ELK into Prelude allows the injection and visualization of third-party logs, received from both system and network components, via TCP/IP messages.
The collection of data can still be combined with the traditional collection and visualization tools of Prelude. For instance, we can keep using Prelude's LML (Log Monitoring Lackey) and third-party sensors, to monitor and process syslog messages generated from different hosts on heterogeneous platforms. LML has two main operation modes:

Watching log files on the host where it is running (e.g., Syslog data feeds).
Receiving UDP Syslog messages from other hosts on the network.

In addition, any other third-party sensors (e.g., Suricata and Snort) can still be registered into Prelude-ELK, following the steps below:

Allocating a unique identity for the sensor.
Creating a directory to be used by the sensor.
Registering to a remote manager, e.g., via signed X509 certificates (to allow secure communication between sensors and managers).

The agent registration process is directed by a single tool, prelude-admin, using the following command-line steps:

$ prelude-admin register <profile name> <requested permission> <manager address> --uid <uid> --gid <gid>

1.6 Correlation of alerts

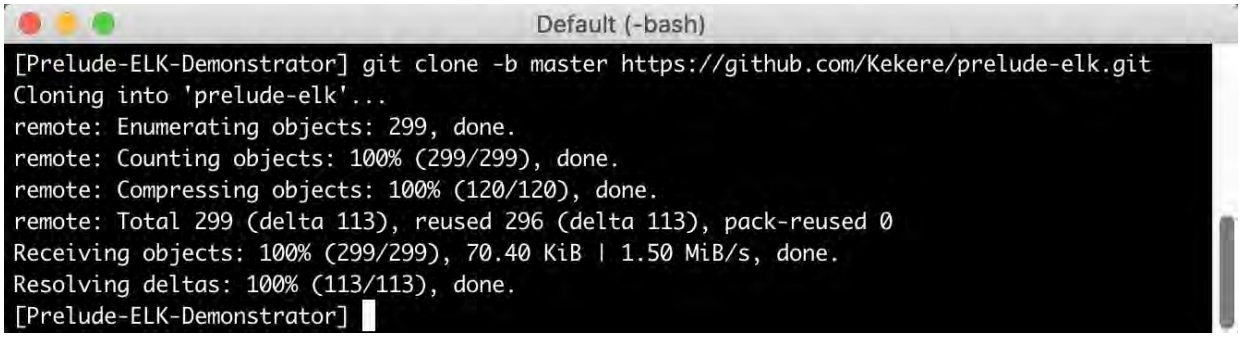
The configuration of LML and third-party sensors allows Prelude-ELK to generate alerts reporting the exploitation of vulnerabilities. Later, a Python script (named prelude-correlator) provides Prelude-ELK with a rule-based correlation engine that connects and fetches alerts from other sensors or managers, providing new alerts (with a higher degree of information, i.e., with information about coordination attacks).

2 Installation of the Prelude-ELK demonstrator

We show next the installation of a dockerized version of Prelude-ELK.

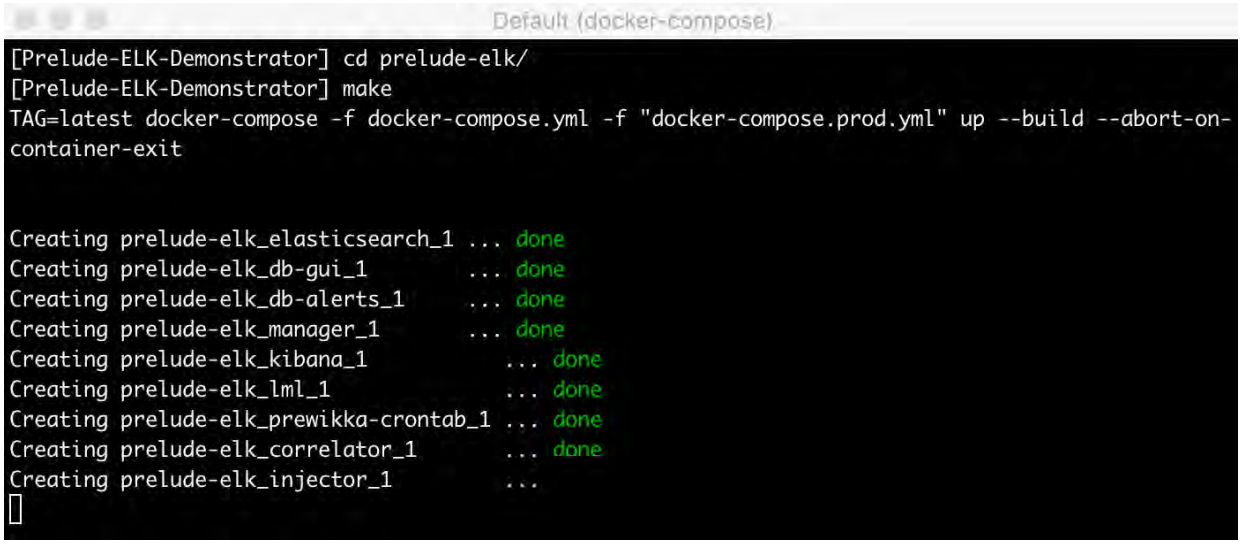
1. Clone the repository using the following command-line step:

$ git clone -b master https://github.com/Kekere/prelude-elk.git



2. Go to the newly created folder and type “make” to build the demonstrator:

$ cd prelude-elk/; make

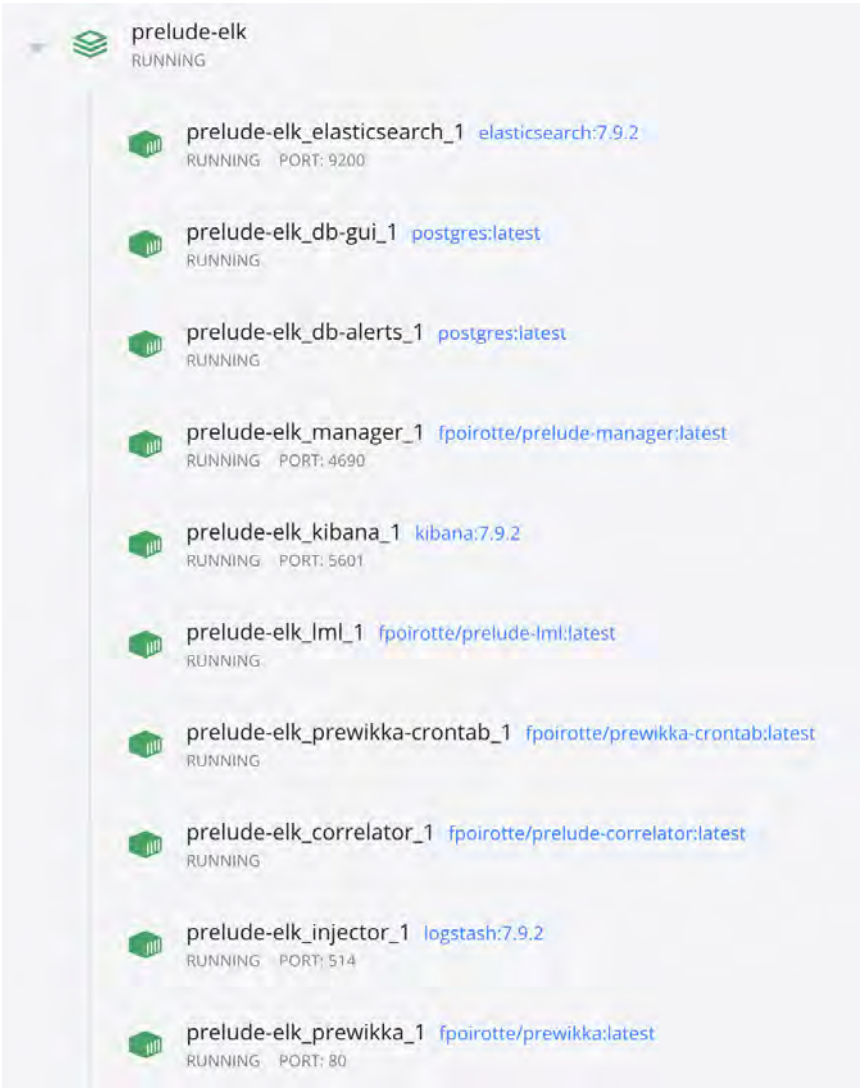


The following containers are created during the installation process:

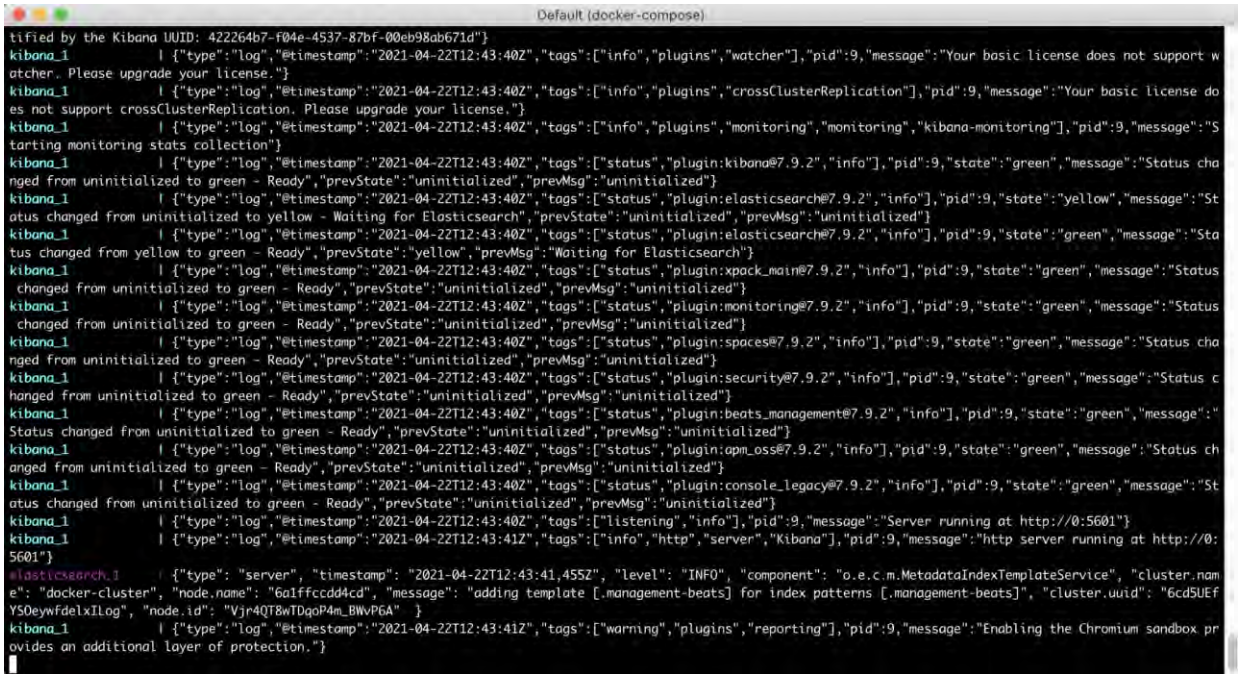
- prewikka: Prelude's web user interface
- prewikka-crontab: periodic scheduler used by prewikka
- manager: Prelude's manager
- kibana: ELK’s data visualization
- elasticsearch: ELK’s log storage
- correlator: alert correlator
- injector: entrypoint for logs
- lml: Prelude's log management servant
- db-alerts: database server for Prelude’s alerts
- db-gui: database server for Prewikka
- logstashalert: ingestor of alerts

- apache: web interface of the attack graph generator
- php: php service

The containers can be displayed and managed using graphical user interfaces such as Docker Desktop for Windows or macOS, as shown next:



The following two screenshots illustrate the startup process of the dockerized version of Prelude-ELK on macOS:



3 Prewikka Dashboards

Prewikka is the official Graphical User Interface (GUI) of Prelude, for the visualization and management of the alerts. Next, we show some representative information, using a Web browser pointing out to the http://localhost url (i.e., prelude-elk_prewikka_1, port 80).

3.1 Overview

Alerts timeline	General	Time distribution of alerts.
	Raw data	Chart bar of alerts based on time distribution.
Threats timeline	General	Time distribution of threats.
	Raw data	Chart bar of threats based on time distribution.
Aggregated alerts	General	Alerts group by selected criteria.
	Raw data	Chart bar of alerts based on selected criteria.
Aggregated threats	General	Threats group by selected criteria.

	Raw data	Chart bar of threats based on the selected criteria;
Alerts table	Date	Time when the alert was generated.
	Classification	Type and description of the attack.
	Sources	Information about the attacker machine.
	Target	Information about the target machine.
	Analyzer	Information about the analyzer that detects the attack.
Threats table	Date	Time when the threat was detected.
	Classification	Type of the threat.
	Sources	Information about the attacker machine.
	Target	Information about the target machine.
	Program	Name of the program.
Heartbeats timeline	General	Time distribution of the heartbeats.
	Raw data	Chart bar of heartbeats based on time distribution.
Heartbeats table	Date	Time when the heartbeat is registered.
	Agent	Agents that register the heartbeat.
	Node address	Ip address of the agent;
	Node name	Identification of the heartbeat.
	Model	Name of the agent;
Agents	Alert listing	Timeline chart bar for the alerts and table listing alerts for the agent.
	Heartbeat listing	Timeline chart bar for the heartbeats and table listing heartbeats for the agent.
	Heartbeat analysis	Heartbeats information for the alerts.

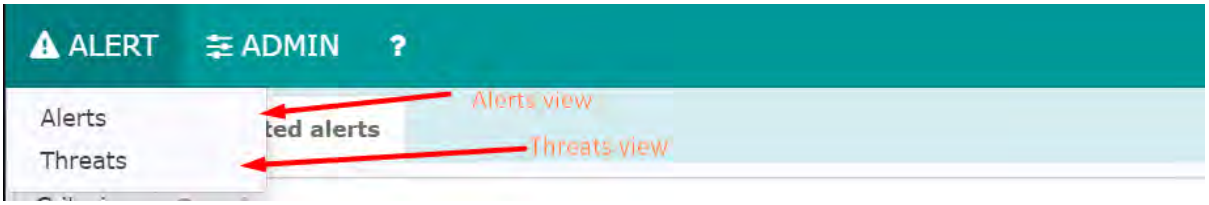
All alerts are shown in the same dashboard for the sake of simplicity including the option to group them by criteria or to search specific alerts by keywords. All threats are shown in the same dashboard for the sake of simplicity including the option to group them by criteria or to search specific alerts by keywords. All heartbeats are shown in the same dashboard for the sake of simplicity including the option to group them by criteria or to search specific alerts by keywords. All agents are shown in the same dashboard with the possibility to see the list of alerts for each agent and the heartbeats of the agents too.

3.2 Content of the dashboards

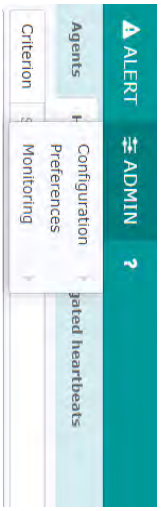
In this section we show some representative dashboards and menu options of prewikka.

3.2.1 Menu bar alerts

Graph	Content	Capabilities
	Alert options	Select if you want to see the alerts dashboard or the threats dashboard.



3.2.2 Menu bar admin

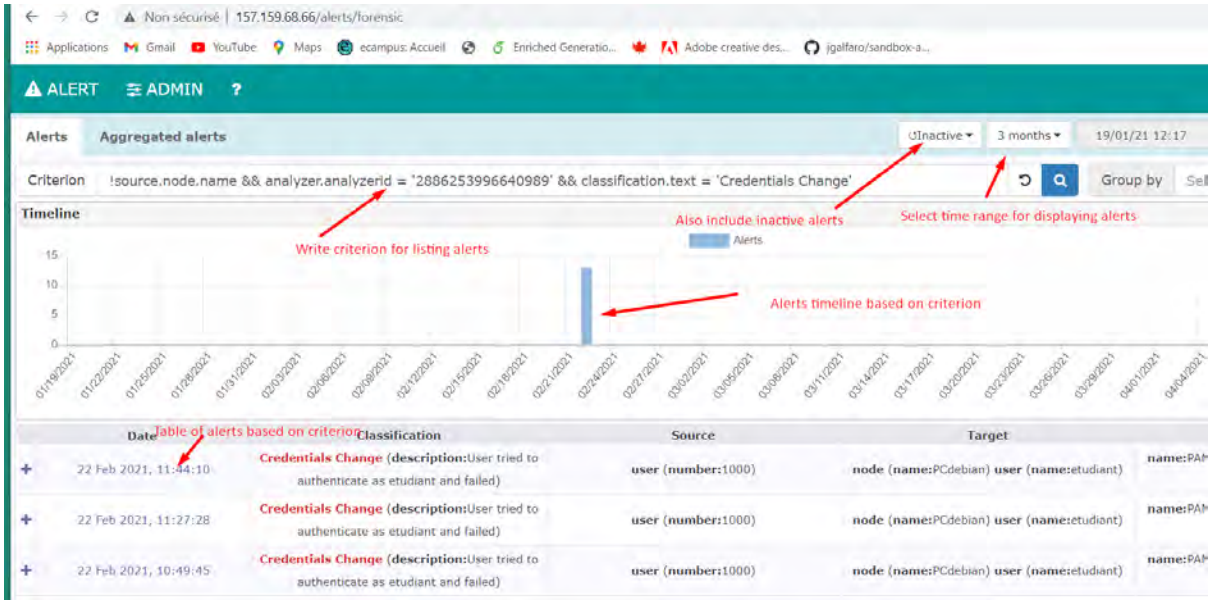
Graph	Content	Capabilities
	Admin options	• In the configuration section, we can schedule alerts. • In the preferences section, we can save a name and an email. • In the monitoring section, we can access the agents dashboard, heartbeats dashboard or aggregated heartbeats dashboard.

3.2.3 General - Alerts

This page contains all the alerts generated on the network, in a table format. A chronologic distribution of the alerts is also shown.

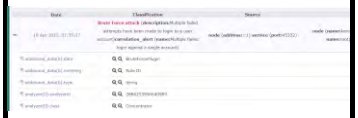
It is useful for searching and filtering across all the information collected and showing them in a table format.

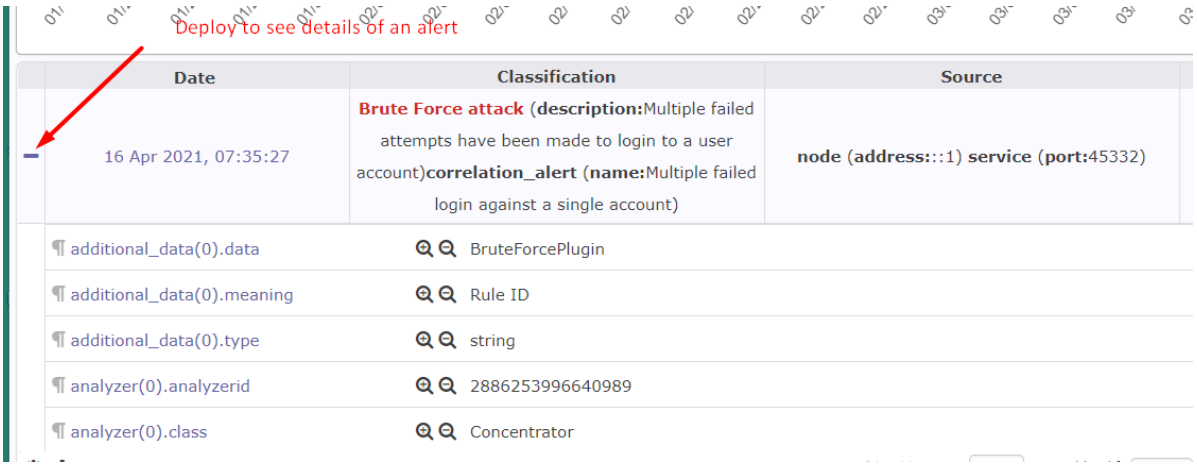
Graph	Content	Capabilities
	• Search area to specify criterion for displaying alerts table • Selection area to select option to group alerts by • Chart bar to visualize alerts distribution on the time • Alerts table	Group by: • Time values (minute, hour, day, month, year) • Alert fields (classification, source, target, analyzer)



3.2.4 Content - Alerts details

Drop down section that contains detailed information about a specific alert.

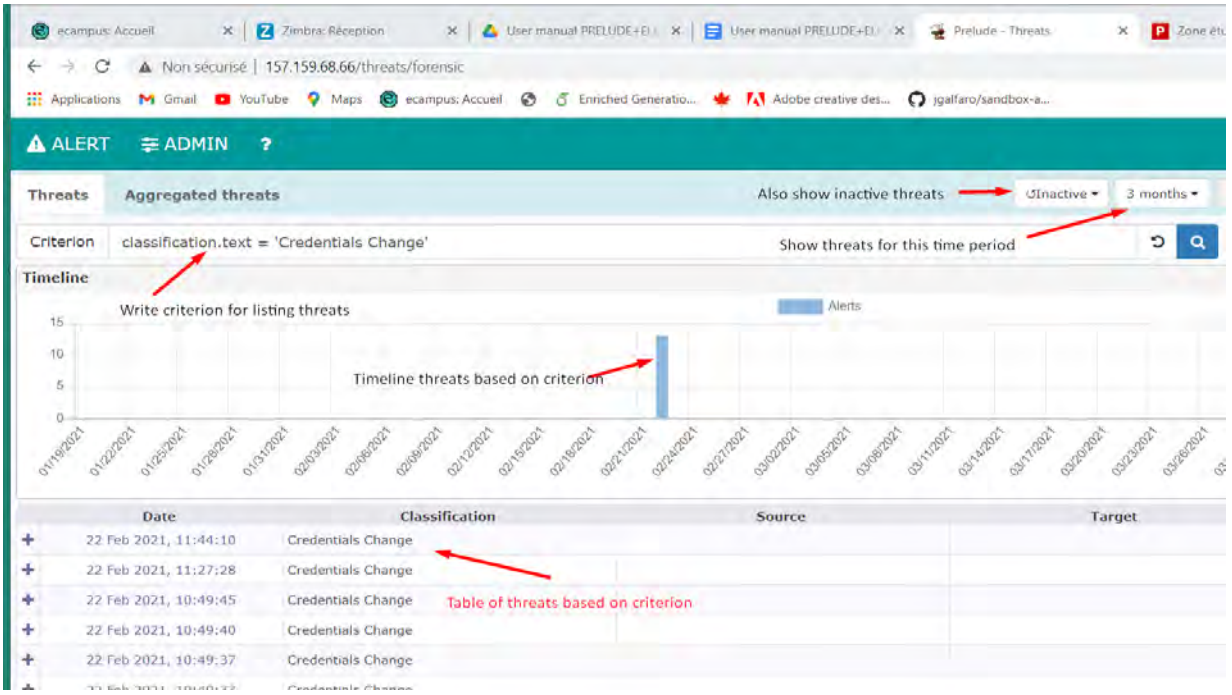
Graph	Content	Capabilities
	Detailed information about a specific alert	Additional information that does not appear in the table is shown.
Hints	Discover all details about an alert	Deploy the plus icon before the date column.



3.2.5 General - Threats

Page contains all the threats generated on the network, in a table format. A chronologic distribution of the threats is also shown. It is useful for searching and filtering across all the information collected and showing them in a table format.

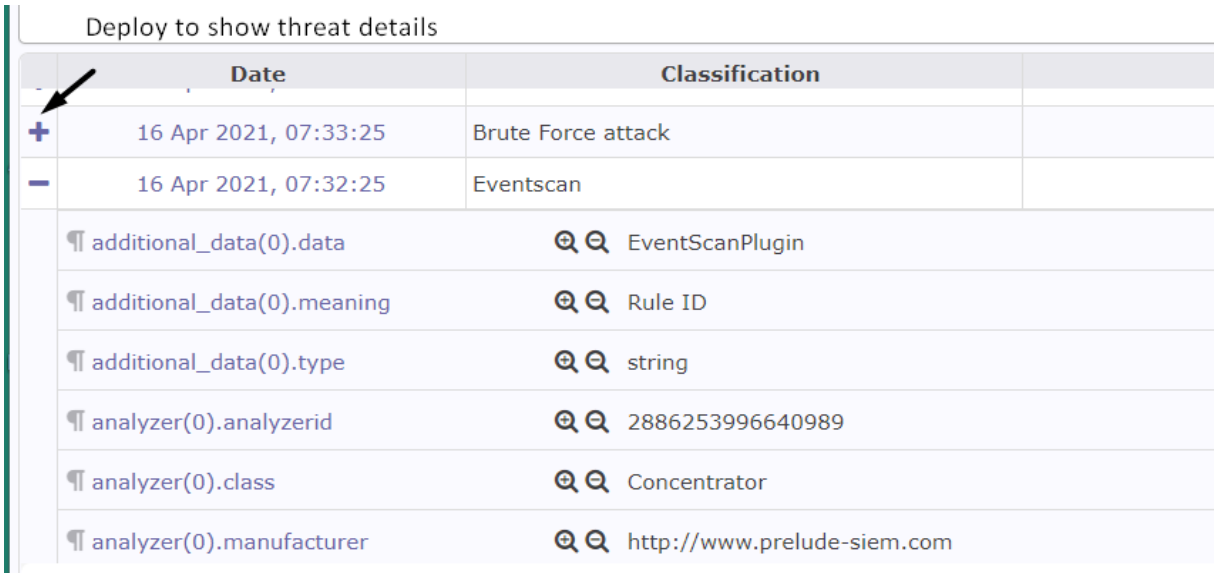
Graph	Content	Capabilities
	• Search area to specify criterion for displaying threats table • Selection area to select option to group threats by • Chart bar to visualize threats distribution on the time • Threats table	Group by: • Time values (minute, hour, day, month, year) • Alert fields (classification, source, target, analyzer)



3.2.6 Content - Threats details

Drop down section that contains detailed information about a specific threat.

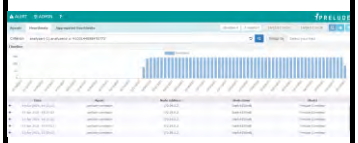
Graph	Content	Capabilities
	Detailed information about a specific threat.	Additional information that does not appear in the table is shown.
Hints	Discover all details about a threat.	Deploy the plus icon before the date column.

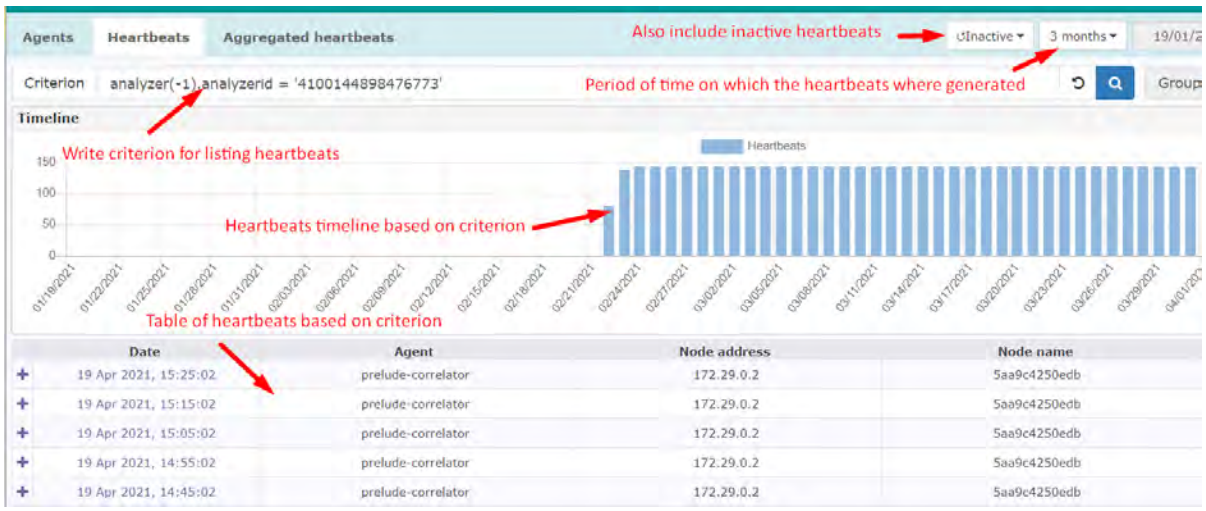


3.2.7 General - Heartbeats

Page contains all the heartbeats generated by the agents, in a table format. A chronologic distribution of the heartbeats is also shown.

It is useful for searching and filtering across all the information collected and showing them in a table format.

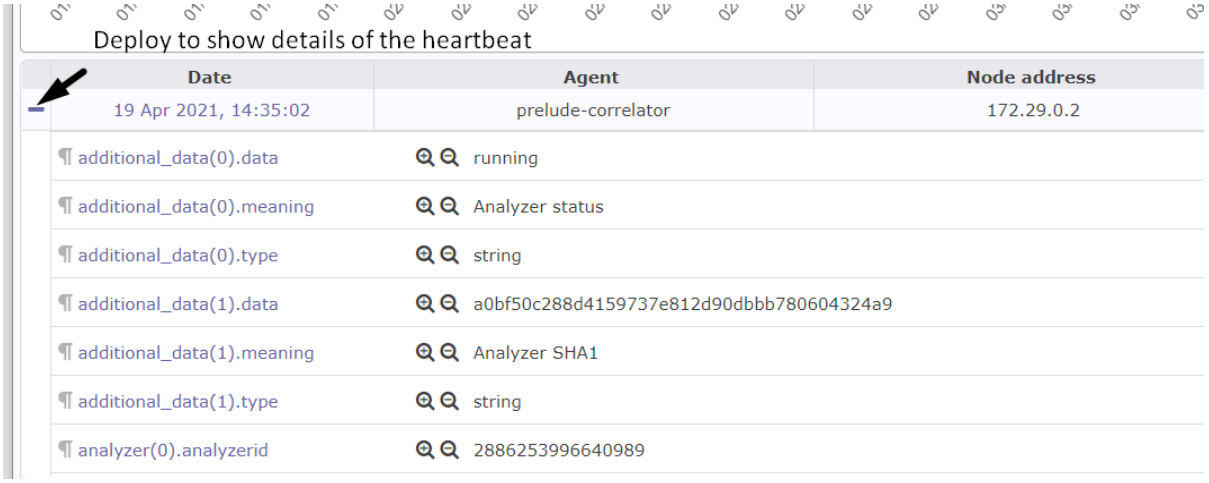
Graph	Content	Capabilities
	• Search area to specify criterion for displaying heartbeats table • Selection area to select option to group heartbeats by • Chart bar to visualize heartbeats distribution on the time • Heartbeats table	Group by: • Time values (minute, hour, day, month, year) • Alert fields (create_time, heartbeat_interval, additional_data, messageid, analyzer)



3.2.8 Content - Heartbeats details

Drop down section that contains detailed information about a specific heartbeat.

Graph	Content	Capabilities
	Detailed information about a specific Heartbeat.	Additional information that does not appear in the table is shown.
Hints	Discover all details about a heartbeat.	Deploy the plus icon before the date column..

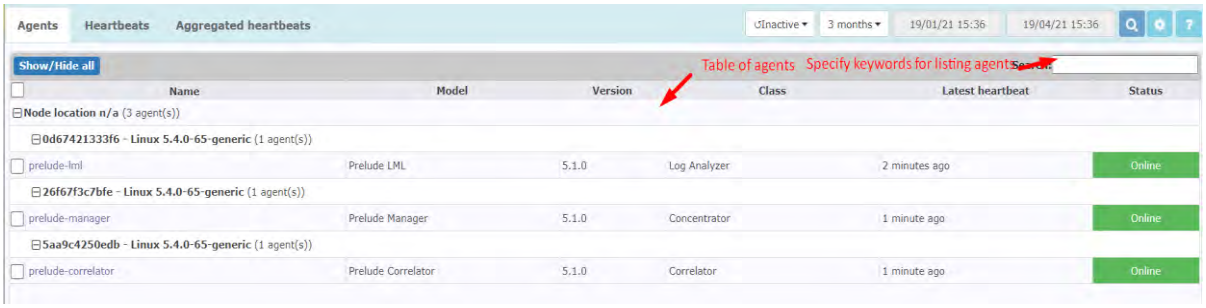


3.2.9 General - Agents

Raw data page contains information about the agents, in a table format.

Graph	Content	Capabilities
	Table with the alerts listed	We have the possibility to list the alerts and heartbeats for each agent. We can also visualize an analysis of heartbeat for each agent.

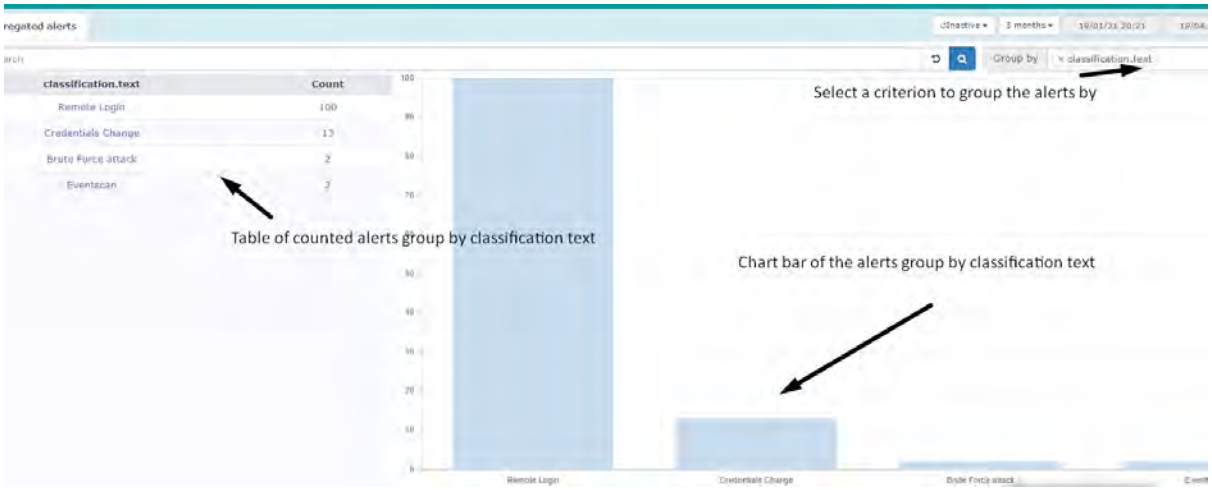
Hints	Discover alerts list group by agent.	Click on the agent name and click on Alerts listing option
	Discover heartbeats list group by agent.	Click on the agent name and click on Heartbeats listing option.
	Discover heartbeat analysis.	Click on the agent name and click on Heartbeat analysis.



3.2.10 General - Aggregated alerts

This page contains a table and a chart bar of the alerts counted, group by the selected option.

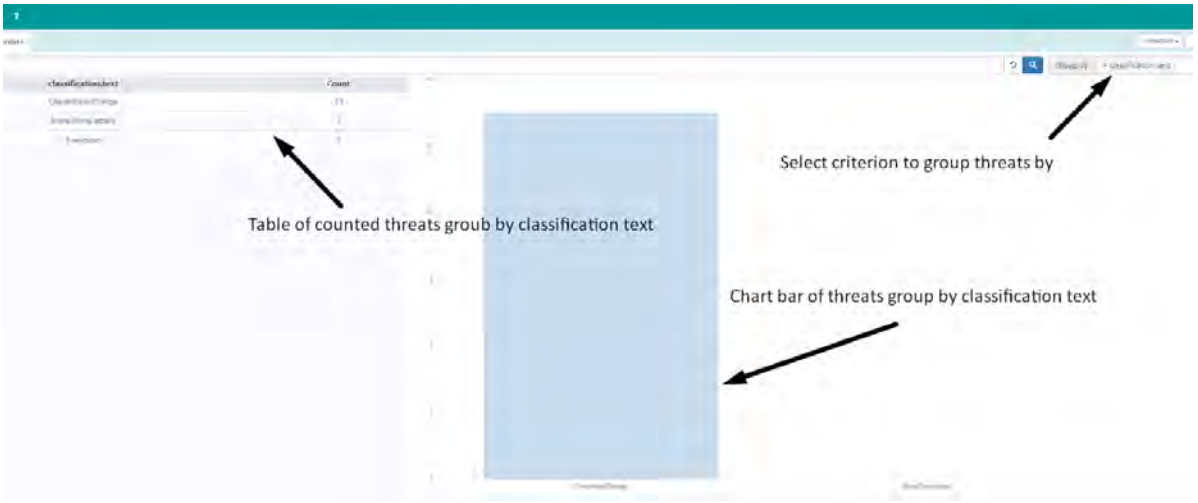
Graph	Content	Capabilities
	• Search area to specify filter for displaying chart bar. • Selection area to select criterion to group alerts by • Table with counted alerts group by filter. • Chart bar of alerts grouped by criterion or filtered alerts	Group by: • Time values (minute, hour, day, month, year) • Alert fields (classification, source, target, analyzer)



3.2.11 General - Aggregated threats

This page contains a table and a chart bar of the threats counted, group by the selected option.

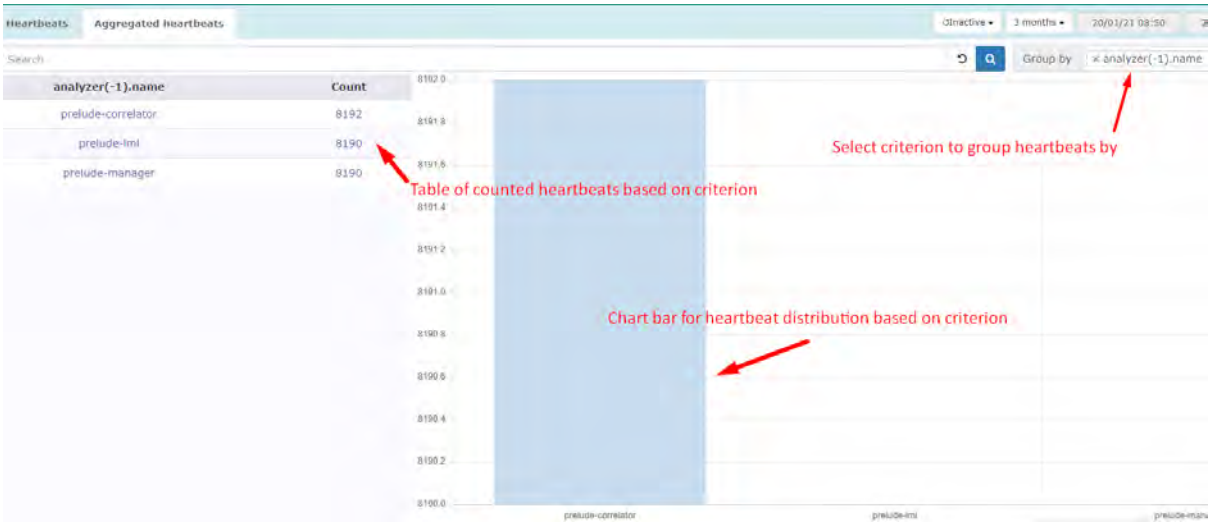
Graph	Content	Capabilities
	• Search area to specify filter for displaying chart bar. • Selection area to select criterion to group threats by • Table with counted threats group by filter. • Chart bar of threats grouped by criterion or filtered threats	Group by: • Time values (minute, hour, day, month, year) • Threat fields (classification, source, target, analyzer)



3.2.12 General - Aggregated heartbeats

This page contains a table and a chart bar of the heartbeats counted, group by the selected option.

Graph	Content	Capabilities
	• Search area to specify filter for displaying chart bar. • Selection area to select criterion to group heartbeats by • Table with counted heartbeats group by filter. • Chart bar of alerts grouped by criterion or filtered heartbeats	Group by: • Time values (minute, hour, day, month, year) • Heartbeat fields (classification, source, target, analyzer)

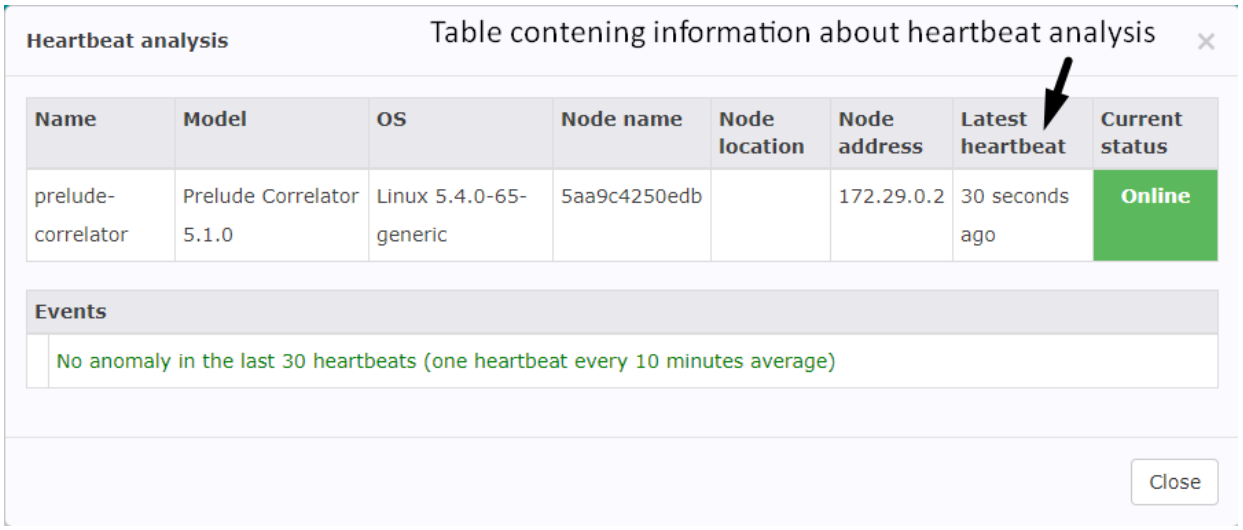


3.2.13 Content - Heartbeats analysis

A modal window contains an analysis of the heartbeats.

Graph	Content	Capabilities

	• Search area to specify filter for displaying chart bar. • Selection area to select criterion to group heartbeats by • Table with counted heartbeats group by filter. • Chart bar of alerts grouped by criterion or filtered heartbeats	Group by: • Time values (minute, hour, day, month, year) • Heartbeat fields (classification, source, target, analyzer)



4 Kibana Dashboards

Kibana is the official Graphical User Interface (GUI) of the ELK stack, for the visualization and management of the Elasticsearch indexes. Next, we show some representative information, using a Web browser pointing out to the http://localhost:5601 url (i.e., prelude-elk_kibana_1, port 5601).

4.1 Overview

Table of logs	Name	Name field
	Type	Type of data
	Format	Format data (it is empty)

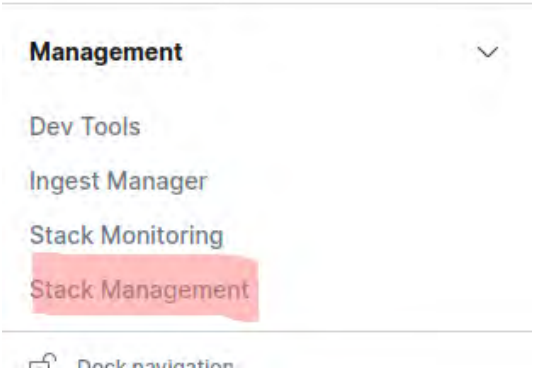
	Searchable	Indicate if the field is searchable.
	Aggregatable	Indicate if the field is aggregatable.
	Excluded	Indicate if the field is excluded.
Discover dashboard	Chart bar	Chart of the logs counted ,group by the time metric specified in the select area.
	Table of logs	Table with all the logs.
	Filter areas	Filter by: • message • received_at • received_from • syslog_message • etc
Goal visualization of logs	Filter areas	Filter by: • message • received_at • received_from • syslog_message • etc
	Goal graphic	Visualize logs counted in a goal graphic format.
	Graphic functionalities options	The style of the graphic can be changed. The metric can also be changed; count can be changed for average.

4.2 Content of the dashboards

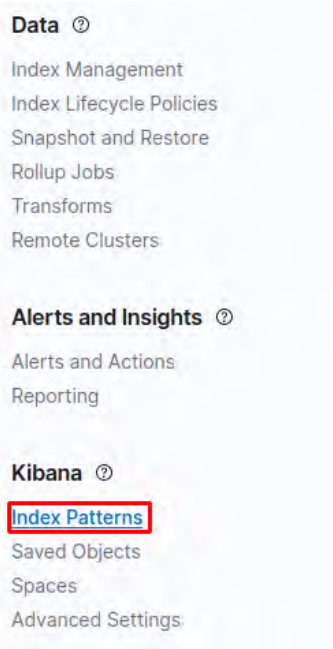
4.2.1 Menu bar

Click on the deployment icon to deploy the menu.

