



Grant agreement number: 883273
Start date of project: 01/10/2020
Revision

Lead contractor: CNR
Duration: 36 months

Project funded by the European Commission within the EU Framework Programme for Research and Innovation HORIZON 2020	
Dissemination Level	
PU = Public, fully open, e.g., web	✓
CO = Confidential, restricted under conditions set out in Model Grant Agreement	
CI = Classified, information as referred to in Commission Decision 2001/844/EC.	
Int = Internal Working Document	

D2.1 – AI4HealthSec Requirements and Research Directives

Editor

Lena Griebel (KLINIK)
Haralambos Mouratidis (UOB)
Mario Ciampi (CNR)
Vasilis Tountopoulos (AEGIS)
Spyridon Papastergiou (FP)

Contributors

Jihane Najar (AEGIS)
Anca Bucur (Philips)
Stefano Silvestri (CNR)
Eftychia Lakka (FORTH)
Manos Athanatos (FORTH)
Marco Fruscione (EBIT)
Stephan Kiefer (Fraunhofer)
Gabriele Weiler (Fraunhofer)
Dmitry Amelin (Fraunhofer)
Vasilis Tountopoulos (AEGIS)
Efsthathios Karanastasis (ICCS)

Reviewers

Eleni-Maria Kalogeraki (FP)
Othonas Soultatos (STS)

VERSION	DATE	REVISION AUTHOR	SUMMARY OF CHANGES
0.1	10/03/2021	Stephan Kiefer, Gabriele Weiler, Dmitry Amelin, Marco Fruscione, Haralambos Mouratidis, Vasilis Tountopoulos, Stefano Silvestri, Manos Athanatos, Efsthios Karanastasis	Initial version
0.2	12/03/2021	Jihane Najar, Anca Buru, Eftychia Lakka, Spyridon Papastergiou	More input to chapter 4 and 5 and report for EAB
0.3	30/03/2021	Lena Griebel, Vasilis Tountopoulos	Changes regarding chapter 5 and conclusion chapter
1.0	31/03/2021	Vasilis Tountopoulos, Eftychia Lakka	Final changes

Executive Summary

At the start of the AI4HealthSec project it was necessary to elicit requirements from different perspectives. We identified three main pillars:

- 1 User challenges/business needs
- 2 Domain requirements
- 3 Technical requirements

For each pillar, methods were defined and the requirements analysis was performed. This deliverable contains description of methods and results concerning the requirements analysis. Furthermore, the basic ideas of AI4HealthSec are presented.

We found that an AI4HealthSec framework needs to work in an environment of different international and national standards such as ISO 90001. Moreover, it needs to take six challenges that express the wishes of potential users into account. 67 technical requirements have been formulated on the basis of the users' wishes, domain analysis and discussions among project partners. Those requirements also need to be the basis for the future development of an AI4HealthSec framework.

Contents

Executive Summary	4
List of acronyms	9
List of tables	11
List of figures	12
1 Introduction	13
1.1 Scope	13
1.2 Background: The AI4HealthSec Framework	13
1.3 Contribution to other work packages and tasks	16
1.4 Structure of the document	17
2 Requirements Elicitation and Analysis Methodology	17
2.1 Security requirements engineering process	17
2.2 Business Needs/User Challenges Elicitation and analysis	19
2.2.1 Objectives of the questionnaires, creation of the questionnaires	19
2.2.2 Questionnaire content and structure	19
2.2.3 Methodology for analysing the user requirements questionnaire	21
2.3 Domain requirements elicitation and analysis	21
2.4 Cybersecurity tools and systems requirements elicitation and analysis	21
2.5 EAB engagement in user requirements elicitation and analysis	23
3 Results: Business needs/user challenges elicitation and analysis	23
3.1 Internal user requirements analysis: Part A – Information on pilot sites security policies ...	24
3.2 Internal user requirements analysis: Part B- Insights into participants background and experience with cybersecurity	24
3.2.1 Vulnerable groups regarding cyber-attacks at the pilot sites	25
3.2.2 Insights: Members of pilot organizations on risk awareness, organization policies and experiences with cybersecurity topics	26
3.3 External user requirements analysis	32
3.3.1 External user requirements analysis: Information on security policies	32
3.3.2 External user requirements analysis: Vulnerable groups	33
3.3.3 Insights: Representative of external organization on risk awareness, organization policies and experiences with cybersecurity topics	34
3.4 Concerns against AI4HealthSec framework	40
3.5 Business Needs/ User Challenges	41
4 Results: Domain requirements elicitation and analysis	41
4.1 Healthcare security management standards and best practices	41
4.1.1 Security Management Standards	42
4.1.2 Health Care domain management standards	46
4.1.3 Best Practices and Guidance	54

4.2	Analysis of healthcare security domain requirements and challenges	67
4.2.1	Security of AI4HEALTHSEC Circles of consideration.....	67
4.2.2	Analysis of healthcare security domain and challenges.....	67
4.2.3	Incidents handling of healthcare security domain.....	69
4.2.4	Risk management of healthcare security domain	71
5	Results: Healthcare security management solutions relevant to AI4HealthSec	73
5.1	Evidence-driven Maritime Supply Chain Risk Assessment (MITIGATE) System	73
5.1.1	Short Description.....	73
5.1.2	Key Features.....	74
5.1.3	Component Advantages	75
5.1.4	Examples Usage Scenario	76
5.1.5	Expected extensions and potential new implementations	77
5.2	Security & Privacy Assurance Platform	78
5.2.1	Short Description.....	78
5.2.2	Key Features.....	78
5.2.3	Component Advantages	79
5.2.4	Examples Usage Scenario	80
5.2.5	Expected extensions and potential new implementations	80
5.3	Cloud - based Intrusion Detection System	80
5.3.1	Short Description.....	80
5.3.2	Key Features.....	81
5.3.3	Component Advantages	81
5.3.4	Examples Usage Scenario	81
5.3.5	Expected extensions and potential new implementations	81
5.4	Data Harmonization & Pattern Recognition	82
5.4.1	Short Description.....	82
5.4.2	Key Features.....	82
5.4.3	Component Advantages	82
5.4.4	Examples Usage Scenario	83
5.4.5	Expected extensions and potential new implementations	83
5.5	Reasoning Engine	84
5.5.1	Short Description.....	84
5.5.2	Key Features.....	84
5.5.3	Component Advantages	84
5.5.4	Examples Usage Scenario	84
5.5.5	Expected extensions and potential new implementations	85
5.6	Advanced Visualization Toolkit	85

