NGSOTI Deliverable D2.1 - requirement analysis

The NGSOTI (Next Generation Security Operator Training Infrastructure) requirement analysis document serves as a critical component in the software development life cycle. It outlines the specifications and expectations for the NGSOTI platform. Those requirements are elaborated in this deliverable.

Classification

- TLP:CLEAR

Document history

- Version 1.0

Introduction

This section describes a brief overview of NGSOTI project and the platform that is build. It also includes the purpose and scope of the document and identifies the stakeholders.

Background

This section describes the contextual information about NGSOTI. It shows relevant history and the interoperability with existing systems.

The project aims to establish an open-source infrastructure tailored for the practical training of Security Operation Center (SOC) operators, with a specific focus on network-related alerts. This training initiative is designed to cover critical areas that are essential for SOC operators to effectively navigate the challenges of cybersecurity. The key training domains include:

- Incident Response: SOC operators will gain proficiency in incident response methodologies, encompassing incident triage, investigation, containment, and remediation.

- Log Management and Analysis: Emphasis will be placed on acquiring knowledge in log collection, analysis, correlation, and the utilization of various log analysis tools.

- Security Operations Center Management: SOC operators will familiarize themselves with SOC processes and procedures, including incident response and threat hunting techniques.
- Communication and Documentation: Good communication and documentation skills, crucial for conveying technical information to non-technical stakeholders and maintaining accurate incident records, will be emphasized.
- Cyber Threat Intelligence: Training will cover understanding various types of cyber attacks, their tactics, techniques, and procedures. It will also involve practicing knowledge management procedures and workflows, including gathering, analyzing, and interpreting facts. The goal is to enable operators to report findings to intelligence consumers and decision-makers, staying ahead of emerging threats and ongoing campaigns.

The significance of training SOC operators using real data is underscored by several benefits over artificial mock-up data:

- Relevance: Exposing trainees to real data allows them to immediately grasp the practical applications of concepts, enhancing their understanding and retention of information.
- Realistic Scenarios and Playbooks: Training with real data provides hands-on experience in dealing with realistic scenarios encountered in a real-world SOC environment, fostering critical thinking and problem-solving skills.
- Familiarity with Real Tools and Workflows: Working with real data familiarizes trainees with actual tools and workflows used in the SOC, instilling confidence in their ability to work effectively with these tools and processes.
- Practice with Real Threats: Training on real data allows trainees to practice identifying and mitigating real-world security threats, aiding in the development of their threat hunting skills.

The project incorporates a range of open-source technologies and tools to build the training platform, leveraging the expertise of consortium partners. Notable components include the D4 Distributed Denial of Services Detection Devices, AIL-project for data leak monitoring, Suricata and Zeek for network intrusion detection, OpenSearch and GrayLog for SIEM, VAST for security data analysis, and the ticketing system RTIR. Additionally, the Cyber Threat Intelligence (CTI) platform MISP and open-source monitoring tools like OpenNMS will play integral roles in the comprehensive SOC training infrastructure. The inclusion of low-interaction honeypots, aligned with initiatives like the honeynet project, further enriches the training environment.

Overall, this initiative strives to create a cutting-edge, open-source SOC training platform, equipping operators with the skills and experience needed to navigate real-world cybersecurity challenges. The identified stakeholders are:

- Engineering students
- Bachelor and master students
- IT security professionals
- Threat analysts
- Open source enthusiasts

Business Objectives

This section includes a clear statement of the business goals and objectives. It also includes identification of key performance indicators (KPIs).

- Development of Next-Generation SOC Operators:
  - Requirement: The system should facilitate the development and training of a new generation of SOC operators.
  - Justification: Address the need for operators equipped to handle evolving challenges including complex vulnerability disclosure processes, AI-related abuse or AI-related positive usage.
- Improvement of Information Sharing Practices:
  - Requirements: Collaboration among various entities is a common practice in incident response and SOC (Security Operations Center) operations.
  - Justification: To address the challenges of information sharing, this involves providing best practices and innovative solutions that facilitate sharing and collaboration both within a SOC and externally. This includes managing data collected from feed providers and other CSIRTs (Computer Security Incident Response Teams).
- Enhancement of Training Tools and Data Feeds:
  - Requirement: The platform must provide advanced toolsets and data feeds to support SOC operators in their training.
  - Justification: Acknowledge the existing gap where potential operators lack good toolsets or data feeds, aiming to bridge this gap and ensure comprehensive training resources.
- Integration with Academic Curricula:
  - Requirement: The system should seamlessly integrate with academic curricula, providing support for students pursuing cybersecurity education in universities.
  - Justification: Recognize the alternative path to SOC careers through academic curricula, ensuring alignment with educational programs and creating a cohesive learning experience.
- Incorporation of Real Operational Infrastructure:

  - Requirement: The proposal should focus on establishing a real operational infrastructure for training SOC operators.
  - Justification: Highlight the importance of hands-on experience in a real SOC environment, emphasizing the need for practical training using actual operational infrastructure usage (such as interoperability between toolsets, processes and practices).
- Dynamic Training Techniques:
  - Requirement: The platform must support dynamic training techniques such as cyber ranges driven by playbooks in the different field required by SOCs (e.g. digital forensic analysis, threat intelligence processes).
  - Justification: Recognize the effectiveness of cyber ranges and playbooks in training SOC operators, emphasizing the need for dynamic and evolving training methodologies.
- Inclusion of Real Data in Training:
  - Requirement: The proposal should emphasize the inclusion of real data in training scenarios.
  - Justification: Acknowledge the limitations of cyber ranges with synthetic data and stress the importance of exposing SOC operators to real-world data and scenarios for more effective training.
- Collaboration with Industry Experts:
  - Requirement: The system should facilitate collaboration with leading industry experts such as as guest lecturers, open source contributors or private organisations operating SOC/CSIRTs.
  - Justification: Recognize the value of industry insights in shaping the curriculum, ensuring that SOC operators are exposed to real-world experiences and challenges.

These business requirements collectively address the objectives outlined in the NGSOTI project and serve as a foundation for the development and implementation of the proposed SOC training platform. The deliverables will be open sourced and accessible to the community at large.

The Key Performance Indicators are listed below

KPI Number	Description	Value
1	How many and how market-ready innovative cybersecurity solutions have been adopted.	3
2	Number of open-source solutions benefited from this action	6
3	Maturity analysis pre and post implementation to measure the change in cybersecurity capacity of the beneficiary(ies).	2

4	Number of SME and public institutions trainings performed	3

User Requirements

This section includes a detailed description of the needs and expectations of end-users. It includes the profiles of personas, use cases and scenarios (such as the ones described by the MISP users).

The NGSOTI platform is primarily designed to cater to the needs of several specific groups of individuals, including:

- Engineering Students: The platform is intended to be a resource for students pursuing engineering degrees. This could include those studying various engineering disciplines such as computer science, electrical engineering, or related fields.
- Bachelor and Master Students: The platform is also targeted at students pursuing both bachelor’s and master’s degrees. This suggests that the content or tools provided by the NGSOTI platform are relevant and beneficial for individuals at different academic levels within the university system.
- IT Security Professionals: The platform is designed to be useful for professionals working in the field of Information Technology (IT) security. This includes individuals with job roles specifically focused on securing digital systems and networks.
- Threat Analysts: The platform caters to the needs of threat analysts. Threat analysts are professionals who analyze and assess potential cybersecurity threats, helping organizations understand and mitigate risks.
- Open Source Enthusiasts: The platform is inclusive of individuals who have an enthusiasm for open-source technologies. This suggests that the NGSOTI platform may leverage or promote the use of open-source tools, software, or methodologies.

In summary, the NGSOTI platform aims to serve a diverse audience, including students in engineering programs, IT security professionals, threat analysts, and those with an interest in open-source technologies. This approach suggests a broad focus on education, skill development, and community engagement within the realm of cybersecurity and threat intelligence. The purpose of this wide audience is to maximize market adoption (see KPI section)

Functional Requirements

This section includes a detailed description of the system’s functionality. It include use cases and scenarios for system operations and functional specifications.

The constraints section outlines essential limitations and restrictions that impact the project's development and sustainability. To ensure the platform's viability beyond the conclusion of EU co-funding, several key constraints have been identified.

Firstly, a critical constraint is the sustainability requirement. The platform must be designed and managed in a way that ensures its continued functionality and relevance even after the conclusion of EU co-funding. This necessitates the development of strategies for ongoing maintenance, updates, and potential future enhancements without relying solely on external financial support.