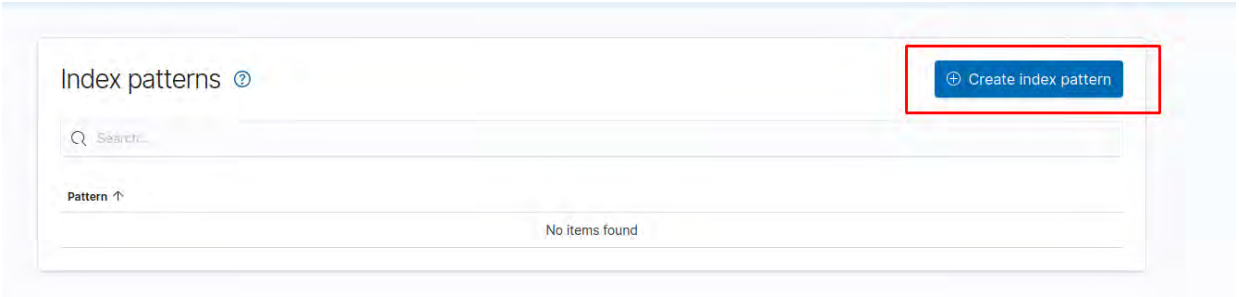
Click on the Stack Management subsection in the Management section to have access to the management menu.



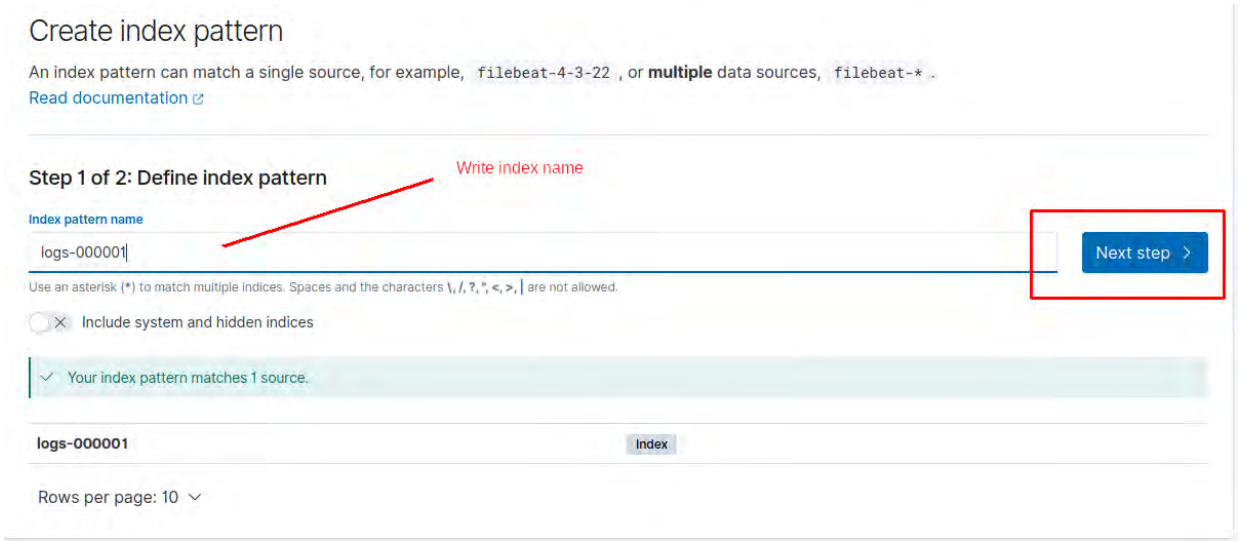
Click on the Index Patterns subsection in the Kibana section to have access to the list of index patterns and to create new index patterns.



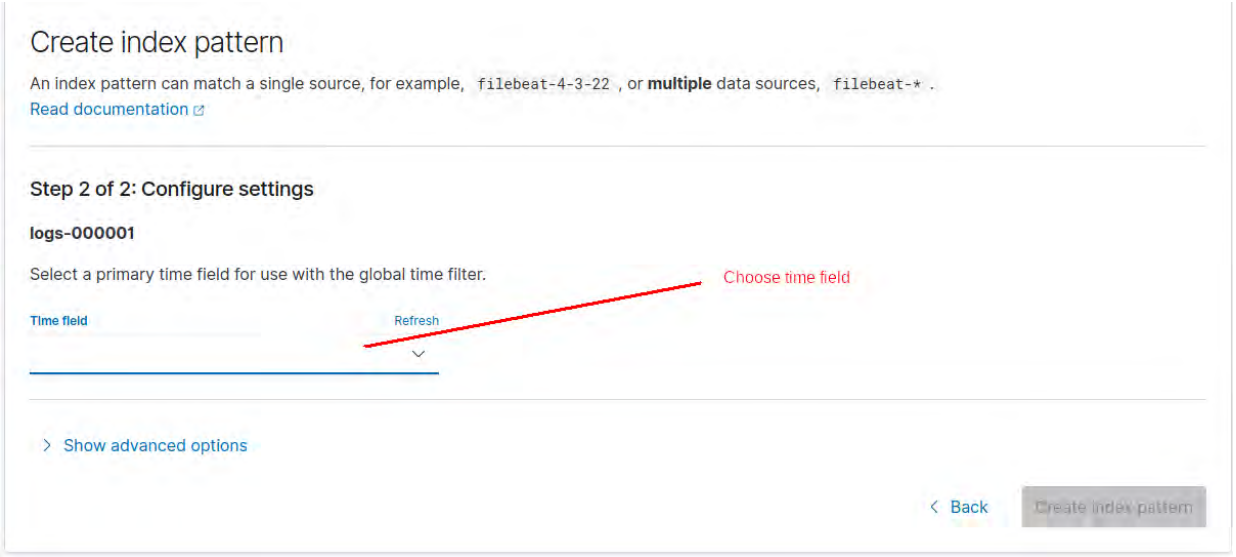
Click on the blue button to create a new index pattern.



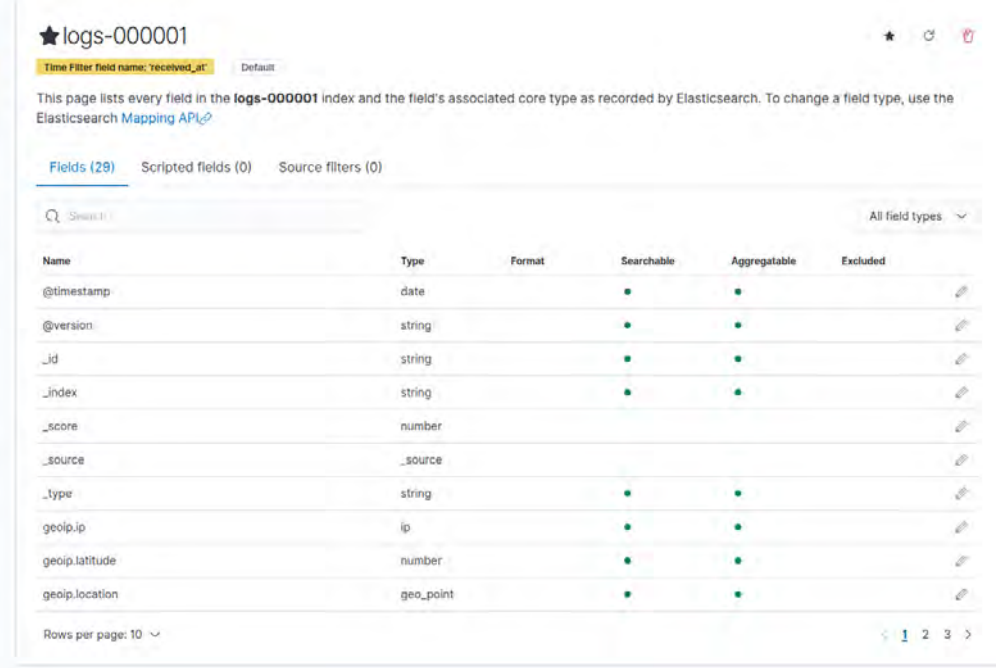
If the index exists in Elasticsearch, click on the blue button to go to the next step.



Choose a time field in the select area as the primary time field.

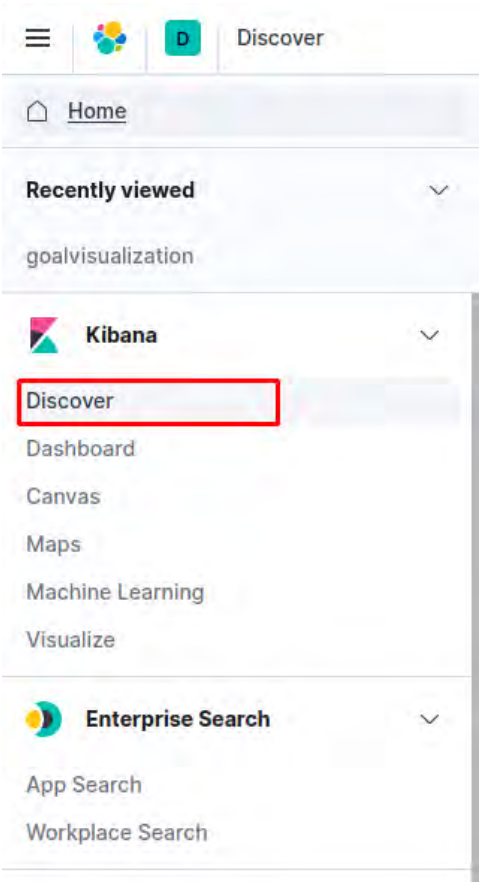


A table will be shown with the name fields and their characteristics.

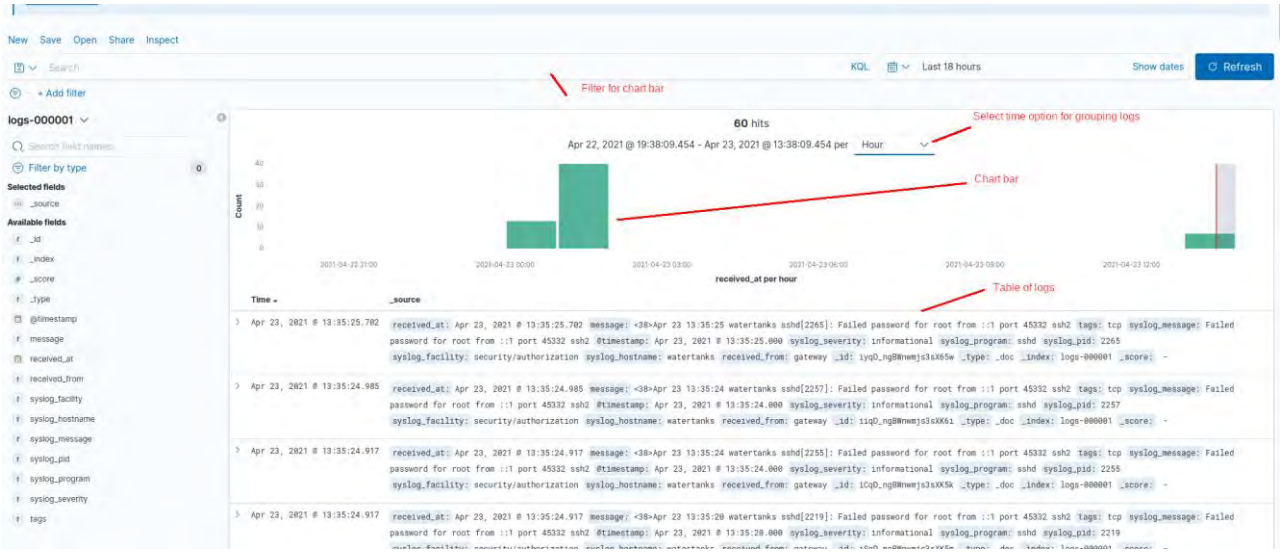


4.2.3 General - Discover dashboard

Click on the Discover subsection in the Kibana section.

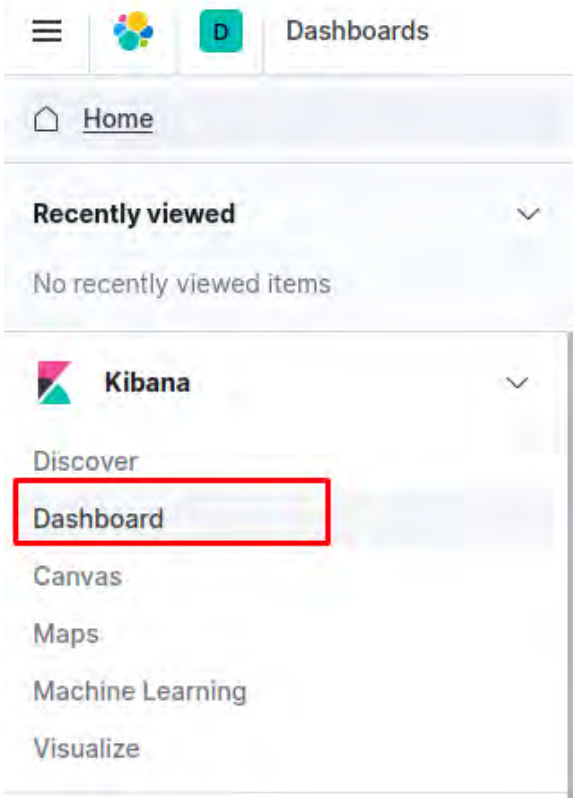


In the Discover dashboard, a chart bar of the counted logs group by the time metric selected is shown. A table with all the details of the logs is also shown. There are filter areas to search for specific logs. The available fields are listed in the left.

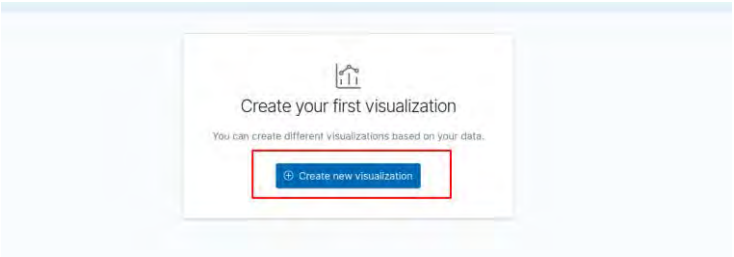


4.2.4 General - Goal graphic of logs

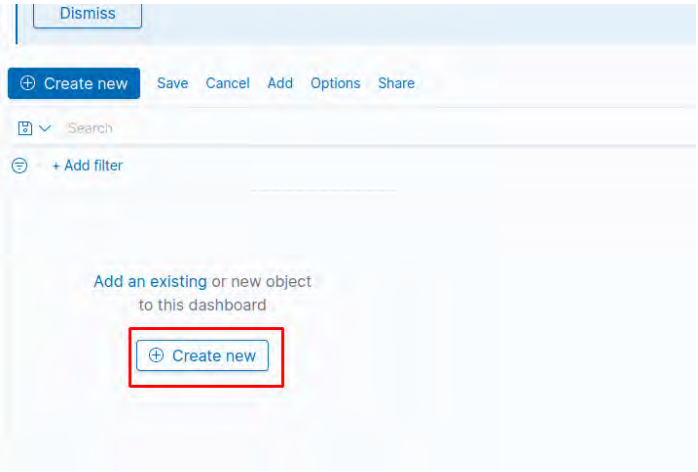
Click on the Dashboard subsection in the Kibana section.



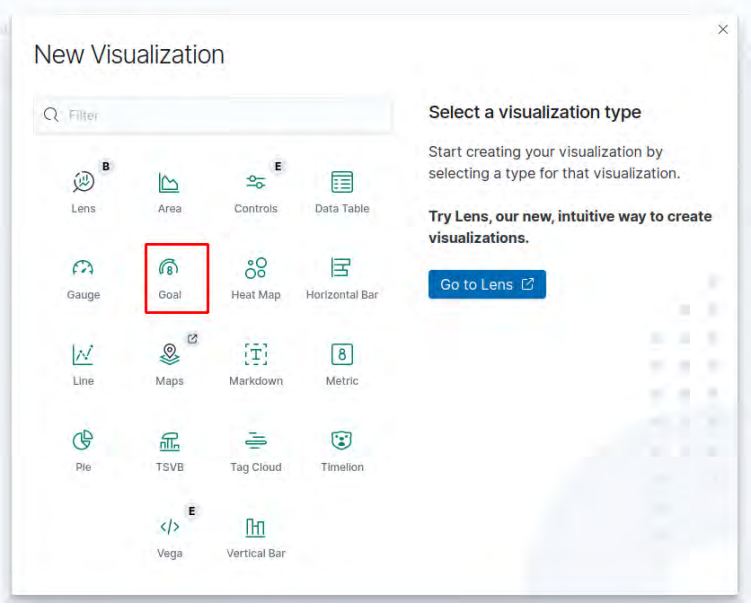
Click on the blue button to create a new visualization.



If you have already saved a visualization, you can click on Add an existing, if not click on the create new button to create a new one.

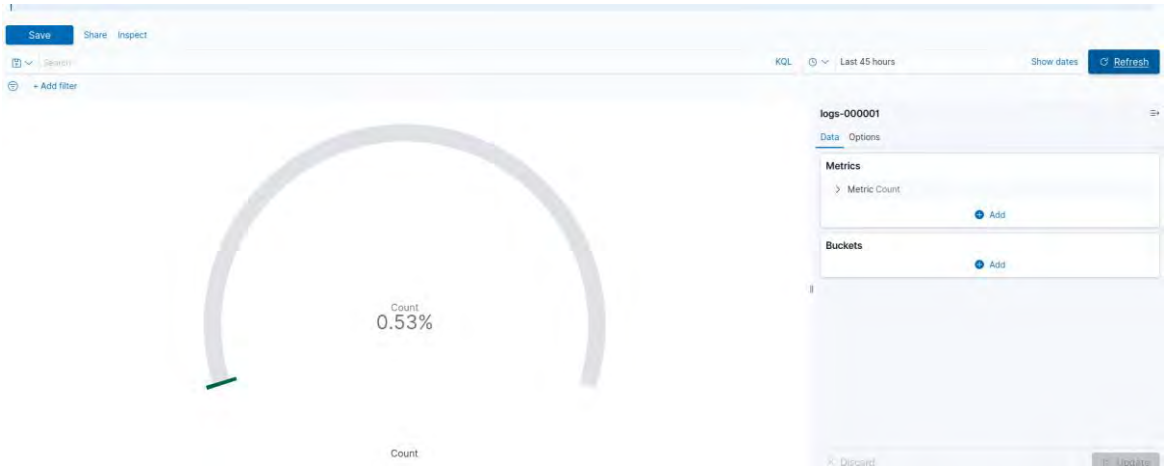


Choose the type of visualization.



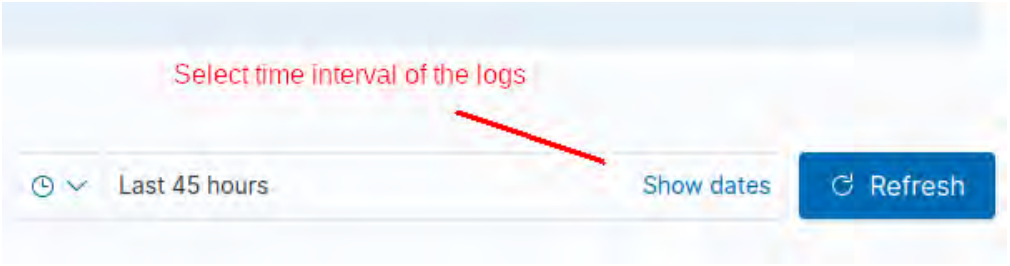
Choose an index to create visualization on.

The dashboard is created, some options are available to customize it.

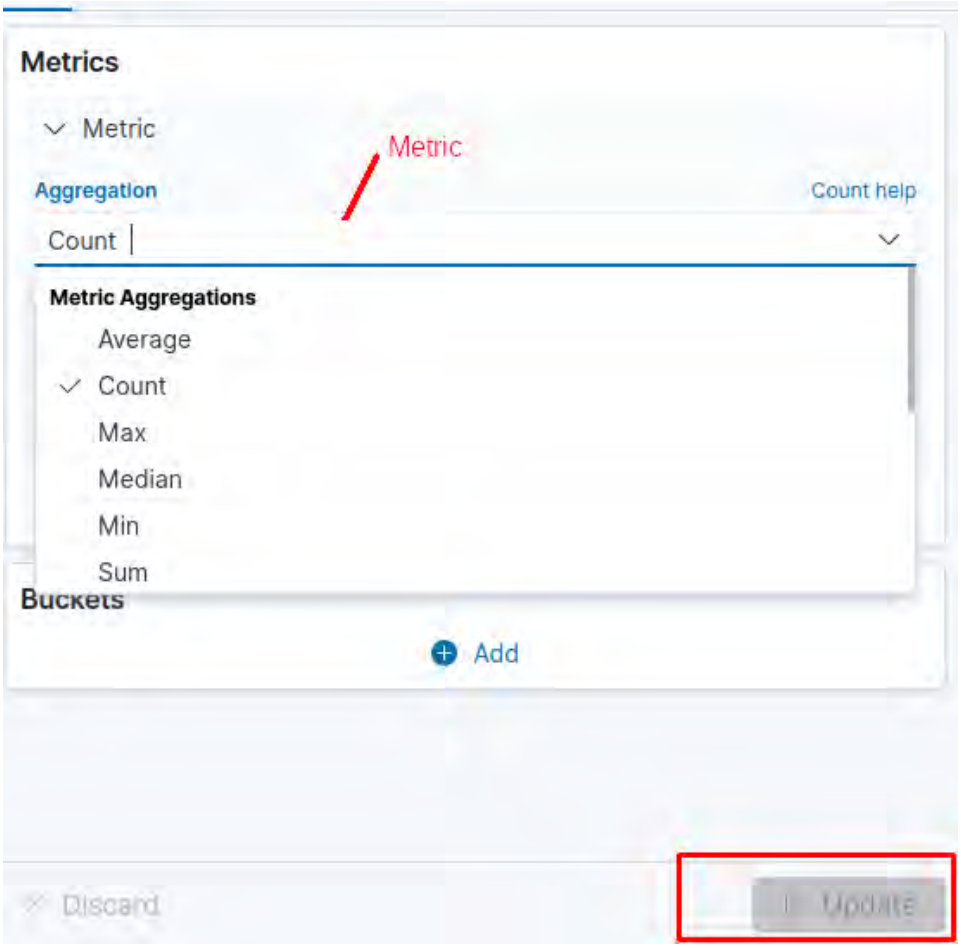


You can select filters to visualize the graphic only for a category of logs.

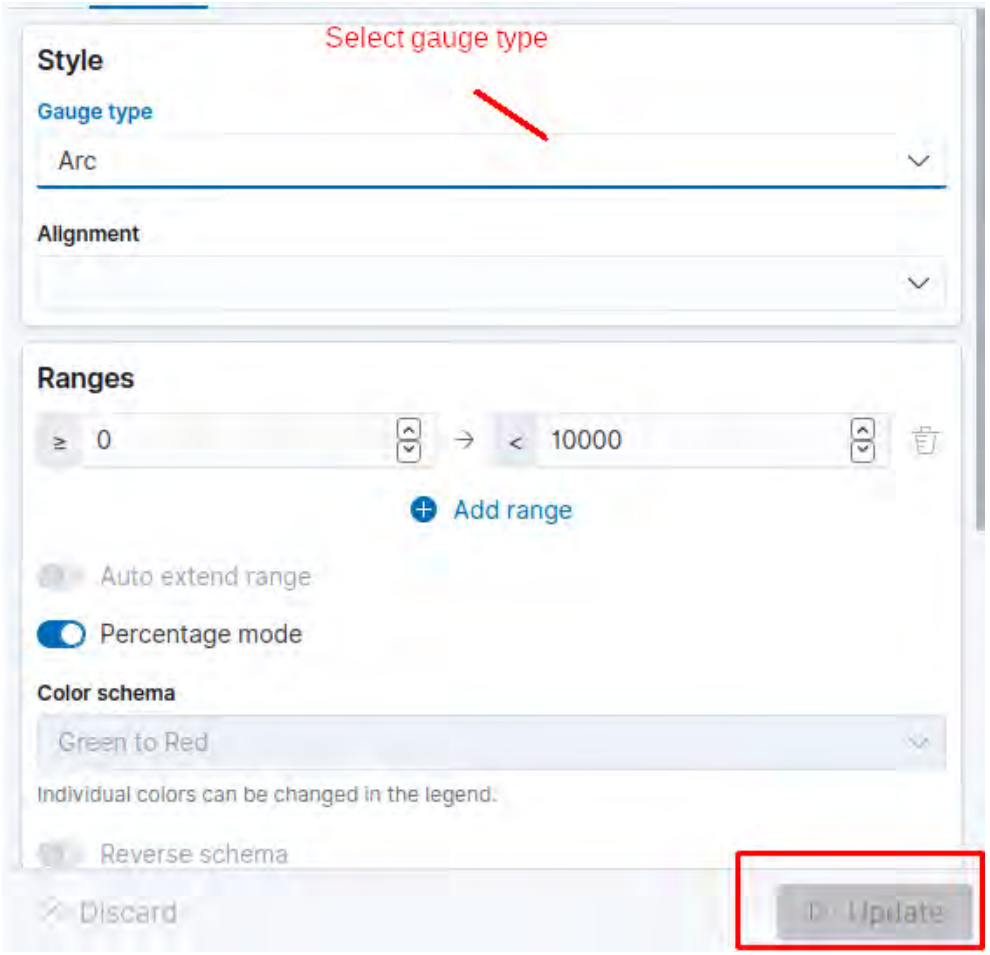
Select the time interval in which the logs that will be represented were generated.



Select the metric for the graphic representation of the logs.



Select the style for the goal graphic representation of the logs.



5 Attack Graph Generator Interface

The Attack Graph Generator Interface is the main Graphical User Interface (GUI) of the CTDR, for the visualization of the attack graph and the attack-defense graph. Next, we show some representative information, using a Web browser pointing out to the http://0.0.0.0:8082 url.

5.1 Overview

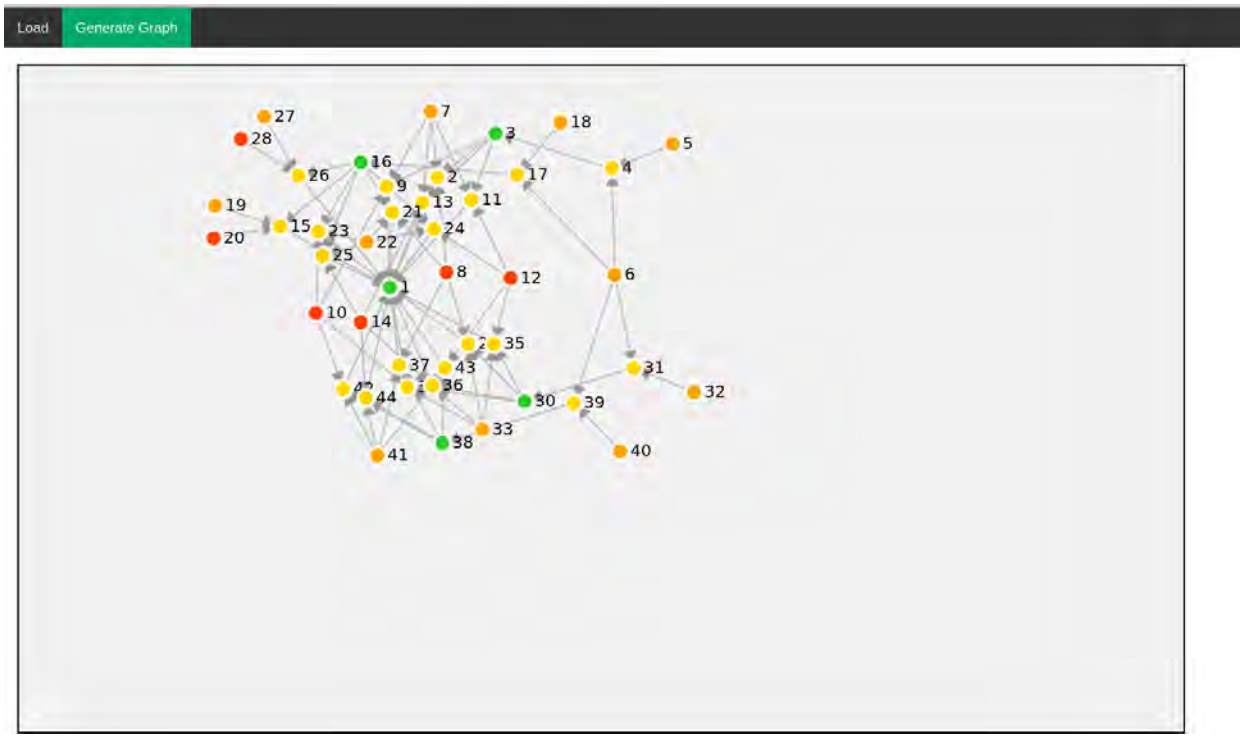
Upload View	Input file	Input file to upload Nessus scan
	Submit button	Button to submit the input for graphs generation
Graph View	Attack Graph	A graphical representation of the step an attacker can take to reach a goal.
	Attack-Defense Graph	A graphical representation of the countermeasures that

		can be applied on the vulnerabilities in order to remediate them.
	Download Graph button	One button to download the attack graph and another one to download

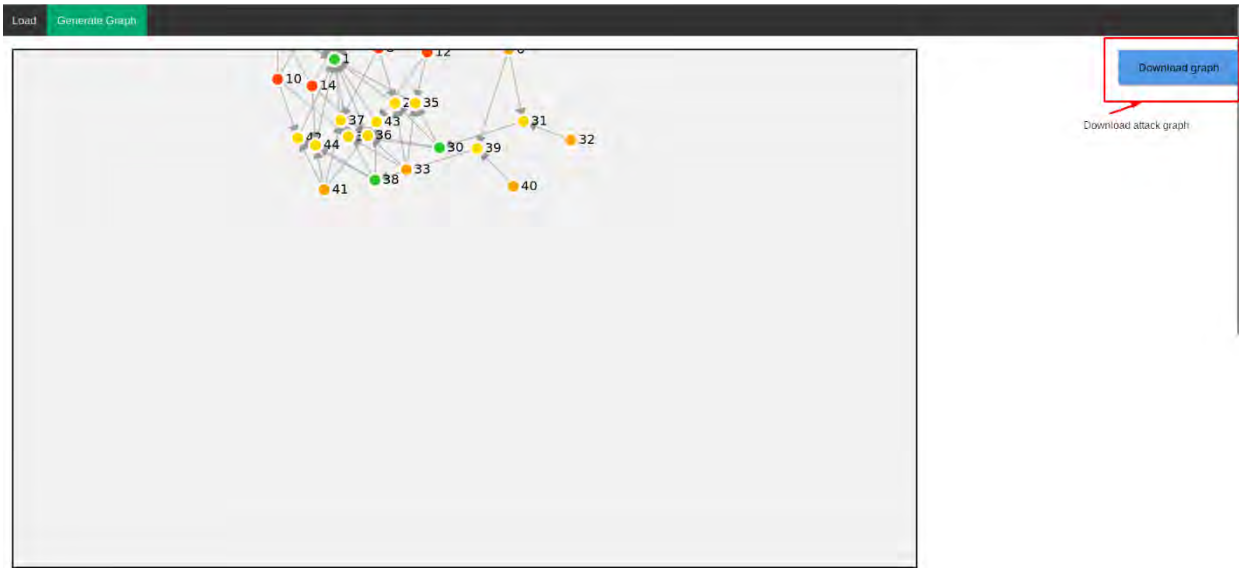
5.2 Content of the Interface

Upload the nessus scan and click on the button Submit.

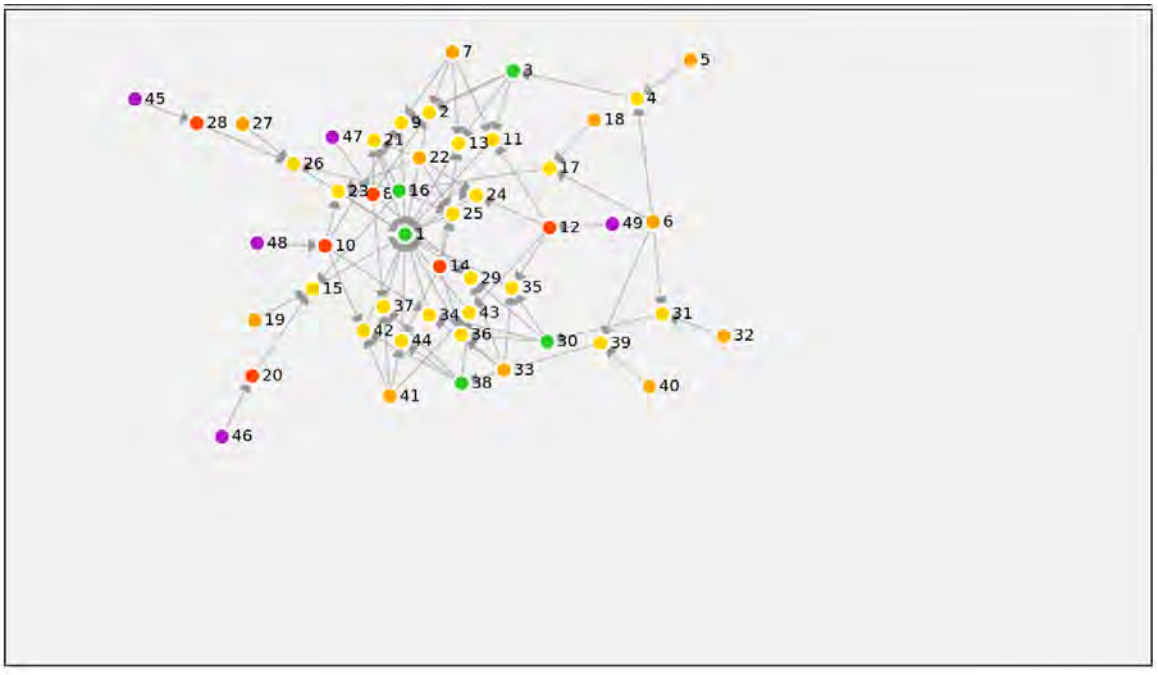
The attack graph is generated and can be visualized.



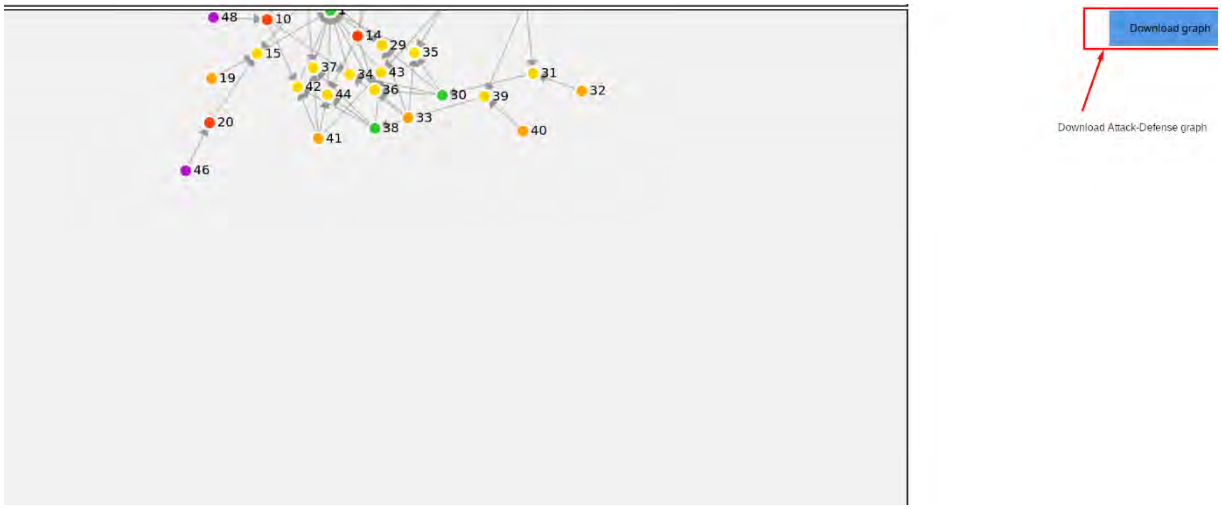
The generated attack graph can be downloaded by clicking on the button Download graph.



The attack-defense graph is generated and can be visualized.



The generated attack-defense graph can be downloaded by clicking on the button Download graph.



6 Conclusion

Prelude is used to generate alerts with the logs received from components of the organization's network. A graphical representation of the distribution of logs is possible on Prewikka, the graphic interface of Prelude. Correlation of alerts is also possible with Prelude. They can also be displayed graphically using Prewikka. In addition, the ELK stack allows Prelude to create additional dashboards for deeper analysis of threats and countermeasures. The attack graph and attack-defense graph are generated based on the output of a Nessus scan. The attack graph is mapped with the alerts from Prelude and a vulnerability ontology in order to update the graphs. Alerts are sent to the IMPETUS platform when a vulnerability existed on the attack graph is exploited.

Members of the IMPETUS consortium

SINTEF	SINTEF, Strindvegen 4, Trondheim, Norway, https://www.sintef.no	Joe Gorman joe.gorman@sintef.no
Institut Mines-Télécom	Institut Mines Telecom, 19 place Marguerite Perey, 91120 Palaiseau, France, https://www.imt.fr	Joaquin Garcia-Alfaro joaquin.garcia_alfaro@telecom-sudparis.eu
UNÎMES	Université de Nimes, Rue du Docteur Georges Salan CS 13019 30021 Nîmes Cedex 1, France, https://www.unimes.fr	Axelle Cadiere axelle.cadiere@unimes.fr
cini consorzio interuniversitario nazionale per l'informatica	Consorzio Interuniversitario Nazionale per l'Informatica, Via Ariosto, 25, 00185 – Roma, Italy, https://www.consorzio-cini.it	Donato Malerba donato.malerba@uniba.it
Università degli Studi di Padova	University of Padova, Via 8 Febbraio, 2 - 35122 Padova, Italy, https://www.unipd.it	Giuseppe Maschio giuseppe.maschio@unipd.it
Entrepreneurship Development Centre for BIOTECHNOLOGY and MEDICINE	Biotehnoloogia ja Meditsiini Ettevõtluse Arendamise Sihtasutus, Tiigi 61b, 50410 Tartu, Estonia, https://biopark.ee	Sven Parkel sven@biopark.ee
SIMAVI	SIMAVI, Complex Victoria Park, Corp C4, Etaj 2, Șos. București – Ploiești, nr. 73 – 81, Sector 1, București, Romania, https://www.simavi.ro	Gabriel Nicola Gabriel.Nicola@simavi.ro Monica Florea Monica.Florea@simavi.ro
THALES	Thales Nederland BV, Zuidelijke Havenweg 40, 7554 RR Hengelo, Netherlands, https://www.thalesgroup.com/en/countries/europe/netherlands	Johan de Heer johan.deheer@nl.thalesgroup.com
CINEDIT INTELLIGENT VIDEO ANALYTICS	Cinedit VA GmbH, Poststrasse 21, 8634 Hombrechtikon, Switzerland, https://www.cinedit.com	Joachim Levy j@cinedit.com
INSIKT INTELLIGENCE	Insikt Intelligence, Calle Huelva 106, 9-4, 08020 Barcelona, Spain, https://www.insiktintelligence.com	Dana Tantu dana@insiktintelligence.com
SIXGILL	Sixgill, Derech Menachem Begin 132 Azrieli Tower, Triangle Building, 42nd Floor, Tel Aviv, 6701101, Israel, https://www.cybersixgill.com	Benjamin Preminger benjamin@cybersixgill.com Ron Shamir ron@cybersixgill.com
Comune di Padova	City of Padova, via del Municipio, 1 - 35122 Padova Italy, https://www.padovanet.it	Enrico Fiorentin fiorentine@comune.padova.it Stefano Baraldi Baraldis@comune.padova.it
Oslo	City of Oslo, Grensen 13, 0159 Oslo, Norway, https://www.oslo.kommune.no	Osman Ibrahim osman.ibrahim@ber.oslo.kommune.no
Institut za sigurnosne politike ISP	Institute for Security Policies, Kruge 9, 10000 Zagreb, Croatia, http://insigpol.hr	Krunoslav Katic krunoslav.katic@insigpol.hr

	International Emergency Management Society, Rue Des Deux Eglises 39, 1000 Brussels, Belgium, https://www.tiems.info	K. Harald Drager khdrager@online.no
	Unismart – Fondazione Università degli Studi di Padova, Via VIII febbraio, 2 - 35122 Padova, Italy, https://www.unismart.it	Alberto Da Re alberto.dare@unismart.it

Grant number: 883286
Project duration: Sep 2020 – Aug 2022
Project Coordinator: Joe Gorman, SINTEF

Horizon 2020: Secure societies
SU-INFRA02-2019
Security for smart and safe cities, including for public spaces
Project Type: Innovation Action

http://www.impetus-project.eu

CTI - Cyber Threat Intelligence

Author: **Ron Ofer**

Version: **1**

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Table of Contents

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

1. Actionable Alerts Module

1.1 Actionable Alerts Module – Introduction

The main module in the Cyber Threat Intelligence (CTI) Tool that is allocated for the IMPETUS IT Operator is the Actionable Alerts module. Through this module, the IT Operator receives notifications in the IMPETUS CTI Dashboard for new incoming alerts and is able to immediately understand his security posture. The alerts are based on the assets that the user has populated beforehand in the asset's configuration page on Cybersixgill's Portal. For an in-depth investigation of the nature of each alert, the IT Operator can easily pivot into the Cybersixgill's Portal and obtain all relevant information to mitigate the risk.

1.2 Assets Configuration

The first step, before you can get any alerts, need to login into Cybersixgill's Portal and inside the setting page, configure your assets. An asset is any data, device, or other components of the environment that supports information-related activities. As such, it's important to configure any asset that your organization owns and potential security risks or gaps can affect it.

By configuring these assets, you can rest assured that every time, each one of your assets will be mentioned in any context in the clear web, deep and dark web forums and markets, instant messaging platform, and more, you will be notified.

The more assets you configure, the more alerts you get. However, it's important to insert relevant assets in an accurate manner, to avoid FP and spam alerts. In each of the assets' category, there is best practices of how to insert the assets (such as format).