5.6.1	Short Description.....	85
5.6.2	Key Features.....	85
5.6.3	Component Advantages	85
5.6.4	Examples Usage Scenario	86
5.6.5	Expected extensions and potential new implementations	87
5.7	Sharing Platform.....	87
5.7.1	Short Description.....	87
5.7.2	Key Features.....	87
5.7.3	Component Advantages	87
5.7.4	Examples Usage Scenario	88
5.7.5	Expected extensions and potential new implementations	88
5.8	CHIMERA.....	88
5.8.1	Short Description.....	88
5.8.2	Key Features.....	89
5.8.3	Component Advantages	90
5.8.4	Examples Usage Scenario	91
5.8.5	Expected extensions and potential new implementations	91
5.9	Asset Explorer tool	92
5.9.1	Short Description.....	92
5.9.2	Key Features.....	92
5.9.3	Component Advantages	92
5.9.4	Examples Usage Scenario	92
5.9.5	Expected extensions and potential new implementations	93
5.10	Data Sharing Management	93
5.10.1	Short Description.....	93
5.10.2	Key Features.....	93
5.10.3	Component Advantages	93
5.10.4	Examples Usage Scenario	94
5.10.5	Expected extensions and potential new implementations	95
5.11	Contribution to the AI4HealthSec Requirements Process	95
6	Results: Input by External Advisory Board	96
7	AI4HealthSec Requirements.....	97
8	Conclusions	104
9	References.....	105
10	Appendix	106
10.1	Introduction Text Questionnaire.....	106
10.1	Part A: Questionnaire Internal Organization	107

10.2	Part B Fraunhofer: Questionnaires Internal Organizations	111
10.3	Part B UoB: Questionnaire Internal Organization	126
10.4	Part B EBIT: Questionnaire Internal Organization	130
10.5	Part B Klinik: Questionnaire Internal Organization	135
10.6	Questionnaire External Organizations	140
10.7	Report for EAB.....	148

List of acronyms

AI	Artificial Intelligence
AICS	Artificial Intelligence Computer System
API	Application Programming Interface
AST	Abstract Syntax Tree
AVT	Aegis Advanced Visualisation Toolkit
CIP	Critical Infrastructure Protection
CSIRT	Computer Security Incident Response Team
CVSS	Common Vulnerability Scoring System
DCS	Distributed Control System
DDoS	Distributed-Denial-of-Service
DMAIC	Define, Measure, Analyze, Improve, Control
DoW	Description of Work
DSAF	Dynamic Situational Awareness Framework
DSPT	Data Security and Protection Toolkit
EAB	External Advisory Board
EHR	Electronic Health Record
ENISA	European Union Agency for Cybersecurity
EUDAMED	European Database for Medical Devices and in-vitro Diagnostics
FDA	Food and Drug Administration (US)
HCII	Health Care Information Infrastructures
HCSCS	Health Care Supply Chains
HIP	Health Information Protection
HIPAA	Health Insurance Portability and Accountability Act
HCSCS	Health Care Supply Chain Services
ICT	Information and Communications Technology
IDS	Intrusion Detection System
IHP	Incident Handling Process
IoCs	Indicators of Compromise
HCSC	Health Care Supply Chain

MDCG	Medical Devices Coordination Group
MDR	Medical Device Regulation
NHS	National Health Service
ORASE	Open Simulation Environment
PDCA	Plan-Do-Check-Act
PHI	Protected Health Information
PMCF	Post-Market Clinical Follow-up
PMS	Post-Market Surveillance
RAAP	Risk Analysis and Assessment Process
SDLC	Software Development Life Cycles
SI	Swarm Intelligence
SOC	Security Operations Centre
SMS	Simple Message Service
VM	Virtual Machine

List of tables

Table 1: Vulnerable groups.....	25
Table 2: Vulnerable groups external user requirements analysis	33
Table 3: Business Needs/ User Challenges.....	41
Table 4: Examples of medical devices' incidents and corresponding severity, security harm and control and safety harm and control	59
Table 5: Technical Challenges linked to Business Needs	98