The timeline constraint stipulates that the platform's setup is scheduled to occur between Q1 2024 and Q3 2024. This timeframe sets clear boundaries for the project's initiation and completion, which includes limitations on the available development time.

Another significant constraint is the prohibition of commercially licensed software for core components on the platform. The use of such licenses would introduce additional costs, compounding the hosting expenses. By opting for open source alternatives, the project aims to control expenses and promote financial sustainability for SOC training facilities. This constraint aligns with the overarching goal of maintaining the platform's affordability and accessibility while avoiding ongoing financial dependencies.

In summary, the identified constraints emphasize the importance of long-term sustainability, adhere to a defined timeline for project execution, and strictly prohibit the incorporation of commercially proprietary licensed software to control costs and enhance financial autonomy. These constraints play a crucial role in shaping the project's strategy and execution on a long-term vision.

Non-functional Requirements

The performance of the platform is crucial to meet the demands of various scenarios. This section encompasses key performance requirements, including the maximum number of users the platform should support and specific usability requirements. Moreover, stringent security measures are imperative to safeguard against potential threats. Compatibility requirements ensure seamless integration with other systems, fostering interoperability.

Furthermore, the platform is designed to be versatile, accommodating new use cases and adapting to best practices identified by the core working group. This adaptability extends to interoperability with existing systems and industry solutions disclosed within the core working group's findings.

In adherence to regulatory and compliance standards, the platform's components are mandated to be open source, governed by a compatible EUPL (European Union Public License)

license. This commitment to open source principles not only promotes transparency but also facilitates collaboration within the community. Additionally, compliance with pertinent European Union regulations, such as the General Data Protection Regulation (GDPR) and the Network and Information Systems (NIS) Directive, is integral. By aligning with these regulations, the platform ensures a robust framework for data protection and cybersecurity, meeting the legal and regulatory expectations set forth by the EU.

System Features

This section includes the specific features or capabilities the system must have. It also describes the feature prioritization.

Based on the provided text, here are some system requirements for the Next Generation Security Operator Training

- Real Data Integration:
    - Requirement: The system should integrate real data, obtained from sources such as the D4 Distributed Denial of Services Detection Devices (https://d4-project.org), to provide practical training scenarios for SOC personnel.
    - Justification: Real data exposure enables trainees to understand and apply concepts, improving the effectiveness of training.
- Low-Interaction Honeypots:
    - Requirement: The system should incorporate low-interaction honeypots, leveraging open-source projects such as the honeynet project (https://www.honeynet.org/), dataplane, shadowserver.org or similar initiatives, to simulate services exposed to the Internet for training.
    - Justification: Simulating realistic scenarios aids in hands-on experience and skill development.
- Data Leak Monitoring Framework:
    - Requirement: The system must utilize the AIL-project (Analysis Information Leak https://ail-project.org) framework for monitoring the Internet and darknet to discover data leaks.
    - Justification: AIL-project provides capabilities to collect, carve, and analyze real unstructured data, enhancing the breadth of scenarios for trainees.
- Network Intrusion Detection Systems (NIDS):
    - Requirement: The system should incorporate open-source NIDS tools such as Suricata (https://suricata.io/) or Zeek (https://zeek.org/) to generate a constant flow of alerts for SOC operators during training.

  - Justification: Realistic models of alert handling prepare SOC operators for actual work scenarios.
- SIEM Technologies:
  - Requirement: The system must use open-source SIEM technologies such as OpenSearch or GrayLog for storing logs, creating alerts, and providing an easily transposable concept for trainees.
  - Justification: SIEM technologies play a crucial role in centralizing and analyzing security-related data.
- VAST/Tenzir for Security Data Analysis:
  - Requirement: The system should utilize the open-source tool VAST (https://vast.io) in conjunction with SIEM for security data analysis, model training, and detection engineering.
  - Justification: VAST/Tenzir enhances the system's analytical capabilities and reflects the state of the art in SOC operator craft.
- Ticketing System for Alerts:
  - Requirement: The system must integrate the ticketing system or case management such as RTIR (https://bestpractical.com/rtir) for centralizing alerts and managing ongoing investigations by trainees.
  - Justification: A centralized system for managing alerts ensures organized and effective handling the communication aspects of incidents.
- Case management for Task assigment and Case handling:
  - Requirement: The system must integrate the case management for centralizing the management (https://github.com/flowintel) of cases performed by trainees.
  - Justification: A centralized tracking the cases to efficiently handle incidents is critical for an operational SOC.
- Cyber Threat Intelligence (CTI) Platform:
  - Requirement: The system should incorporate the MISP platform (https://www.misp-project.org/) as the Cyber Threat Intelligence (CTI) platform for information sharing internally but also externally with other partners.
  - Justification: MISP provides essential capabilities for contextualization, reporting, and interconnection with other tools, enhancing the intelligence-sharing aspect.
- Monitoring Tools:
  - Requirement: The system should use open-source monitoring tools such as OpenNMS (https://www.opennms.com/) to monitor infrastructure responsiveness.
  - Justification: Monitoring tools contribute to network monitoring best practices and aid in problem remediation.
- Flexibility for Tool Integration:

- Requirement: The system must be designed to easily integrate additional tools from the MeliCERTes tooling-set, allowing for the extension of NGSOTI with new open-source tooling.
- Justification: Ensures adaptability and future-proofing of the training infrastructure.

These system requirements collectively define the technical specifications necessary for the successful implementation and operation of the NGSOTI platform.

Constraints

Ensuring the long-term sustainability of the platform is imperative, extending beyond the conclusion of EU co-funding. The platform is scheduled for setup from Q1 in 2024 to Q3 2024, emphasizing a timely and efficient development timeline. A critical constraint is the avoidance of commercially proprietary licensed software for core components, as this would introduce additional costs alongside hosting expenses, undermining financial sustainability.

Furthermore, a non-negotiable constraint is the imperative alignment with existing curricula of schools and their planning. This necessitates careful integration with educational frameworks, ensuring that the platform complements and supports the ongoing academic activities of the targeted institutions. This alignment is essential for the platform’s successful adoption and integration into educational systems.

Dependencies

This section summarizes the external factors or components that the project depends on.

The primary external dependency is closely tied to the alignment with school schedules, curriculum timelines, and the availability of students considering their constraints, including commitments to master’s programs and schedules for thesis defense, among other factors. Additionally, the success of the project relies on the availability of other NGSOTI stakeholders involved in the training process.

Risks and mitigation

This section summarizes the identification of potential risks to the project and strategies for mitigation.

The risks and their corresponding mitigation strategies are listed in the table below:

Risk Number	Description
1	[LOW] Unavailable key personnel: Key personnel. Contributions in this project will be open source. Hence, in case of leave, sickness, or work contract termination, other personnel could take over.
2	[MEDIUM] Software licenses incompatibility: NGSOTI relies on open source software, some of them being quite mature. The software license of MISP, for instance, has not changed for 10 years. However, other open source components or dependencies such as OpenSearch might change and render the usage or interoperability impossible. CIRCL will clone the third-party software used by NGSOTI and do an evaluation of the last released compatible open source software that could be maintained, and decide whether it should be substituted with an alternative piece of software.
3	[HIGH] Procurement delays; Hardware equipment cannot be delivered at estimated time period. Make quick order decisions to order hardware at an early stage, negotiate fixed prices for hardware (to counter price fluctuations). Try to avoid large amounts of hardware to avoid the necessity of national/european procurement legislation processes. In case of long delays, plan to have already existing hardware on site to deploy (recycle existing hardware if possible until new hardware is delivered).
4	[LOW] Illegal data processing: Data processing procedures in place and regular data control. Proposed Mitigation Measures are put in place.

Approach

This section describes the methodology and approach to be used for development.

The development of open-source components within the project will notably encompass the creation of the “Programming Methodology Framework,” commonly known as PMF. This framework is envisioned to be a comprehensive and versatile set of tools and guidelines designed to enhance programming methodologies. By adopting an open-source approach, PMF seeks to promote collaboration, transparency, and community engagement, allowing developers and organizations to leverage and contribute to its evolution. PMF is anticipated to provide a structured and adaptable framework that facilitates efficient programming practices. Its development aligns with the principles of openness, enabling users to access, modify, and distribute the framework freely. The goal is to foster a community-driven ecosystem where diverse perspectives contribute to the continuous improvement of programming methodologies. This open-source initiative reflects a commitment to innovation, knowledge sharing, and the collective advancement of programming practices. Developers, educators, and organizations are encouraged to explore, adopt, and contribute to PMF, thereby promoting a collaborative and dynamic environment within the broader programming community.