Figure 1 – Login page

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

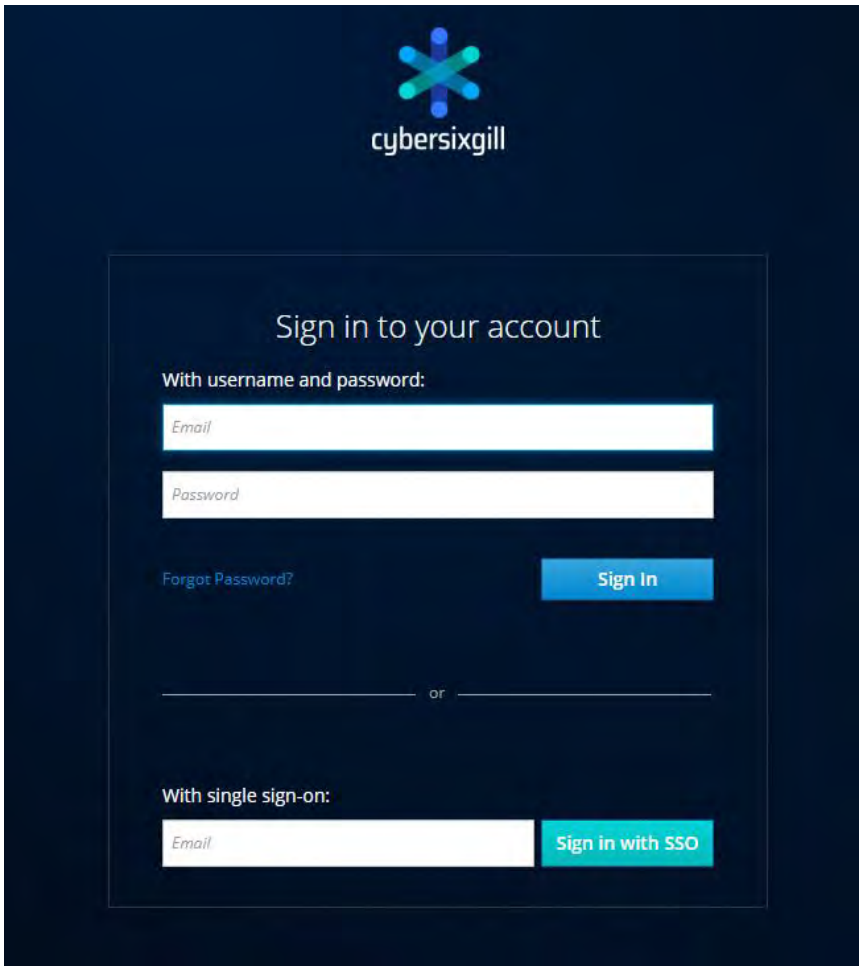


Figure 2 – Settings page

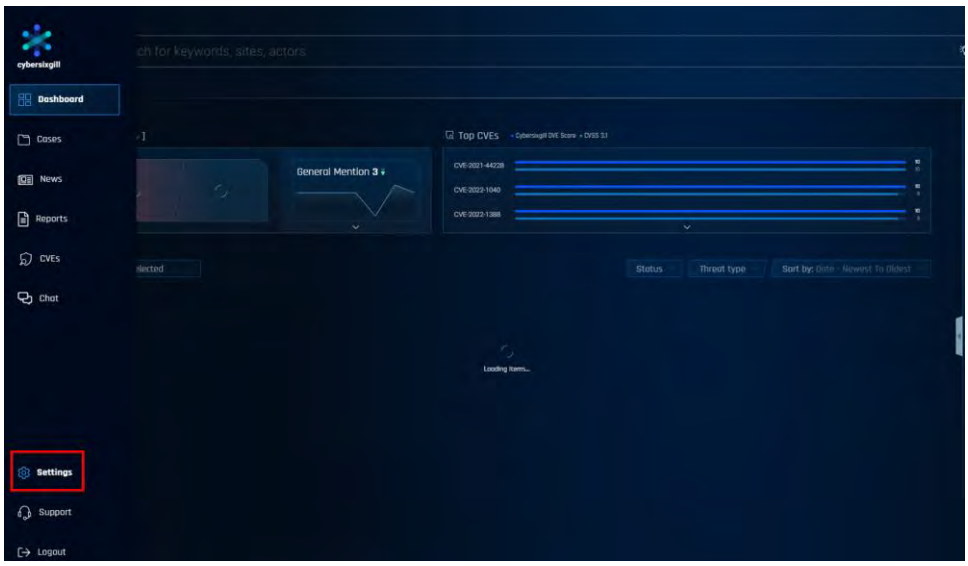


Figure 3 – Asset's configuration page

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

1.3 IMPETUS Platform

Once you are done with configuring all assets in the Portal, you will start getting the alerts. Now you need to login to IMPETUS platform (figure 4), navigate into the CTI dashboard (figure 5) and wait for the alerts notifications to appear (figure 6).

Figure 4 – IMPETUS Platform

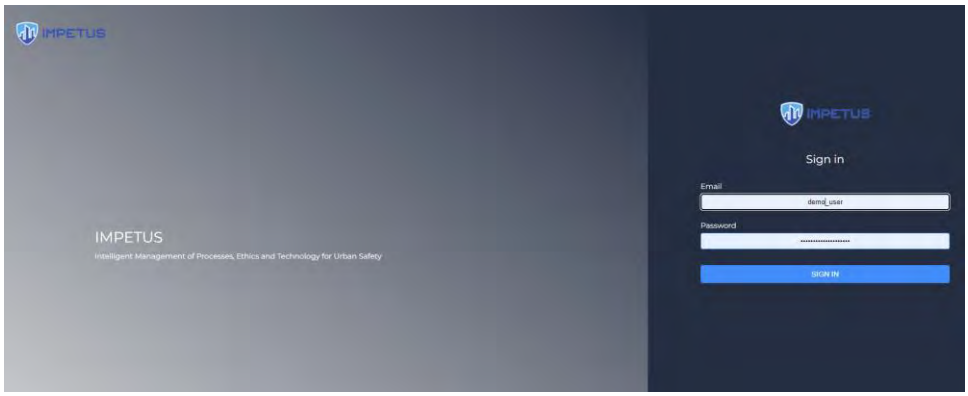


Figure 5 – CTI tool

Figure 6 – CTI Dashboard

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

1.4 Redirect into Cybersixgill's Portal

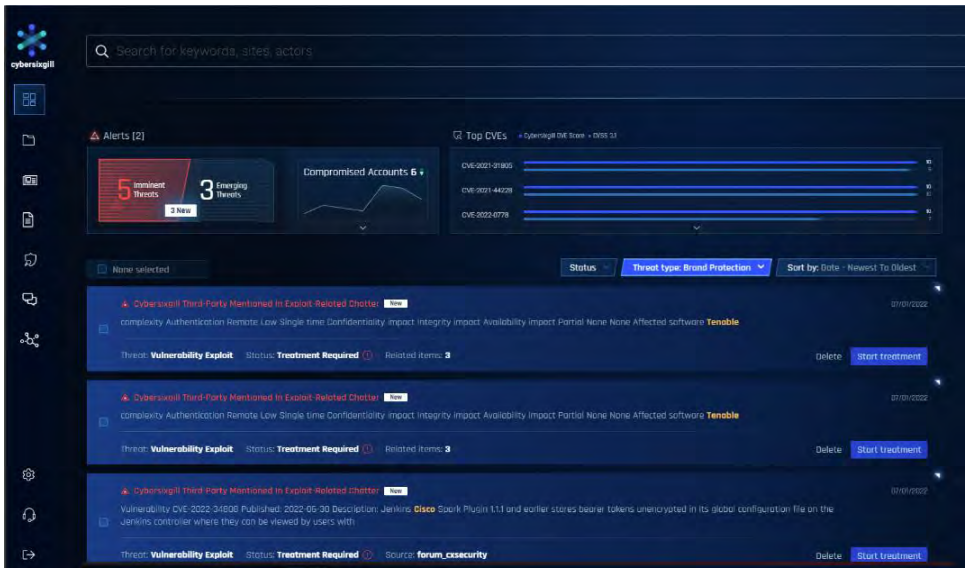
When the user presses the redirect button on the CTI dashboard in the IMPETUS platform, he is being redirected to Cybersixgill's Portal. The page he lands at is already being filtered by the threat type he pressed on the dashboard.

For example, pressing the redirect button on the "brand protection" threat type, in the list of "Imminent threats" will bring you into the Cybersixgill's Portal where the alerts queue is already filtered accordingly (figure 8).

Figure 7 – CTI Dashboard handling – redirect button

Figure 8 – Redirected into Cybersixgill's Portal

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.



1.5 Actionable Alerts Queue

The alerts are divided into 2 types, which you can filter the queue by:

1. Imminent - The asset is mentioned in a prominent forum or there is a clear attack vector. Alert that is posing a concrete or real threat to the customer, and the customer can act upon it. These alerts usually include a clear modus operandi (DDoS attack, Data leak, etc.), will be related to a specific asset (alias, domain etc.) and will be contextual (for example, include actor attribution).

2. Emerging - Alert that is not posing a concrete or real threat to the customer, but may emerge to such in the future. These alerts include sector alerts (campaigns against vertical or geography), threats with missing context (Customer's name was mentioned, but no modus operandi present) or threats which we cannot validate that are targeting the customer. These alerts will usually be less actionable.

Important Note: If you use any filter, remember to disable it once you want to see all results. The filter will still be on until you disable it (refreshing the page won't help).

Figure 9 – Alerts Queue

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

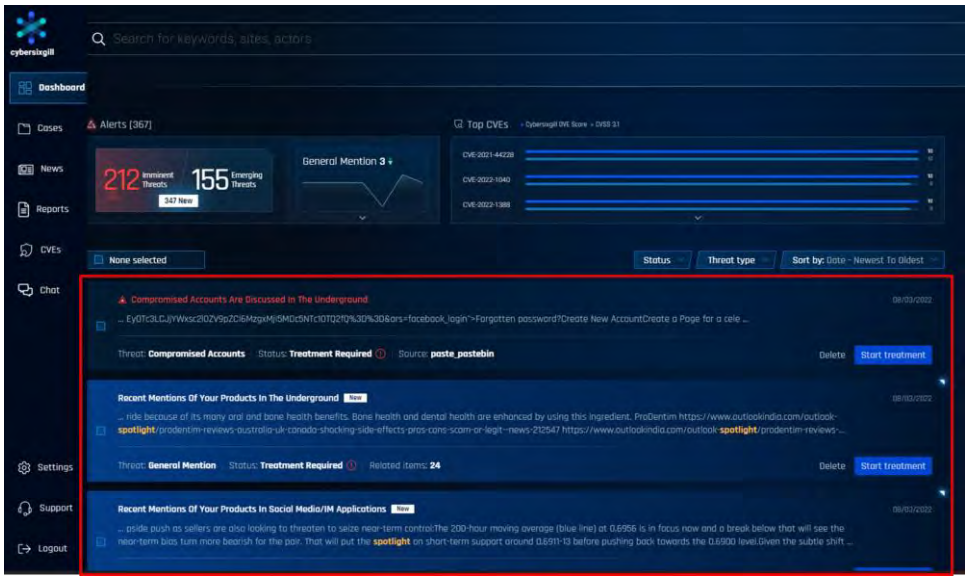


Figure 10 – Alerts Filters

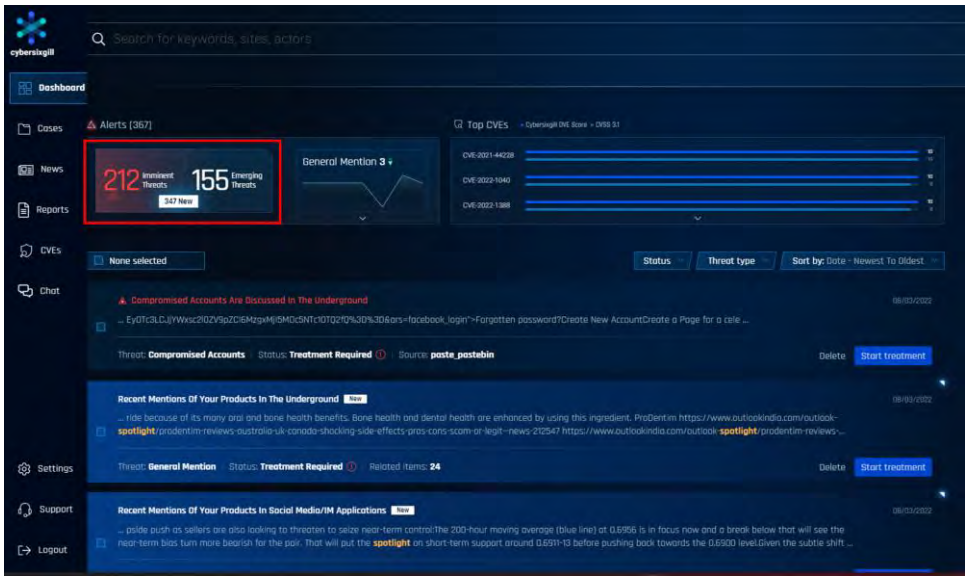
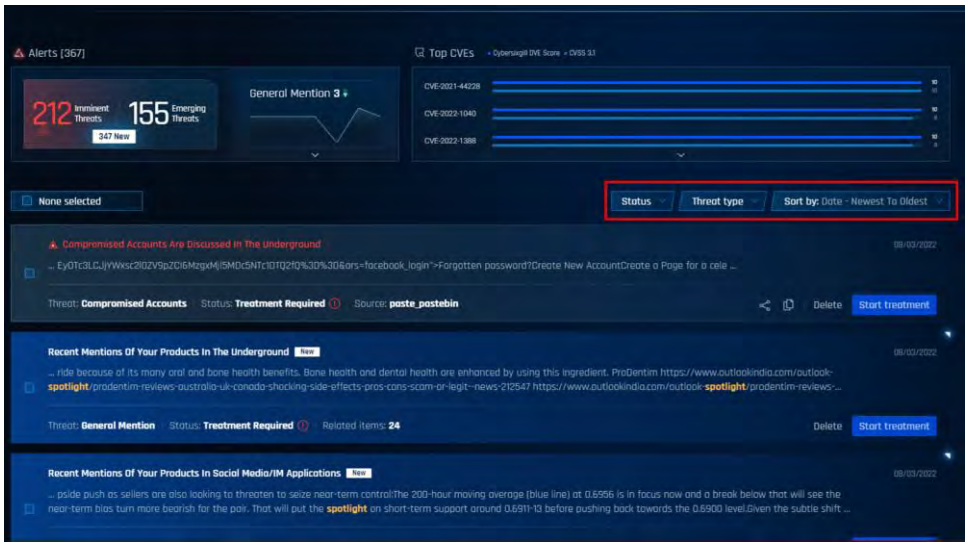


Figure 11 – Alerts Filters continue

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.



1.6 Consuming the Actionable Alerts

By clicking any alert in the queue, you will open the alert and start your investigation. There are 2 kinds of alerts:

1. Single Alert – only one asset triggered an alert of a specific threat type. Opening this alert will show the user only one post that is relevant to a specific asset that triggered the alert.
2. Aggregated Alert – one or more assets triggered more than one alert of a specific threat type. Opening this alert will show the user a few "sub-alerts" each for a specific asset that triggered a sub-alert.

Figure 12 – Single Alerts Structure

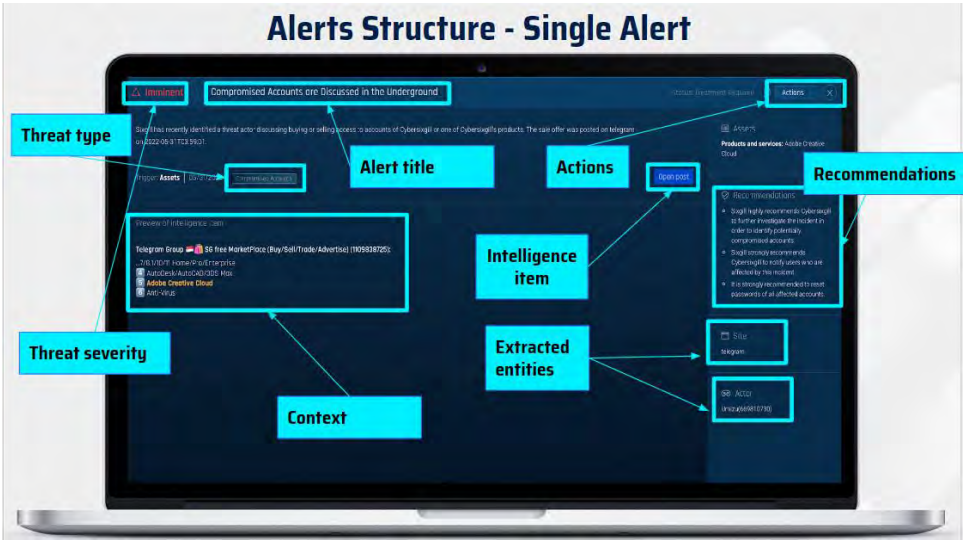


Figure 13 – Aggregated Alerts Structure

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Figure 14 – Alerts handling

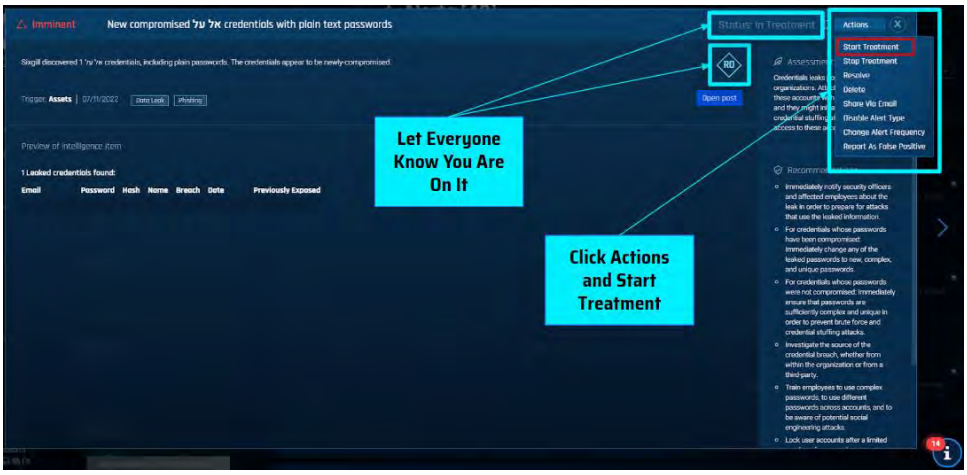


Figure 15 – Alerts handling continue

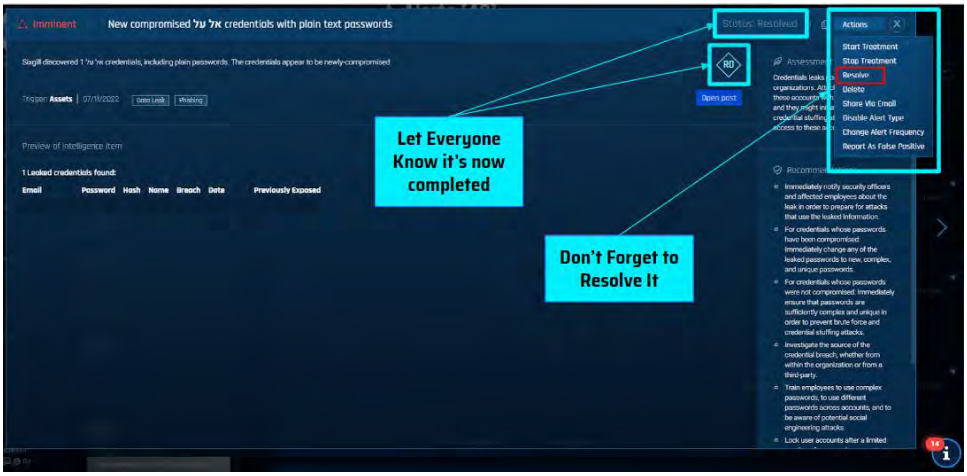
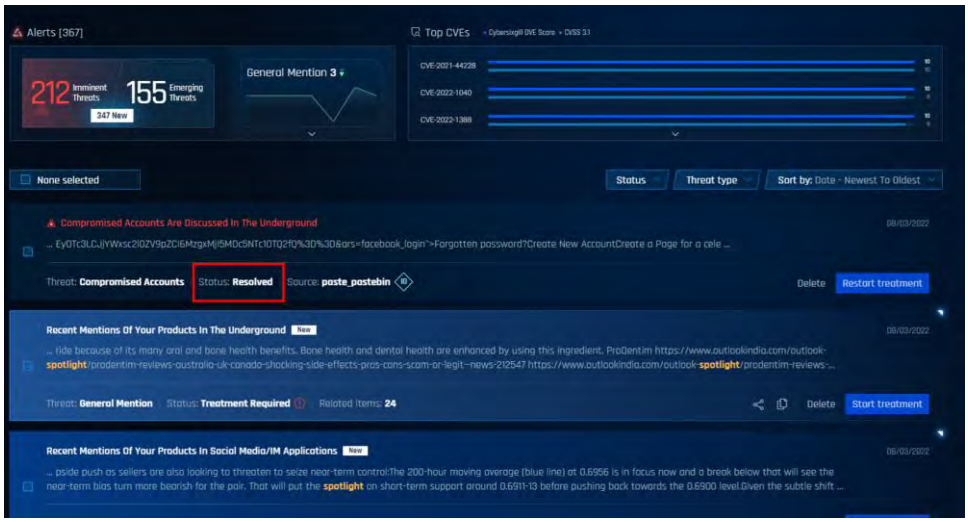


Figure 16 – Alerts handling continue

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

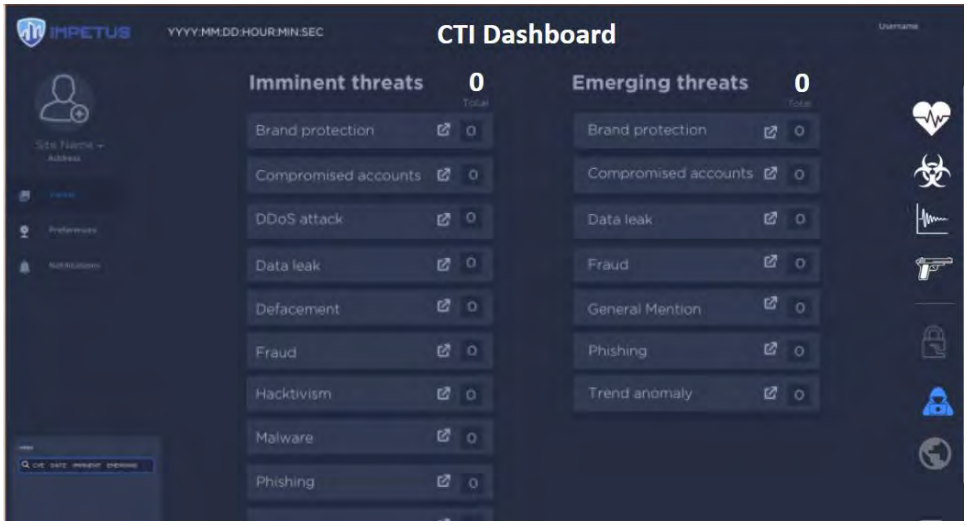


1.7 Alerts management

When you are done resolving all alerts in the queue, your CTI dashboard will be updated accordingly and will show no further notifications for new alerts.

Now it's time to rest.

Figure 17 – Clean CTI Dashboard



The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

2. Investigation Module

2.1 Introduction to Investigation Module

Deep dive into any escalation in real-time and understand the context. Research threat actor's profile, modus operandi, and history. Review and analyze across languages, sites, timeframes, types of products, topics, entities, and more.

The search bar allows you to search Cybersixgill's large-scale data lake for any sort of information we have collected with a simple query language. All items we collect are being indexed and saved in our database, and are still searchable even if later on they are deleted from the original source.

Figure 18 – Search Bar

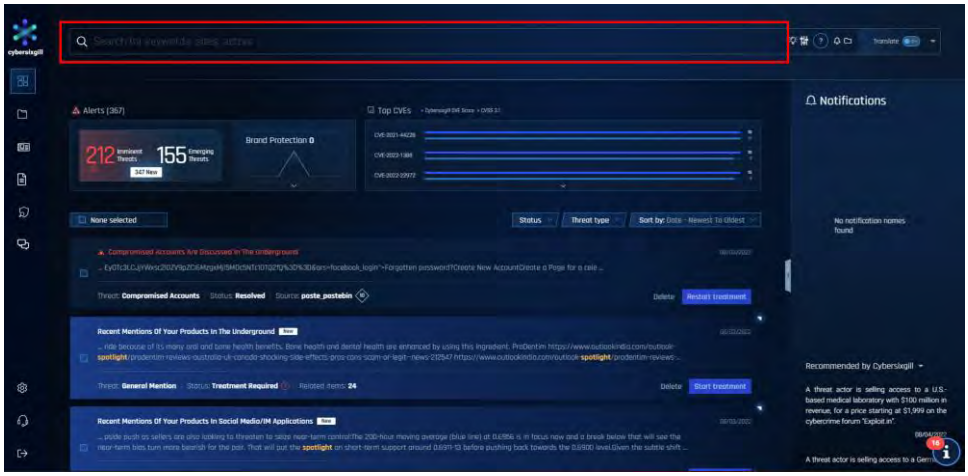


Figure 19 – Search Assistant button

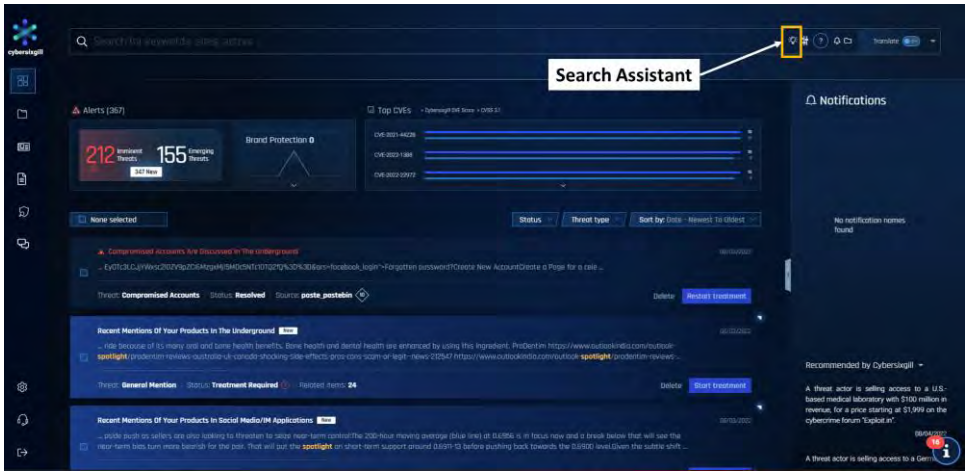


Figure 20 – Search Assistant page

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

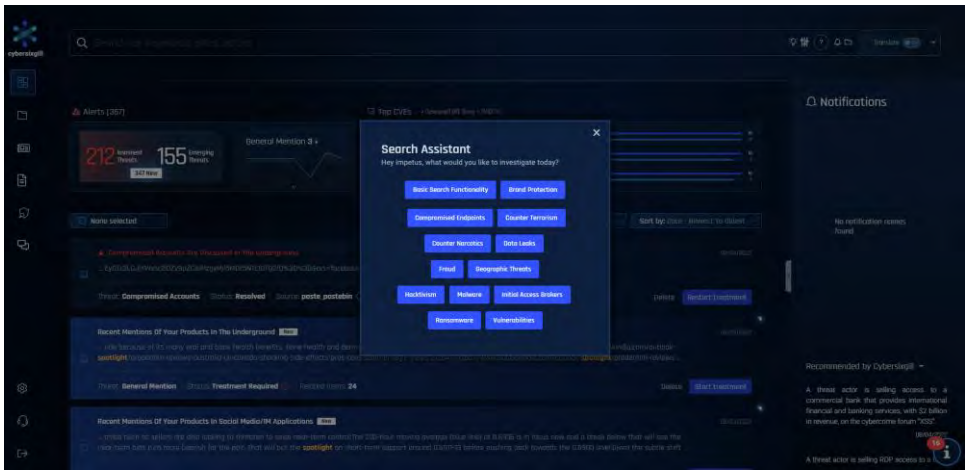


Figure 21 – Advanced Search button

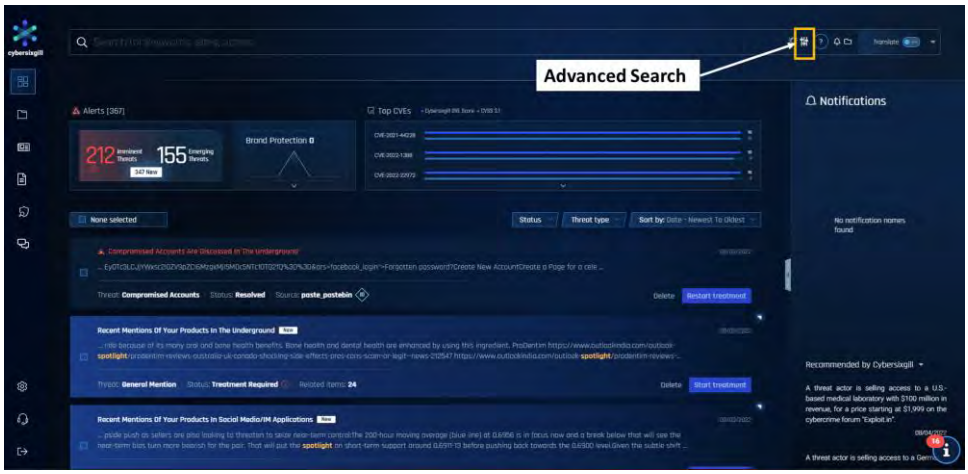


Figure 22 – Advanced Search page

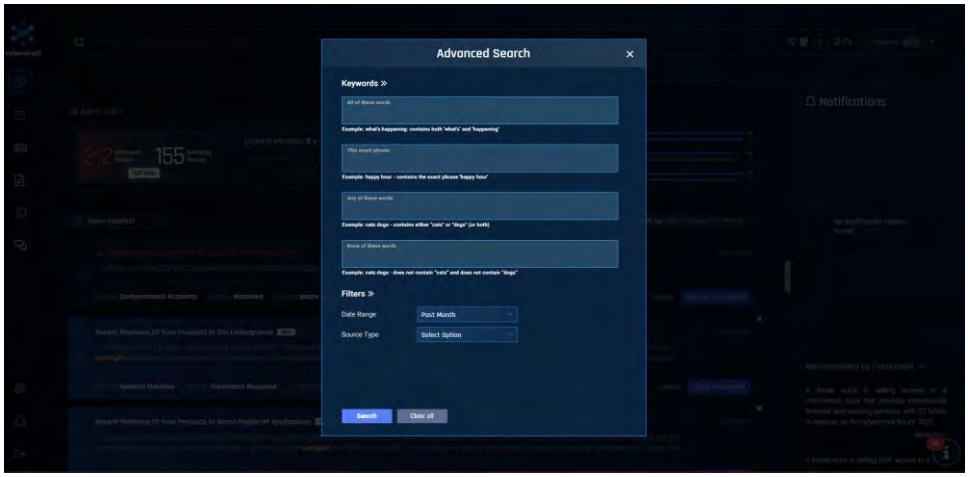


Figure 23 – How to Query button

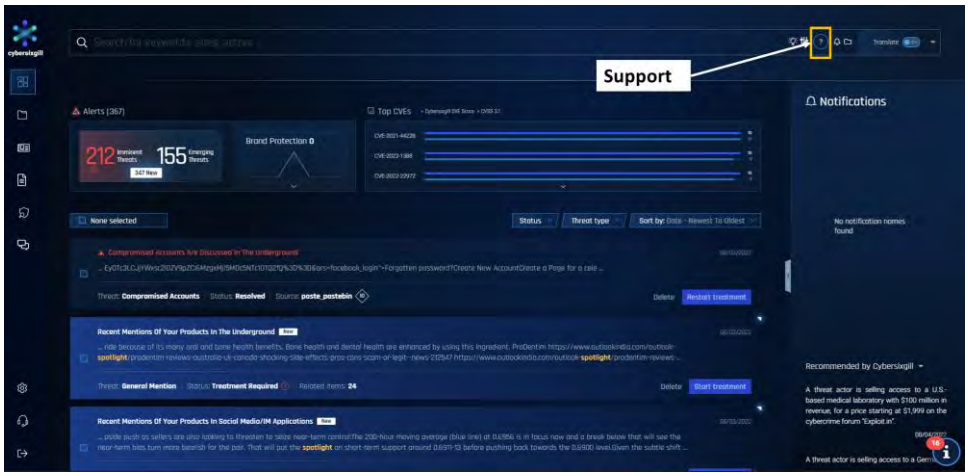
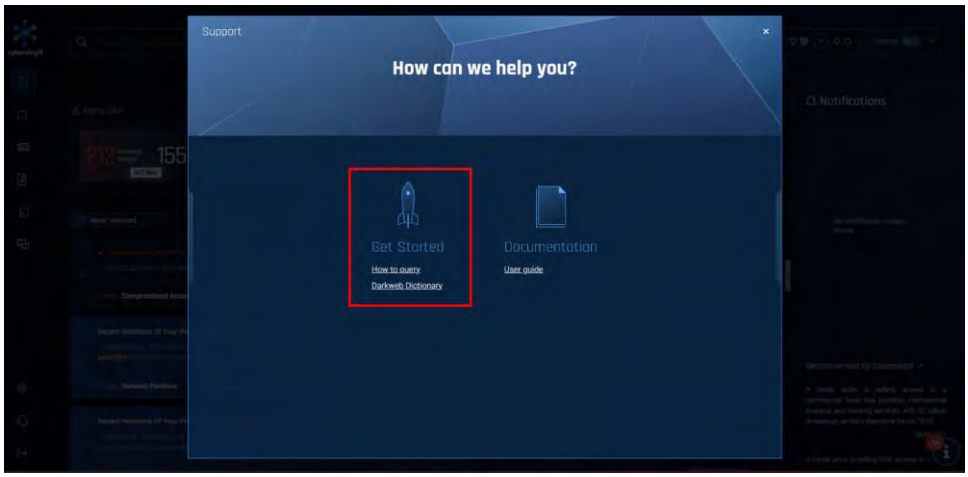


Figure 24 – How to Query page



The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

3. CVE Module

3.1 Introduction to CVE Module

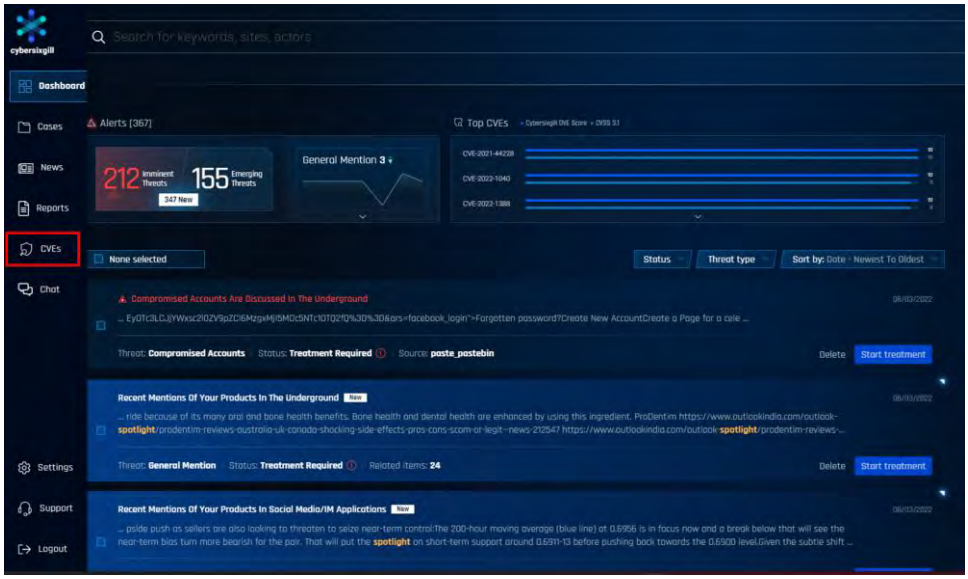
CVE identification and prioritization are a critical part of every vulnerability management tool, and an integral component in any risk assessment. With over a thousand new CVEs published every month, this is an exponential challenge to cybersecurity professionals.

To overcome this challenge, Cybersixgill has developed the CVE Dynamic Rating, bringing together the worlds of vulnerability management and cyber threat intelligence. The CVE Dynamic Rating is comprised of state-of-the-art machine learning models that automatically predict the probability of a CVE being exploited. The Rating is derived from automated AI analysis of underground discourse on deep and dark web forums, and combined with intelligence from other sources -- such as code repositories and technical know-how – empowers you to track threats from CVEs that most others define as irrelevant or obsolete, but have a higher probability of being exploited in the near term by active threat actors in the cyber underground.

The mechanism is powered by Cybersixgill's best-in-class proprietary collection methodology and the cutting-edge technology that supports it. Our real-time automated collection provides users with the broadest, deepest, and most comprehensive access to threat intelligence from underground sources (we collect all items without filtering). These include deep and dark restricted access web forums and markets, IRC channels, underground paste sites, as well as closed-access groups on Telegram, QQ, and Discord.

The Sixgill CVE Dynamic Rating provides a complementary approach to NVD's static CVSS score, adding the dimension of probability. Ultimately, Sixgill's CVE solution helps you answer that critical question - how likely is this CVE to be exploited in the near term - so you can stay ahead of the threats that are out there. The solution is available as an add-on module in Sixgill's SaaS portal, as well as via API integrations.

Figure 25 – Navigate to CVE Module



The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Figure 26 – CVE Module page (before configuring your CVEs)

Figure 27 – Configuring CVEs in the assets page

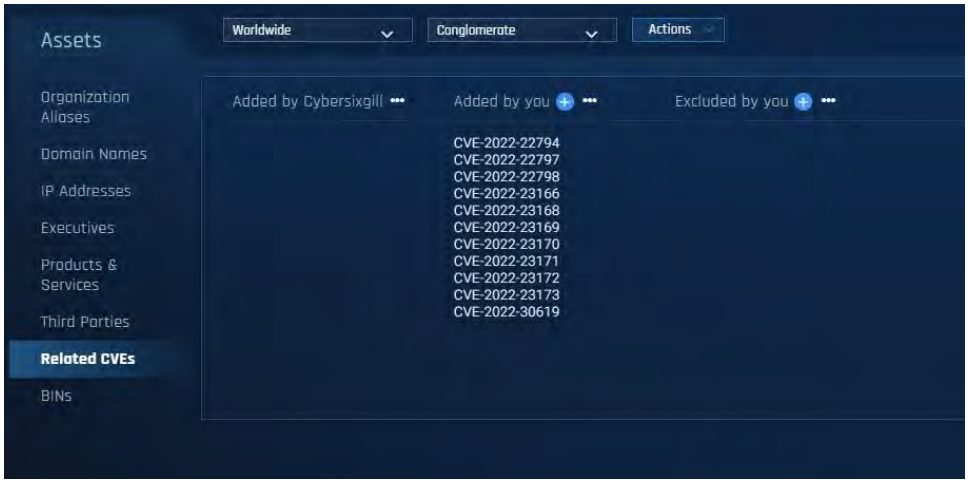


Figure 28 – CVE Module page after inserting your CVEs

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4. Support

Figure 29 – Support Center button

Figure 30 – Support Center

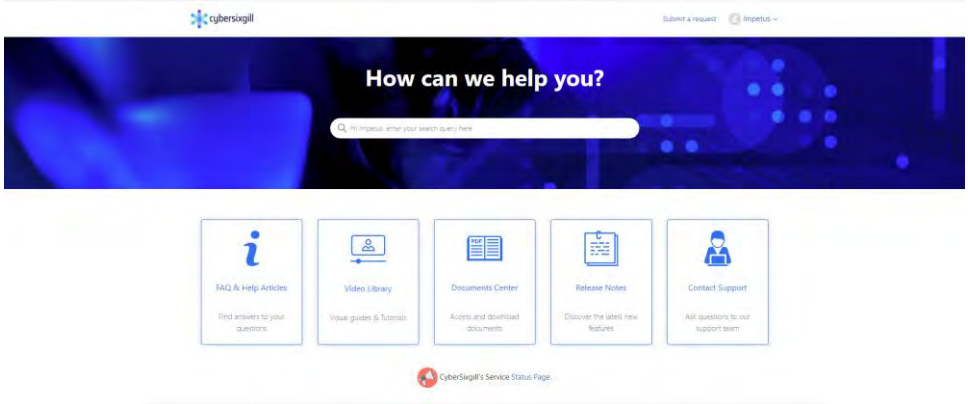


Figure 31 – Resource Center button

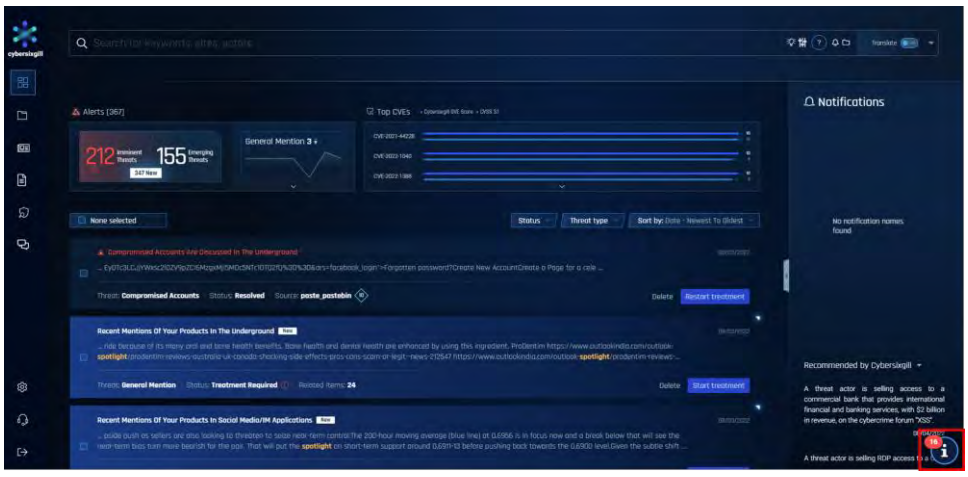
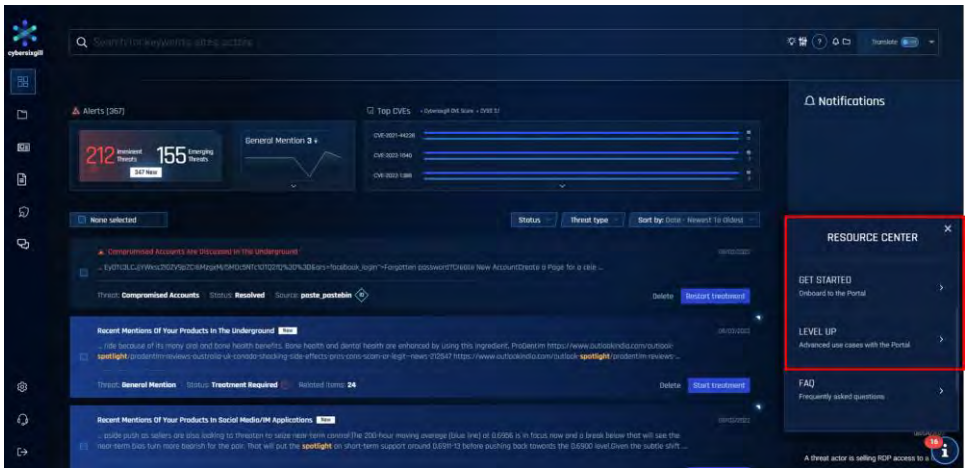


Figure 32 – Resource Center Interactive Guides

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.



4.1 Important Note:

For more in-depth explanations we suggest you take a look at our complete Portal User Guide (inside the documents center) and for any unclear matters, please contact us directly through the support center (figure 30).