List of figures

Figure 1. AI4HEALTHSEC Circles of Consideration	14
Figure 2. Main Aspects and Principles of the AI4HEALTHSEC framework	15
Figure 3: The process for collecting user requirements in the AI4HealthSec project.	18
Figure 4: The overall conceptual elements of the AI4HealthSec framework	22
Figure 5: Self-assessed knowledge on cybersecurity topics	26
Figure 6: Encountered cybersecurity incidents over the past 3 years	27
Figure 7: Knowledge what to do in case of cybersecurity incident	28
Figure 8: Personal involvement in risk management process.....	28
Figure 9: Personal involvement in cyber-security tasks.....	29
Figure 10: Policy allows to improve situational awareness concerning cybersecurity	30
Figure 11: Opinion of organization's security officers towards engagement in CIP program	30
Figure 12: Invisible vs. visible framework	31
Figure 13: Interaction with external cybersecurity framework	31
Figure 14: External_self-assessed knowledge on cybersecurity	34
Figure 15: External_trained on cybersecurity by own organization.....	35
Figure 16: External_encountered cybersecurity incidents	35
Figure 17: External_knew what to do in case of incident.....	36
Figure 18: External_Policy allow to improve situational awareness concerning cybersecurity	37
Figure 19: External_Opinion security officers towards CIP engagement	37
Figure 20: External_Invisible vs. visible framework.....	38
Figure 21: External_Interaction with external cybersecurity framework.....	38
Figure 22 PHD-to-gateway Communication Model.....	48
Figure 23: ENISA cloud architecture model for medical device (extracted from [3])	65
Figure 24: Incident Response Life Cycle [8]	69
Figure 25: Lifecycle of CSIRT [10]	71
Figure 26: Evidence-driven Maritime Supply Chain Risk Assessment (MITIGATE) System.....	74
Figure 27: The high-level architecture of the Security & Privacy Assurance Platform.....	79
Figure 28 - Chimera Component Diagram.....	89
Figure 29: Normal workflow of example usage scenarios	91
Figure 1. AI4HEALTHSEC Circles of Consideration	149
Figure 2. Main Aspects and Principles of the AI4HEALTHSEC framework	151

1 Introduction

1.1 Scope

This document describes the process used for eliciting the requirements. It includes the synthesis of information obtained from the AI4HealthSec user-representative questionnaires and the approaches taken for the specification of technology solutions, enriched by requirements from literature analysis.

The first step of identifying user needs, expectations, and concerns is enriched by the description of technology solutions that will be part of the AI4HealthSec framework and by requirements excluded from the literature on cybersecurity in healthcare. For the elicitation of requirements, the pilot partners of AI4HealthSec (Fraunhofer, Ebit, UoB, KLINIK) provided representatives of typical end-users or persons in positions to enable them to give more details on the user perspective. External organizations were included as well. Those organizations were not limited to the healthcare sector, but also included other domains potentially endangered by cybersecurity attacks, such as the energy sector. Furthermore, project partners of the pilot sites gave more information on the respective organization concerning the company size, and the way of handling with cybersecurity issues. The requirements were collected by means of a questionnaire by e-mail or in bilateral interviews. The data obtained from the questionnaires were analysed and the requirements were formulated. Finally, AI4HealthSec's External Advisory Board (EAB) reviewed the requirements so that this deliverable is able to give a validated insight into the requirements towards an AI4HealthSec framework.

The document has the following scopes:

1. To provide requirements for the AI4HealthSec framework from a typical user's perspective by including potential end-users, representatives of pilot organizations and external experts
2. To provide an overview on technical requirements that should be met when designing the AI4HealthSec framework
3. To provide a domain analysis on requirements from a literature analysis.

1.2 Background: The AI4HealthSec Framework

In the digital era the healthcare ecosystem in Europe has turned into a complex mosaic, composed by large health systems and institutes, single physician practices, device developers, etc. This ecosystem can be defined as a widely distributed, interconnected set of entities (i.e., organizations, individuals or/and CIs), processes and services that relies upon interconnected ICT infrastructures, establishing a dynamic Health Care Supply Chain (HCSC). The established interconnections reflect the relationships that exist between the involved entities.

In this context, these HCSCs are characterized by a high degree of complexity and interconnectivity of the ICT systems. As depicted in Figure 1, the health care ecosystem can be represented as being composed by four **circles of consideration** that puts the patient at the centre of attention. The **first inner circle**, our starting point, includes health components that are very close to the user (e.g., implants, sensors). The **second circle** encapsulates the previous one as well as all the medical equipment and devices (e.g., pathology scanners and servers) used in health institutes. The **third circle** encloses the two previous ones and incorporates the **individual Health Care Information Infrastructures (HCIIIs)**. Finally, the **fourth and outer circle** contains all the above circles and represents the **interdependent HCIIIs** composing the whole health ecosystem, including the supporting **Health Care Supply Chain Services (HCSCS)**.

However, the evolving digital interconnectivity of medical ICT systems has also changed the threat landscape, as the digitalization of patient data is attracting more attention from cybercriminals, producing a wide range of security and privacy challenges and increasing the danger of potential cybersecurity attacks in Healthcare Infrastructures. Thus, there is an urgent need to ensure that these identified four distinct areas of consideration are all properly secured. However, despite the fact that these areas have their own unique characteristics, they are not independent from each other. Inner circles can be seen as the building blocks of the external ones, meaning that the security of the external circles is directly affected by the inner ones. Thus, the security of the **interdependent HCIIIs** and the **HCSCS**, is directly affected by the security of the **individual HCIIIs** that compose it. However, it should be noted that the overall system is not secured by simply securing its “building blocks”. There are interdependencies between the different layers that have their own specificities and require cross layer coordination.

AI4HEALTHSEC’s aim is to enhance the security and resilience of the modern digital healthcare ecosystems and the provided medical supply chain services through the provision of a novel **Artificial Intelligence Dynamic Situational Awareness Framework (DSAF)**. The main goal of the proposed approach is to improve, intensify and coordinate the overall security efforts for the effective and efficient identification, evaluation, investigation and mitigation of realistic risks, threats, and multi-dimensional attacks within the cyber assets in the four distinct **areas of consideration** (Figure 1). The proposed approach seeks to support, prepare and help the **Interdependent HCIIIs** participating in different types of **HCSCS** to: (i) thoroughly assess the vulnerabilities of all cyber assets; (ii) continuously forecast and evaluate the probability of cyber-attacks; (iii) access/receive warnings for

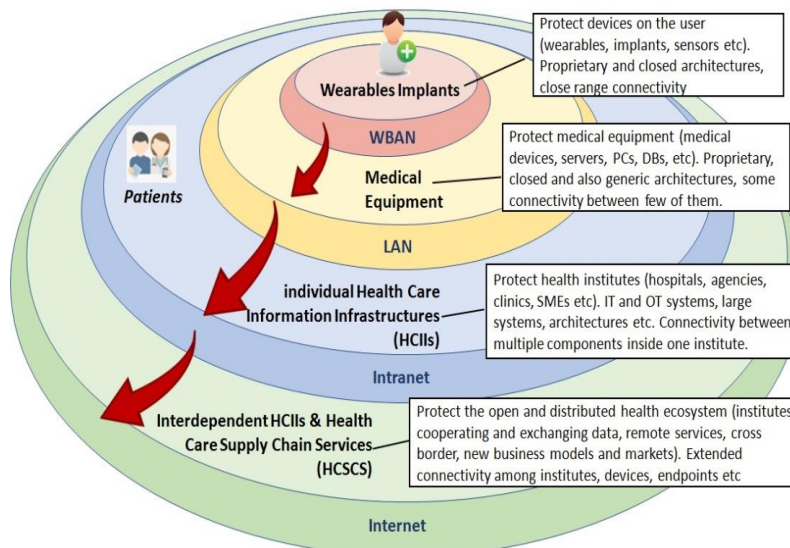


Figure 1. AI4HEALTHSEC Circles of Consideration

upcoming attacks and vulnerabilities; (iv) see the continuum between indicators of compromise, advanced persistent threats, cyber alerts and adversaries (v) easily recreate, visualize and forecast propagation and cascading effects of attacks in their **Interdependent HCIs** and anticipate how these attacks propagate across the **HCSCS**; (vi) follow a targeted step-by-step framework providing timely technical assistance and guidance on investigating and handling complex, interrelated cyber security incidents and data breaches and extracting all relevant information; (vii) combine and analyse all security incident-related information and proofs in an effective and accurate manner; and (viii) receive guidelines and, share information and warnings with all **HCIs**.

In order for **DSAF** to meet its objectives, it consists of consists of 7 main conceptual layers, 4 horizontals (“**Risk and Privacy management & Cyber-Attack Forecasting**”, “**Incident Identification**”, “**Security Events Evaluation**” and “**Analysis and Decision-Making**”) dealing with the situational awareness process and three vertical, the “**Information Sharing & Individualised Autonomous Networking**” responsible to distribute, disseminate, self-publish, broadcast or circulate the security-related information, the “**Security & Privacy**” incorporating a set of security, privacy and data protection features and the “**Context-Rich/Analytical Exploration**” providing environment that allows the **HCIs**’ operators to have a better understanding of the cyber environment.

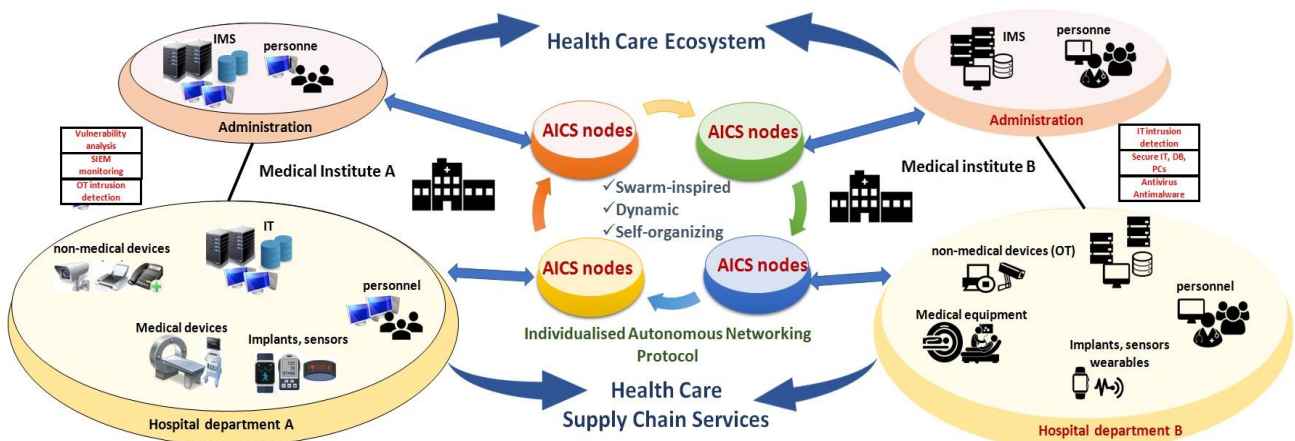


Figure 2. Main Aspects and Principles of the AI4HEALTHSEC framework

In addition, the proposed framework will be built upon a new type of Swarm Intelligence (SI), self-organizing and dynamic collaboration approach implemented through an **individualised Autonomous Networking protocol** (Figure 2) that provides autonomic deployment, cluster formulation and hierarchical communication in **HCIs**. This protocol, will connect the four circles of the health ecosystem grouping individual ICT elements, systems, and components into a population of simple or group of nodes, named **AICS nodes** (group of ICT assets or individual **HCIs**), allowing them to interact locally with one another and with their Interdependent Health Care environment. In this way, the proposed protocol will build networking infrastructures that manage the effective coordination of the **AICS nodes** of **Interdependent HCIs** by defining and leveraging the actions that should be performed by them. These agents are linked together and cooperate with each other through local interactions to achieve distributed optimization of the risk analysis and incident handling in real time. The continuous diffusion of security-related information across the network

enables the agents to optimize the evaluation and mitigation of the interdependent threats and risks as well the investigation of complex security events and data breaches.

1.3 Contribution to other work packages and tasks

This deliverable D2.1 is the result of Task T2.1 which is part of Work Package (WP) 2 “Refinement of pilot requirements, evaluation metrics and AI4HEALTHSEC Architecture”.

This WP contributes to others in the project:

It interacts with WP3 “Design of self-organized swarm intelligence framework”, WP4 “Design of dynamic cyber situational awareness system”, and WP5 “Development of dynamic situational awareness system”.