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Members of the consortium

SINTEF	SINTEF, Strindvegen 4, Trondheim, Norway, https://www.sintef.no	Joe Gorman joe.gorman@sintef.no
Institut Mines-Télécom	Institut Mines Telecom, 19 place Marguerite Perey, 91120 Palaiseau, France, https://www.imt.fr	Joaquin Garcia-Alfaro joaquin.garcia_alfaro@telecom-sudparis.eu
UNÎMES	Université de Nimes, Rue du Docteur Georges Salan CS 13019 30021 Nîmes Cedex 1, France, https://www.unimes.fr	Axelle Cadiere axelle.cadiere@unimes.fr
cini consorzio interuniversitario nazionale per l'informatica	Consorzio Interuniversitario Nazionale per l'Informatica, Via Ariosto, 25, 00185 – Roma, Italy, https://www.consorzio-cini.it	Donato Malerba donato.malerba@uniba.it
Università degli Studi di Padova	University of Padova, Via 8 Febbraio, 2 - 35122 Padova, Italy, https://www.unipd.it	Giuseppe Maschio giuseppe.maschio@unipd.it
Entrepreneurship Development Centre for BIOTECHNOLOGY and MEDICINE	Biotehnoloogia ja Meditsiini Ettevõtluse Arendamise Sihtasutus, Tiigi 61b, 50410 Tartu, Estonia, https://biopark.ee	Sven Parkel sven@biopark.ee
SIMAVI Software Imagination & Vision	SIMAVI, Complex Victoria Park, Corp C4, Etaj 2, Șos. București – Ploiești, nr. 73 – 81, Sector 1, București, Romania, https://www.simavi.ro	Gabriel Nicola Gabriel.Nicola@simavi.ro Monica Florea Monica.Florea@simavi.ro
THALES	Thales Nederland BV, Zuidelijke Havenweg 40, 7554 RR Hengelo, Netherlands, https://www.thalesgroup.com/en/countries/europe/netherlands	Johan de Heer johan.deheer@nl.thalesgroup.com
CINEDIT INTELLIGENT VIDEO ANALYTICS	Cinedit VA GmbH, Poststrasse 21, 8634 Hombrechtikon, Switzerland, https://www.cinedit.com	Joachim Levy j@cinedit.com

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

INSIKT INTELLIGENCE	Insikt Intelligence, Calle Huelva 106, 9-4, 08020 Barcelona, Spain, https://www.insiktintelligence.com	Dana Tantu dana@insiktintelligence.com
SIXGILL	Sixgill, Derech Menachem Begin 132 Azrieli Tower, Triangle Building, 42nd Floor, Tel Aviv, 6701101, Israel, https://www.cybersixgill.com	Benjamin Preminger benjamin@cybersixgill.com Ron Shamir ron@cybersixgill.com
Comune di Padova	City of Padova, via del Municipio, 1 - 35122 Padova Italy, https://www.padovanet.it	Enrico Fiorentin fiorentine@comune.padova.it Stefano Baraldi Baraldis@comune.padova.it
Oslo	City of Oslo, Grensen 13, 0159 Oslo, Norway, https://www.oslo.kommune.no	Osman Ibrahim osman.ibrahim@ber.oslo.kommune.no
Institut za sigurnosne politike ISP	Institute for Security Policies, Kruge 9, 10000 Zagreb, Croatia, http://insigpol.hr	Krunoslav Katic krunoslav.katic@insigpol.hr
T S I M E	International Emergency Management Society, Rue Des Deux Eglises 39, 1000 Brussels, Belgium, https://www.tiems.info	K. Harald Drager khdrager@online.no
UniSMART Fondazione Università di Padova	Unismart – Fondazione Università degli Studi di Padova, Via VIII febbraio, 2 - 35122 Padova, Italy, https://www.unismart.it	Alberto Da Re alberto.dare@unismart.it

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Grant number: 883286
Project duration: Sep 2020 – Aug 2022
Project Coordinator: Joe Gorman, SINTEF

Horizon 2020: Secure societies
SU-INFRA02-2019
Security for smart and safe cities, including for public spaces
Project Type: Innovation Action

http://www.impetus-project.eu

FD – Firearm Detector

Author: **Joachim Levy**

Version: **1**

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Table of Contents

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

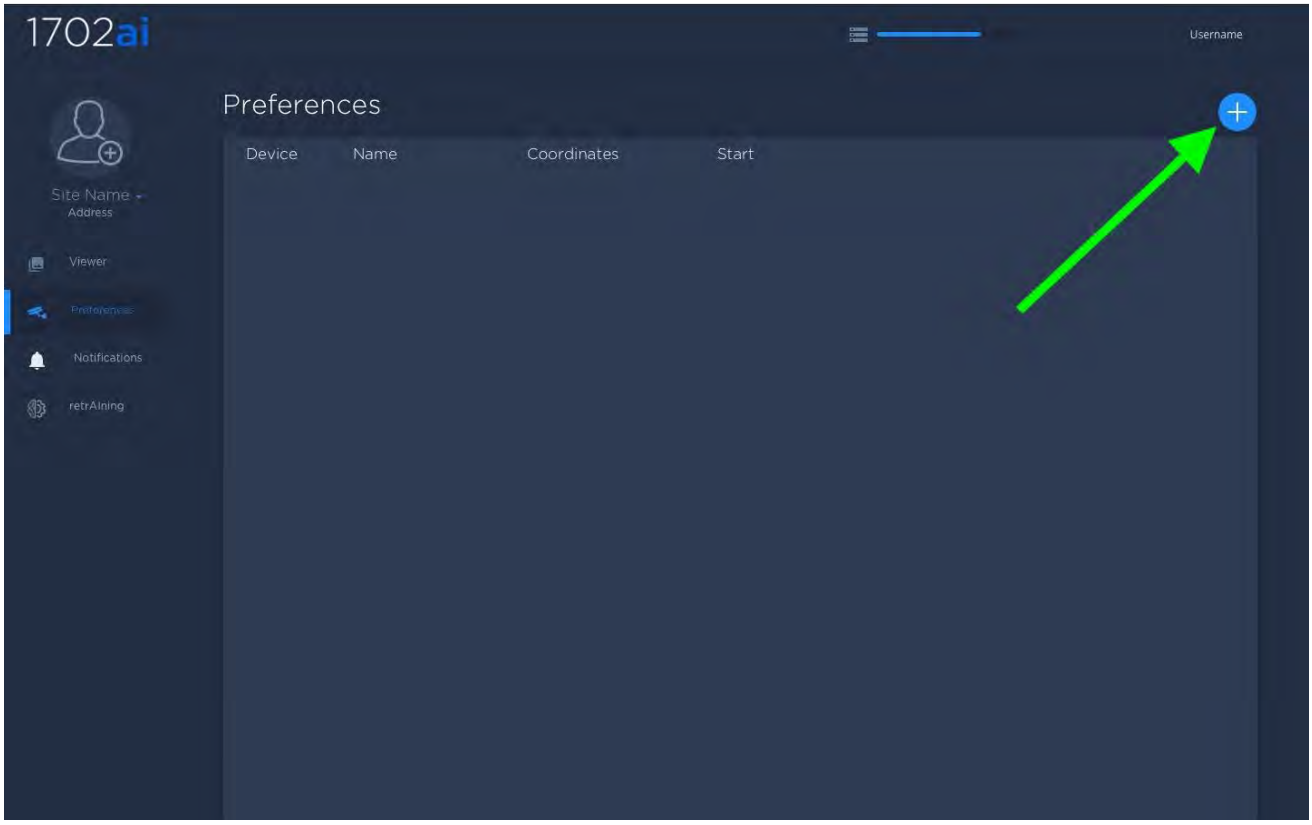
1. Alerts Module

1.1 Alerts Module – Introduction

The main module in the Firearm Detector (FD) Tool is aimed for the IMPETUS SoC Operator. Through this module, the SoC Operator receives notifications in the IMPETUS FD Dashboard for new incoming alerts and is able to immediately validate the alerts. The alerts are based on the surveillance camera streams that the user has provided beforehand in the configuration.

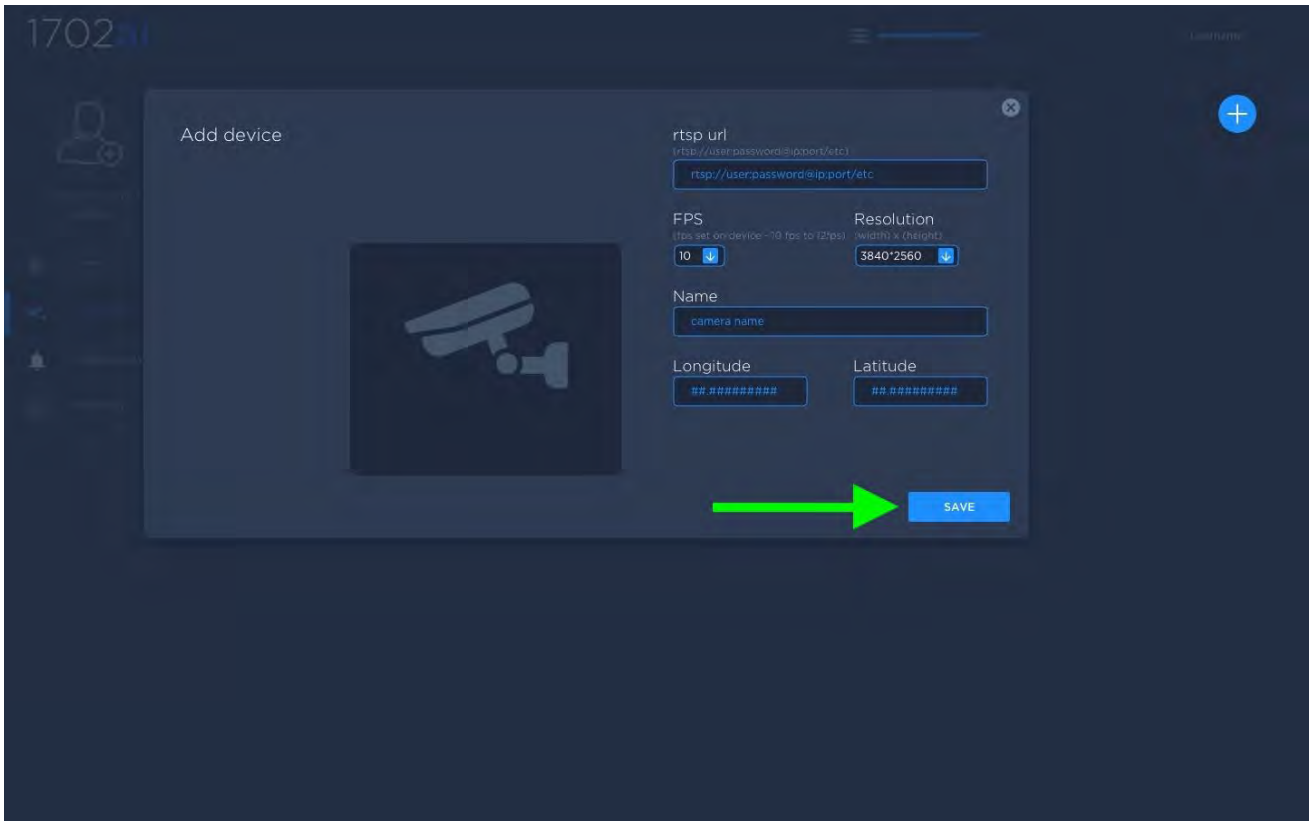
1.2 Surveillance Camera Configuration

The first step, before you can get any alerts, is to provide the surveillance cameras rtsp url into the FD tool preference page. Click the + button in the dashboard in order to add a surveillance camera.

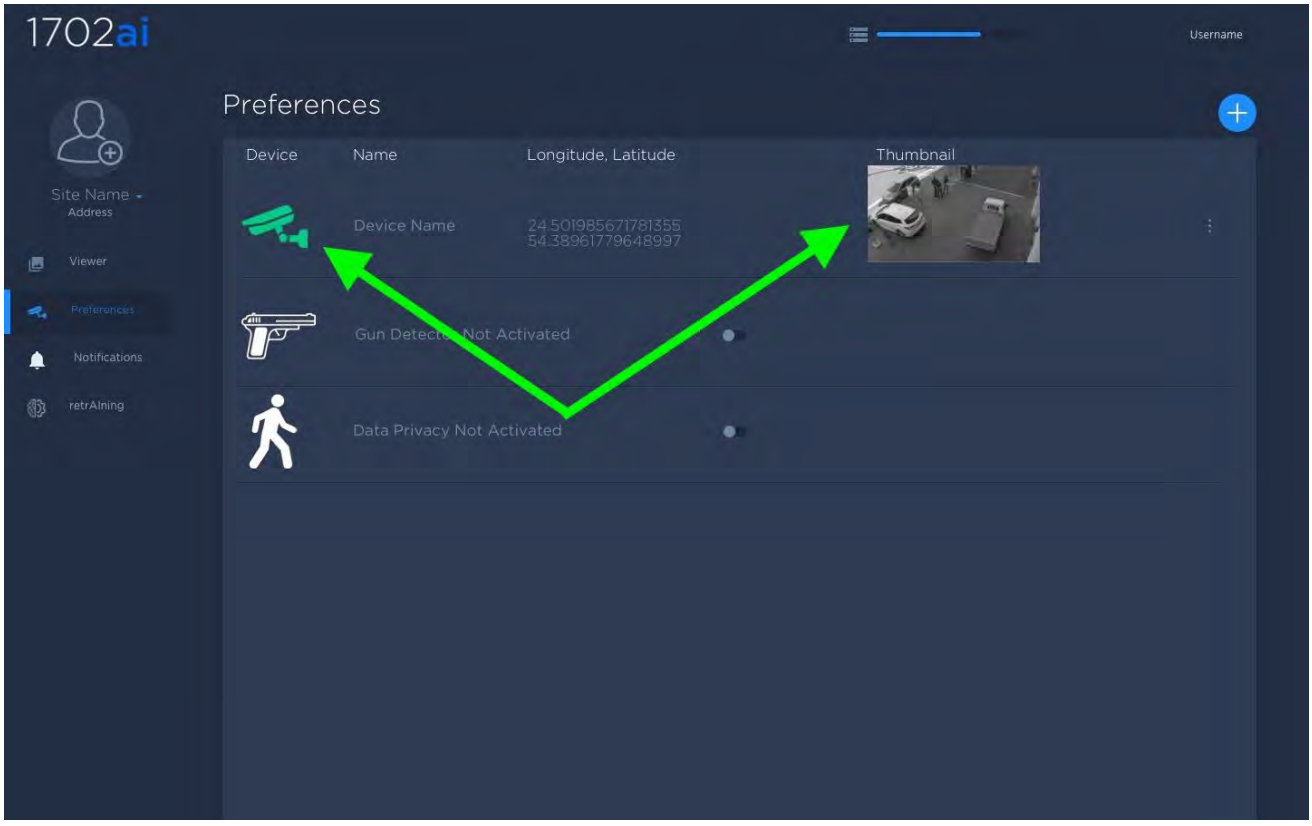


Then add the proper rtsp url of the camera along with the longitude and the latitude of the device and the name of the device. Using the dropdown menu, select the proper resolution and the fps (frame per seconds). The resolution and the fps must match your current surveillance camera settings.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

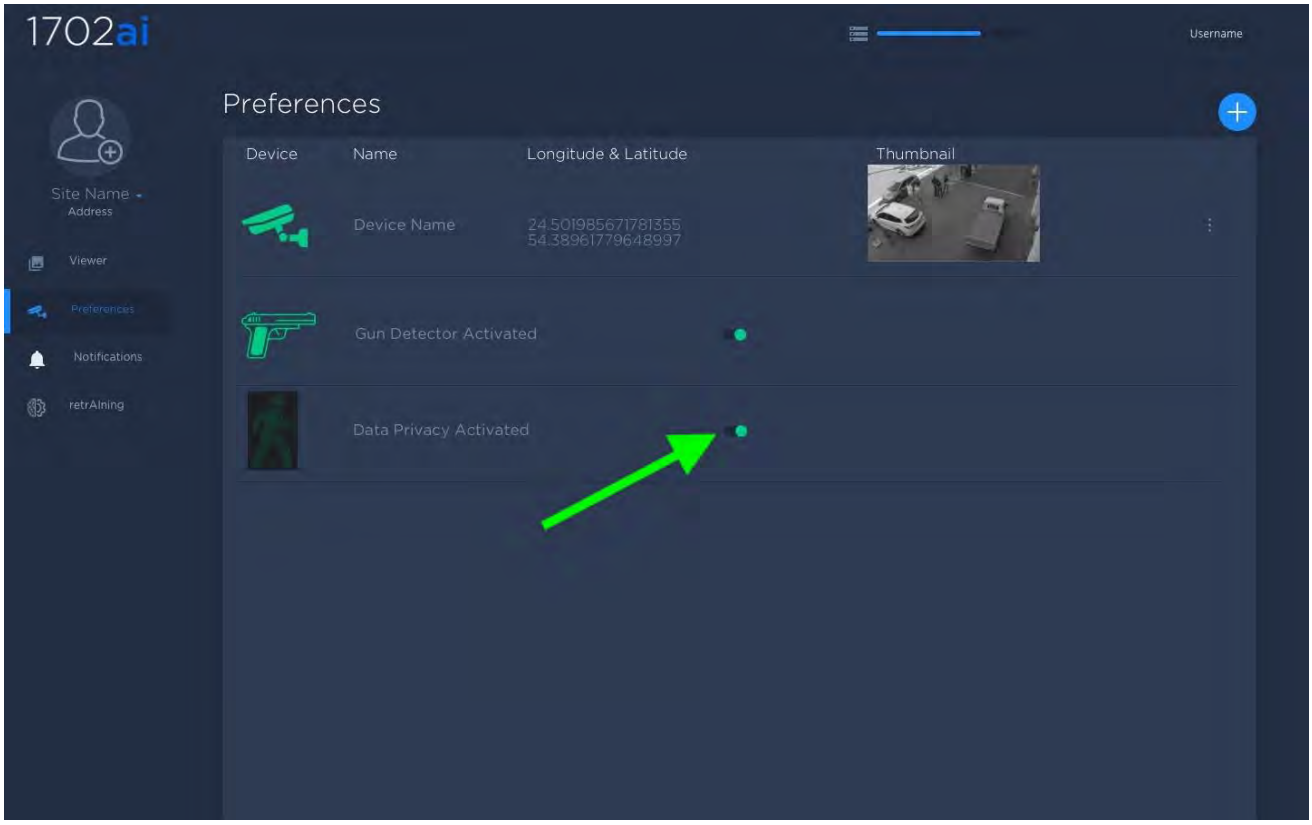


Click on the blue SAVE button. If the correct rtsp and camera settings are provided the camera icon turns green and a low resolution thumbnail of the stream will appear on the dashboard.



The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

By default the FD will not run until the user activates the Data Privacy slider. Once activated the weapon detector is automatically running.



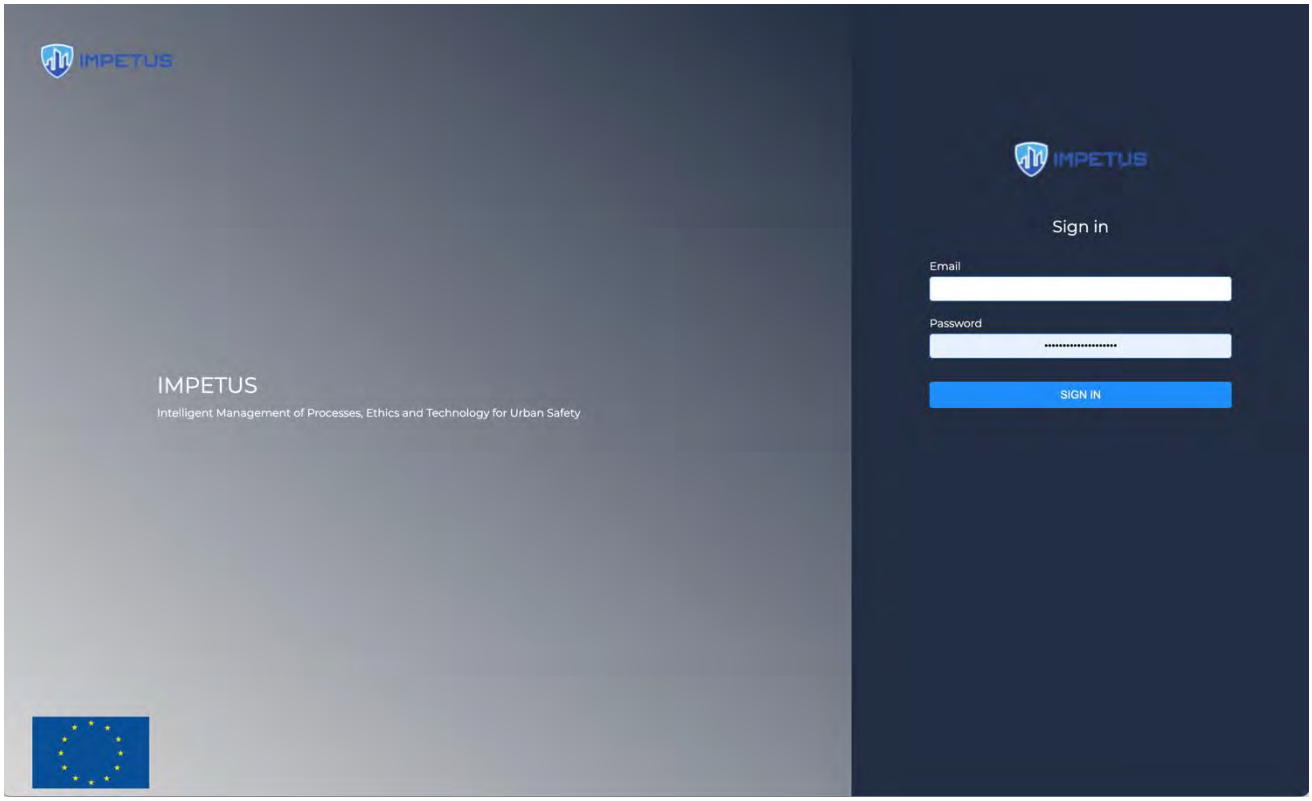
The FD tool is now running, the users can now login to the IMPETUS URL. Please refer to 1.3 IMPETUS Platform for further details.

1.3 IMPETUS Platform

Once you are done configuring all assets in the preference pane, you can monitor the alerts. The next step requires the end user to login to the IMPETUS platform (figure 4) using a web browser. We recommend firefox of chrome as a web browser. The IMPETUS platfrom url is to be provided by your integrator.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Figure 4 – IMPETUS Platform



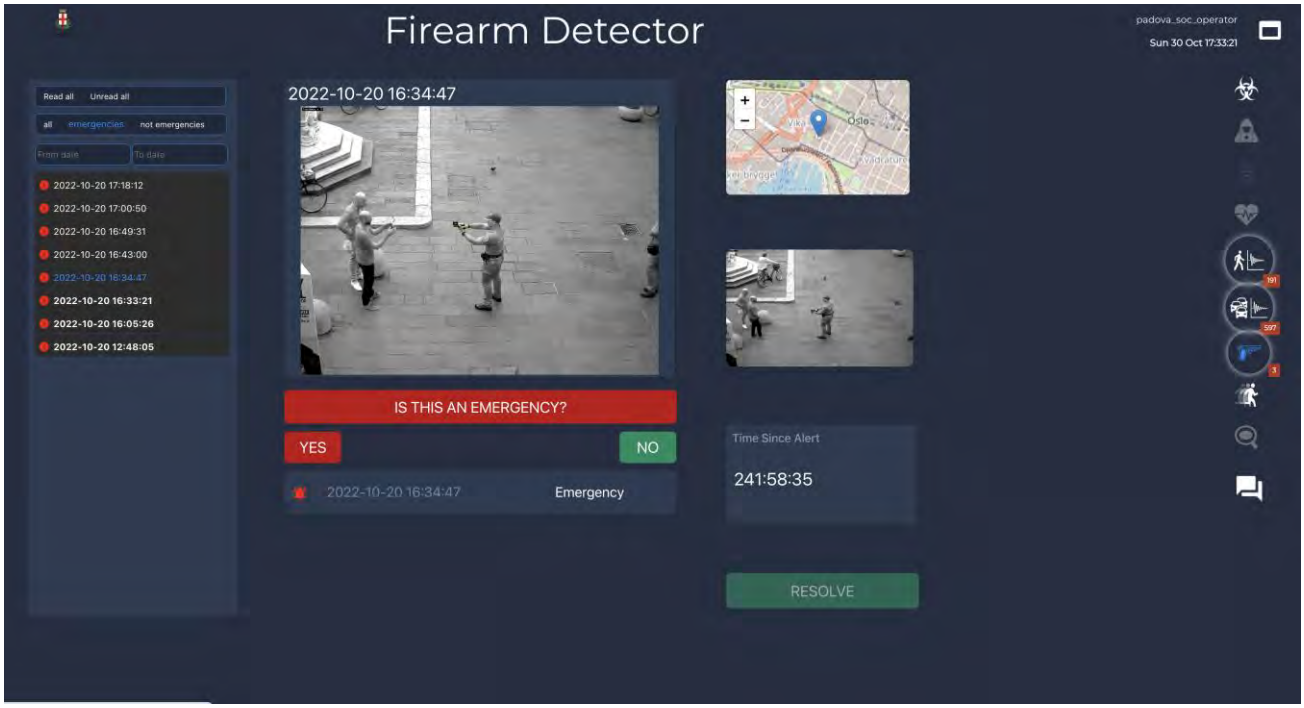
Log in using the email or user name along with the password you were provided with by your integrator. Once logged in, navigate into the FD dashboard (figure 5) by either selecting the dashboard icon or the sidebar icon (displayed using a green rectangle and circle in the figure below)

Figure 5 – FD tool

By selecting the FD icon, it will redirect you the FD tool Dashboard (figure 6).

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Figure 6 – FD Dashboard



The Firearm Dashboard (Figure 7 below – FD Dashboard Detailed) is composed of the following sections:

index: displays contextual alerts and allows for certain search functionalities

canvas: in this section the video alerts are displayed

action area: the action buttons can share an alert with law enforcement using the YES button or label an alert as false alert using the NO button.

map viewer: displays the GPS coordinates of the surveillance camera position. The + and – transport controls allow the end user to zoom in or out of the area being displayed.

jpeg viewer: displays a still image of the alert as a .jpeg

timer: displays the elapsed time since the alert was detected by the artificial intelligence.

resolve button: once an alert is treated as either an emergency or not an emergency, it must be resolved.

tool bar: Displays the other tool and their alerts available in the IMPETUS Platfrom.

Figure 7 – FD Dashboard Detailed

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

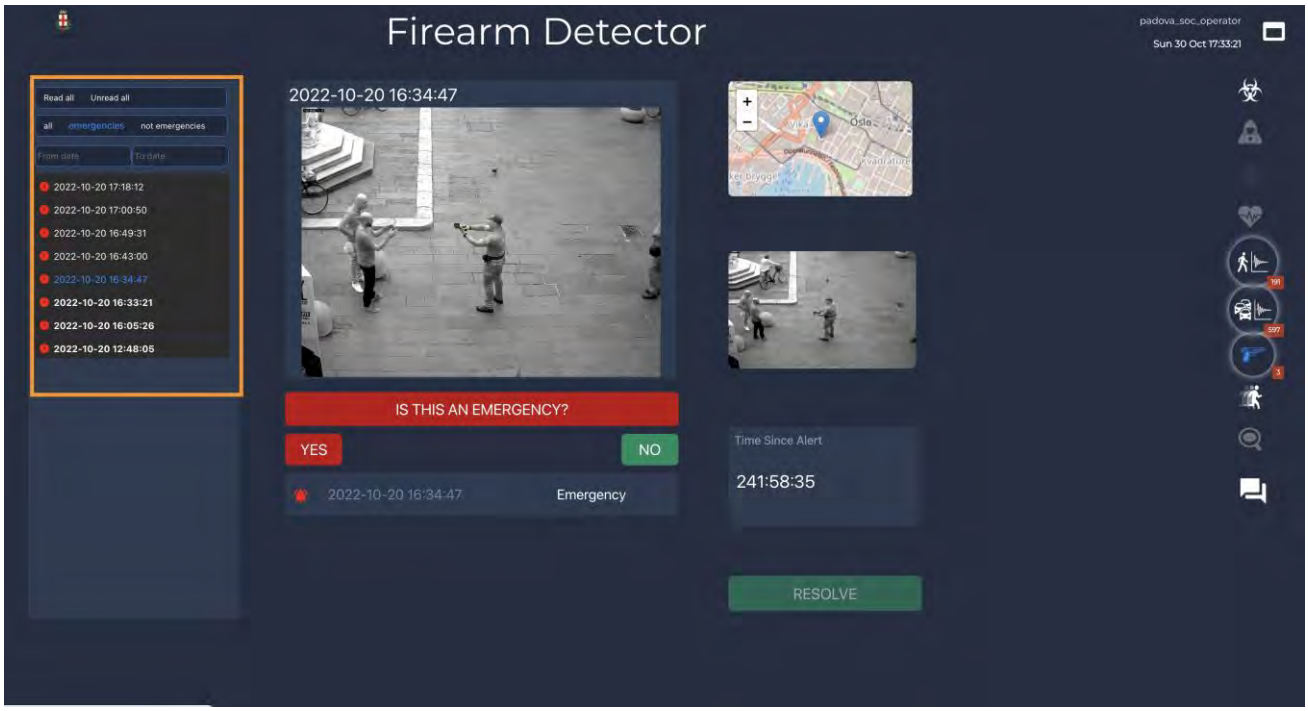
1.4 How to process alerts

Alerts are automatically populated in the FD dashboard in the *index* (figure 7 - marked in orange). The *index* allows the end user to filter alerts by dates and by type such as: emergencies, not emergencies or display all.

Note: A powerful feature is the *Read all* or *Unread all* selection. When selected this function resets all displayed alerts as untreated alerts as long as these were not previously resolved by the end-user.

Figure 7 – FD Index

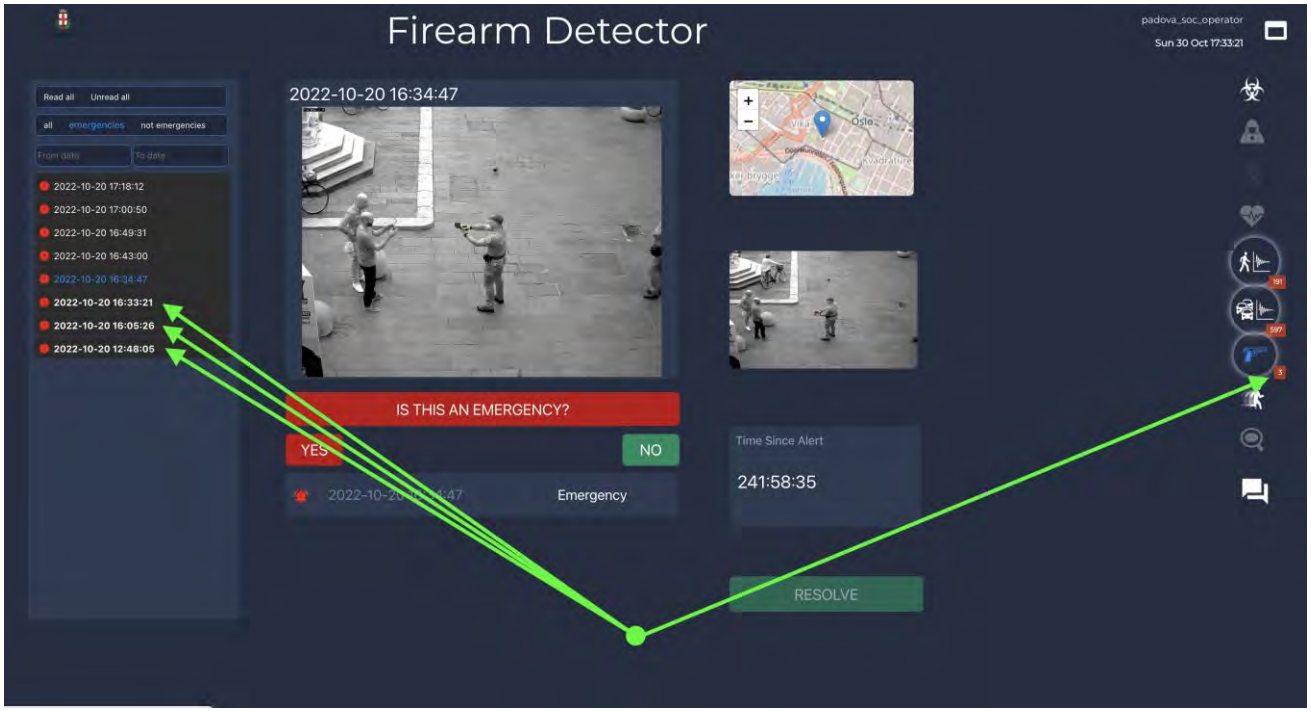
The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.



In the *index* the unread alerts are displayed in **bold**. Additionally the amount of new alerts are also displayed as a decimal number located next to the Firearm Detector icon in the *Tool Bar*. (shown using green arrows in the Figure 8 below).

Note: The number of alerts appearing in **bold** and preselected by a red dot always corresponds to the number displayed next to the firearm icon located in the *Tool Bar*.

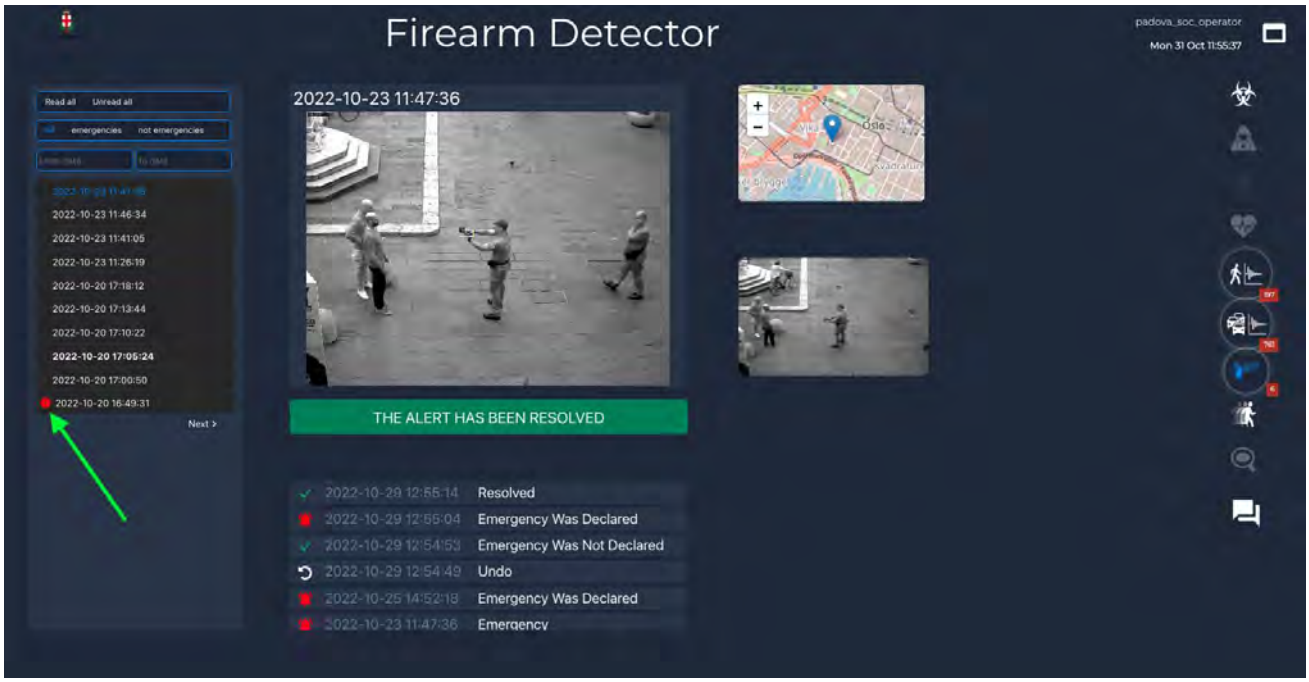
Figure 8 – FD Tool Bar



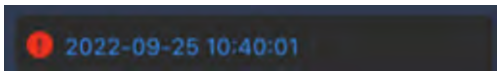
The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

In the *index* the unresolved alerts are also displayed using a red dot. (shown here using a green arrow in the figure 9 below).

Figure 9 – Unresolved alert.



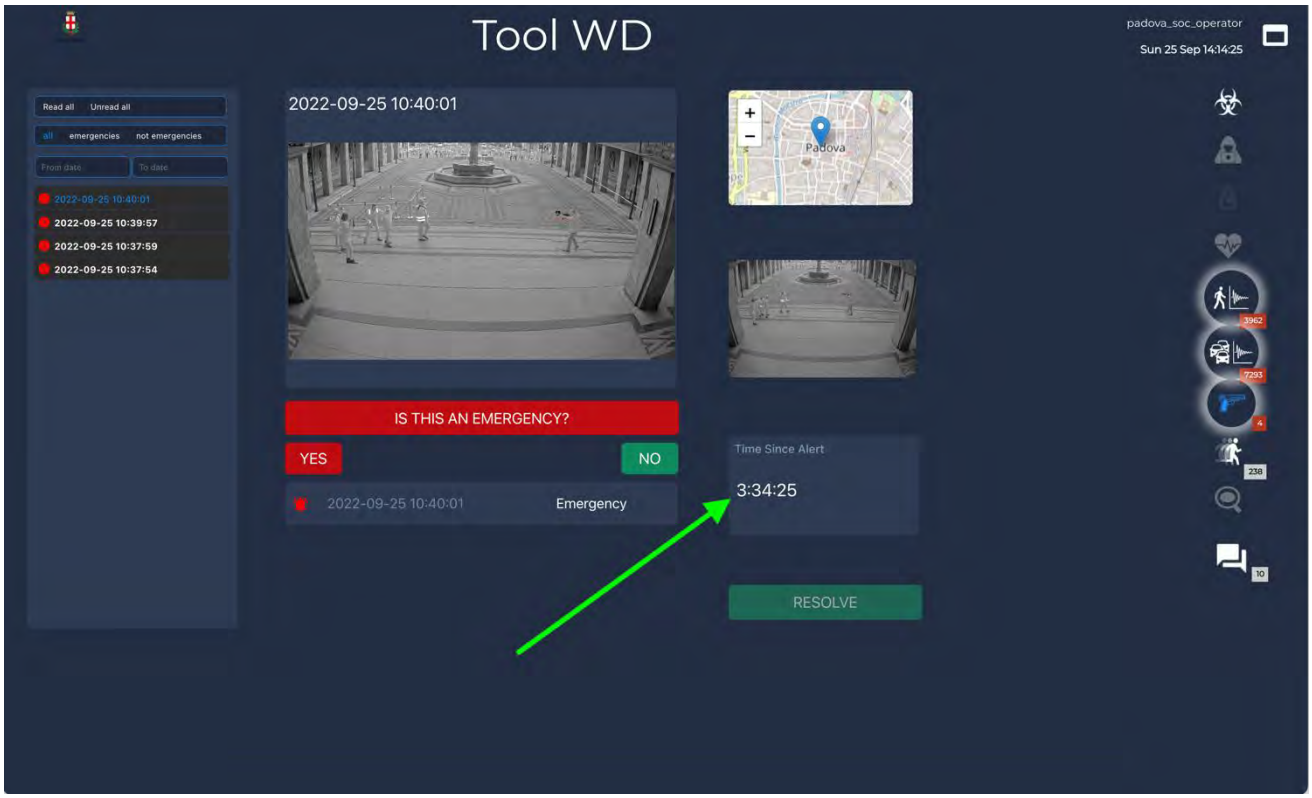
When an alert is selected in the *index* the time stamp is changed from white to blue.



The Timer is displayed until the Resolve button is clicked.

Figure 10 – Timer

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.



Each alert (Figure 11) displays a video sequence of the weapon anomaly in the *canvas* along with a still image in the *jpeg viewer.* Both display a red bounding box around the weapon anomaly.

Figure 11

For better UX (User Experience) as seen in Figure 12 below, it is possible to view the video and the still image in full screen mode by clicking on the still image inside the *jpeg viewer.* The same can be done inside the *canvas* by selecting the full screen option next to the transport controls located below the video.

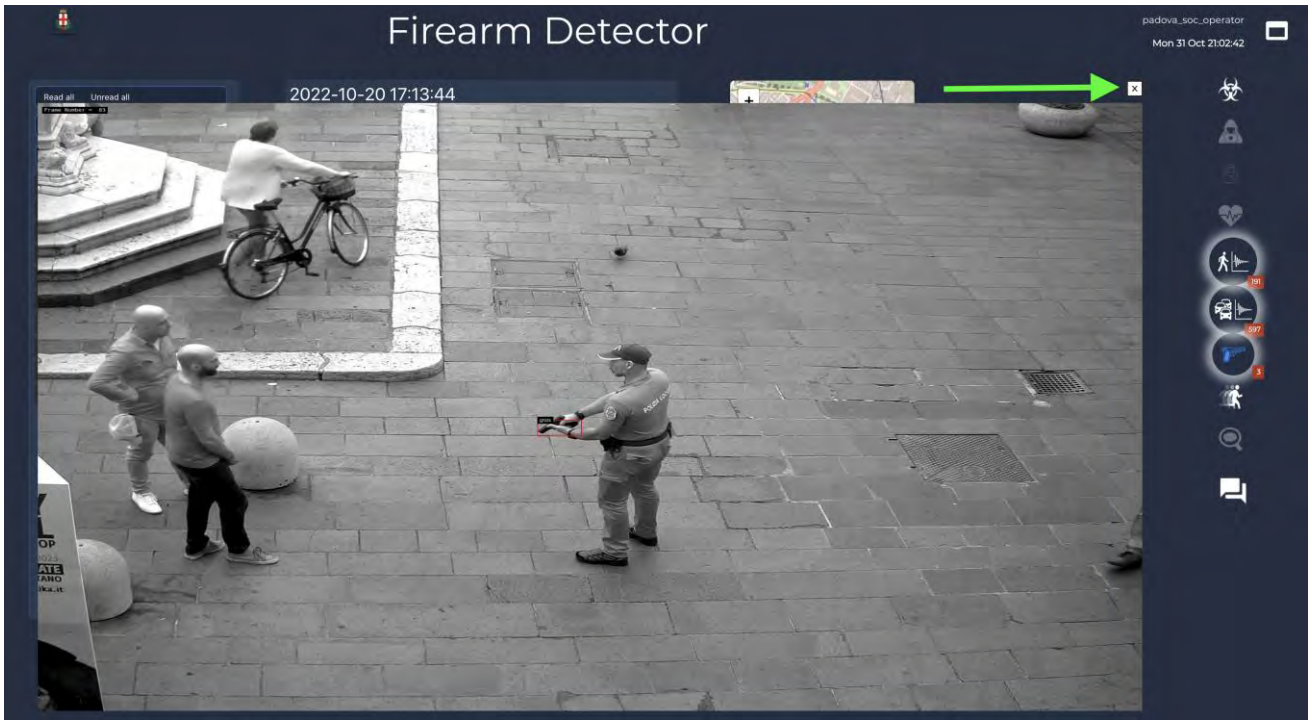
Note: The video transport control position may vary according to the web browser the end user is using.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Figure 12

When the *jpeg viewer* is clicked, the jpeg image populates the majority of the dashboard. To close the image, the x icon located above the image on the right must be clicked.

Figure 13

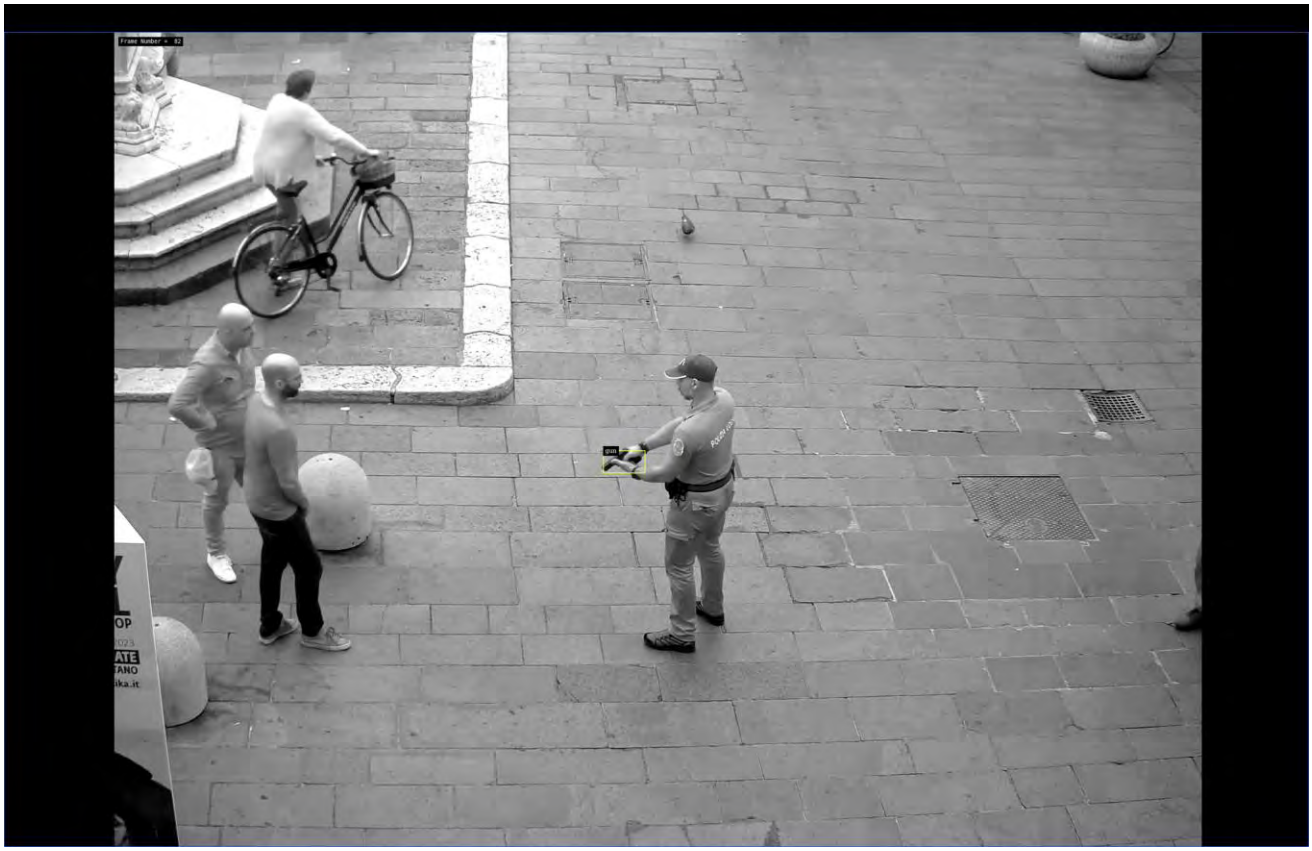


The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Figure 14 illustrates when the *canvas* full screen option is clicked.