The objectives of WP2 are:

- To elicit and analyse requirements associated with the needs of the digital healthcare environments, including and other sectors as well.
- To specify the real-life pilot scenario of the project
- To entail a preliminary analysis of the legal and ethical framework applicable to AI4HealthSec
- To provide the specifications of the AI4HealthSec architecture and interfaces and delineate the implementation process to be undertaken within the project
- To identify the high-level legal and ethical requirements associated with the technological innovation of the project and
- To define the appropriate evaluation methodology and corresponding metrics for the demonstration of the unique characteristics of AI4HealthSec

the requirements will be considered in WP3 and WP4.

This deliverable D2.1 is the basis by providing the broader context an AI4HealthSec framework should take into account.

Moreover, WP2 provides input for WP6 “Pilots development of the AI4HealthSec system”.

Task 2.1 will provide the basis for T2.3 where a methodology and certain metrics (specified in the form of Key Performance Indicators) for the qualitatively and quantitatively evaluation of the identified requirements will be developed. D2.3 will detail pilot scenarios and user requirements according to the pilots.

Task 2.1 will furthermore provide user requirements as input for Task 2.4 in order to produce a set of functional and non-functional requirements provided and validated by the AI4HealthSec Health Care operators, which will describe in detail what functionalities will be implemented and how.

Requirements defined in this deliverable will have to be transferred to the technical perspective which will be presented in D2.4.

1.4 Structure of the document

This document is structured in seven main sections:

After the introduction chapter, the methods used for user requirements analysis, for the domain requirements elicitation and analysis and for the identification of cybersecurity tools and system requirements, and for the input by the EAB are described.

The third chapter contains results from the user requirements analysis. Afterwards, results from the literature analysis, for the identification of cybersecurity tools and from the validation by the EAB are included. Finally, a conclusion is drawn.

2 Requirements Elicitation and Analysis Methodology

2.1 Security requirements engineering process

The security requirements engineering process entails the way that key project objectives will be materialised into concrete expectations of the intended end users from the AI4HealthSec framework. Such users belong to teams with discrete roles in the cyber security arena, like a community emergency response team (CERT), a Security Operations Centre (SOC) and a computer security incident response team (CSIRT). The roles in these teams in an interconnected hospital environment are entitled with responsibilities to collect information about cyber-attacks, monitor and analyse potential incidents, evaluate the identified events, and eventually propose and apply actions in response to these events.

The main objectives of the AI4HealthSec project are summarized in the following lines:

- Detection and analysis of cyber-attacks and threats on Health Care Information Infrastructures (HCIIIs)
- Knowledge awareness on cyber security and privacy risks
- Reaction capabilities in case of security and privacy breaches
- Exchange of reliable and trusted incident-related information

To achieve these objectives, the project will define, develop and validate a framework that supports the implementation of two main processes, namely the Risk Assessment Process (RAP) and the Incident Handling Process (IHP). The establishment of the respective framework needs to stand on top of solid user requirements that express the expectations of the teams in the cyber security domain for support in managing cyber-attacks and minimising the impact from their existence in the HCIIIs through the implementation of relevant preventive, detective, and corrective mechanisms. To this end, this deliverable presents and implements a well-defined methodology for the elicitation of security related requirements in the AI4HealthSec project for the design and development of the relevant framework.

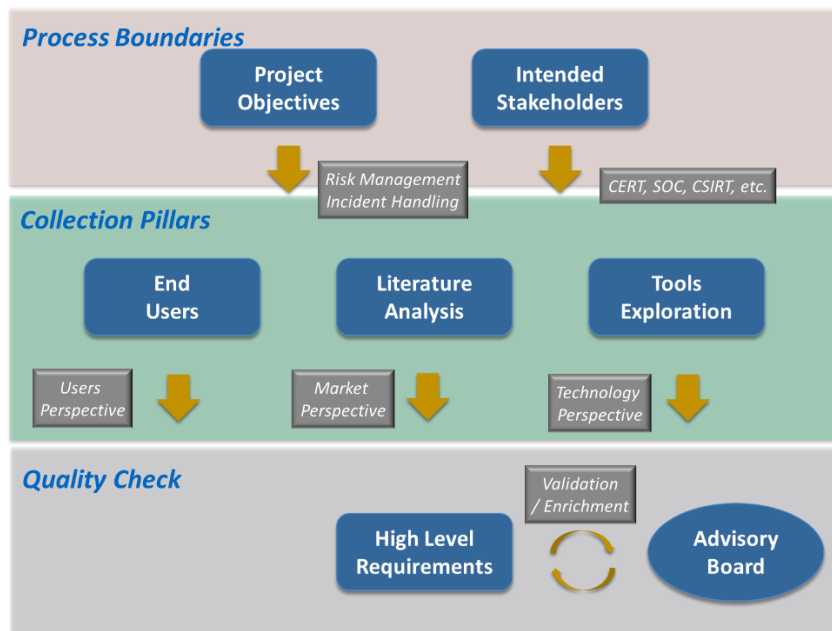


Figure 3: The process for collecting user requirements in the AI4HealthSec project.

In this methodology the requirements elicitation process unfolds in three parallel pillars, which are presented in Figure 3 and are analysed below:

The end-users' pillar: in this pillar, we include the activities for engaging representatives from the intended stakeholders to come up with high level requirements on the way that the AI4HealthSec framework will assist them in exercising their everyday activities for managing risks and handling incidents. The objectives of this pillar will be analysed in Section **Errore. L'origine riferimento non è stata trovata..**

The cross-domain literature analysis pillar: in this pillar, we aim to depict domain agnostic trends in the cyber security field with respect to the implementation of the RAP and IHP processes, by analysing the literature for best practices and guidelines in a variety of business domains, like digital health and healthcare, finance, logistics, etc. This pillar showcases the close link of the activities in task T2.1 with the other tasks in WP2, and especially task T2.2. The objectives of this pillar will be analysed in Section **Errore. L'origine riferimento non è stata trovata..**

The tools exploration pillar: in this pillar, we introduce the analysis of existing tools to operate the two processes (RAP and IHP) as the baseline for identifying new challenges and specifying additional functions and features to be delivered in the AI4HealthSec framework. The objectives of this pillar will be analysed in Section **Errore. L'origine riferimento non è stata trovata..**

As shown in Figure 3, these three pillars present a set of independent streams for commencing the work in the elicitation of high-level user requirements for the AI4HealthSec framework. As an additional quality check before releasing this list, we have already identified the importance of requirements enrichment and validation by the project Advisory Board, as an external and subjective board of experts in the cyber security field not only on the healthcare domain, but also in additional

business sectors that would essentially raise similar challenges in their risk assessment and incident handling processes. More about this step in the methodology is presented in Section 2.5.

2.2 *Business Needs/User Challenges Elicitation and analysis*

The first large analysis area dealt with the elicitation and analysis of business needs. Those needs should represent challenges that have to be met when further designing the AI4HealthSec framework. For this purpose, special questionnaires were developed. The questionnaires can be found in the appendix of this deliverable.

2.2.1 Objectives of the questionnaires, creation of the questionnaires

All questionnaires were developed iteratively with all WP2 project partners. The objective for the internal questionnaires is threefold:

1. To elicit organizational characteristics concerning cybersecurity policies and training
2. To elicit wishes and expectations towards a cybersecurity framework
3. To elicit the knowledge and involvement of different potential user groups into cyber-security issues at their organization

The external questionnaires did focus on the elicitation of organizational characteristics and on the analysis of wishes and expectations towards a cybersecurity framework.

Before creating the questionnaires, each pilot partner was asked to define the typical user groups in their organization; they filled in a form which asked the following questions:

1. **Whom** will you hand the questionnaire?
2. Are those **people also the possible end users** of an AI4HealthSec system?
 - If those people are **not** the end users: Why did you choose them to fill out the questionnaire?
3. Who are the **actual end users** in your use case? (if they are not the same people that will answer the questionnaires) people that will answer the questionnaires)
4. What did you think were the **possible advantages an AI4HealthSec system could offer** to the persons that will fill out the questionnaire when you created your use cases?
5. Are there any **difficulties when it comes to the conduction of the questionnaires**? E.g., are there strict time constraints of the potential participants? Do the potential participants have only low motivation to participate?
6. In what **setting** can the participants answer the questionnaires? E.g., in their workplace, at home.
7. Will the possible participants get an **instruction on the project** before they fill out the questionnaire or will they have never heard of the project AI4HealthSec before?
8. Will the possible participants have **heard of cybersecurity topics** before or are they completely blank on this topic?

Based on the answers to the questions for each defined user group an internal questionnaire was designed which fit to the expected motivation

2.2.2 Questionnaire content and structure

In total, ten questionnaires for the internal user requirements analysis (i.e. analysis within project's pilot partners) were created (internal questionnaires) – one questionnaire for each pre-defined use

scenario. In addition, one questionnaire for the external use outside of AI4HealthSec partners was developed (external questionnaire).

As the internal questionnaires should ask certain possible user groups for their wishes and expectations as well as for organizational details concerning personal experiences with cyber-attacks and cybersecurity and also hopes and wishes for AI4HealthSec components and possible fears and objections, they were designed to fit to each group of persons that are part of the user scenarios. Fraunhofer provided three different pilot scenarios; one of them was separated between users with decent and users with less experience on cybersecurity topics. UoB, EBIT and KLINIK provided together a pilot scenario. UoB and EBIT received one internal questionnaire to fit all pilot users, KLINIK three different questionnaires so that they fit to distinct potential end users at their site.

All AI4HealthSec partners used the same external questionnaire to hand it to organizations outside of the project consortium from different domains (not only healthcare, but e.g., energy, logistics).

Both internal and external questionnaires consisted of two main parts: One part focusing on organizational details, and one wishes and expectations for a cybersecurity framework. For the internal questionnaires only one partner had to answer the part on organizational details as we needed those input only once. The second part of the internal questionnaire was to be fulfilled by several potential end users provided by each pilot partner.

External organizations were all asked to fill in both parts of the questionnaire. All questionnaires contain both closed and open questions. Open questions were mainly included to ask for wishes and expectations towards a cybersecurity framework.

The first part (part A) of the internal questionnaire was filled in by only one member of each pilot organization of the AI4HealthSec project contained the following question categories:

- Details of organization (Public/private organization; size of organization)
- Details on security management (Outsourced or in-house security and incident management; security management standards; incident response teams, procedures to cover cyber-attacks; response capabilities; procedures to estimate cascading effects of security events; cooperation and exchange with external entities to share incident information; automated mechanisms to support incident handling process; collection of security-related data; performance of cyber risk assessments)
- Details on training of staff regarding cyber-security (drills provided)
- Employment of solution to centralize incident information for organization-wide perspective

Part B of the internal questionnaire consisted of the following topics:

- Vulnerable groups in the organization
- Preferred features of AI4HealthSec framework
- Concerns against AI4HealthSec framework
- Main possible benefits of AI4HealthSec framework
- Knowledge on cybersecurity and situational awareness of the staff
- Form of interaction with AI4HealthSec framework (interactive vs. autonomous system; invisible vs. visible system)

- Experience with cyber-security incidents

The External Questionnaires contained questions from both parts A and B of the internal questionnaires.

2.2.3 Methodology for analysing the user requirements questionnaire

The answers given in the closed questions were analysed in a descriptive manner using MS Excel. We did not intend to get statistically significant answers, but insights into potential end users' preconditions and expectations. Open questions were analysed using response categories. Those categories evolved while reading the given answers and by grouping them according to similar answer aspects.