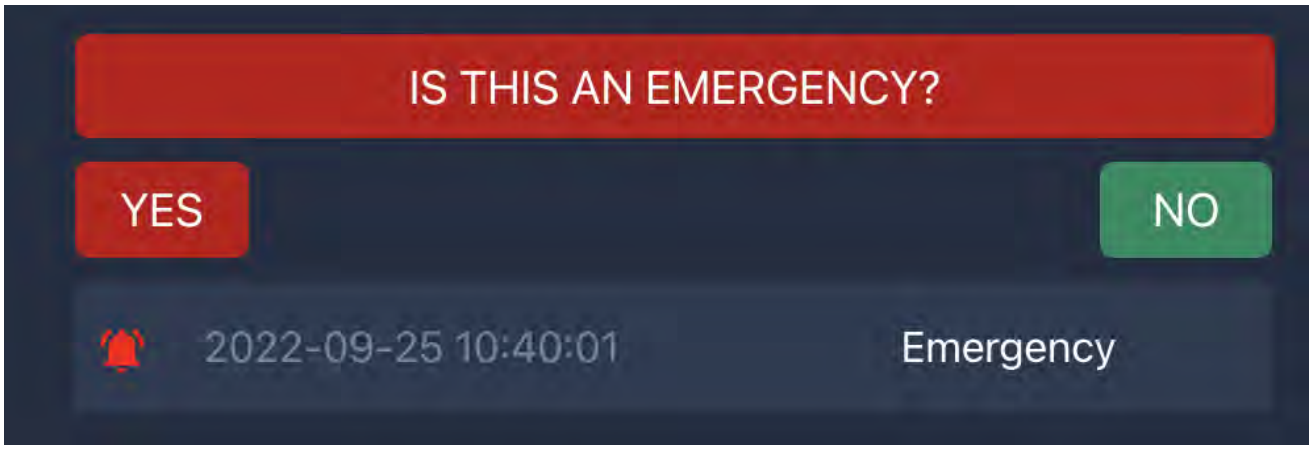
Note: the entire dashboard is masked.

Figure 14 – *Canvas* full screen mode



1.5 Validating Alerts

In red alert mode, the UI (User Interface) will populate an alert with the following fields and buttons inside the *action area*



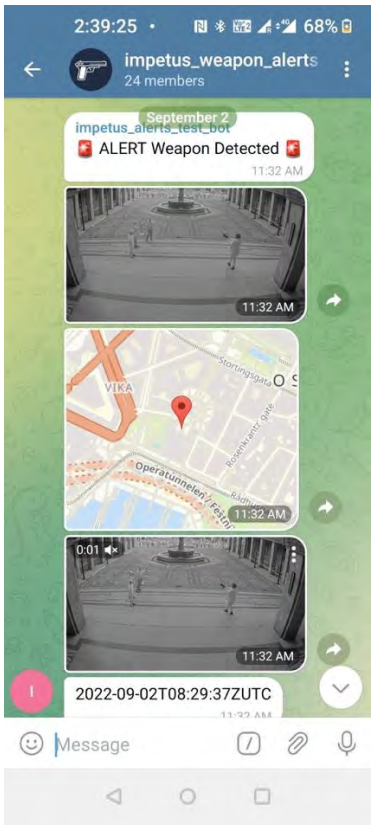
The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

The SoC Operator has two choices and can either confirm if an alert is YES an emergency or NO not an emergency.

Note: When selecting YES, the FD tool will share an instant message to the partner city telegram Channel (Figure 15 – *Telegram*)

When selecting NO, the FD tool artificial intelligence is updated.

Figure 15 – *Telegram*

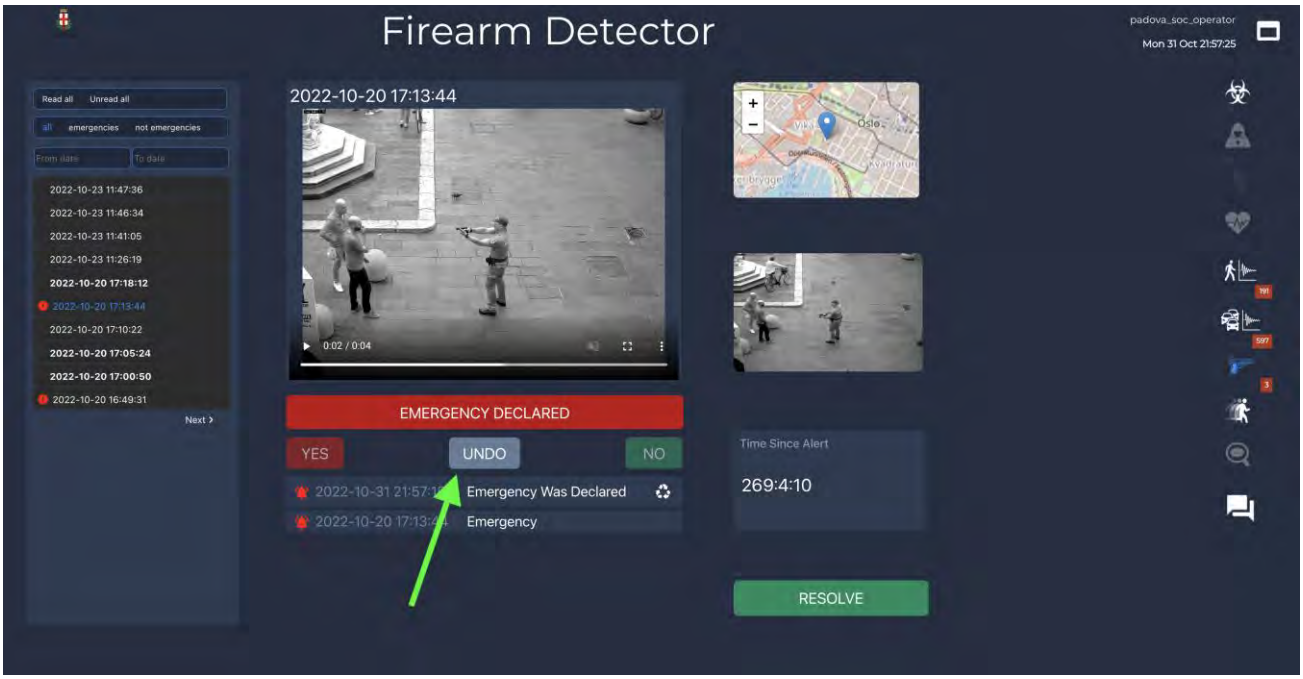


The data is shared with the partner city Telegram Channel and provides instant situation awareness to the first responders who are registered members of the telegram channel. The telegram channel is configured outside of the IMPETUS platform and can be easily integrated using the FD Tool support channel (See chapter 2.0 for details).

When the SoC Operator validates an Alert an "EMERGENCY DECLARED" message is displayed in the *action area*. The SoC operator has then the possibility to cancel the alert by selecting UNDO. See Figure 16.

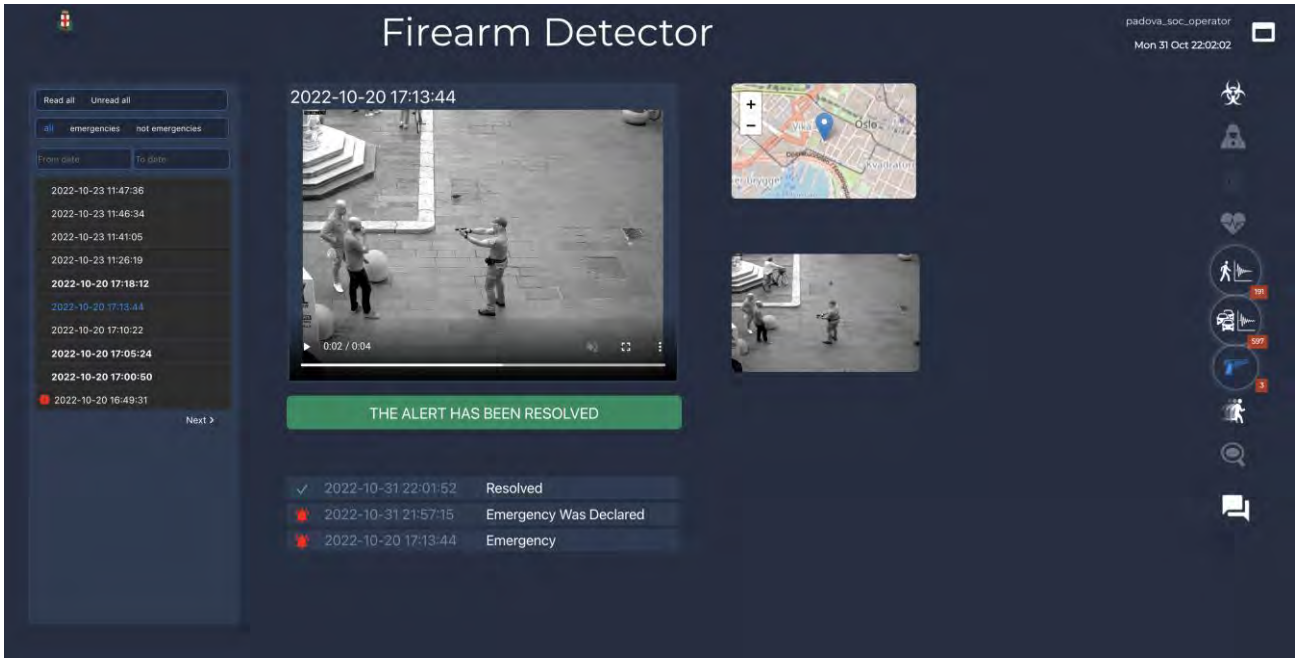
Figure 16 – *Undo*

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.



When the alert is treated by the first responders, the SoC Operator clicks the *resolve button* and the alert can no longer by shared. (Figure 17 – *Resolve*)

Figure 17 – *Resolve*



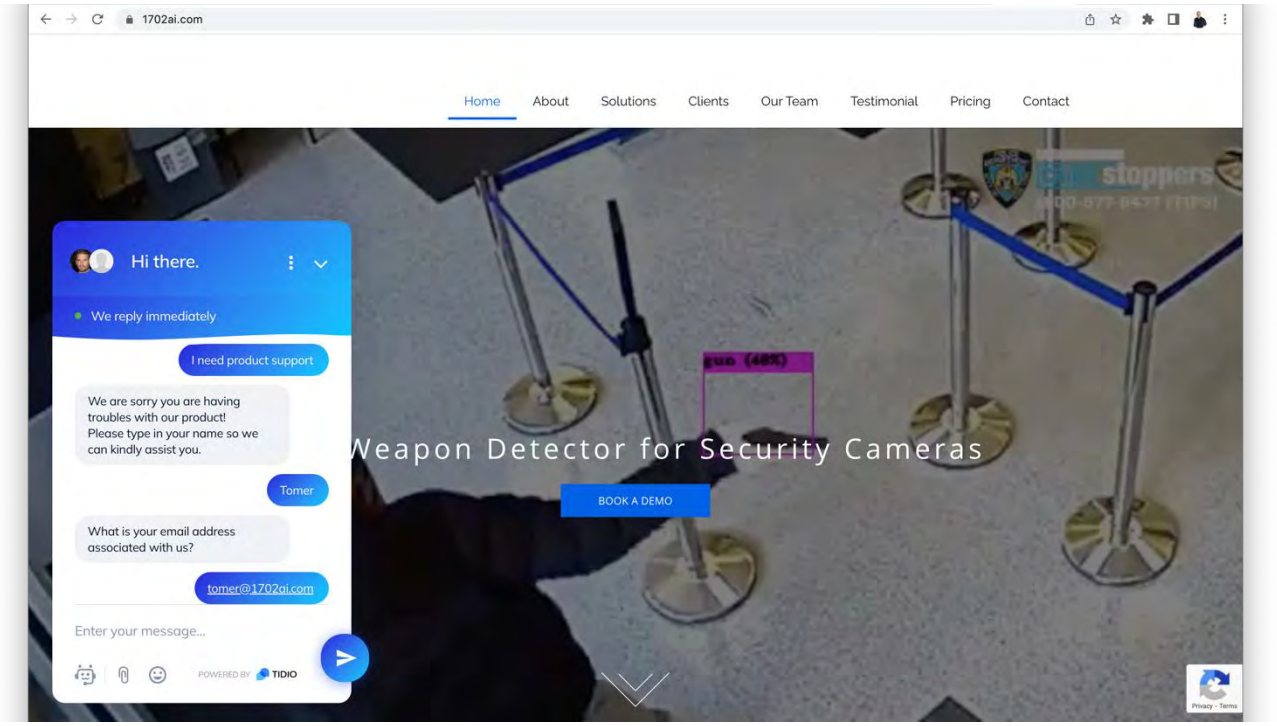
1.6 Alerts management

When you are done resolving all alerts in the queue, your FD dashboard will be updated accordingly and will show no further notifications for new alerts.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

2. Support

If you need support, our team is ready to chime in. Log on to www.1702ai.com and talk to our support bot for a fully automated experience.



2.1 Trouble shooting:

A) *The canvas is greyed out and does not display video alerts.*

Make sure you are using IMPETUS Platform using either chrome or firefox as a web browser.

B) *The UI is not responsive, there is lag.*

Make sure you are using the IMPETUS Platform using a high bandwidth internet connection (50Mb down 10Mb up) and avoid using internet tethering at all costs or a modem using a SIM card.

C) *How can people be added to the Telegram channel ?*

The admin or the owner of the telegram channel can add anyone, they can also make you admin as it is possible to have multiple admins on a telegram channel.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Members of the consortium

SINTEF	SINTEF, Strindvegen 4, Trondheim, Norway, https://www.sintef.no	Joe Gorman joe.gorman@sintef.no
Institut Mines-Télécom	Institut Mines Telecom, 19 place Marguerite Perey, 91120 Palaiseau, France, https://www.imt.fr	Joaquin Garcia-Alfaro joaquin.garcia_alfaro@telecom-sudparis.eu
UNÎMES	Université de Nimes, Rue du Docteur Georges Salan CS 13019 30021 Nîmes Cedex 1, France, https://www.unimes.fr	Axelle Cadiere axelle.cadiere@unimes.fr
cini consorzio interuniversitario nazionale per l'informatica	Consorzio Interuniversitario Nazionale per l'Informatica, Via Ariosto, 25, 00185 – Roma, Italy, https://www.consorzio-cini.it	Donato Malerba donato.malerba@uniba.it
Università degli Studi di Padova	University of Padova, Via 8 Febbraio, 2 - 35122 Padova, Italy, https://www.unipd.it	Giuseppe Maschio giuseppe.maschio@unipd.it
Entrepreneurship Development Centre for BIOTECHNOLOGY and MEDICINE	Biotehnoloogia ja Meditsiini Ettevõtluse Arendamise Sihtasutus, Tiigi 61b, 50410 Tartu, Estonia, https://biopark.ee	Sven Parkel sven@biopark.ee
SIMAVI Software Imagination & Vision	SIMAVI, Complex Victoria Park, Corp C4, Etaj 2, Șos. București – Ploiești, nr. 73 – 81, Sector 1, București, Romania, https://www.simavi.ro	Gabriel Nicola Gabriel.Nicola@simavi.ro Monica Florea Monica.Florea@simavi.ro
THALES	Thales Nederland BV, Zuidelijke Havenweg 40, 7554 RR Hengelo, Netherlands, https://www.thalesgroup.com/en/countries/europe/netherlands	Johan de Heer johan.deheer@nl.thalesgroup.com
CINEDIT INTELLIGENT VIDEO ANALYTICS	Cinedit VA GmbH, Poststrasse 21, 8634 Hombrechtikon, Switzerland, https://www.cinedit.com	Joachim Levy j@cinedit.com

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

INSIKT INTELLIGENCE	Insikt Intelligence, Calle Huelva 106, 9-4, 08020 Barcelona, Spain, https://www.insiktintelligence.com	Dana Tantu dana@insiktintelligence.com
SIXGILL	Sixgill, Derech Menachem Begin 132 Azrieli Tower, Triangle Building, 42nd Floor, Tel Aviv, 6701101, Israel, https://www.cybersixgill.com	Benjamin Preminger benjamin@cybersixgill.com Ron Shamir ron@cybersixgill.com
Comune di Padova	City of Padova, via del Municipio, 1 - 35122 Padova Italy, https://www.padovanet.it	Enrico Fiorentin fiorentine@comune.padova.it Stefano Baraldi Baraldis@comune.padova.it
Oslo	City of Oslo, Grensen 13, 0159 Oslo, Norway, https://www.oslo.kommune.no	Osman Ibrahim osman.ibrahim@ber.oslo.kommune.no
Institut za sigurnosne politike ISP	Institute for Security Policies, Kruge 9, 10000 Zagreb, Croatia, http://insigpol.hr	Krunoslav Katic krunoslav.katic@insigpol.hr
T S I M E	International Emergency Management Society, Rue Des Deux Eglises 39, 1000 Brussels, Belgium, https://www.tiems.info	K. Harald Drager khdrager@online.no
UniSMART Fondazione Università di Padova	Unismart – Fondazione Università degli Studi di Padova, Via VIII febbraio, 2 - 35122 Padova, Italy, https://www.unismart.it	Alberto Da Re alberto.dare@unismart.it

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Grant number: 883286
Project duration: Sep 2020 – Aug 2022
Project Coordinator: Joe Gorman, SINTEF

Horizon 2020: Secure societies
SU-INFRA02-2019
Security for smart and safe cities, including for public spaces
Project Type: Innovation Action

IMPETUS

http://www.impetus-project.eu

SMD – Social Media Detection Tool

INSIKT SPOTLIGHT

Uncovering Hidden Intelligence

User Manual V10.5

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Table of Contents

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

1 QuickStart

1.1 INSIKT Spotlight

SMD tool also known as INSIKT Spotlight is a powerful tool for analysing online content for detecting criminal activities, misbehaviour, and misconduct.

The system's core is analysing text content (with NLP, Natural Language Processing). It has additional features that analyses relationships between users, with SNA (Social Network Analysis).

1.2 Limitations of the current version

Please note that SMD tool as part of the Impetus project has some limitations, both technical and from a feature point of view:

Limitation	Description
Number of projects	The user can create up to 20 different projects.
Allowed areas of analysis	The SMD version includes the following areas of analysis: ● Engagement Spotlight also has available other areas such as: ● Jihadism ● Extreme right ● And others…
Languages	The SMD version includes the following languages of analysis: ● English ● Norwegian ● Italian The system only allows selecting one language per project.
Additional analysis available	The further analysis module available is: ● Hate Speech detection. Spotlight has available other types: ● Psychoprofile analysis (called Personality insights in the dashboard). ● Social Network Analysis.
Type of projects	● CSV → A file with specific columns and format. ● Keywords → Search for words or phrases.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

	Spotlight also has the type: • User → Search for specific Users, Pages, or Groups.
Limitation of project type CSV	• CSV file with a maximum of **20000** rows. If the file has more rows, it will be disposed of. • CSV file with a maximum size of **10Mb**. A larger file will not be uploaded.
Limitations for project	The system will acquire a maximum of **150** messages per keyword (the last 150 messages). The maximum number of keywords or users per project will be **3**.
Project editing	The projects can be edited to change the configuration. The projects can be deleted at any time.
Active projects at the same time	The SMD version doesn't allow you to create a new project when there is another project active, i.e., analysing new data.

1.3 Human-in-the-Loop

Human-in-the-Loop in AI in general, and Machines Learning in particular, introduces the need for human interaction with machine learning systems to increase its accuracy.

INSIKT Spotlight requires human input to reach its full potential. Especially the results in the domain scores demand a final human understanding to understand them fully.

The context is always essential in NLP, and the data collected from social media usually lacks context. So, take with caution the results and check everything in their circumstances.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

2 The science behind INSIKT Spotlight

2.1 Overview

The current chapter includes describing the features extracted from the Natural Language Processing engine, referred to as the NLP engine, from here onwards.

2.2 2.2 NLP Features

The core of Spotlight is an NLP engine that analyses all the messages and extracts the following information:

Feature	Definition	Values
Concept	Terms included in the message	Words
Entity	People, organizations, and places are included in the message.	Words
Topic	Message topic	List of topics
Sentiment	Author's attitude towards a person, product, topic, etc. (positive, negative, or neutral).	-5 (very negative) to +5 (very positive)
Hashtags	Hashtags included in the message.	Hashtags

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

2.3 Significance Score

This Score evaluates the relevance of the message regarding the selected domain. Each domain includes its model in each one of the languages currently in Spotlight.

2.3.1 Engagement Score

For the SMD tool, the Engagement Score is the only available domain. The Engagement Score evaluates the impact of the messages and their virality, i.e., if other users reply, mention, comment, or like the content.

It ranges from 0 to 1, meaning 0 no engagement and one a high Engagement Score.

The way to calculate it depends heavily on each platform; some examples:

Platform	Factors to Calculate the Engagement
Twitter	Tweet • Number of mentions • The number of likes.
Facebook	Posts • Reactions (Likes, haha, etc.). • The number of comments. • The number of positive comments.
Instagram	Posts • The number of comments. • The number of positive comments.
YouTube	Video • The number of comments. • The number of positive comments. • The number of likes.

This table is not an exhaustive list of all the variables; this score is based on many other factors.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

2.4 Hate Speech Score

The Hate Speech score detects:

- Personal attacks: threats against people or organizations.
- Bigotry, hate speech against minorities, groups, countries, etc.

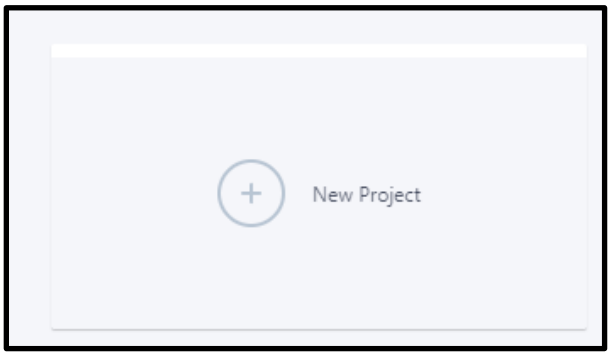
Values	Severity
0.0 - 0.6	No hate speech
0.6 - 0.75	Light
0.75 - 0.9	High
0.9 - 1	Extreme

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

3 How to start: create a project - delete - edit - search

3.1 Create a new project

Click New Project

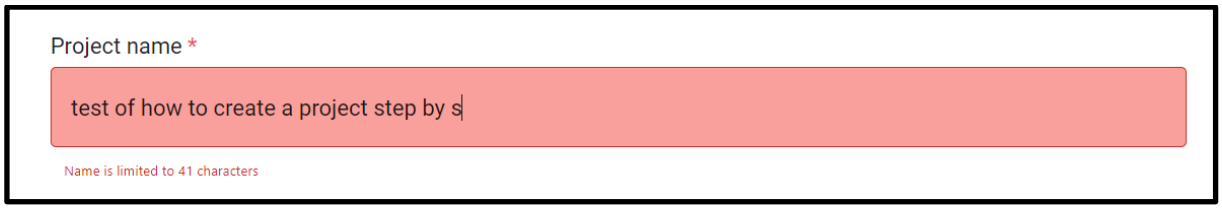


You will see the following dashboard for creating the project.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Step 1. Introduce the project name.

The name is limited to 41 characters.



Also, the project name must be unique.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Step 2. Write a project description (optional).

The project description helps describe the project in detail.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Step 3. Choose a project type.

Project type is related to the way to collect data (from users, from keywords, or uploading a CSV file)

There are three types of projects:

Type	Description
User	The user introduces a list of users (see more detail later), and the system collects and analyses all the messages of these users.
Keywords	The user introduces a list of keywords (see later in what format), and the system collects and analyses all the messages containing these keywords.
CSV	The user introduces a CSV (see later in what format), and the system ingests it and analyses its content.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Step 4. Select an area of analysis

The user selects which analysis model will be applied during the data analysis.

These NLP classification models extract information at a message level from different perspectives (see methodology section for a detailed description of all the models). In particular:

Area of analysis	Description
Jihadism	Model to detect if the message contains content related to jihadism ideology.
Extreme right	Model to detect if the message contains content related to extreme-right ideology.
Engagement	Model to evaluate the impact of a message, proportional to the potential audience and its reactions (likes, shares, comments).

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

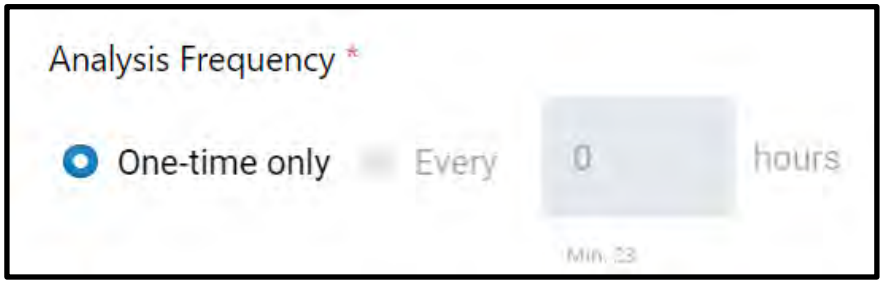
Step 5. Choose the selected sources (only for User and Keyword type of project).

The user selects the source/s of information where the system will collect the data. New additional sources can be added by requested.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Step 6. Choose a frequency (only for User and Keyword project).

The user defines the frequency of the data collection.



Among the following options:

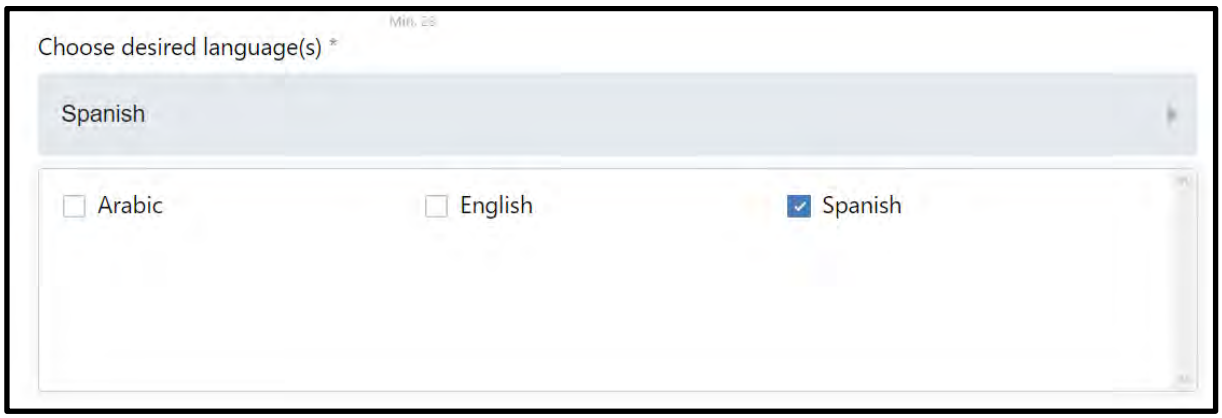
Option	Description
One-time only	There is no data acquisition in the future, i.e., the system collects data from a specific date range defined by the user.
Every X hours	The system collects data from the exact keywords or users each X hours, where the user defines the value of X.

There is no option to choose the date range. The system collects the last messages containing the keyword.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Step 7. Choose the desired languages.

This option is only available for User and Keyword projects. For CSV projects, the system includes a language detector to analyse each message in the correct language.



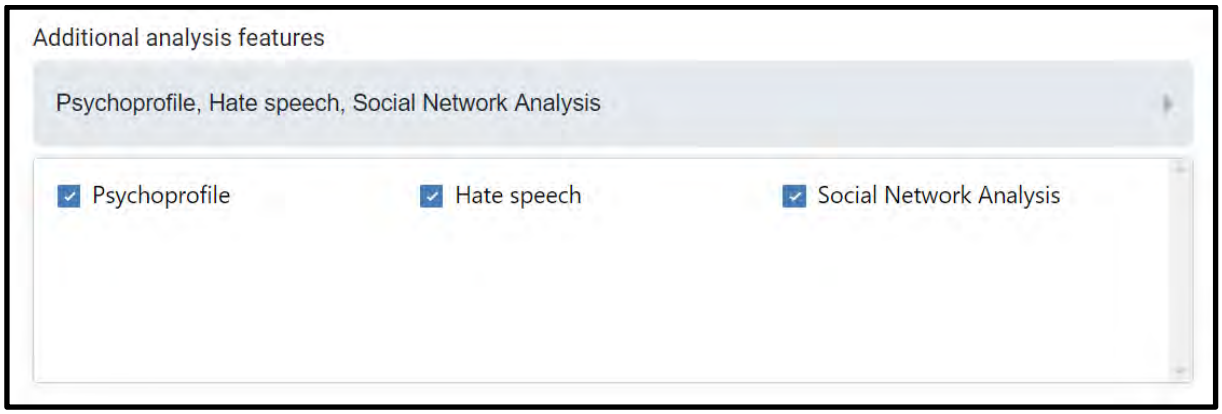
Insikt Spotlight will only show the data in your selected language, the rest of the data acquired during the keyword search or user search is discarded.

T*he user can select only one language by the project.*

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Step 8. Additional features.

Some of the features included in the analysis are optional.



Specifically:

Feature	Description
Psychoprofile	If activated, the project includes an analysis of the users' personality traits.
Hate speech	If activated, the project includes the detection of Hate Speech in all the messages.
Social Network Analysis	If activated, the project includes the Social Network Analysis to analyse the connections between users.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Step 9. Definition of keywords or user names (User and Keyword type of project only).

The last step is to define the keyword or user names.

The system will collect the messages containing these options.

The user can define different type of keywords:

Keyword	Description
term	Term to search messages containing it.
“several words”	Search for the whole expression in the messages.
#hashtag	Hashtags just including the character #.

Regarding user names, they should be added without @ (in Twitter), for example:

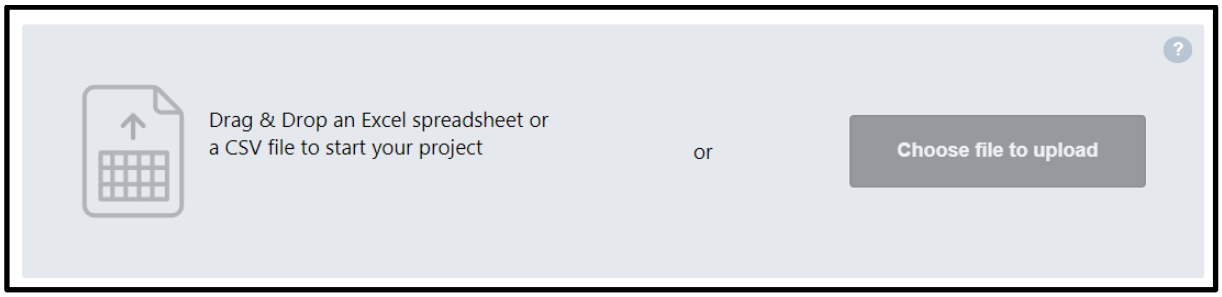
The SMD version allows the inclusion of a maximum of 3 keywords or users.

The option of excluded keywords is not available in the SMD version.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Step 5 bis. Upload the CSV file (only for CSV project type).

In projects based on the CSV file, the user must upload the file by drag and drop it or choose the file to upload.



The user must upload the file in the correct format:

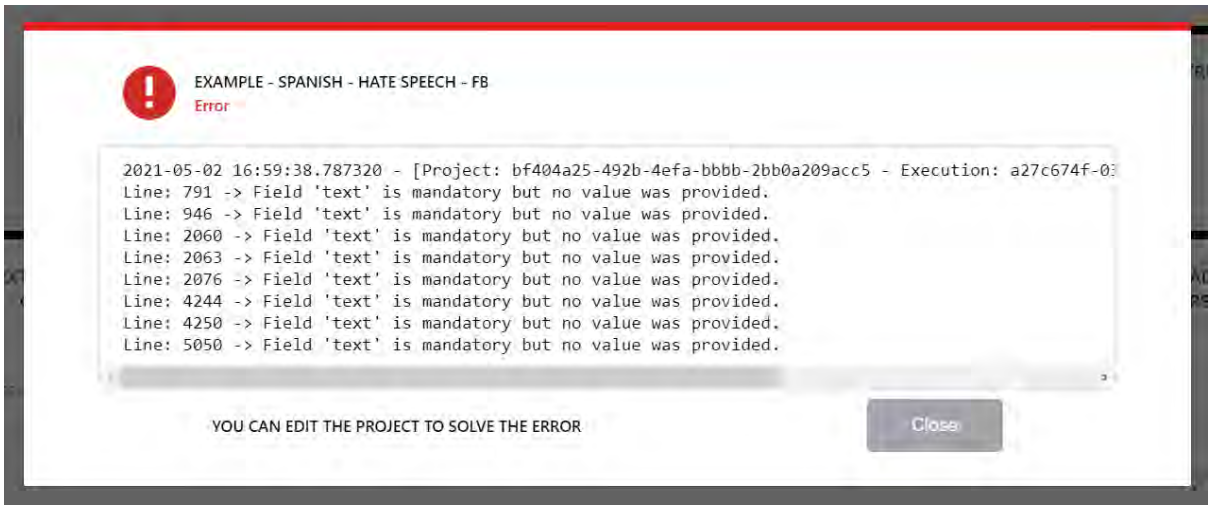
Field name	Type and format	Mandatory
Field separator	; (semicolon)	Yes
User	String with double quotes (" ")	Yes
Timestamp	Epoch Milliseconds (Integer)	Yes
Text	String with double quotes (" ")	Yes
Nshares	Integer	No
Ncomments	Integer	No
Nlikes	Integer	No
Nfriends	Integer	No
Replytoauthor	String with double quotes (" ")	No

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

If some of the field values are not formatted correctly, the system will display an error message.

It is possible to check the incorrect lines and the type of error by clicking on the error icon:

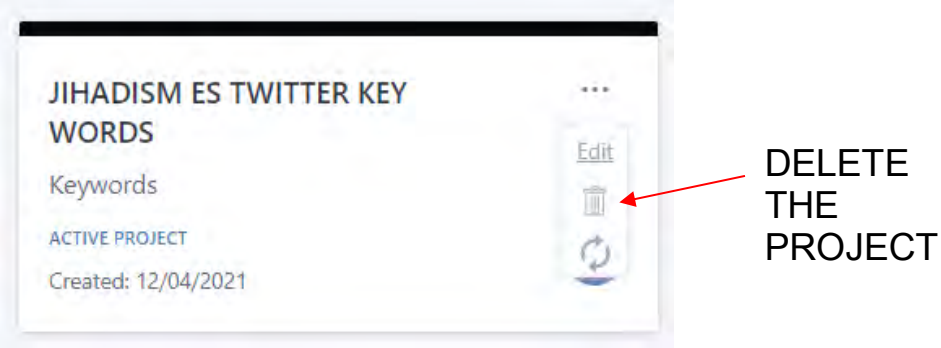
Example of error message:



The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

3.2 Delete a project

To delete a project is straightforward:



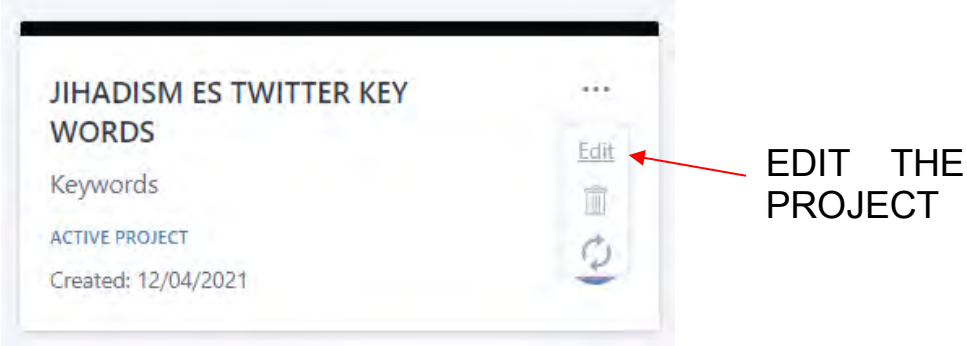
Just click the three dots "…" at the top of the project card, and then the garbage symbol.

We just have to accept by clicking "Yes," and the system will remove the project.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

3.3 Edit a project

To edit a project is pretty straightforward:



Just click the three dots … at the top of the project card and then the "Edit" symbol.

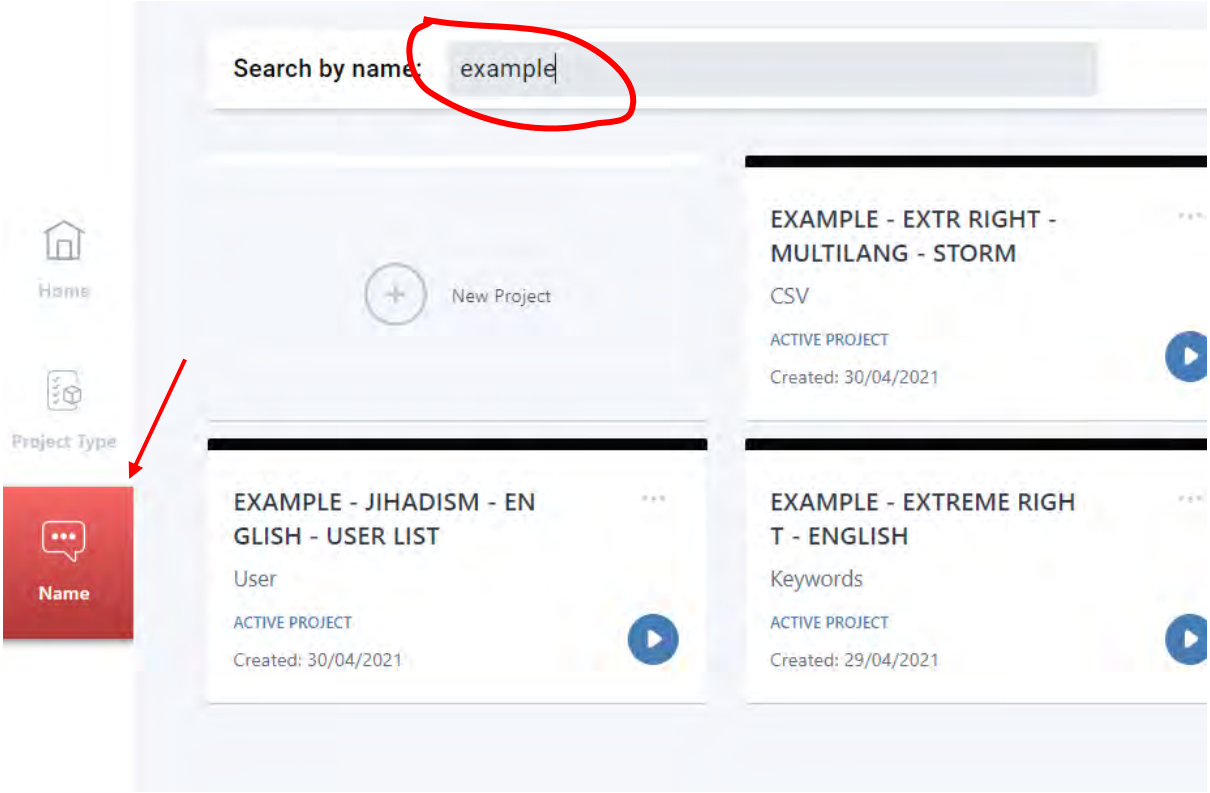
The user can change any of the options defined to create the project.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

3.4 Search for a project

It is helpful to search by project name if the users created a lot of projects.

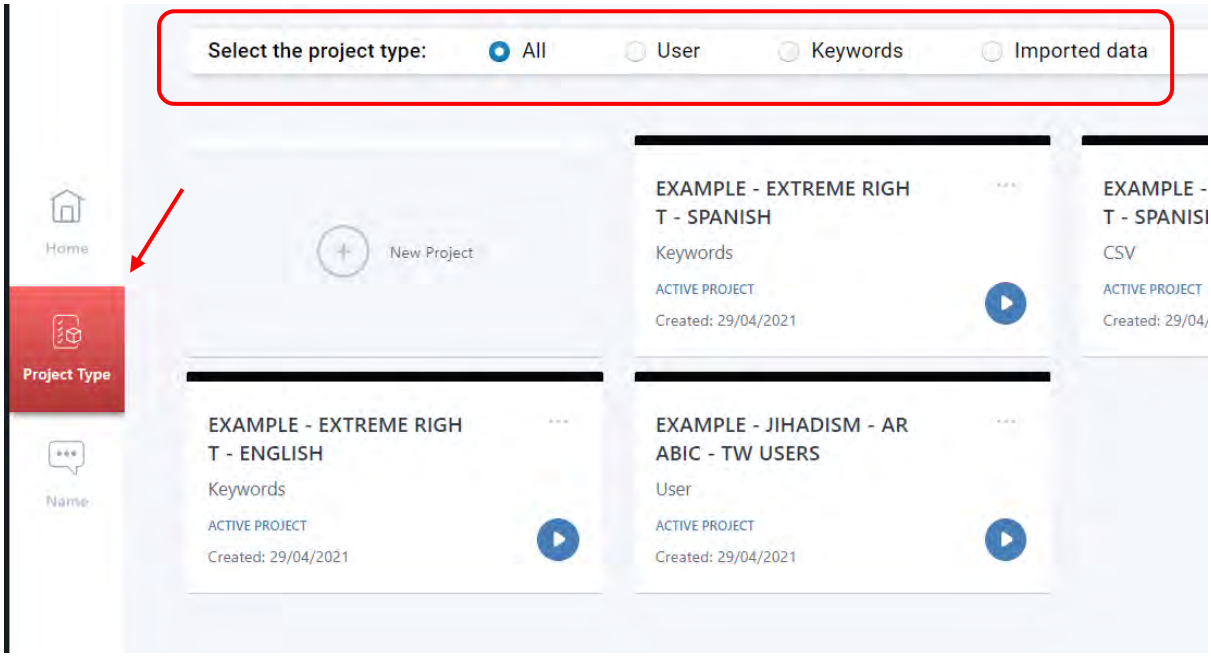
At the left tab, just select Name.



Then, write the term you want to search for finding the project.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

It is also possible to search by project type:



All the options appear at the top, User, Keyword, and Imported data project type.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4 The results: a dashboard tour

4.1 Overview

The INSIKT Spotlight dashboard structure is as follows:

Summary	General
	Raw data
Content	Messages
	Concepts
	Entities
	Topics
	Key-ideas
	Hashtags
Users*	General
Networks*	Interactions
Location*	Messages
	Users

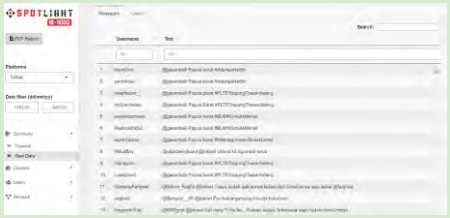
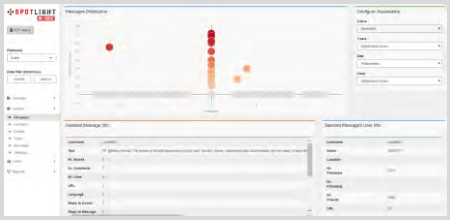
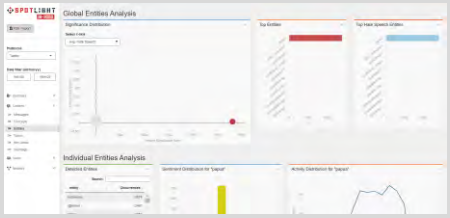
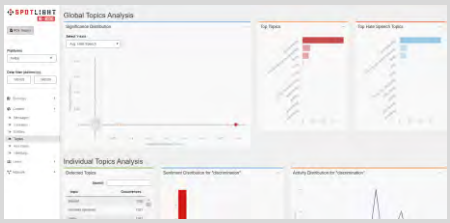
*Not included in SMD version.

All the social media data sources are shown in the same dashboard for simplicity, including dashboards to analyse the message contents and extract information about the users.

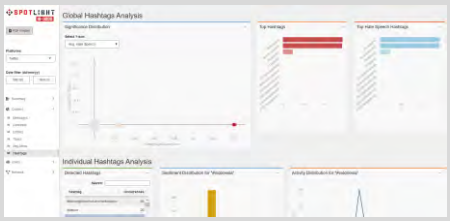
The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4.2 Objectives

The main objectives of each dashboard are the following:

Summary	General	Summary of all the content.	
	Raw data	Table with all the data analysed in the dashboard.	
Content	Messages	Scores at a message level.	
	Concepts	Statistics of the message concepts (terms).	
	Entities	Content linked to entities.	
	Topics	Message topics (defined by the user).	
	Key-Ideas	Message topics (not pre-	

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

		defined).	
	Hashtags	Statistics of the message hashtags.	