2.3 Domain requirements elicitation and analysis

This section provides information on how the project partners have approached the domain requirements elicitation and analysis task. It focuses only on the approach and it does not discuss the results of the activity. Results are presented in Section 4.

The methodology for the domain requirements elicitation is based on a detailed literature review that focuses on the identification of the state of the art related to security standards, regulations and best practices for digital security of the healthcare sector. As part of this we consider standards such as ISO27001, ISO27005, ISO28000, and the CEN/TC 251 Committee; regulations such as the General Data Protection Regulation (GDPR); and recommended best practices, such as the Technical Safeguards for Data Security. As part of this, our approach identifies and analyses the healthcare market based on the AI4HealthSec circles of consideration, i.e. health components (first circle), medical equipment (second circle), individual HCIs (third circle) and interconnected HCIs (fourth circle).

2.4 Cybersecurity tools and systems requirements elicitation and analysis

As we mentioned in sections 1.2, and 2.1 and as we have introduced in the DoW, the AI4HealthSec framework defines two methodological processes and implements a set of tools to support stakeholders in the healthcare ecosystem to realise security and privacy risks and address related implications arising from the detection and analysis of cyber-attacks on the respective HCIs. The relevant mechanisms that the project will develop are to be deployed across the four levels of circles of consideration, ranging from devices within the patient personal space (i.e. wearables and implants) or used in the medical professionals' offices and in related hospital departments (medical equipment and devices, client side software, etc.) to the integrated hardware and software solutions comprising individual and interconnected HCIs (like laboratory and hospital information systems, PACS, etc.).

The envisaged contribution of AI4HealthSec spans across the specification and implementation of the tool supported methodologies for privacy and risk assessment and cyber-attacks related incident handling. The project provides the corresponding mechanisms and software components for the development of these methodological processes. These are categorised into four horizontal and three vertical layers, as shown in Figure 4 and they are summarised into the following high-level functions that the AI4HealthSec framework should address:

- Risk and privacy assessment: implement mechanisms for assessing the performance of risk and privacy management practices applied to interconnected assets found in HCIIIs.
- Incident management and realisation: implement mechanisms for integrating and correlating security and risk-related information and detecting anomalies with respect to cyber-attacks.
- Cyber-attacks forecasting and implications: implement mechanisms for constructing the path for the impact of detected anomalies across all the assets in the interconnected HCIIIs.
- Response and knowledge sharing: implement mechanisms for supporting decision making for the enactment of mitigation actions and establishing and sharing a knowledge base with lessons learnt.

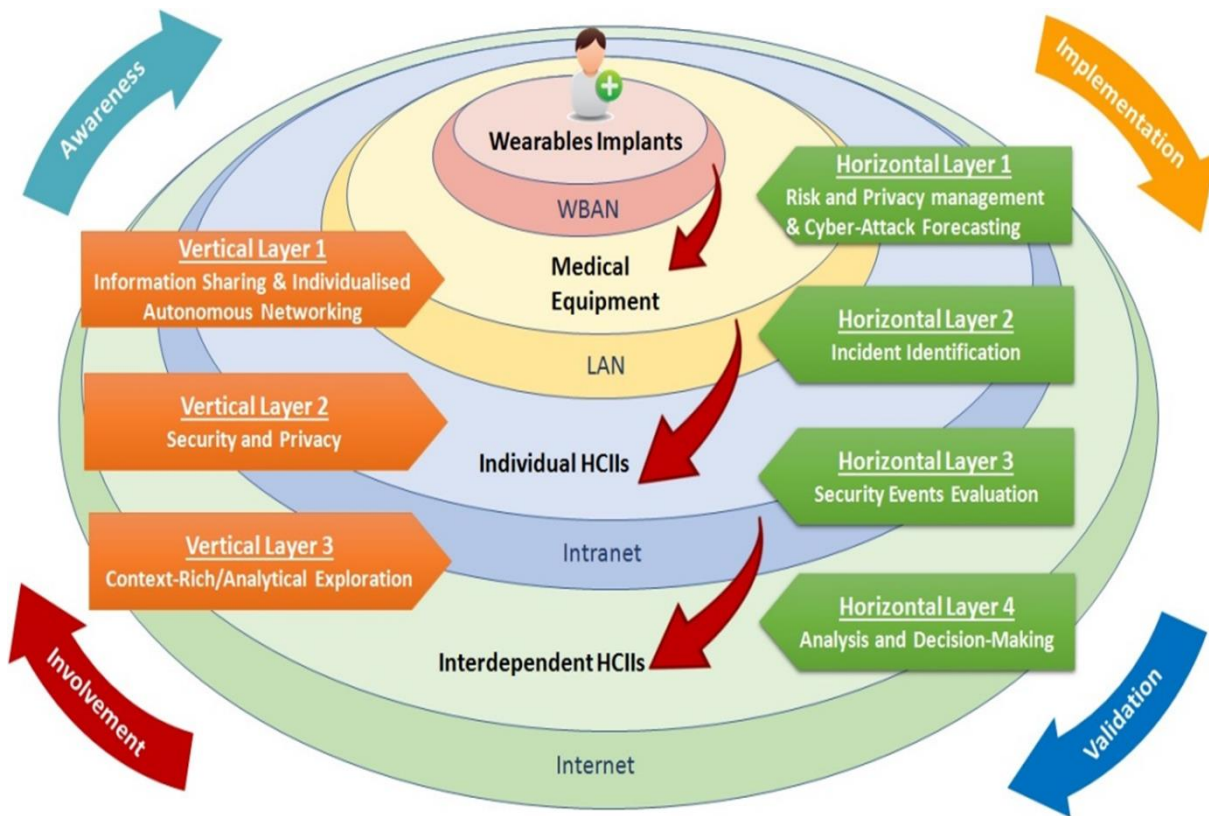


Figure 4: The overall conceptual elements of the AI4HealthSec framework

The mechanisms that the framework needs to develop have been partially addressed in existing solutions and approaches that most of the technical partners in the Consortium have already introduced into the market and need to be further developed and extended, subject to the research activities that the project foresees in WP3 – WP5. The respective tools and services will be analysed in Section 5 of this document with the aim to present the current maturity of relevant technical solutions and identify additional requirements that the intended users of the framework may have towards building a chain of technical tools and services that develop the mechanisms of the risk and privacy assessment and incident handling processes.