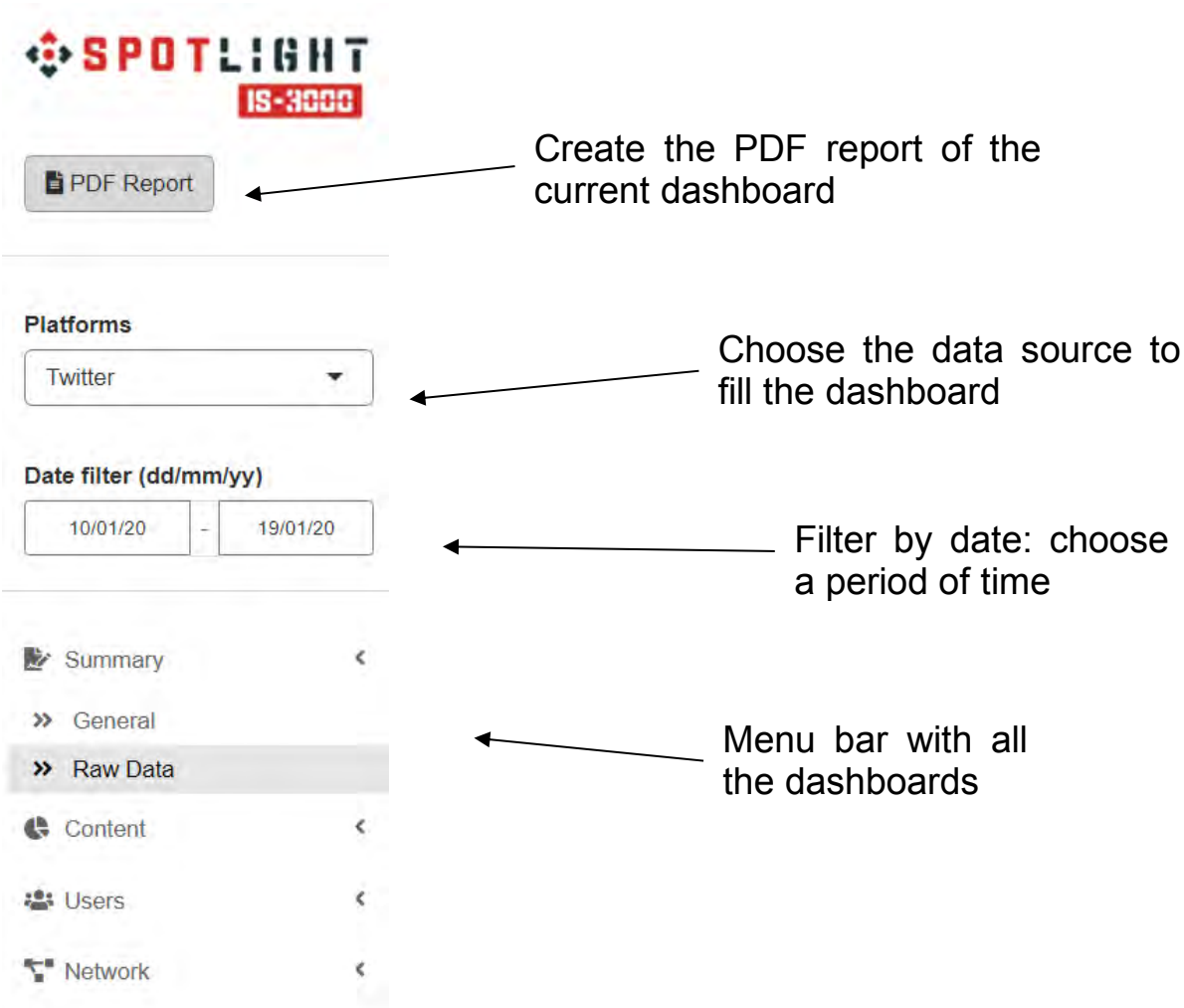
The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4.3 Content of the Dashboards

4.3.1 Menu Bar

Graph	Content	Capabilities
	• Summary of all the data retrieved and analysed	Filter all the dashboards by: • Date • Data source Choose the dashboard tab.
Hints	Daily updates.	• Choose a time of the current day to review the daily updates.
	Significance score threshold.	
	Specific data source.	• Filter by the data source to show all the data of a specific platform.
Advanced version (not in the SMD version)	Download data.	• Download all data in different formats (CSV, Excel, etc.).

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.



The PDF report creation is not available in the SMD version.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

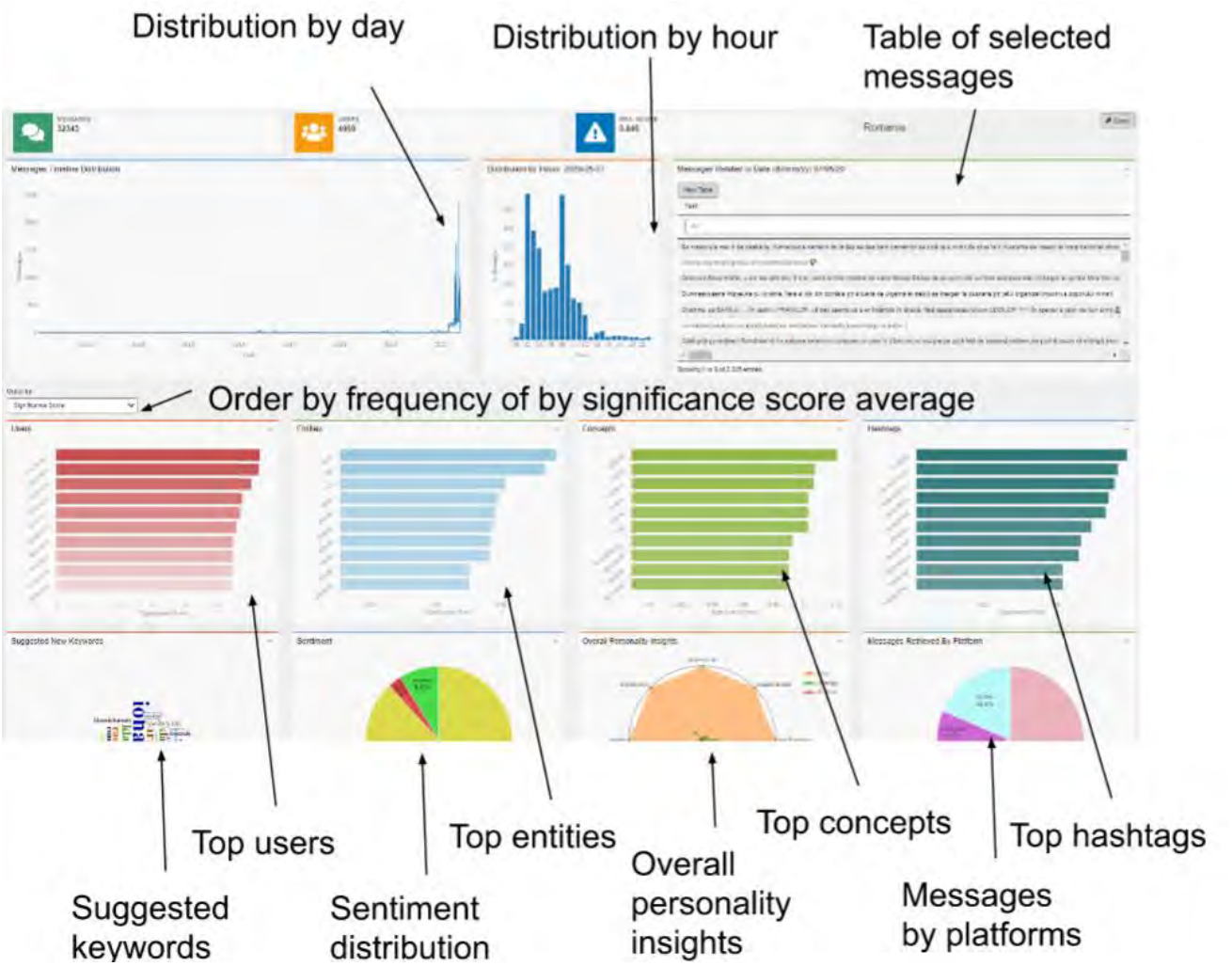
4.3.2 General - Summary

4.3.2.1 Overview of the features

Graph	Content	Capabilities
	• Summary of all the data retrieved and analysed	• The number of messages by day. • The number of messages by the hour. • Table of messages by day. • Top values of users, entities, concepts, and hashtags. • Suggested keywords. • Sentiment distribution. • Overall personality insights. • Messages by platforms.
Hints	Daily updates.	• Click over one day to get all the messages of this day.
	Order values by frequency	• Choose order by frequency to show the most frequent terms and more active users.
	Order by Significance Score	• Choose order by Significance Score to show terms and users most frequent in messages with high Significance Score.
	Messages containing a term	• Click on a bar to show all the messages with this content.
	Suggested keywords	• Terms that can be used for future searches
Advanced version (not in the SMD version)	Download data.	• Download all data in different formats (CSV, Excel, etc.).

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4.3.2.2 Description of the Dashboard options



Ordered by frequency:

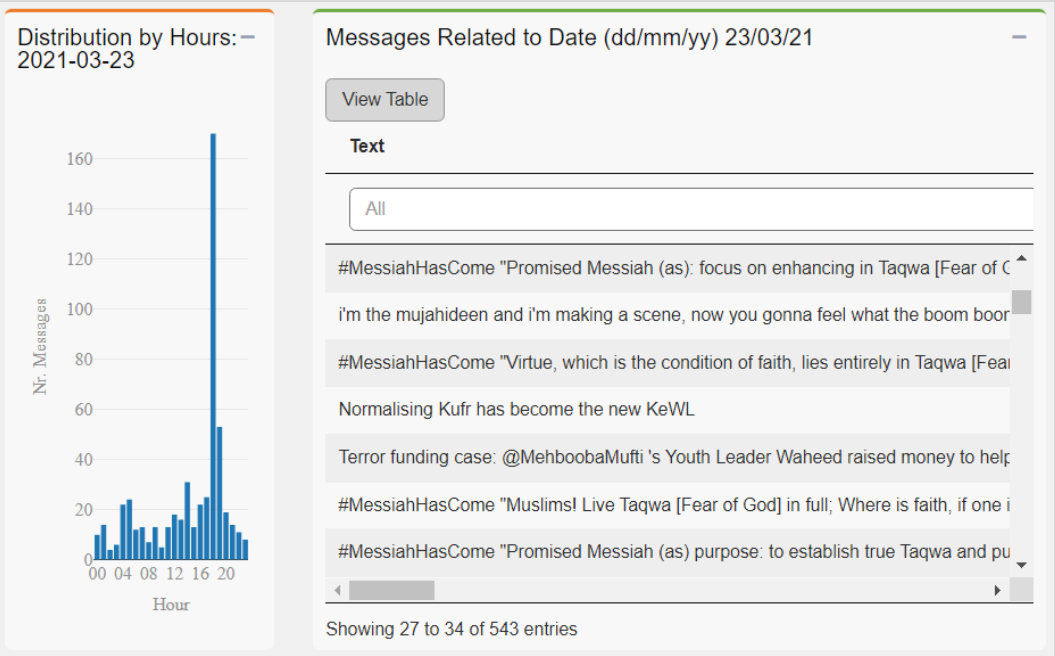
- The dashboard shows the most common terms and more active users.

Ordered by Significance Score:

- Most frequent terms in messages with high Significance Score.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4.3.2.3 Examples

Goal	What are the messages of a specific day? Ex: 23/03/2021
Steps	• Date filter: select from 23/03/2021 • Click the day in the timeline to fill the message table • Check the table of Messages Related
Result	

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4.3.3 General - Raw Data

4.3.3.1 Dashboard options

This tab contains all the data analysed in the project in a table format.

It helps search and filter across all the information collected and showing them in a table format.

Graph	Content	Capabilities

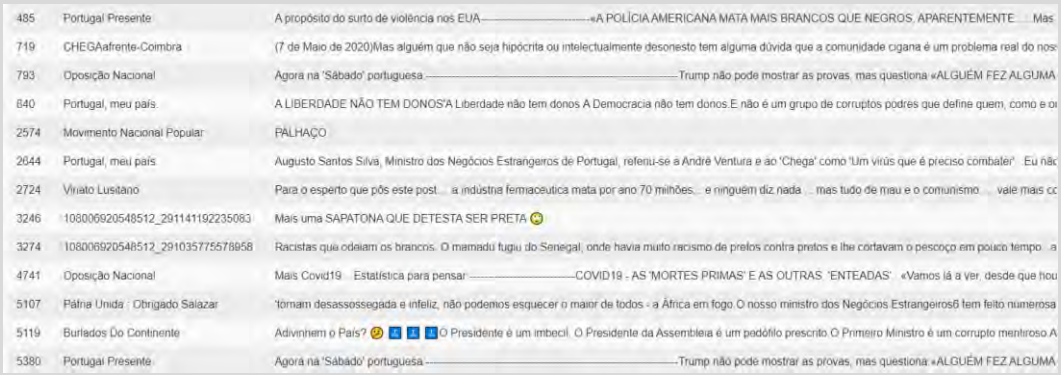
	• Table with all the data.
	• Table with the content of a specific message.
	• Filters.
		Filter by:
		• NLP features
		  ○ Topics
		  ○ Concepts
		  ○ Hashtags
		  ○ Entities
		  ○ Sentiment
		• Impact
		• Engagement
		• Date
		• Keyword

Hints	Discover influence content.	• Filter by Significance Score above 0.5 to discover messages with high relevance.
	Daily updates.	• Filter by time to check the daily new content.
	Discover negative-positive messages	• Filter by Sentiment below -0.5 (above +0.5) to discover negative (positive) messages.
Advanced version (not in the SMD version)	Download data.	• Download the filtered table in different formats (CSV, Excel, etc.).

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4.3.3.2 Examples

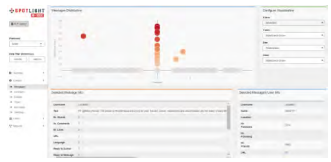
Goal	What are the most negative messages?
Steps	• Go to table headers. • Order by Sentiment, from low to high. • Check the messages with Sentiment -5.
Result	Top negative messages.



485	Portugal Presente	A propósito do surto de violência nos EUA ---------- «A POLÍCIA AMERICANA MATA MAIS BRANCOS QUE NEGROS, APARENTEMENTE ... Mas
719	CHEGAafrente-Coimbra	(7 de Maio de 2020)Mas alguém que não seja hipócrita ou intelectualmente desonesto tem alguma dúvida que a comunidade cigana é um problema real do nos
793	Oposição Nacional	Agora na 'Sábado' portuguesa ---------- Trump não pode mostrar as provas, mas questiona «ALGUÉM FEZ ALGUMA
840	Portugal, meu país	A LIBERDADE NÃO TEM DONOS'A Liberdade não tem donos A Democracia não tem donos.E não é um grupo de corruptos podres que define quem, como e o
2574	Movimento Nacional Popular	PALHAÇO
2644	Portugal, meu país	Augusto Santos Silva, Ministro dos Negócios Estrangeiros de Portugal, referiu-se a André Ventura e ao 'Chega' como 'Um vírus que é preciso combater'. Eu não
2724	Viriato Lusitano	Para o esperto que pôs este post ... a indústria farmacêutica mata por ano 70 milhões... e ninguém diz nada ... mas tudo de mau e o comunismo ... vale mais co
3246	1080069205485 12_291141192235083	Mais uma SAPATONA QUE DETESTA SER PRETA
3274	1080069205485 12_291035775578958	Racistas que odeiam os brancos. O mamadú fugiu do Senegal, onde havia muito racismo de pretos contra pretos e lhe cortavam o pescoço em pouco tempo ...a
4741	Oposição Nacional	Mais Covid19 ... Estatística para pensar ---------- COVID19 - AS 'MORTES PRIMAS' E AS OUTRAS 'ENTEADAS' «Vamos lá a ver, desde que hou
5107	Pátria Unida : Obrigado Salazar	'tornam desassossegada e infeliz, não podemos esquecer o maior de todos - a África em fogo.O nosso ministro dos Negócios Estrangeiros6 tem feito numerosa
5119	Burlados Do Continente	Adivinhem o País? O Presidente é um imbecil. O Presidente da Assembleia é um pedófilo prescrito.O Primeiro Ministro é um corrupto mentiroso.A
5380	Portugal Presente	Agora na 'Sábado' portuguesa ---------- Trump não pode mostrar as provas, mas questiona «ALGUÉM FEZ ALGUMA

4.3.4 Content - Messages

4.3.4.1 Features

Graph	Content	Capabilities
	Significance Score, Hate speech, and Sentiment of all the messages	Filter by: • Type • Country • User • Language
Hints	Discover top messages in terms of impact and positive/negative sentiment.	Select all the top-right messages.
	Discover messages with very positive (or negative) content.	Select all the top messages.
	Discover messages with high impact.	Select all the messages to the right.

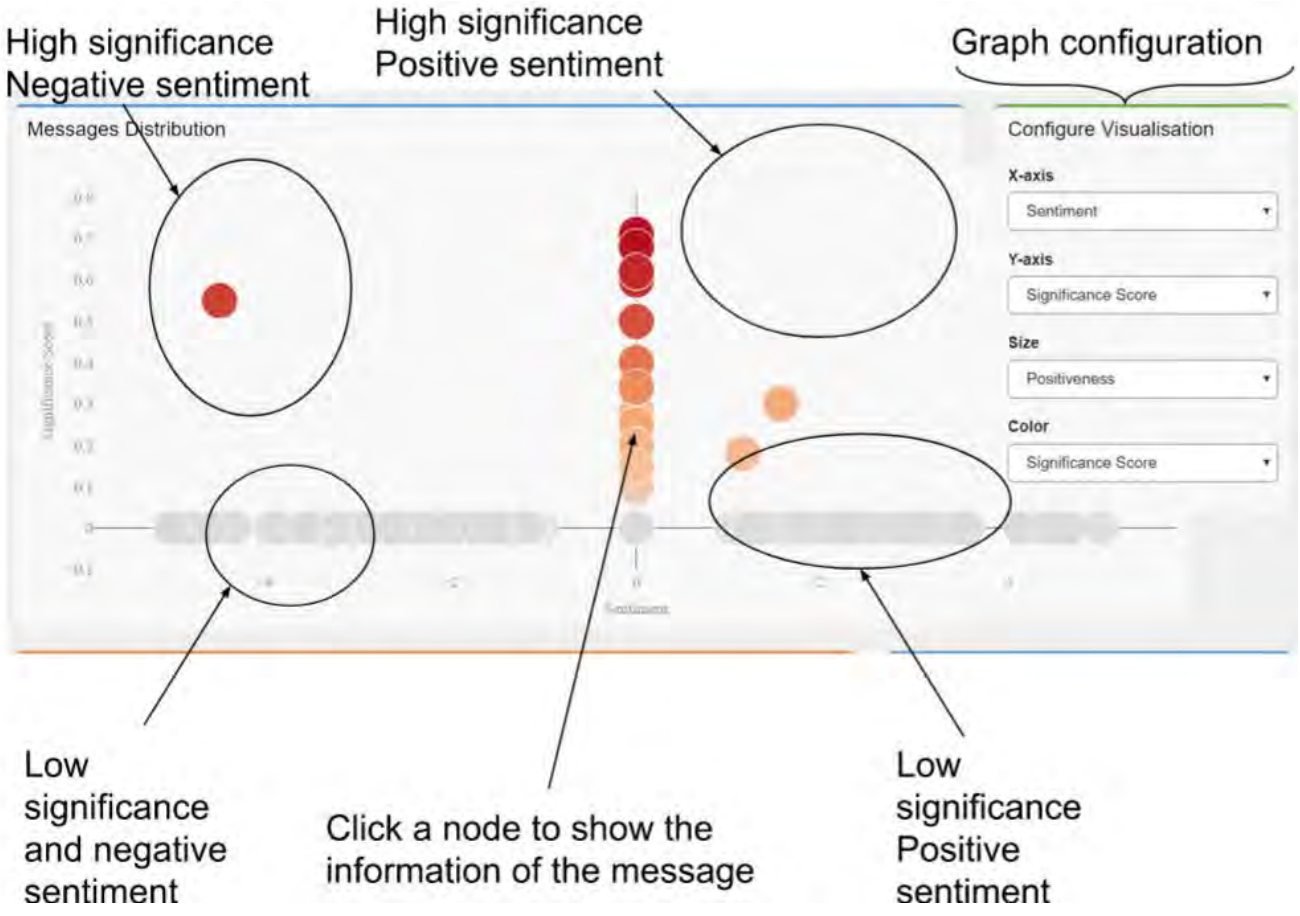
The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Advanced version (not in the trial version)	Download	Download all the selected messages as Excel, CSV, txt. Download all the graphs as PDF.
	Specific features.	Definition of specific targets to track.

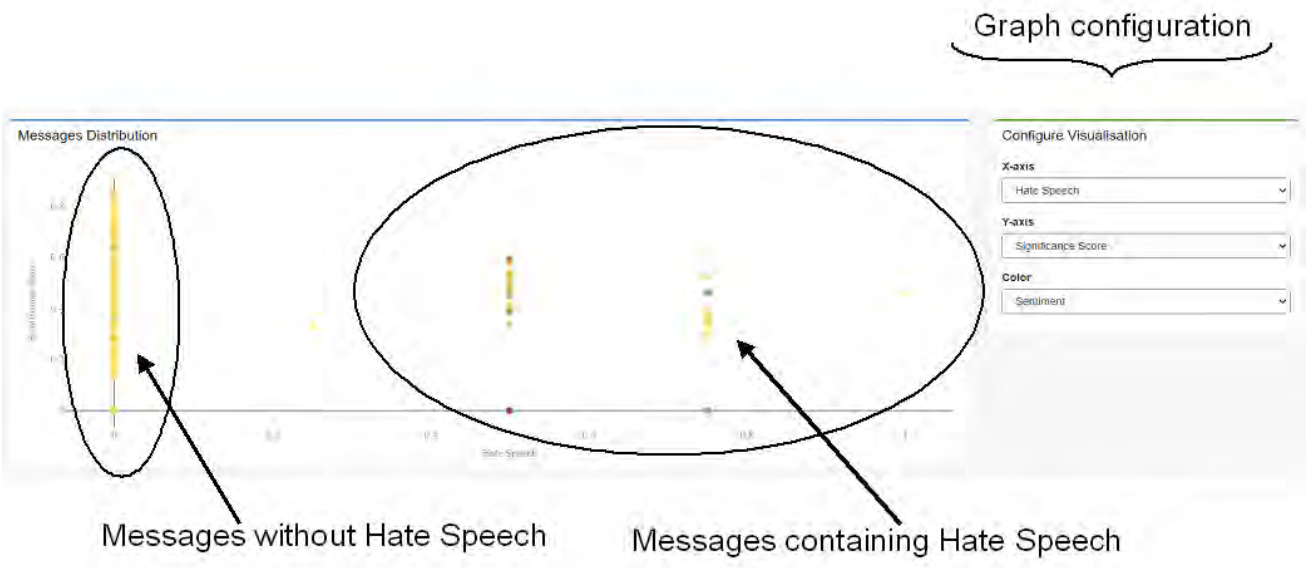
In terms of the Significance Score equal to impact.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4.3.4.2 Dashboard options

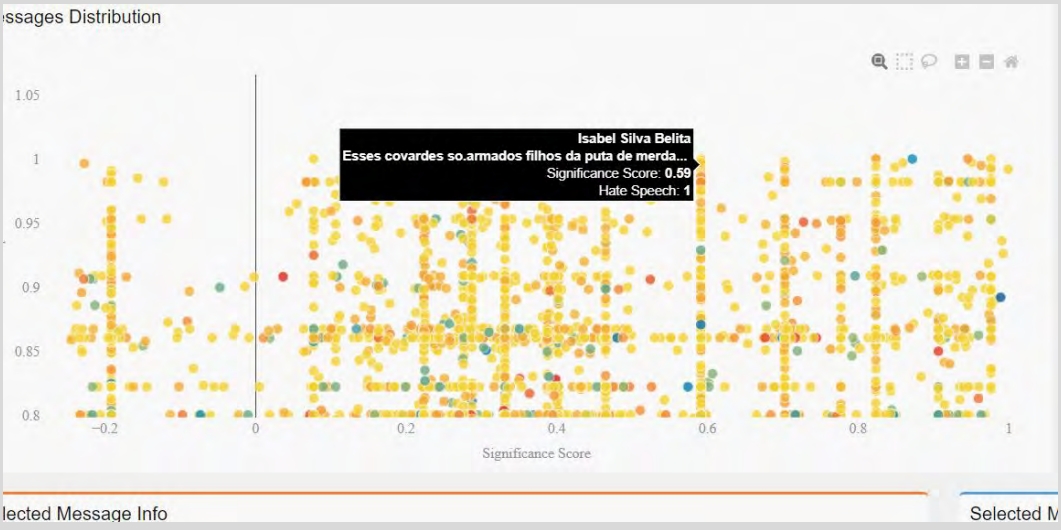


If we choose Hate Speech and Significance Score:

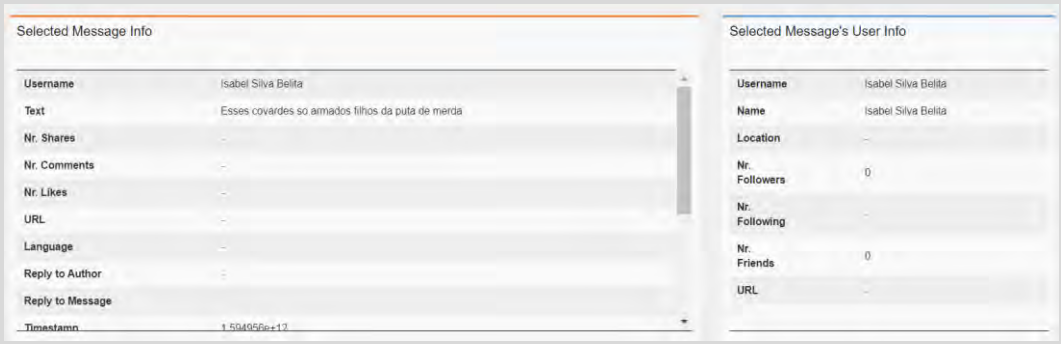


The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4.3.4.3 Examples

Goal	What are the messages containing Hate Speech?
Steps	• In Configure Visualisation, select Hate Speech as Y-axis. • Zoom the top of the graph.  • Select some of the nodes to check the message. 
Result	Example: click the message to check all the information about it.

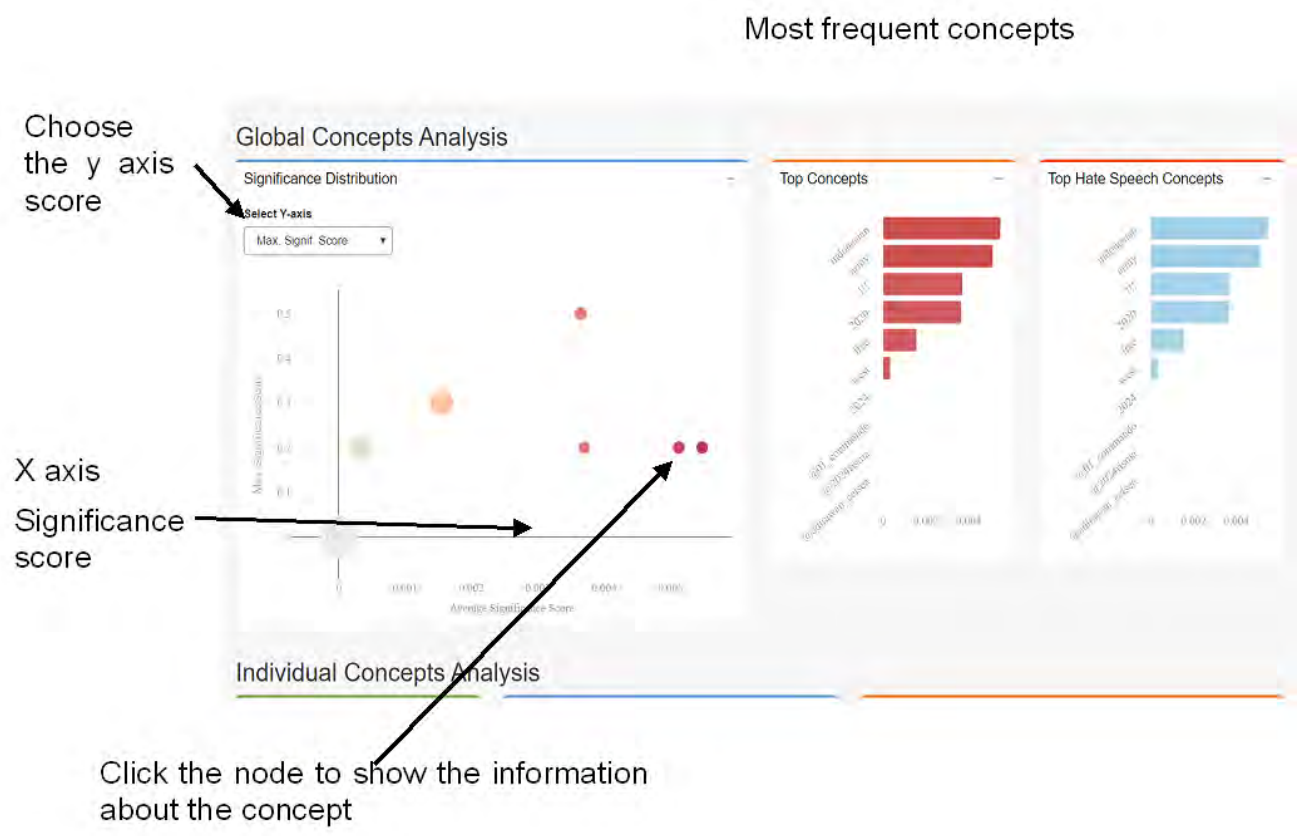
The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.



Goal	What are the messages containing Threats or Supporting Violence?
Steps	• In Configure Visualisation, select Significance Score as X-axis. • Zoom the top of the graph. • Select some of the nodes to check the message.
Result	Example: click the following message “Cortem a cabeça e este covarde corrupto de merda e grande traidor a Portugal.” 

4.3.5 Content - NLP features (concepts, entities, topics, key ideas, hashtags)

Graph	Content	Capabilities
	• Graphs with the most common features: ○ Topics ○ Concepts ○ Hashtags ○ Entities ○ Key ideas	Select attributes to visualize evolution.
Hints	Discover trends in the conversation.	• Choose different periods to check the trends in these periods.
Advanced version (not in the Demo)	Download data.	• Download all data related to this feature in different formats (CSV, Excel, etc.).

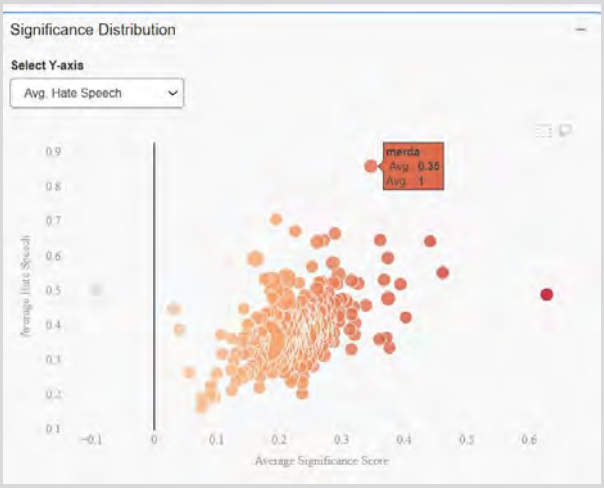
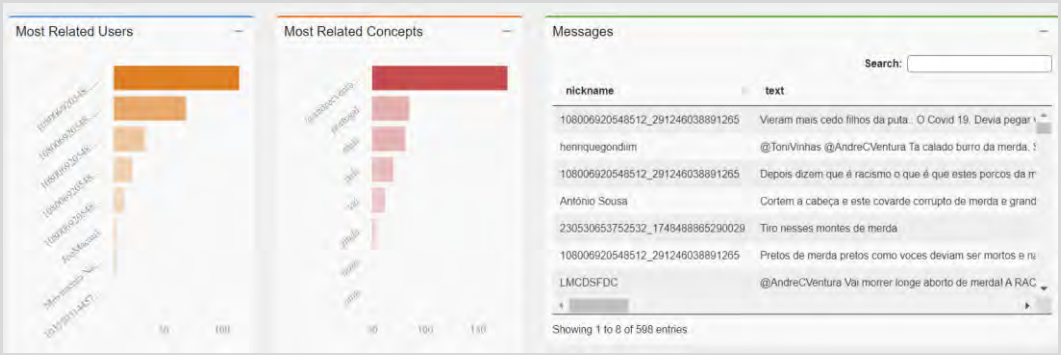


With the Significance distribution graph, we can discover the following information:

Dashboard	Meaning	Useful for
Concepts	Concepts used in suspicious messages.	Discover new keywords.
Entities	Entities that appear in suspicious messages.	What are they talking about in terms of people, locations, and organizations?
Topics	Topics	Understand the main topics of interest of suspicious users.
Key Ideas	Key ideas used in suspicious messages.	A way to detect trending topics in suspicious messages.
Hashtags	Hashtags used in suspicious messages.	Discover hashtags that can be used as keywords.

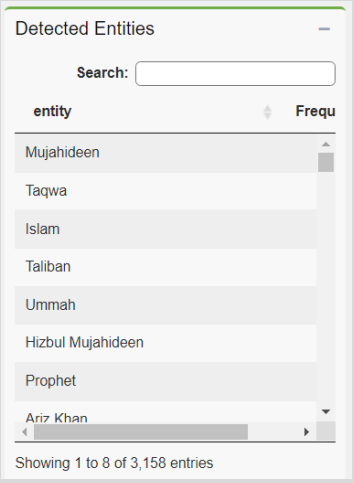
Information about each term:

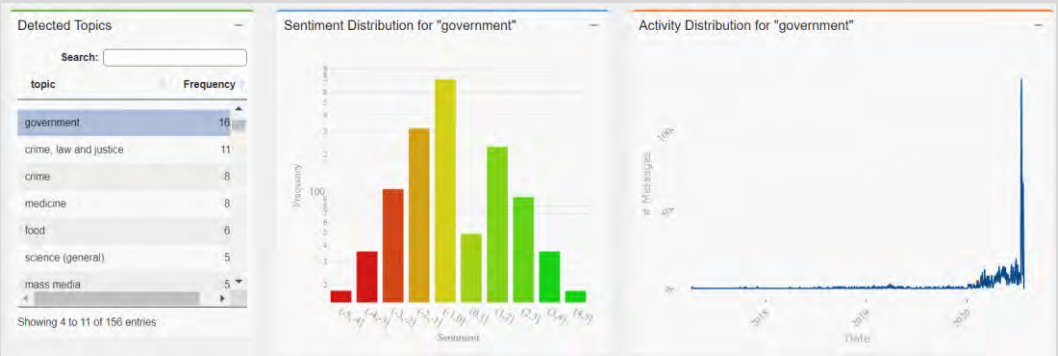
4.3.5.1 Examples

Goal	What are the terms most frequently used in Hate Speech messages?
Steps	• In Global Concepts Analysis, select x-axis as Avg. Hate Speech. • Hover over top nodes to check the most frequent terms in Hate Speech messages. 
Result	Example: check all the messages that include “merda” in their message.  • Top users who use it. • Related concepts (terms in the same message). • Messages (author and message).

Goal	What are the entities (people, organizations, location) most frequently cited in the messages?

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Steps	• Go to the Entities tab. • Check the table at the left that includes all the entities, ordered by frequency.
Result	Top entities by frequency 

Goal	Can we know what messages are related to the government?
Steps	• Got to the Topics tab. • Select “government” into the list or the graph. • Check all the information and messages related to this topic.
Result	Example: Timeline of messages with government as the main topic. 

Goal	Can we discover other hashtags to search for potential relevant content?

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

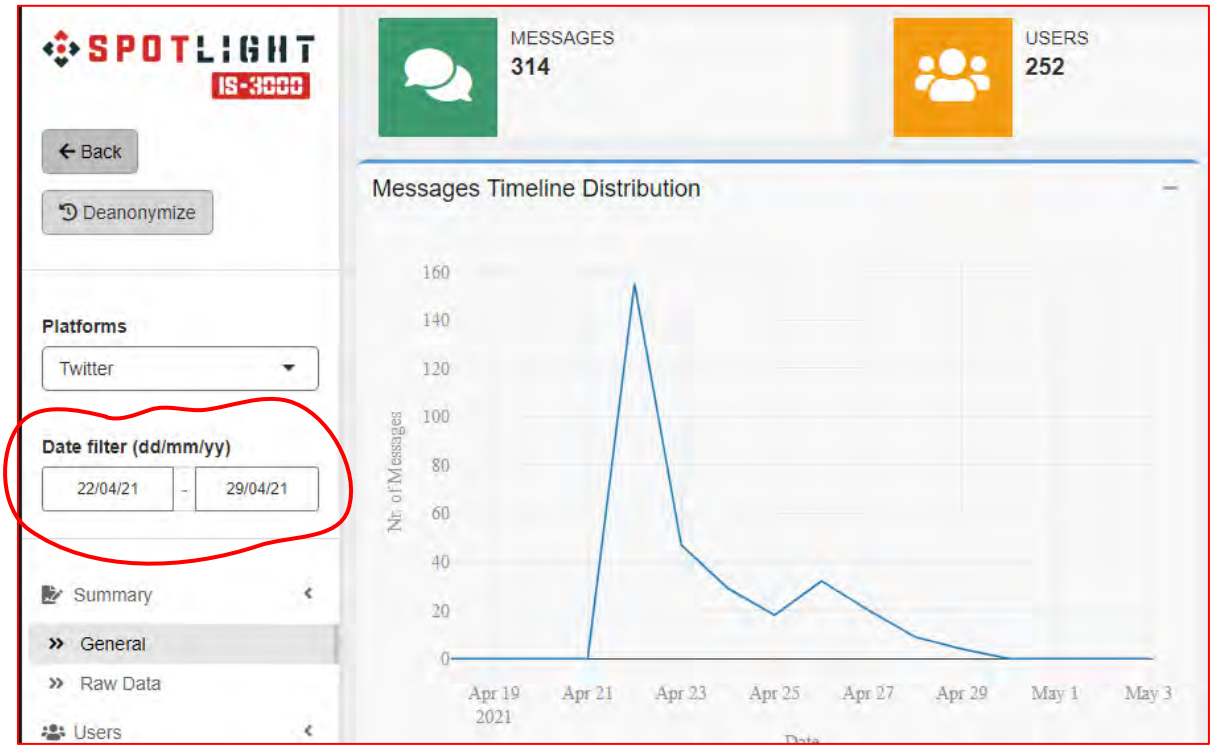
Steps	• Go to the Hashtags tab. • Select the x-axis as Avg. Hate Speech. • Hover over top-right nodes to check the most frequent terms in Hate Speech and Threats messages. • These Hashtags are potential ones for searching.
Result	Examples of this type of Hashtags: • #comunonazis • #hitleralilas • #mediascum • #emigrantes

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

5 Tip and tricks

5.1 How to detect radical content posted in the last week?

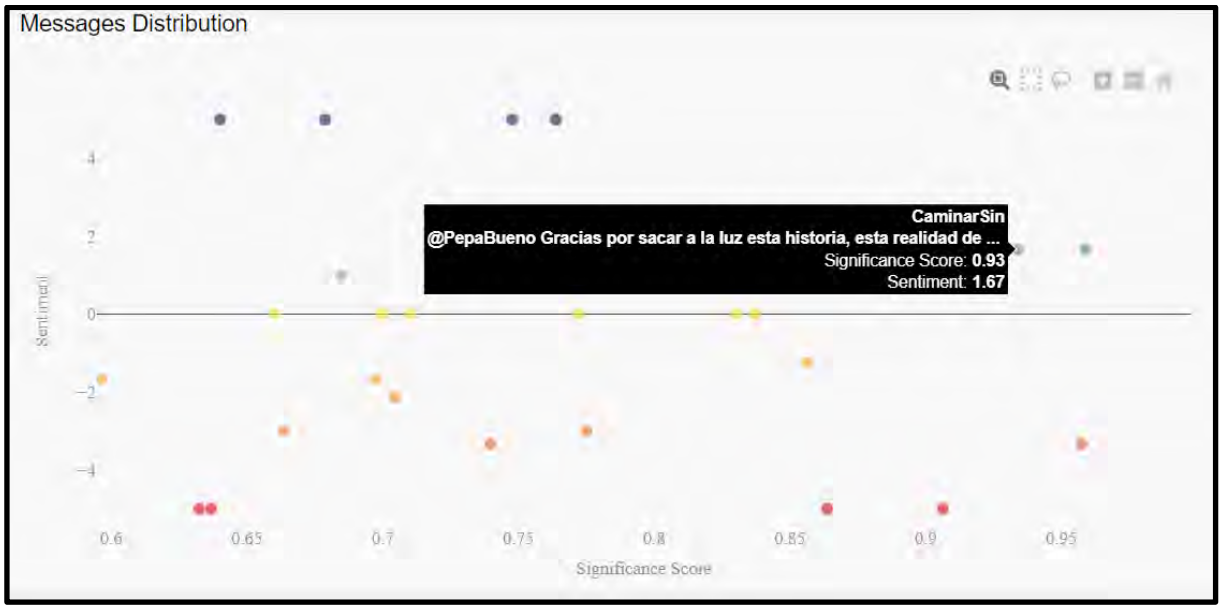
Step 1. Filter all the data by date.



Step 2. Go to the Content / Messages tab and zoom the messages with a Significance Score above 0.6.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Step 3. All the messages with this score are viewed as containing extreme right topics.



The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

5.2 How to discover what people are talking about?

There are many different ways to understand what is the conversation about the selected topic, for example, through the topics.

Step 1. Go to the tab Content / Topic.

We can have a clear picture of the topics of the conversation in Detected Topics.

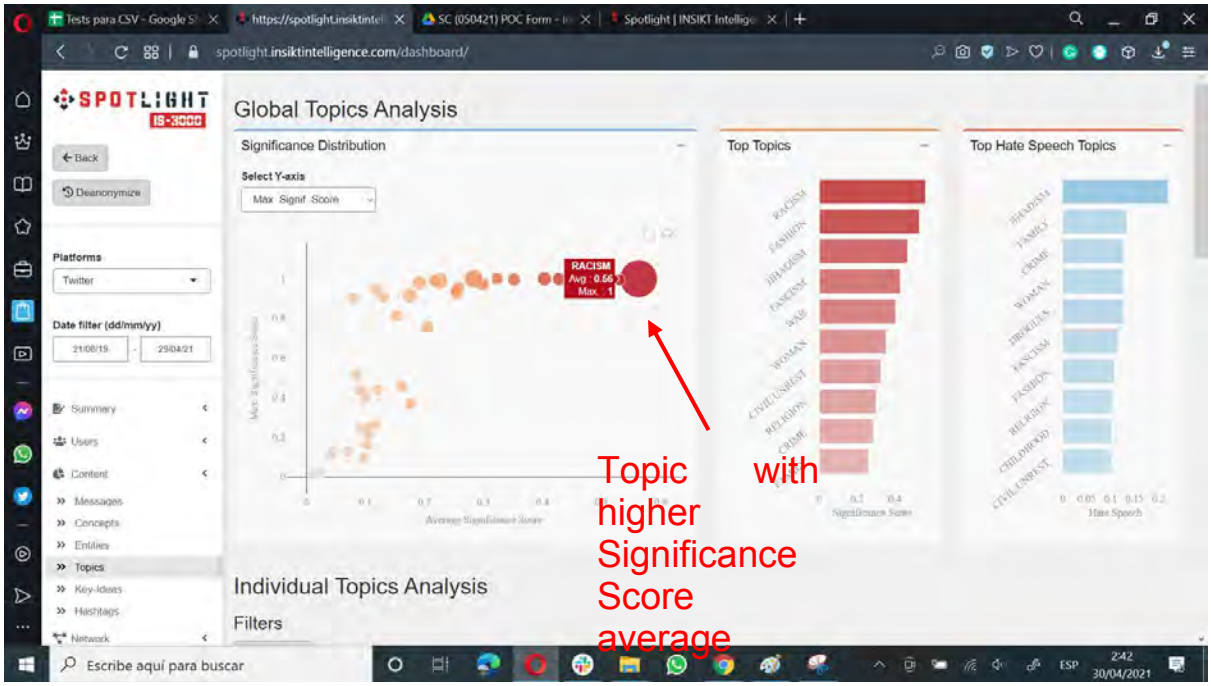
It is easy to get more insight about that just by clicking any of them and check the positive/negative messages related to this topic.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

For example, we select “immigration.”

We get a lot of information about the messages related to this topic.

It is also possible to discover which topic is most frequently mentioned in extreme right messages.



Racism is the most mentioned topic in the conversation.

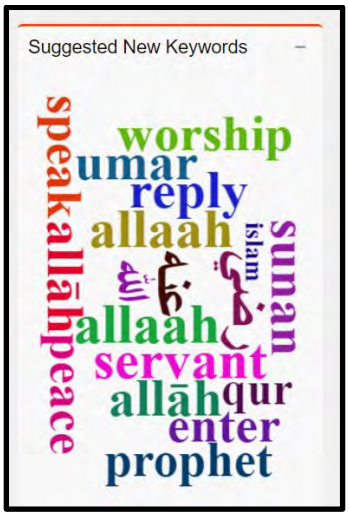
The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

5.3 Could I discover potential new keywords?

Yes, the system suggests some keywords for enlarging the search.

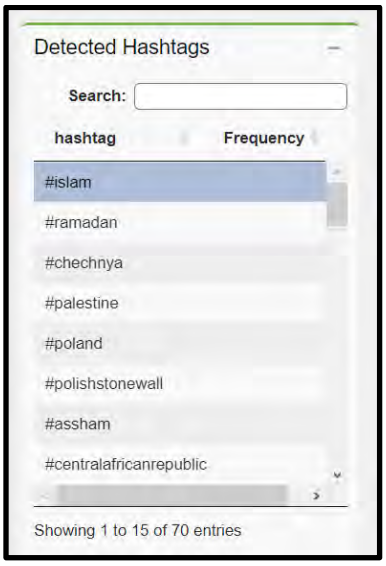
Step 1. Go to Summary / General.

Step 2. Scroll down, and there is a graph containing suggested keywords.



Another interesting way is to check Hashtags.

Step 3. Go to tab Content / Hashtags and check the Hashtags looking for some interesting ones for continuing searching.



The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

6 FAQs

6.1 7.1 From how many data sources is Spotlight capable to extract information?

For the Impetus SMD version, the only available data sources are:

- Online Sources from Social Media:
  - Twitter
  - YouTube
  - TikTok
- Online Sources from Local Press:
  - mattinopadova.gelocal.it
  - document.no
  - reset.no
  - vg.no
  - dagbladet.no
- CSV - external file

For Insikt Spotlight, it depends on the user needs, but, in general, the available sources are:

- Social media platforms
- Blogs
- Forums
- Websites
- News

6.2 What Significance Score is?

It is a score that measures the relevance of a message or user related to the domain you are analysing.

For example, if the selected domain is Extreme Right, the Significance Score will evaluate if the message contains content related to this ideology. I.e., if the message expresses ideas that are usually expressed in extreme right messages.

6.3 7.3 Is it possible to export the raw data?

Yes, but not in the SMD version.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Grant number: 883286
Project duration: Sep 2020 – Aug 2022
Project Coordinator: Joe Gorman, SINTEF

Horizon 2020: Secure societies
SU-INFRA02-2019
Security for smart and safe cities, including for public spaces
Project Type: Innovation Action

http://www.impetus-project.eu

Urban Anomaly Detector (UAD) Tool

SparkGHSOM & DENCAST

Authors: Edoardo Cavallo, Michelangelo Ceci, Paolo Mignone, Claudio Ardagna

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Table of Content

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

1 SPARK-GHSOM- Anomaly Detection

Spark-GHSOM is a machine learning algorithm supposed to automatically identify anomalies from geo-referenced sensors data. This type of task is implemented in several real-world applications, such as the detection of anomalies in car traffic, air pollution, pedestrian trajectories, and so on.

In a geometrical point of view, an anomaly is an outlier data that is far away from the rest of the other instances, and detecting those, help to raise alarms into monitoring system in an urban context.

1.1 Background

Anomaly detection is usually performed by training a model (henceforth anomaly detector) that is capable of catching anomalies from data. For the purposes of the IMPETUS project, three possible phases to train an accurate anomaly detector were identified: the i) initial phase, ii) update phase, and iii) identification phase. In the first two, the model is trained on data, using a batch-learning approach for starting and then an updating approach (using micro-batch too) to keep upgraded on future data the pre-trained model. In the last stage, the anomaly detector is capable to discover the anomalies event, directly from data collected by sensors.

Spark-GHSOM can handle the available sensor data containing spatial (e.g., GPS coordinates) and temporal information (e.g., timestamp) and a set of descriptive variables that are acquired by the specific sensor for the monitoring of the city. For instance, independently of the type of the sensor (e.g., traffic or air pollution), the anomaly detector acts with the same approach. Indeed, the anomaly detector works with values usually indicating the level of different things: pedestrians' concentration, traffic level, temperature, humidity, level of PM2.5, level of CO2, and so forth, that are automatically captured and transmitted by the sensor network. Therefore, in the real situation, the anomaly detector will analyze the data coming from different sensors and it will be able to judge the data as normal or anomalous.

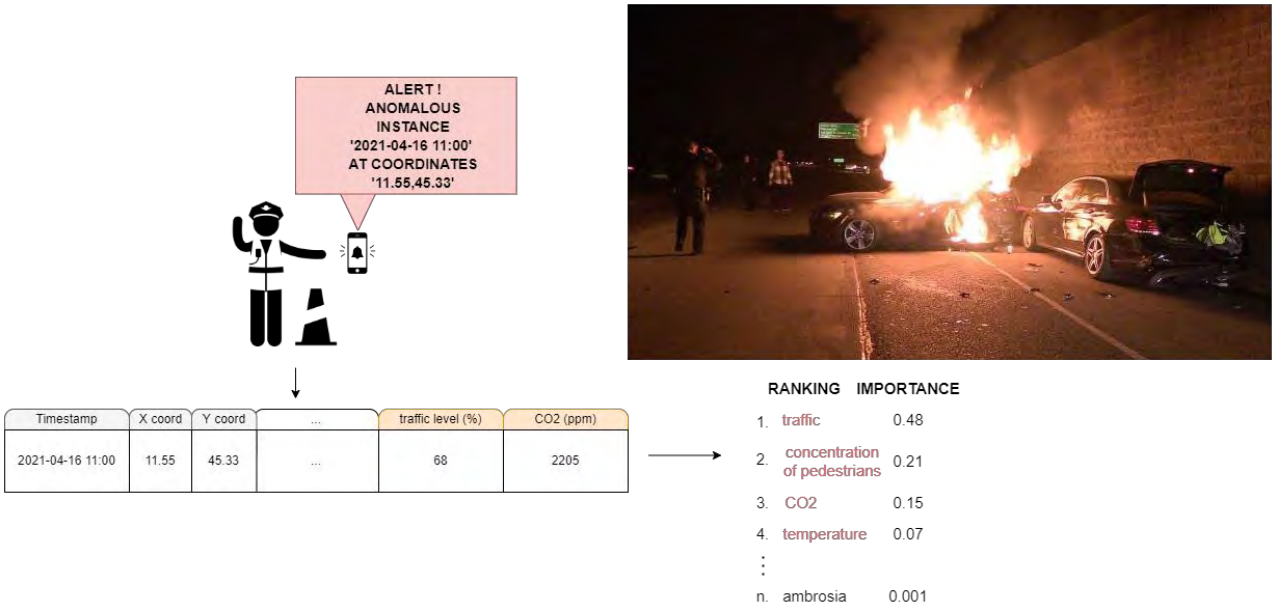


Figure 1 When an anomaly is detected, the system raises an alert and provides a feature ranking according to the features importance in detecting the anomaly.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

1.2 Input parameters

SparkGHSOM takes as input a set of parameters written in a 'parameters.txt' file where are defined:

- HIVE connection (used for fetching models and data)
  - `datasetName`: name of the current running experiment
  - `datasetPath`: path to the dataset used for training phase
  - `testPath`: path to dataset used for evaluating phase

- Kafka queues address of where the output is written to
  - `kafkaAddr`: IP address of host server
  - `writeJson`: Flag boolean (1 = write on file, 0 = don't)
  - `writeKafka`: Flag boolean (1 = write on Kafka queue, 0 = don't)

- Method parameters used for training phase (settings referred to analyst):
  - τ_1 (`tau1`): parameter regulating the growth of a single SOM Layer
  - τ_2 (`tau2`): parameter regulating the hierarchical growth
  - `epochs`: number of reading iterations of the training dataset

1.3 Usage

- Command to run

```
spark-submit --driver-memory 12g --executor-memory 12g --master spark://spark-master:6066 --deploy-mode cluster -class com.sparkghsom.main.utils.AnomDetRunner hdfs://namenode:9000/user/bda/FinalAnalytic/spark_ghsom_final.jar hdfs://namenode:9000/user/bda/FinalAnalytic/parameters.txt
```

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

2 DENCAST – Event Classification

DENCAST is a tool capable of receiving geo-referenced data collected by sensors and return as output the nature of events given in input. It relies on two cluster models, trained by analyst in pre-release state upon data selected and collected for this goal. One model can classify events defined as 'normal' those measurements that represent routine or regular state; the other model can detect the category of those anomalous events that diverge from the normal historical data.

The main difference with the anomaly detection task is that usually the classification of an instance is guided by the learning of a predictive model in a supervised setting. This means that the data set used for the training of the model must be annotated by describing the possible threats for the real scenario. However, this could be demanding to obtain, since it is necessary to hire data annotators to make the data fall into specific categories.

To overcome this problem, unsupervised algorithms could be also considered for the classification task as for the anomaly detection. These algorithms are usually less accurate than the supervised ones since they exploit less informative data avoiding considering predefined classes.

2.1 Background

DENCAST is built upon the concept of DBSCAN algorithm that used a density-based clustering approach to perform the task. DBSCAN starts with an arbitrary object o and, if this is a core object, retrieves all the objects which are density-reachable from it w.r.t. eps and minPts, returning a cluster. The algorithm then proceeds with the next unclustered object.

The algorithm identifies density peaks according to two indicators: a local density indicator, which corresponds to the concept of eps-neighbourhood in DBSCAN, and the maximum similarity (or minimum distance) indicator, computed between the current object and any object with higher local density.

DENCAST model relies on neighbourhood graph, a low dimensional representation of the observed object, built from high-dimensional data thorough the use of the LSH method. This model it's capable of conducting unsupervised learning by identifying density-based clusters.

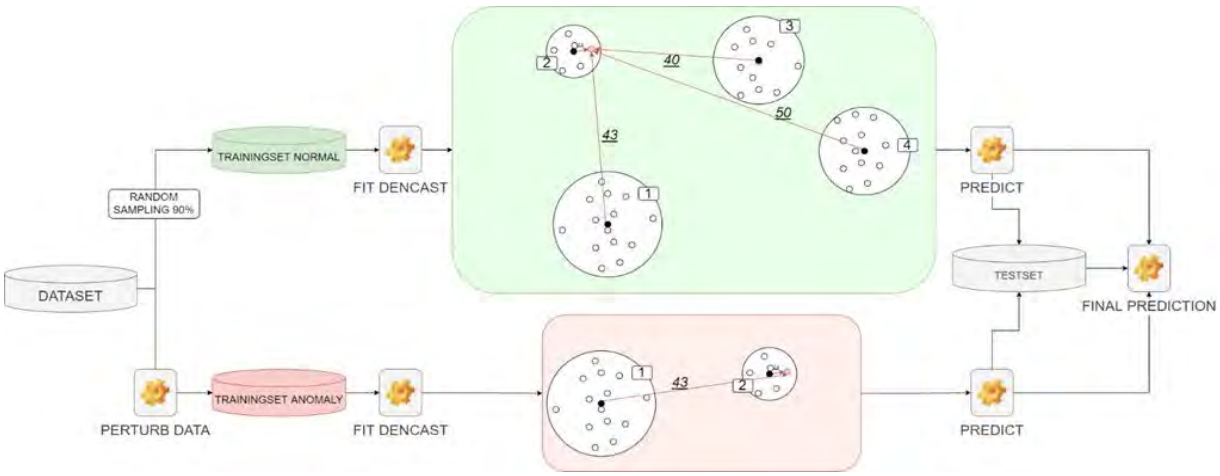


Figure 2 Graphic representation of the clustering approach

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

2.2 Input parameters

DENCAST take in input a set of parameters written in a `'parameters.txt'` file where are defined:

- HIVE connection (used for fetching models and data)
  - `datasetName`: name of dataset to take in input
  - `datasetPath`: path to dataset used for training phase
  - `testPath`: path to dataset used for evaluating phase

- Kafka queues address of where the output is written to
  - `kafkaAddr`: IP address of host server
  - `writeJson`: Flag boolean (1 = write on file, 0 = don't)
  - `writeKafka`: Flag boolean (1 = write on Kafka queue, 0 = don't)

- Method parameters used for training phase (settings referred to analyst):
  - `minPts`: number of minimum of instances that a cluster can store
  - `minCosine`: minimum grade of similarity for comparing instances
  - `dimensions`: number of random vectors (hyperplanes) to generate bit vectors
  - `numNeighbors`: minimum number of neighbors to compare
  - `numPermutations`: number of times bitsets are permuted
  - `partitions`: number of partitions
  - `numThreads`: number random split of dataset

For the training phase, DENCAST take in input two dataset, one is used for training the normal instances (attribute 'anomaly' set to 0), the other one is used for training those instances where the attribute 'anomaly' is set to 1.

2.3 Usage

- Command to run

```
spark-submit --master local[*] --driver-memory 26G --class EventEvaluation
dencast-with-dependencies.jar <path-to-parameters.txt>
```

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

3 Dataset

Spark-GHSOM and DENCAST tool use both same dataset as input data, sharing the same schema.

- Oslo transports (dataset 3): it contains information from a traffic monitoring service of public transportations, aggregated by their position that fall into a specific area (cluster).

Name attribute	Type attribute	Description
TimeWindow	timestamp	Timestamp of measurement (window)
ClusterLatitude	double	Latitude of cluster
ClusterLongitude	double	Longitude of cluster
Numberofvehicles	integer	Number of vehicles passing into the cluster
AvgDelay	double	Average delay of arrival into the cluster
PercInPanic	integer	Percentage of buses in panic into the cluster
PercInCongestion	integer	Percentage of buses in congestion into the cluster
Cluster	integer	ID of cluster
AvgMonitoredCall_VisitNumber	double	Average of buses' route progress (percentage) into the cluster
DateTimedayofTheWeek	integer	Day of the week
DateTimeDay	integer	Day of the month
DateTimeHour	integer	Time hour of the window
OriginAimedDepartureTimeHour	integer	Origin aimed departure time hour
HeadwayService_False	integer	Field defining whether the service is a headway service
Anomaly	integer	Normal = 0, Anomaly = 1

- Padova traffic-control (dataset 1): this dataset contains information received by the car plate readers at the gates of plate-number monitoring system. The sensors recognize vehicles passing through the gates and record their license plate and timestamp along with other information.

Name attribute	Type attribute	Description
Start_Time	timestamp	Timestamp of window
X	double	Latitude of cluster
Y	double	Longitude of cluster
Cameraname	string	ID of camera
No_approaching	integer	Number of cars approaching to sensor

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

No_leaving	integer	Number of cars leaving from sensor
No_unknown	integer	Number of cars not detected
DateTimedayofTheWeek	integer	Day of the week
Month	integer	-
Day	integer	-
Hour	integer	-
Minute	integer	-
Anomaly	integer	Normal = 0, Anomaly = 1

- Padova people counter (dataset 2): this dataset contains information about the people counters installed in Piazza Dei Signori. The data are aggregated by a 5-minute time window and indicate the people who entered and left a specific entrance during a given time window.

Name attribute	Type attribute	Description
start_Time	timestamp	Timestamp of window
X	double	Latitude of cluster
Y	double	Longitude of cluster
Sensor_Id	integer	ID of sensor
Year	integer	-
Month	integer	-
Day	integer	-
Hour	integer	-
Minute	integer	-
Week_Day	integer	Day of the week
Count_People_In	integer	Number of people inside of area
Count_People_Out	integer	Number of people outside of area
Holiday	Integer	Not holiday = 0, Holiday = 1
Anomaly	integer	Normal = 0, Anomaly = 1

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

4 Output

Here is shown a schema in JSON format as result of the evaluation phase given in output by both tools. Output file contains a field 'ranking' where are listed into an array the features taken in input by the dataset itself (feature ranking process).

Name attribute		Type attribute	Description
Timestamp		Timestamp	Timestamp ISO 8601
Coordinates		Pair of double	Latitude and longitude of the area where the event happened
Anomaly		Boolean	Normal = false, Anomaly = true
Ranking		Array of object of features	List of all features and their 'importance' in describing the nature of event (why it is an anomaly or not).
ARRAY OF FEATUREs	**Feature**	**Type of attribute**	**Description**
	Feature	String	Name of the feature
	Value	String | Double | Integer | Boolean | Timestamp	Value of the feature
	Importance	Double	Percentage of how much a specific variable influences the nature of the event.

For instance, is reported an example of output from Padova traffic-control (dataset 1) for a single sensor data measurement.

```
{
"timestamp" : "2022-06-28T10:50:00+01:00",
"coordinates" : [45.427622, 11.879039],
    "ranking" : [
    {
          "feature" : "no_approaching",
          "value" : 150.0,
          "importance" : 0.3
    },
    {
          "feature" : "no_unknown",
          "value" : 150.0,
          "importance" : 0.28
},…
anomaly = true}
```

The output generated by both tools is sent via Apache Kafka that enable asynchronous authentication between users and platform, buffering of the data sent to the platform, and scalability.

Tools communicate in secure encrypted channel not giving access to data by third parties, so an authentication with username and password is required. Once the output is sent to the platform, it could process such output in order to show it within a graphical interface

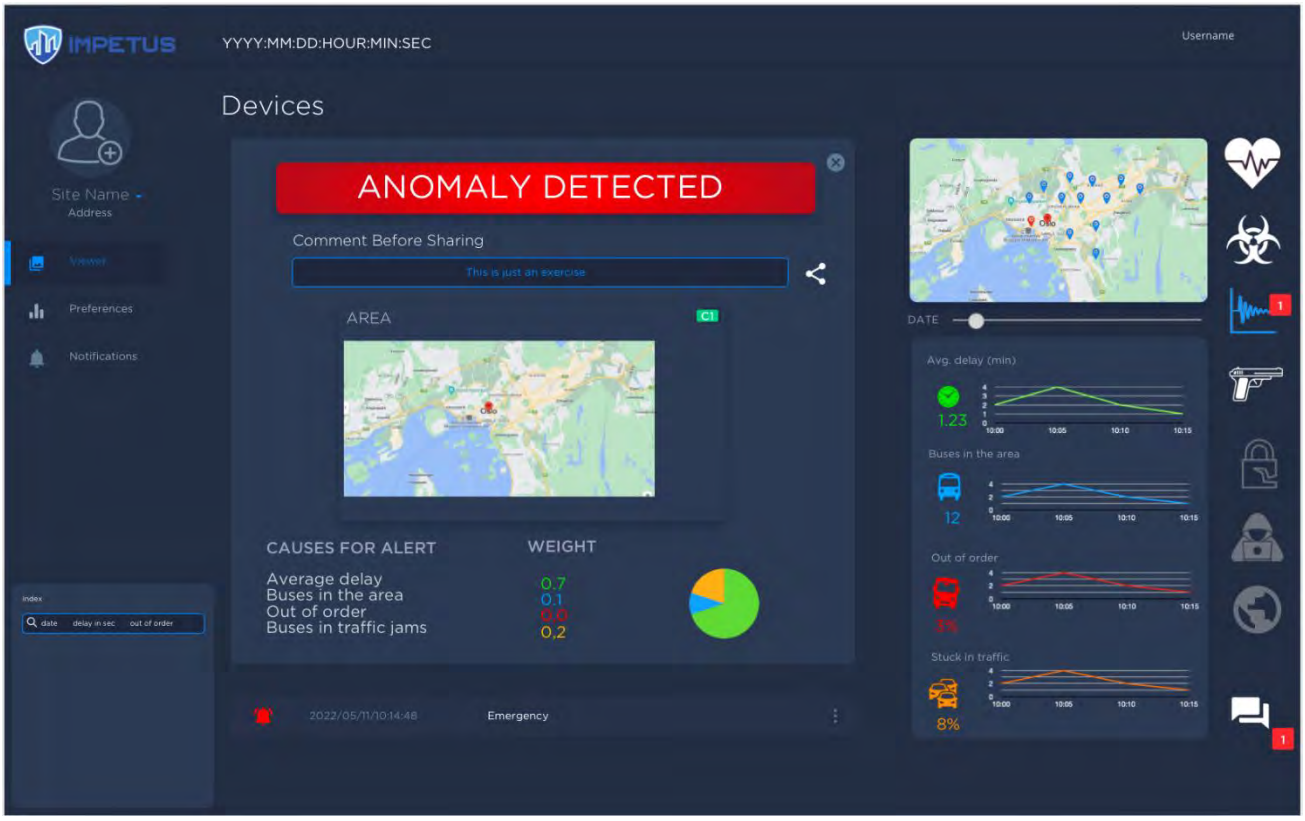


Figure 3: Example of interface that exploited the provided output of the Urban Anomaly Detector tool. The cause for alert" represents the motivation in terms of variables that contributed to raise an alert; the "weight" represents the importance of the feature that feature.

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

Members of the IMPETUS consortium

SINTEF	SINTEF, Strindvegen 4, Trondheim, Norway, https://www.sintef.no	Joe Gorman joe.gorman@sintef.no
Institut Mines-Télécom	Institut Mines Telecom, 19 place Marguerite Perey, 91120 Palaiseau, France, https://www.imt.fr	Joaquin Garcia-Alfaro joaquin.garcia_alfaro@telecom-sudparis.eu
UNÎMES	Université de Nimes, Rue du Docteur Georges Salan CS 13019 30021 Nîmes Cedex 1, France, https://www.unimes.fr	Axelle Cadiere axelle.cadiere@unimes.fr
cini consorzio interuniversitario nazionale per l'informatica	Consorzio Interuniversitario Nazionale per l'Informatica, Via Ariosto, 25, 00185 – Roma, Italy, https://www.consorzio-cini.it	Donato Malerba donato.malerba@uniba.it
Università degli Studi di Padova	University of Padova, Via 8 Febbraio, 2 - 35122 Padova, Italy, https://www.unipd.it	Giuseppe Maschio giuseppe.maschio@unipd.it
Entrepreneurship Development Centre for BIOTECHNOLOGY and MEDICINE	Biotehnoloogia ja Meditsiini Ettevõtluse Arendamise Sihtasutus, Tiigi 61b, 50410 Tartu, Estonia, https://biopark.ee	Sven Parkel sven@biopark.ee
SIMAVI Software Imagination & Vision	SIMAVI, Complex Victoria Park, Corp C4, Etaj 2, Șos. București – Ploiești, nr. 73 – 81, Sector 1, București, Romania, https://www.simavi.ro	Gabriel Nicola Gabriel.Nicola@simavi.ro Monica Florea Monica.Florea@simavi.ro
THALES	Thales Nederland BV, Zuidelijke Havenweg 40, 7554 RR Hengelo, Netherlands, https://www.thalesgroup.com/en/countries/europe/netherlands	Johan de Heer johan.deheer@nl.thalesgroup.com
CINEDIT INTELLIGENT VIDEO ANALYTICS	Cinedit VA GmbH, Poststrasse 21, 8634 Hombrechtikon, Switzerland, https://www.cinedit.com	Joachim Levy j@cinedit.com

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

INSIKT INTELLIGENCE	Insikt Intelligence, Calle Huelva 106, 9-4, 08020 Barcelona, Spain, https://www.insiktintelligence.com	Dana Tantu dana@insiktintelligence.com
SIXGILL	Sixgill, Derech Menachem Begin 132 Azrieli Tower, Triangle Building, 42nd Floor, Tel Aviv, 6701101, Israel, https://www.cybersixgill.com	Benjamin Preminger benjamin@cybersixgill.com Ron Shamir ron@cybersixgill.com
Comune di Padova	City of Padova, via del Municipio, 1 - 35122 Padova Italy, https://www.padovanet.it	Enrico Fiorentin fiorentine@comune.padova.it Stefano Baraldi Baraldis@comune.padova.it
Oslo	City of Oslo, Grensen 13, 0159 Oslo, Norway, https://www.oslo.kommune.no	Osman Ibrahim osman.ibrahim@ber.oslo.kommune.no
Institut za sigurnosne politike ISP	Institute for Security Policies, Kruge 9, 10000 Zagreb, Croatia, http://insigpol.hr	Krunoslav Katic krunoslav.katic@insigpol.hr
TIEMS	International Emergency Management Society, Rue Des Deux Eglises 39, 1000 Brussels, Belgium, https://www.tiems.info	K. Harald Drager khdrager@online.no
UniSMART Fondazione Università di Padova	Unismart – Fondazione Università degli Studi di Padova, Via VIII febbraio, 2 - 35122 Padova, Italy, https://www.unismart.it	Alberto Da Re alberto.dare@unismart.it

The research leading to these results has received funding from Horizon 2020, the European Union's Framework Programme for Research and Innovation (H2020) under grant agreement n° 883286.

IMPETUS

Workload Monitoring System

(Human-Computer Interaction tool)

User Manual

v3

Table of Contents

1 Overall workflow

Here, we describe the overall idea and the purpose of the tool, in the broader context of Human-Machine teaming.

1.1 The loop

The purpose of the tool is to measure bio-signals of the Human operators, who are interacting with their equipment (the machines) and each other, while performing their given tasks. Then, using customized and personalized machine-learning models, to assess the operators' momentary mental workload and emotional stress levels. This assessment, made by the tool, is then presented (in a graphic form) as feedback to the operators and their supervisor.

Measuring of the bio-signals is done continuously, in real-time and as unobtrusively as possible. Specific configuration and functions of the sensors that are used for the measurement of the bio-signals are explained later in this document.

The Tool processes the bio-signal data-streams in real-time and uses (personalized) models, created prior to the use of the tool, to produce an assessment of the mental, emotional, and physical workloads of each individual operator. These assessments are updated continuously and frequently.

The assessment can be shown as feedback in configurable amount of detail, on individual and aggregated (team) levels, to a person or persons of choosing, by means of a (digital) dashboard.

The ultimate goal of the tool is to provide timely feedback and assure that the operators can perform their tasks without being overloaded or overstressed, by removing these unwanted effects which can impede their work and introduce unwanted fatigue and stress. This feedback can then be used to increase their well-being and effectiveness.

The following graphic summarizes the steps of the continuous loop of the tool: biosensing, analysis, assessment, and feedback.

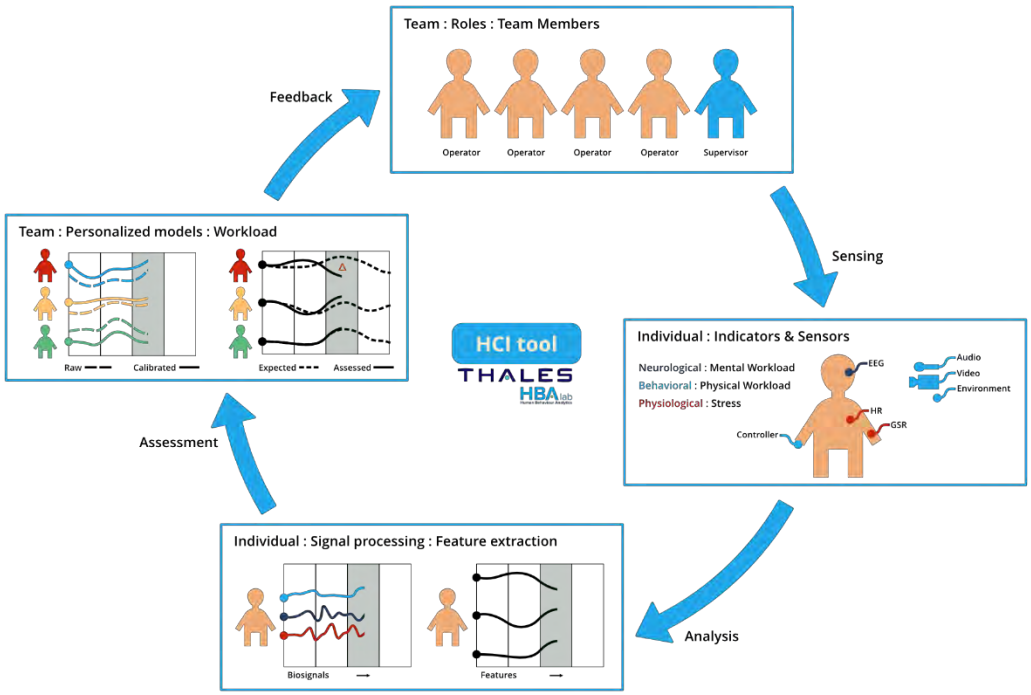


Figure 1: The Workload Monitoring System loop

1.2 The assessment

The workload assessment provided by the tool comprises three individual components or indicators, namely: the Mental, Emotional and Physical.

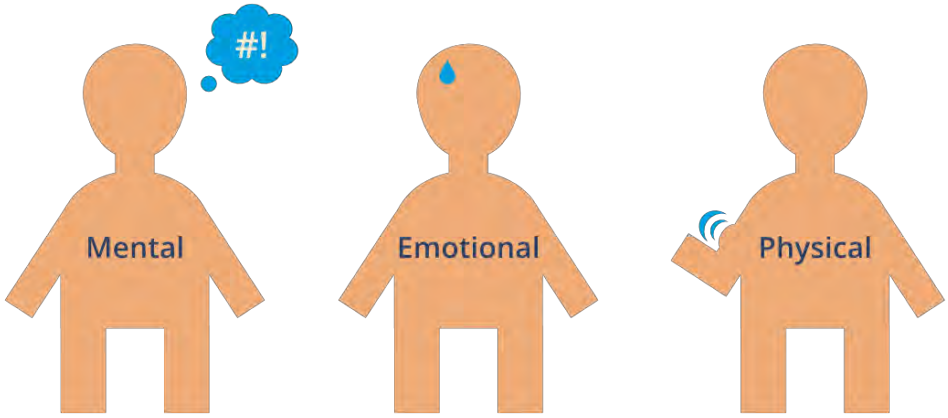


Figure 2: The tree Workload indicators

1.2.1 Mental Workload

This indicator is a measure of the cognitive effort, exerted by our brains. It involves information processing, mental calculation, (short term) planning and memory tasks. In most simple terms it can be summarized as the effort of (rational) thinking.

1.2.2 Emotional Workload

This indicator captures the emotional component to workload, that is the more basis physical arousal or (short term) stress. It can be thought as the irrational counterpart of the mental workload.

1.2.3 Physical Workload

This indicator reflects the work involving muscle movement or tension. Even though the work of most operators using the Workload Monitoring System has a relatively limited physical component, it is prudent to monitor this as increased levels of stress, repetitive or prolonged movement or tension can lead to undesirable effects on the whole body and thus in turn, affect the other mental and emotional indicators.

2 Intended use of the tool

The intended users of the tool are primarily the operators and their supervisors. The Tool measures the operators' bio-signals and assesses their workload. This information is shared with the supervisors via a graphical interface. Both users have their own dedicated hardware and software, which they can use in a specific way. The following sections explains this in more detail.

2.1 Operator

This section first explains the hardware and software and subsequently explains it in the context for the operator.

2.1.1 The Sensor

The Muse sensor is a wearable brain-sensing headband. The device measures electrical brain activity via four electroencephalography (EEG) sensors (electrodes). The Muse is manufactured by InteraXon, a company based in Toronto, Ontario, founded in 2007. We use the Muse 2 and Muse S, which also measures the pulse (via photoplethysmography, or PPG), and head movement (rotation with accelerometers and translation with gyroscopes).

Muse is worn over the ears and wirelessly connects to a computer via Bluetooth. The data connection is made using the Dashboard Streams communicating the current EEG, PPG, and accelerometer values at a refresh rate of 256 hertz.

Figure 3: The Muse 2 and Muse S Sensors

2.1.2 Sensor Instructions

2.1.2.1 The Muse

The Muse sensor measures electrical brain signals (using ElectroEncephaloGraphy, or EEG) and heart rate (using PhotoPlethysmoGraphy, or PPG). The Muse sensor has two different versions: the 'glasses' and the 'headband' version. Both sensors use Bluetooth to wirelessly send and receive data. Both operate on a battery. The batteries are large enough to last a full day.

2.1.2.2 Muse S (headband variant)

The Muse S integrates EEG and PPG sensors in one, wireless headband. The clasp can be used to adjust the band to fit snugly on your head. The "Pod" houses the main sensors, electronics and battery. Additional sensors are the gold metallic strips on either side of the head, resting on the ears.

Figure 4: The Muse components

How to wear the Muse sensor is explained step-by-step here:

Figure 5: How to wear the Muse S

1. To fit your Muse S appropriately to your head, ensure the Muse logo is right side up and the soft beige color sensors that are placed alongside your ears are facing down. These sensors need to be flush against your skin so if you have long hair you might choose to tie it back. Fasten the clasp at the back of your neck and then slide the band up to your forehead. Ensure the pod sits at the center of your forehead, pod facing forward.
2. The headband should run across the middle of your forehead – not too high or too low. Just above the eyebrows.
3. Be careful not to stretch the band near the sensors as this can damage the sensors.
4. Run your finger over the top of the pod to feel for the power button. If you feel the subtle raise of the power button, you will know your headband is right-side up.

Additional Tips:

- It is important that both ear sensors make direct contact with the skin behind your ears. Move any hair behind your ears out of the way as you adjust the fit to ensure good signal quality.
- If you have long hair, it is recommended that you put your hair up in a ponytail to get a good fit and optimal signal quality.

2.1.2.3 Muse 2 ("glasses" variant)

This variant has the same sensors (EEG and PPG) but in a slightly different format. The sensors are placed in the middle of the band, making contact with the forehead. Additional sensors are in the rubbery loops resting on and behind the ears. The length of the band can be adjusted by sliding the end pieces in or out – much like with many adjustable headphones.

Figure 6: How to wear the Muse 2

1. Expand your Muse to its largest size and place the rubber ear sensors behind your ears.
2. Adjust the headband to tighten it back up for a snug fit that feels comfortable.
3. The headband should run across the middle of your forehead – not too high or too low. Just above the eyebrows.
4. Ensure all sensors have good skin to sensor contact.
5. Make sure that there is no hair between sensors and your skin, as this will prevent the Muse from getting a good signal.

Note:

If you have long hair, it is recommended that you put your hair up in a ponytail to get a good fit and optimal signal quality.

2.1.3 The Data Acquisition Unit (DAU)

Figure 7: The Workload Monitoring System - Data Acquisition Unit (DAU)

2.1.4 Connection and Signal acquisition view

This screen provides the user with the information on basic functioning, like the connection status and the battery level of the sensor. Furthermore, the data-streams that are continuously captured, are plotted. This screen thus provides the first overview and enables the user to check if the sensors operate properly and assess the data quality.

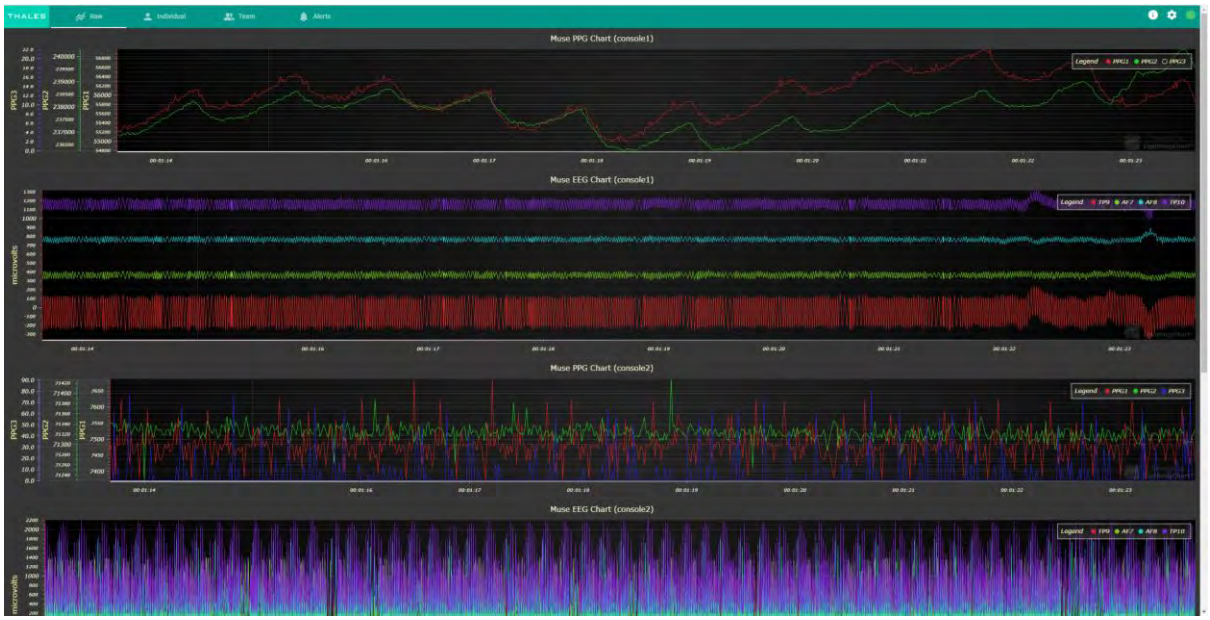


Figure 8: Connection and Signals view

2.1.4.1 Sensor Link

To connect the DAU to the sensor in order to send the data from the sensor to the DAU for analysis, the sensor link needs to be established. To link the sensor, click the Bluetooth icon, in the upper right corner of the interface.

Next turn on the sensor by clicking the power button. You can click the “scan now” button to list the available Bluetooth Devices. Select the sensor you want to connect to and click “Pair”.

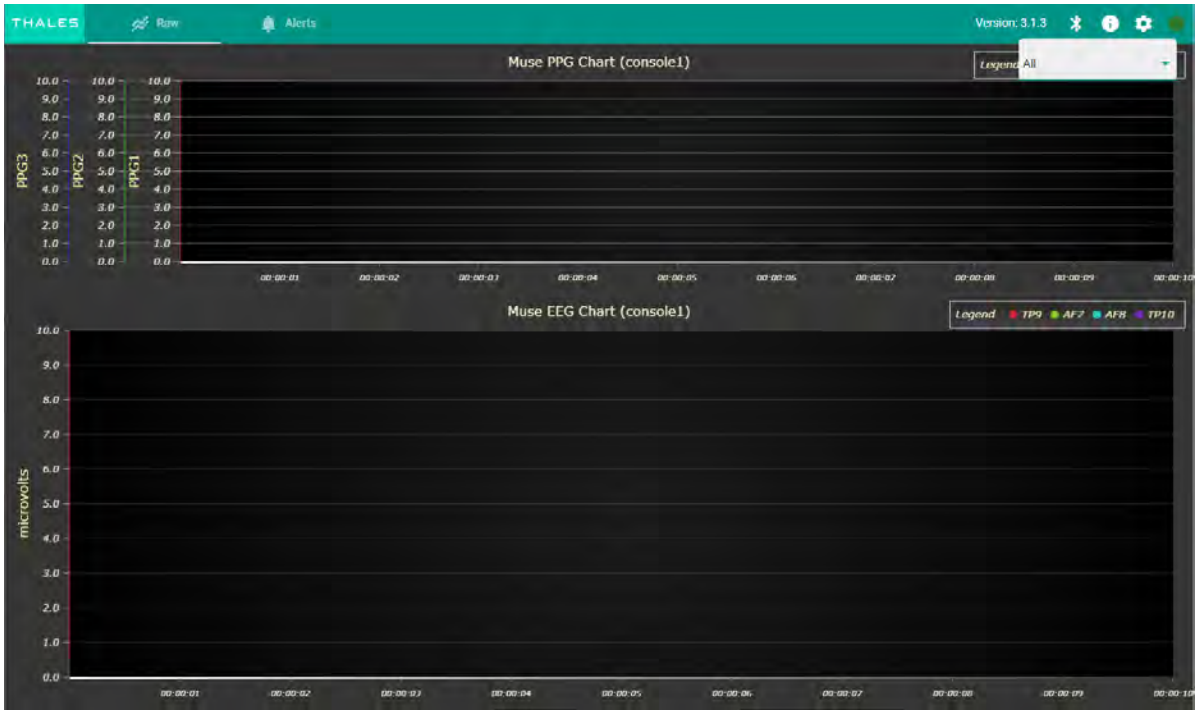


Figure 9: Connecting the sensor to the DAU

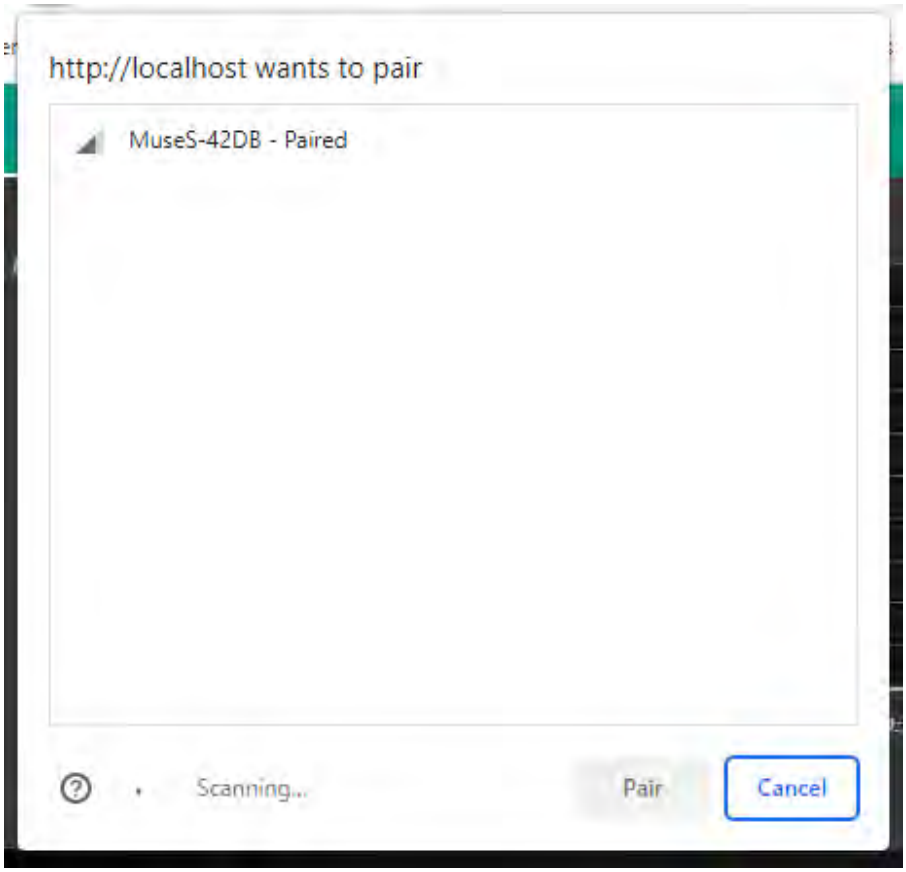


Figure 10: Sensor link confirmation pop-up window

Figure 11: Link established

Once the Sensor link is established, additional user interface elements will appear on the menu bar, in the top-right corner. The first is the battery indicator icon, which shows the available battery level of the connected sensor. The second is the "disconnect" button. This can be pressed to disconnect the sensor (break the sensor link). Finally, once the sensor link is established and the data is being sent to the DAU, the graphs will start to be drawn and will be continuously updated.

2.1.4.2 Alerts view

This view shows notifications of possible sensor connection errors. The function of these notifications is to alert the user of possible issues the sensor, either related to the battery power or connection status. Additionally, workload assessment notifications are also shown here. The newest alerts are added at the top of the list and all the previous alerts (the history), are displayed below.

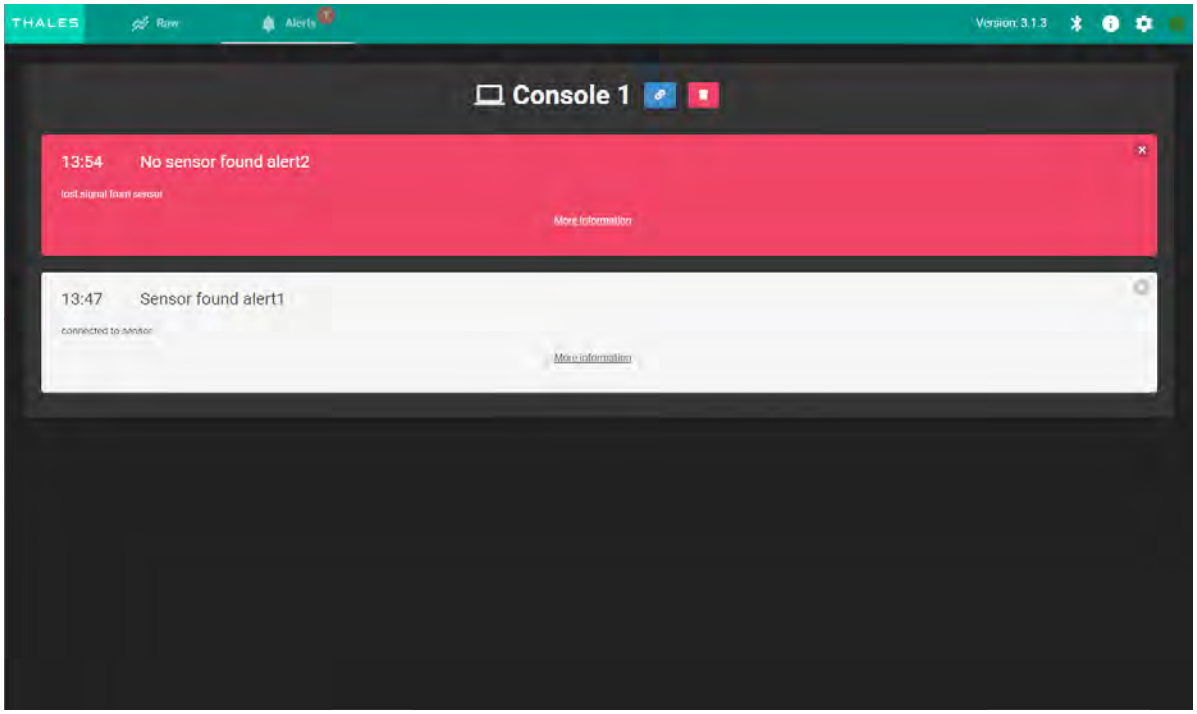


Figure 12: Alerts view

Click “x” to delete one alert.

Click the “trashcan” icon to delete all alerts for the console.

Click the “link” icon to go to the data for the console.

Possible alert codes:

‘001’ : “no sensor found”,

‘002’: “sensor found”,

‘010’: “bad signal”,

‘011’: “average signal”,

‘012’: “good signal”,

‘060’: “unexpected assessment”,

‘061’: “expected assessment

2.2 Supervisor

This section explains the software functionality of the Workload Monitoring System, aimed at the supervisor.

The Dashboard consists of several views. The contents of these views are described here.

2.2.1 Team-view

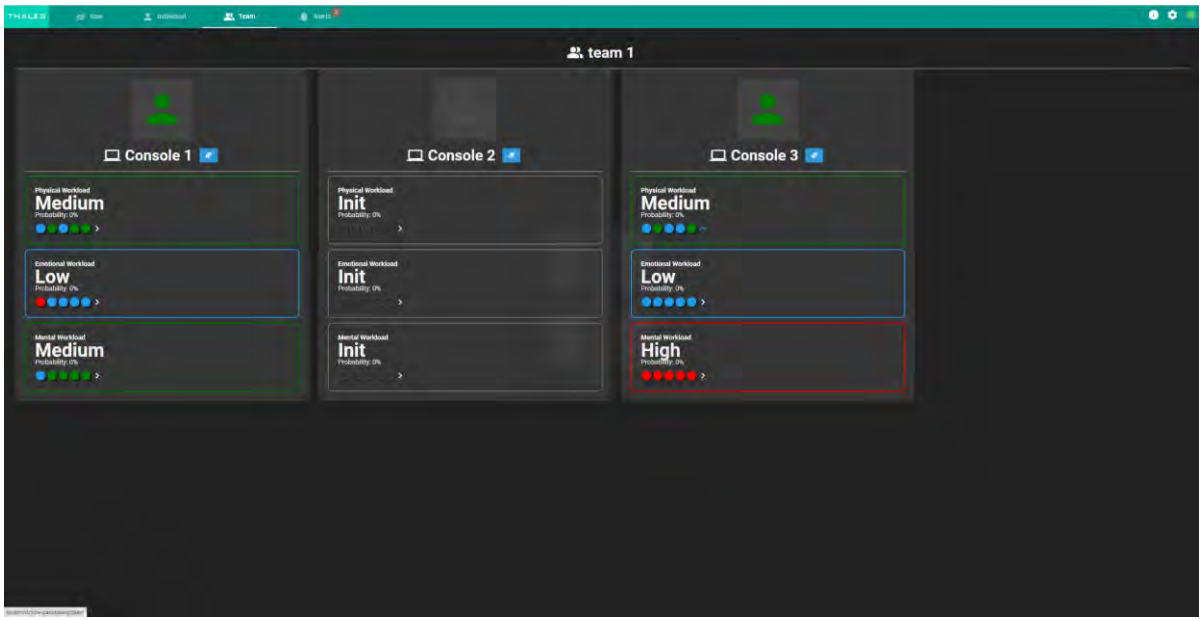


Figure 13: The Team dashboard view

This view shows all connected operators (consoles) and their current workloads. All three workload assessments are shown for each operator, that is the mental, emotional and physical workload. The current workload level is shown as "Low", "Medium" or "High". Additionally, the recent history is displayed below (the previous five workload levels), depicted by colored dots, with the most recent one to the right. This view enables a quick overview of the current status and trend of the team's workload.

2.2.2 Individual operator view

2.2.2.1 Biodata based features

Here, you can keep track of the features of the individual operator. In the upper part of the interface, the most relevant EEG and PPG features are shown. They are continuously updated every 5 minutes. This update interval is configurable in the settings.

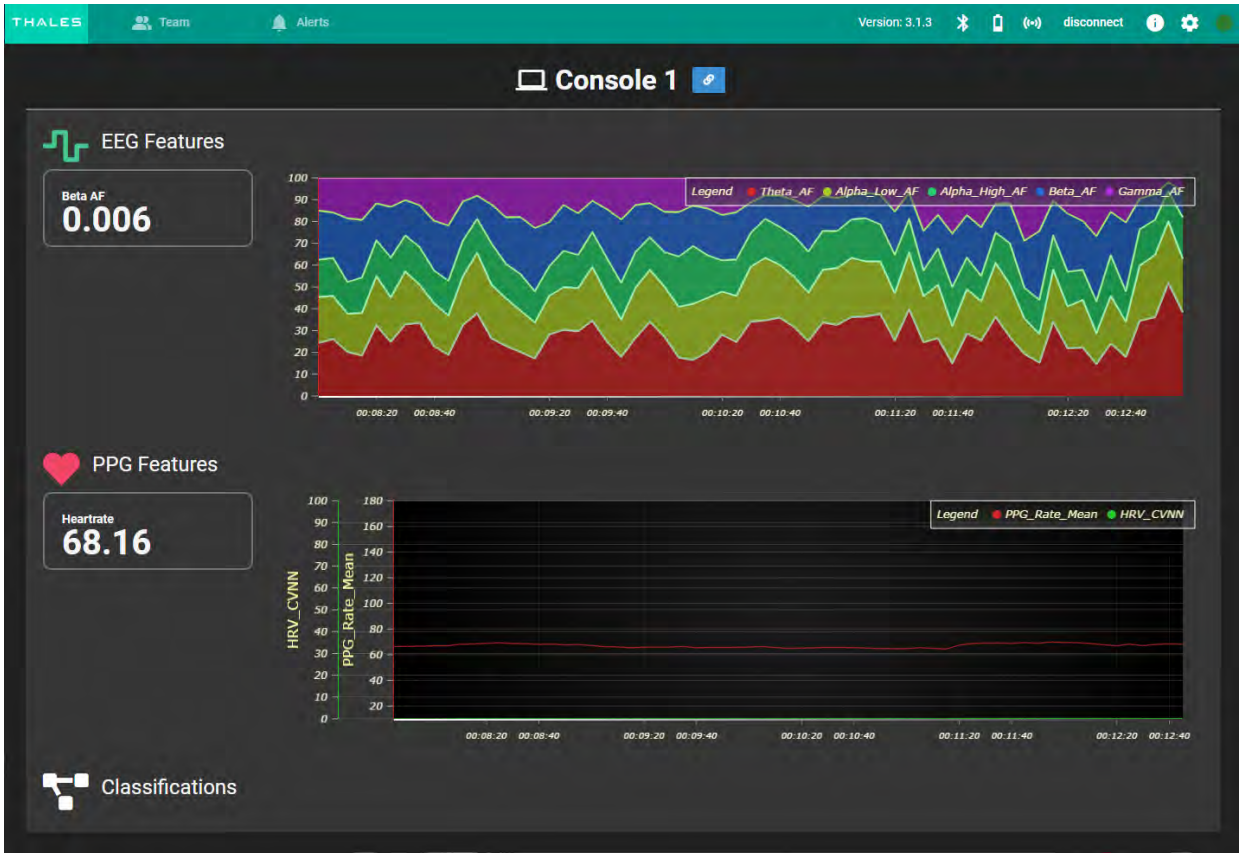


Figure 14: Individual operator view

2.2.2.2 Classifications

Below the feature graphs, the interface shows the workload assessment. The Classifications are provided for the three indicators of workflow, namely Mental, Emotional and Physical. These are described earlier in this document.

Classifications of each of these indicators are updated at every interval and reflect the state of the operator. The current workload assessment of the operator is shown in one of three levels, as low, medium, or high workload. The confidence of this classification, shown as a percentage, reflects the predicted accuracy of the classification.

The most recent history of classifications is shown in five steps from left to right, with the one furthest to the right being the most recent one. The levels of workload are color-coded: blue is low, green is medium and red indicates a high workload. The trend is shown using an arrow to indicate a rise, stabilization or decline in workload classification, over a longer period.

2.2.3 Alerts view

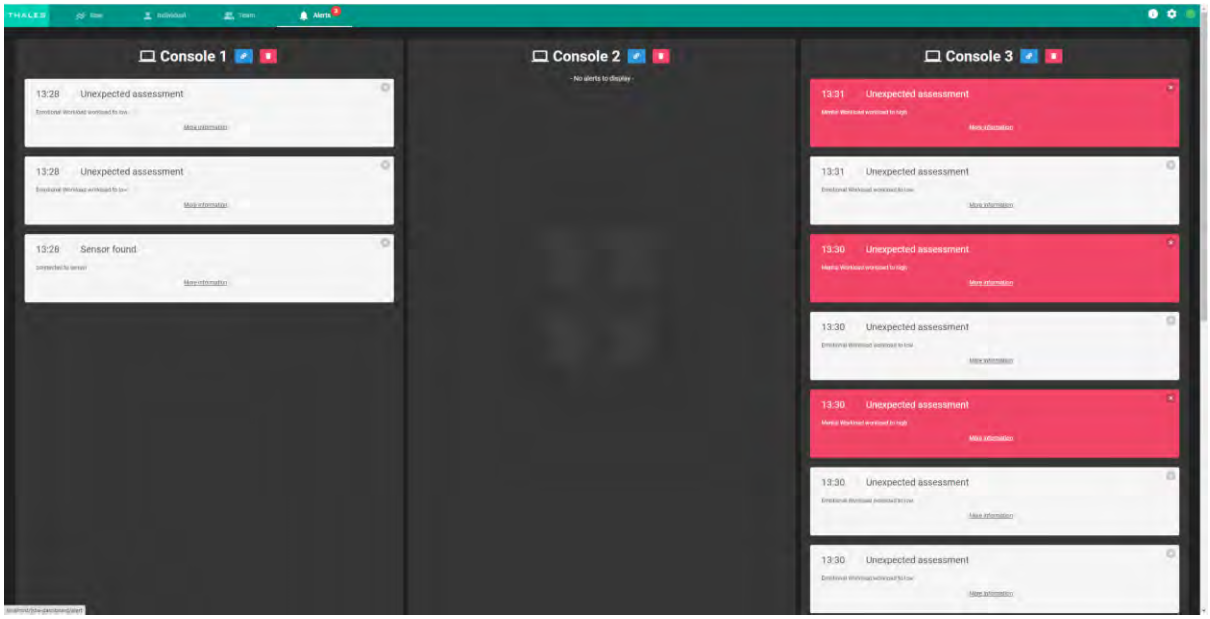


Figure 15: The Alerts view of the dashboard

This view shows only the notifications of workload that exceeds an expected or desired level. The function of these notifications is to alert the user of possible issues with too high or too low workload, of a specific operator. All alerts are displayed, with the most current one at the top, and all the previous alerts (the history), are displayed below.

2.2.4 Connection and Signal acquisition view

This screen provides the user with the information on basic functioning, like the connection status and the battery level of the sensor. Furthermore, the data-streams that are continuously captured, are plotted. This screen thus provides the first overview and enables the user to check if the sensors operate properly and assess the data quality.

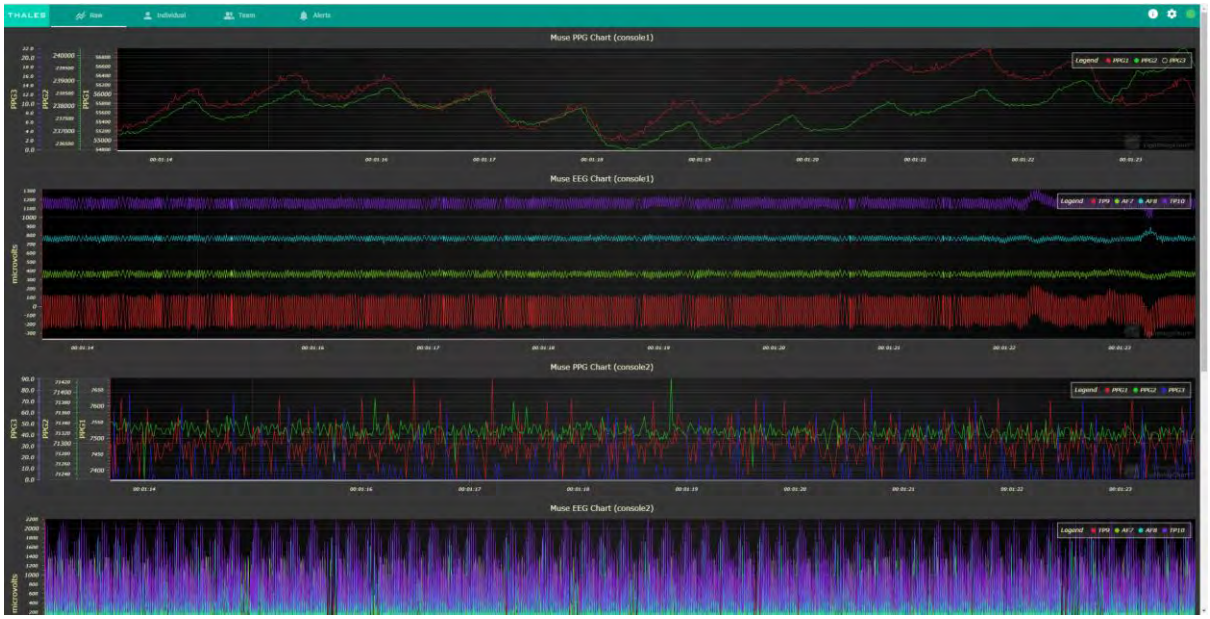


Figure 16: Connection and Signals view

You can click the plots to reset the data visualization. You can click the items in the legend to toggle whether they are drawn in the plot.

3 In-depth Information

3.1 Brain-Computer Interface (BCI)

The sensor used (the Muse) to measure the electrical brain activity, is a so-called 'passive Brain-Computer Interface (BCI)'. A generic definition of a passive BCI is: "A passive BCI derives its outputs from arbitrary brain activity without the purpose of voluntary control, for enriching a human – computer interaction with implicit information". In the specific case of the Muse senor, the brain activity that is measured is the electrical activity of the cortex (most outer layer of the brain, right under the skull) in a number of locations. That electrical signal is then processed and analyzed, and the relevant features of that signal are extracted and used as information. That is the output of the passive BCI, and is shared with the user of the system, in an accessible, explicit and understandable way, via the graphical interface.

Figure 17: The Muse S sensors

3.2 Bio-signals: EEG

Our brain controls and processes all functions of our body, including muscle movement, sensory functions such as touch, hearing and sight, and more complex functions, such as memory, emotion and cognition. The brain consists of around 86 billion neurons, which form a complex messaging and processing system [1]. Neurons are able to pass along and process messages in the form of electrical pulses. Hereby, the brain can be divided in an inner layer (white matter) and an outer layer (gray matter) (see Figure 4). The main function of neurons within the inner layer is to transport the messages within the brain and from and to the brain through the spinal cord. The neurons within the outer layer function as the processing center of the messages.

Figure 18 White and grey matter of the brain (source: https://www.technologynetworks.com/neuroscience/articles/gray-matter-vs-white-matter-322973)

With an electroencephalogram (EEG), we can study the electrical activity of the outer layer of the brain. Sensors in the form of small metal discs/plates are placed on the scalp. These sensors are sensitive to any electrical change in the outer layer of the brain. Consequently, electrical pulses (i.e., messages) that propagate through neurons within the outer layer of the brain will be measured by the sensors and can be visualized as signals. Within EEG signals we distinguish different frequency bands in which signal pulses propagate through neurons. A few of these bands are shown in figure below.

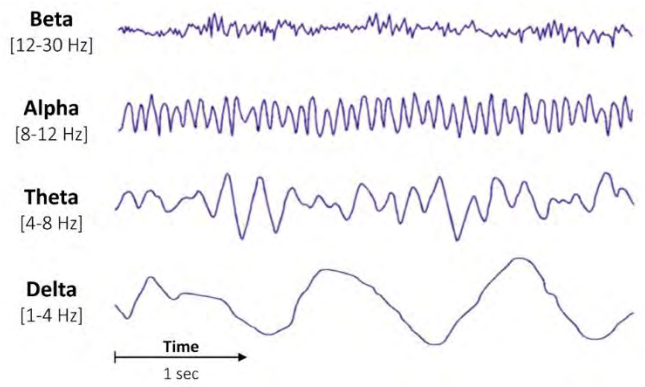


Figure 19:EEG frequency bands (source: https://raphaelvallat.com/bandpower.html)

3.3 Bio-signals: PPG

A photoplethysmogram (PPG) is an optically obtained plethysmogram that can be used to detect changes in blood volume in the tissue microvascular system. A PPG is often obtained by using a pulse oximeter that illuminates the skin and measures changes in light absorption. A conventional pulse oximeter monitors the perfusion of blood to the dermis and the subcutaneous tissue of the skin.

With every heart cycle, the heart pumps blood through the body. Although this pressure pulse is slightly dampened by the time it reaches the skin, it is enough to set up the arteries and arterioles in the subcutaneous tissue. The change in volume caused by the pressure pulse is detected by illuminating the skin with the light from a light-emitting diode (LED) and then measuring the amount of light transmitted or reflected to a photodiode. Each heart cycle appears as a peak, as seen in the figure below. The shape of the PPG waveform varies from individual to individual and varies with the location and manner in which the pulse oximeter is attached. A schematic representation of a typical pulse waveform is depicted below, in Figure 20:

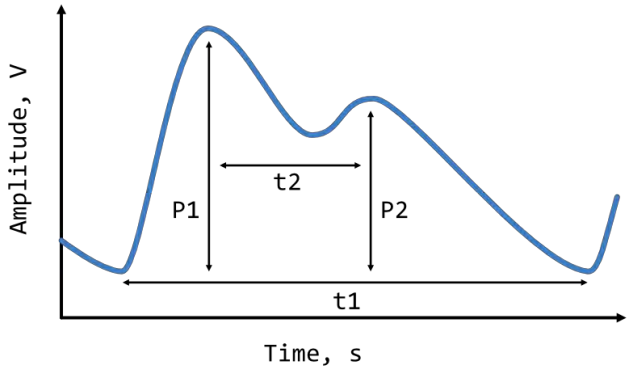


Figure 20: A typical PPG waveform and its major features

From this pulse waveform, many characteristic features can be extracted, such as the Pulse Wave Duration (t1), the amplitudes of the Systolic (P1) and Diastolic (P2) peaks or the Pulse Propagation Time (t2). The time between two subsequent systolic peaks (not depicted above), or the Peak-to-Peak Interval (PPI), is the equivalent of the Inter Beat Interval (IBI), which can be measured using the electrocardiogram (ECG) method. Therefore, the same algorithms which are usually used for the ECG signal analysis, can be applied to the PPG signal to estimate heart rate and heart rate variability (HRV) features.

Heart rate variability (HRV) is the physiological phenomenon of variation in the time interval between heartbeats. It is measured by the variation in time between heartbeats, also known as the beat-to-beat interval (R-R). One of the most popular HRV calculation methods is the Root Mean Square of successive differences in the beat-to-beat interval (RMSSD). It is a measure of how much variation there is in the heart rate. In a healthy heart, there is a natural variation, which is due to a balance between the sympathetic nervous system (SNS) and the parasympathetic parts (PSNS) of the autonomic nervous system. When your body is under stress, the sympathetic system is activated to prepare you for fighting or flight behavior and your heart rate will increase.

The parasympathetic control of your body. Rest and digestion are associated with recovery. Parasympathetic activation saves energy, narrows heart vessels, aids digestion, and slows your heart rate. These two parts of the nervous system are normally in a healthy balance, creating a natural variation in the heart. If this balance is disturbed for any reason, this variance changes. A lower HRV is associated with stress.

All HR(V) related measures are computed from a time window of 60 seconds. This period is long enough to give robust indicators of short-term changes in HR and HRV, and short enough to distinguish the changes in stress that occur during the task. This is also a sufficient frequency (once per minute) to update the feedback to the users to be able to respond to these changes quickly enough, and not too frequent to not react unnecessarily to insignificant momentary changes in physiological measurements.

3.4 Additional Dashboard views

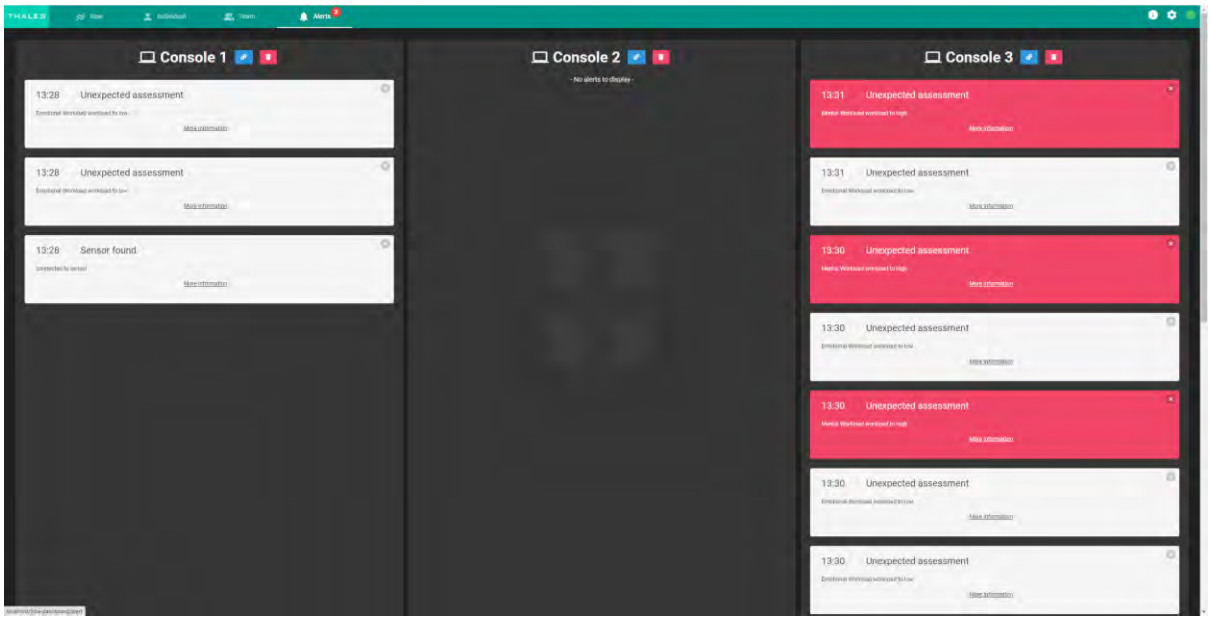


Figure 21: Alerts

List of Possible alerts

'001' : "no sensor found",

'002': "sensor found",

'010': "bad signal",

'011': "average signal",

'012': "good signal",

'060': "unexpected assessment",

'061': "expected assessment

3.4.1 Settings

There are several global settings in the Dashboard, which are user configurable. These settings are accessible upon clicking on the cogwheel icon, in the upper right corner of the interface.

Figure 22: Settings menu icon

The settings open in a pop-up window and are divided over a number of tabs. These are shown in the figures below.

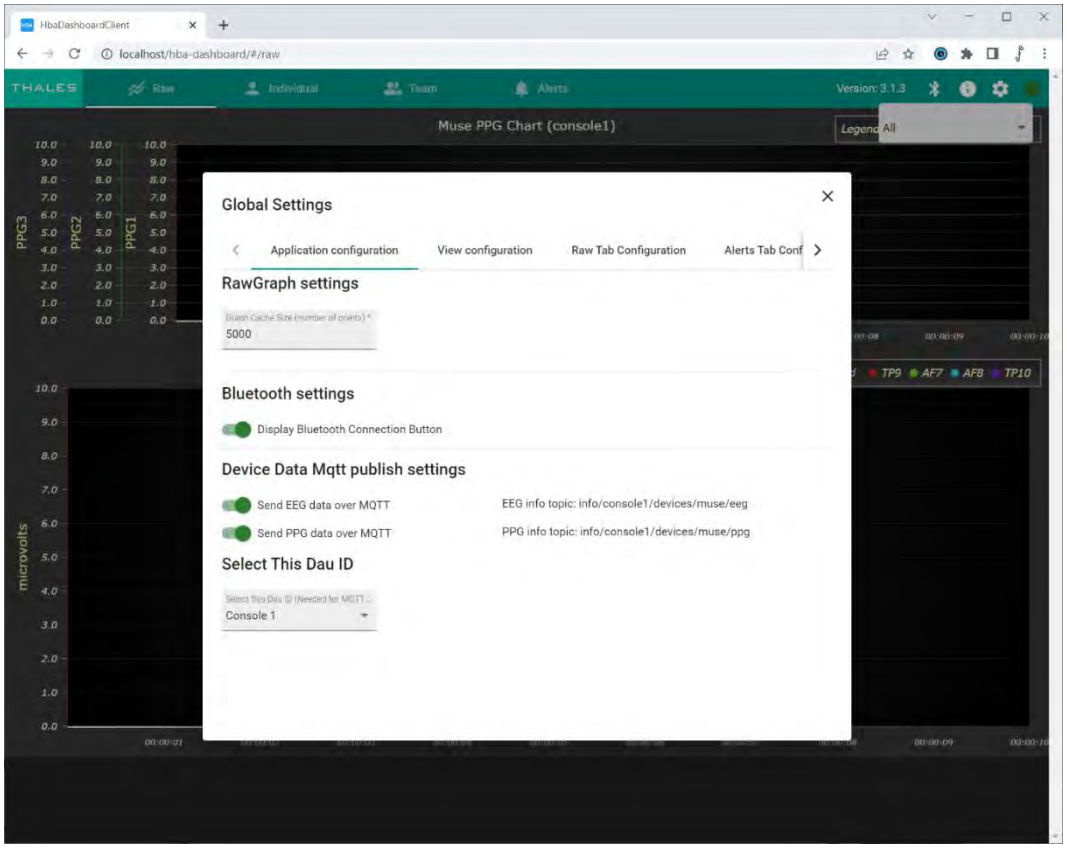


Figure 23: The dashboard global settings - application configuration

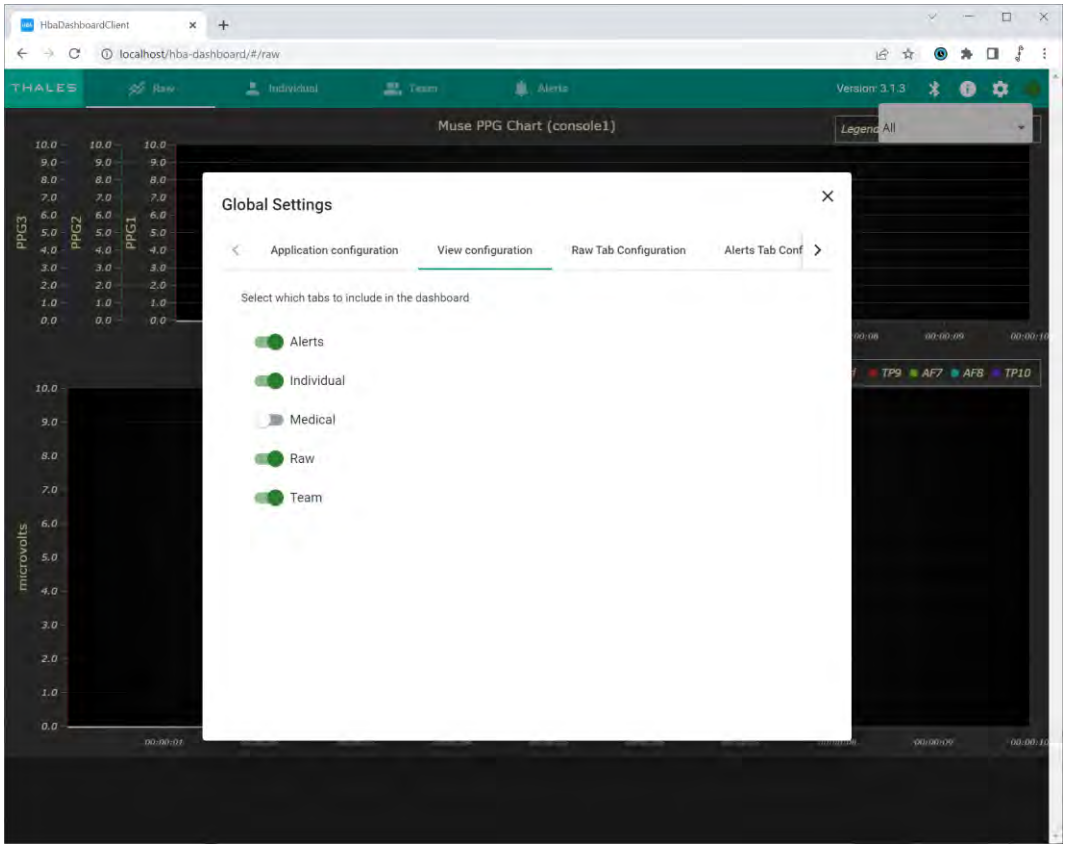


Figure 24: The dashboard global settings – view configuration

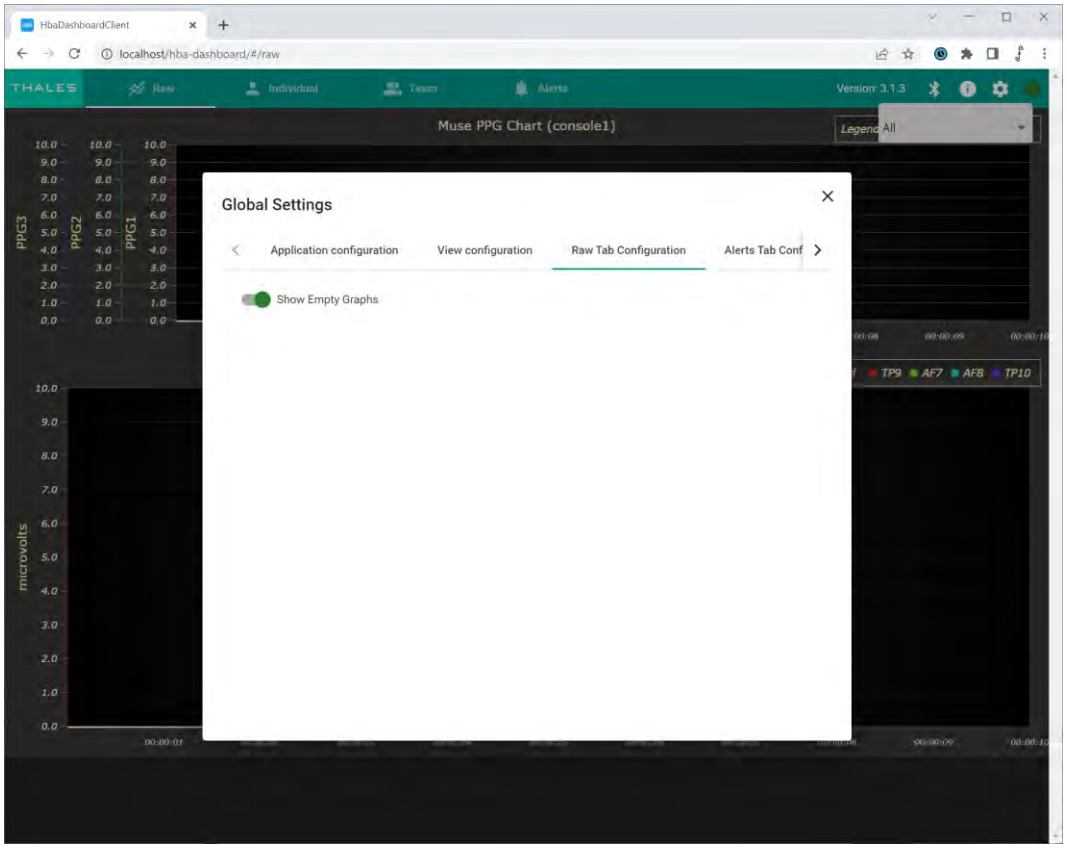


Figure 25: The dashboard global settings – Raw data configuration

3.5 The basic technical architecture

The following schematic depicts the steps and the data flow in the Workload Monitoring System.

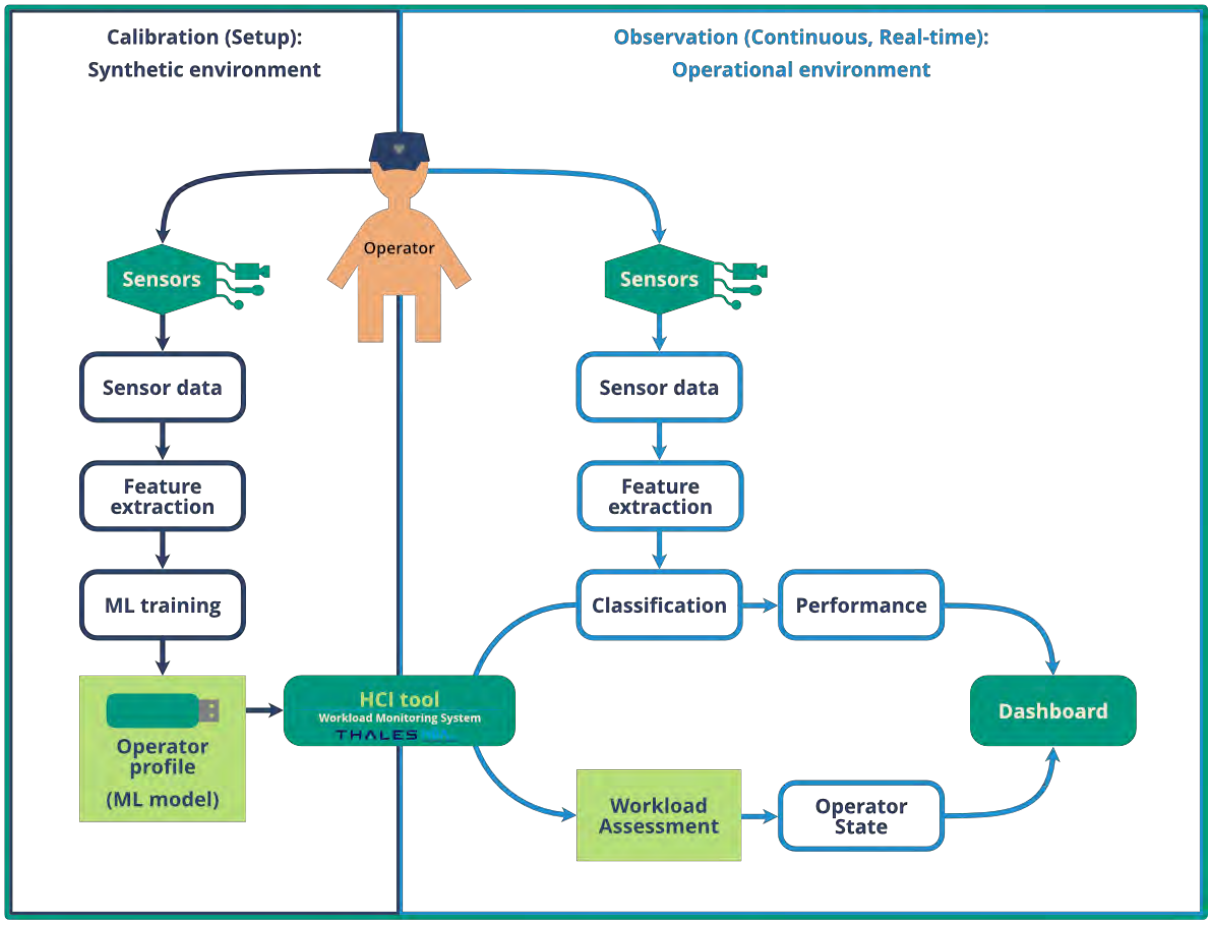


Figure 26: Data flow in the Workload Monitoring System

Explanation of the schematic: There are two steps in the process of use of the Workload Monitoring System tool. The first is shown on the left (depicted in darker shade of blue). It is the setup and calibration phase. Here, the operator performs a series of tasks, using a synthetic environment. At the same time, the sensors are used to collect the bio-data. This data is then processed, the relevant features are extracted from it and, based on these, a Machine Learning (ML) model is trained. This results in a personalized profile, able to interpret the biodata from the specific person (operator) whose bio-data has been collected. This ML model is stored on an portable data storage device (USB stick), along with any user preferences, which may be relevant during the use of the tool.

The second step in the process is the continuous use of the tool, during normal operations. Here, the sensors collect the bio-data, from the operators (while they perform their daily tasks). This data is processed by the tool, in a similar way as it has been in the first step. This processing happens in near-real time, which is referred to as ”online operation”. The same relevant features are extracted from the data. Then, the personalized operator profile is used to perform the classification. The classification of the operator’s workload, is visualized in the Dashboard. There, the current and recent operator workload is shown, along the classification performance. Together, these give an insight in the operator’s state and the tool’s operation.

3.6 Bio-signal Features

The following section lists and describes the features extracted from the bio-signals.

3.6.1 EEG features

3.6.1.1 Frequency domain features

1. Frequency band power for selected bands (Theta, Delta, Alpha, Beta, Gamma) and ratios (Beta over Alpha, Theta over Beta)
2. Phase Lag Index (PLI) computed for each electrode pair

3.6.2 Heartrate features

1. HR_mean: the mean heartrate per minute within the window

3.6.2.1 Time domain HRV features:

2. RMSSD: The square root of the mean of the sum of successive differences between adjacent RR intervals. It is equivalent (although on another scale) to SD1, and therefore it is redundant to report correlations with both (Ciccone, 2017).
3. MeanNN: The mean of the RR intervals.
4. SDNN: The standard deviation of the RR intervals.
5. SDSD: The standard deviation of the successive differences between RR intervals.
6. CVNN: The standard deviation of the RR intervals (SDNN) divided by the mean of the RR intervals (MeanNN).
7. CVSD: The root mean square of the sum of successive differences (RMSSD) divided by the mean of the RR intervals (MeanNN).
8. MedianNN: The median of the absolute values of the successive differences between RR intervals.
9. MadNN: The median absolute deviation of the RR intervals.
10. HCVNN: The median absolute deviation of the RR intervals (MadNN) divided by the median of the absolute differences of their successive differences (MedianNN).
11. IQRNN: The interquartile range (IQR) of the RR intervals.
12. pNN50: The proportion of RR intervals greater than 50ms, out of the total number of RR intervals.
13. pNN20: The proportion of RR intervals greater than 20ms, out of the total number of RR intervals.
14. TINN: A geometrical parameter of the HRV, or more specifically, the baseline width of the RR intervals distribution obtained by triangular interpolation, where the error of least squares determines the triangle. It is an approximation of the RR interval distribution.
15. HTI: The HRV triangular index, measuring the total number of RR intervals divided by the height of the RR intervals histogram.

3.6.2.2 Contains frequency domain HRV features:

16. ULF: The spectral power density pertaining to ultra-low frequency band i.e., .0 to .0033 H by default.
17. VLF: The spectral power density pertaining to very low frequency band i.e., .0033 to .04 Hz by default.
18. LF: The spectral power density pertaining to low frequency band i.e., .04 to .15 Hz by default.
19. HF: The spectral power density pertaining to high frequency band i.e., .15 to .4 Hz by default.
20. VHF: The variability, or signal power, in very high frequency i.e., .4 to .5 Hz by default.
21. LFHF:
22. LFn: The normalized low frequency, obtained by dividing the low frequency power by the total power.
23. HFn: The normalized high frequency, obtained by dividing the low frequency power by the total power.
24. LnHF: The log transformed HF.

3.6.2.3 Contains non-linear HRV features:

Characteristics of the Poincaré Plot Geometry:

25. SD1: is a measure of the spread of RR intervals on the Poincaré plot perpendicular to the line of identity. It is an index of short-term RR interval fluctuations, i.e., beat-to-beat variability. It is equivalent (although on another scale) to RMSSD, and therefore it is redundant to report correlations with both (Ciccone, 2017).
26. SD2: is a measure of the spread of RR intervals on the Poincaré plot along the line of identity. It is an index of long-term RR interval fluctuations.
27. SD1SD2: the ratio between short- and long-term fluctuations of the RR intervals (SD1 divided by SD2).
28. S: Area of ellipse described by SD1 and SD2 (pi * SD1 * SD2). It is proportional to SD1SD2.
29. CSI: The Cardiac Sympathetic Index (Toichi, 1997), calculated by dividing the longitudinal variability of the Poincaré plot (4*SD2) by its transverse variability (4*SD1).
30. CVI: The Cardiac Vagal Index (Toichi, 1997), equal to the logarithm of the product of longitudinal (``4*SD2``) and transverse variability (4*SD1).
31. CSI_Modified: The modified CSI (Jeppesen, 2014) obtained by dividing the square of the longitudinal variability by its transverse variability.

3.6.2.4 Indices of Heart Rate Asymmetry (HRA), i.e., asymmetry of the Poincaré plot (Yan, 2017):

32. **GI**: Guzik's Index, defined as the distance of points above line of identity (LI) to LI divided by the distance of all points in Poincaré plot to LI except those that are located on LI.
33. **SI**: Slope Index, defined as the phase angle of points above LI divided by the phase angle of all points in Poincaré plot except those that are located on LI.
34. **AI**: Area Index, defined as the cumulative area of the sectors corresponding to the points that are located above LI divided by the cumulative area of sectors corresponding to all points in the Poincaré plot except those that are located on LI.
35. **PI**: Porta's Index, defined as the number of points below LI divided by the total number of points in Poincaré plot except those that are located on LI.
36. **SD1d** and
37. **SD1a**: short-term variance of contributions of decelerations (prolongations of RR intervals) and accelerations (shortenings of RR intervals), respectively (Piskorski, 2011).
38. **C1d** and
39. **C1a**: the contributions of heart rate decelerations and accelerations to short-term HRV, respectively (Piskorski, 2011).
40. **SD2d** and
41. **SD2a**: long-term variance of contributions of decelerations (prolongations of RR intervals) and accelerations (shortenings of RR intervals), respectively (Piskorski, 2011).
42. **C2d** and
43. **C2a**: the contributions of heart rate decelerations and accelerations to long-term HRV, respectively (Piskorski, 2011).
44. **SDNNd** and
45. **SDNNa**: total variance of contributions of decelerations (prolongations of RR intervals) and accelerations (shortenings of RR intervals), respectively (Piskorski, 2011).
46. **Cd** and
47. **Ca**: the total contributions of heart rate decelerations and accelerations to HRV.

3.6.2.5 Indices of Heart Rate Fragmentation (Costa, 2017):

48. **PIP**: Percentage of inflection points of the RR intervals series.
49. **IALS**: Inverse of the average length of the acceleration/deceleration segments.
50. **PSS**: Percentage of short segments.
51. **PAS**: IPercentage of NN intervals in alternation segments.

3.6.2.6 Indices of Complexity:

52. **ApEn**: The approximate entropy measure of HRV.
53. **SampEn**: The sample entropy measure of HRV.

3.7 FAQ

This section contains the frequently asked questions.