To process the input from all partners and facilitate the elicitation of user requirements are presented in the tools' exploration pillar of Figure 3 we define a template for the description of the proposed components, which can take the form of a tool, a software solution or a service. This template consists of the following sections:

- Short description: an overview of what the proposed component can bring into the project, the intention of use and the scientific and/or business problems it solves with respect to the aforementioned high-level functions.
- Key features: a short analysis of the key functions of the component.
- Component advantages: a brief introduction of the strong points of the component and the potential weak aspects that need to be considered, as well as the advantages and disadvantages of the use of this component in the context of the AI4HealthSec project.
- Example usage scenario(s): a set of user-driven scenarios detailing on the steps that a business stakeholder needs to follow to realize the key features of the component, presenting the information that need to be fed into the component and the expected output data.
- Expected extensions / new implementations: a summary of the functionalities that can be delivered within the scope of the AI4HealthSec project and their position with respect to the horizontal and vertical layers of the Framework, as presented in Figure 4.

As a result of this process, we will be able to identify the challenges that the AI4HealthSec framework will have to address, in order to allow the intended stakeholder to integrate the proposed components into the methodologies for privacy and risk assessment and cyber-attacks related incident handling.

2.5 EAB engagement in user requirements elicitation and analysis

All user requirements elicited by the methods described above were then presented to the AI4HealthSec External Advisory Board (EAB). For this task, the project consortium organized an online video call with three EAB experts from medical informatics, the finance sector and biomedical engineering.

The objective of the EAB engagement was not to validate the requirements but to get feedback from the experts regarding further project steps and hints on how to deal with the requirements.

3 Results: Business needs/user challenges elicitation and analysis

The following chapter presents the findings from the first pillar of requirements analysis as presented in Figure 3. The user perspective is the basis for the further elicitation of more concrete technical requirements. In total, we collected 31 internal user requirements questionnaires. Most questionnaires have been filled in completely.

3.1 Internal user requirements analysis: Part A – Information on pilot sites security policies

Most pilots adopted or plan to adopt several security management standards, including ISO9001, ISO/IEC 27001 or more specifically B3S which is a hospital-specific security standard. There are no automated mechanisms to support the incident handling process employed at the pilot sites. Security-related data (e.g., logs, attacks) are collected by the hospital site which also stores them in files. There is also some exchange between the pilots and other organizations regarding attack-related data, e.g., the hospital site collects data from other hospitals. The hospital site has disaster recovery policies, malicious software policies, and network access policies and/or procedures in place. Fraunhofer's pilot site furthermore provides incident handling, information security incident management, disaster recovery, access control, network access and identification and authentication policies and/or procedures. In all internal organizations there is trained and non-trained personnel in terms of cyber-security. Numerous organizations are offering (or plan to offer) training programs for the employees.

3.2 Internal user requirements analysis: Part B- Insights into participants background and experience with cybersecurity

Part B of the questionnaire was filled in by several members of the pilot organizations. From all project pilot partners, the following person groups answered the internal questionnaires:

- Biobank operators
- Biologists
- Medical wearables app developers
- Medical wearables backend developers
- Developers of biobank applications
- Software developers of implantable medical devices
- Hardware developers of implantable devices
- Living Lab researcher
- Hospital's data security officer
- Nursing manager
- Hospital's controller Human Resources
- Administration of laboratory IT in a hospital
- Disaster concept developer hospital
- Product manager healthcare
- Project engineer healthcare
- Post Sales manager healthcare IT
- Help desk technical support manager hospital
- Pre-Sales manager healthcare IT
- R&D manager healthcare IT
- Test manager healthcare IT

- Installation manager healthcare IT
- Integration manager healthcare IT

The findings from the internal organization's analysis are presented in the following.

3.2.1 Vulnerable groups regarding cyber-attacks at the pilot sites

The most vulnerable groups in regard to cyber-security incidents were found to be patients, followed by physicians. For more groups see Table 1.

Table 1: Vulnerable groups

Vulnerable Groups	Reason of vulnerability
Patients/Residents of the Living Lab	At risk of suffering consequences from cyber- attacks.
Physicians	At risk causing dangerous cybersecurity situations. / At risk of suffering consequences from cyber-attacks.
Hospital Managers	At risk of suffering consequences from cyber-attacks.
Researchers depending on biomaterial	At risk of suffering consequences from cyber-attacks.
Other hospital staff	At risk of suffering consequences from cyber-attacks. / At risk causing dangerous cybersecurity situations.
Hospital as a whole organization	At risk of suffering consequences from cyber-attacks.

Patients resp. residents of the Living Lab clearly were seen as the most vulnerable group as the negative consequences evolving from cybersecurity breaches would affect them most directly, potentially even resulting in the loss of life of patients (e.g., when important medical information for the treatment of a patient is compromised or lost or when there is a malfunctioning implantable medical device).

Physicians are characterised as vulnerable as well in terms of their tendency to have to work under time constraints in combination with a potentially low level of cybersecurity awareness. Moreover, physicians (representing the end user of software in healthcare IT) identified as are considered the most relevant gateway for malware and cyberattacks. Of course, physicians are also the ones to suffer from consequences if they are not able to access patient data needed for their daily work as well.

Hospital managers characterised are at risk as well, because they could potentially be victims of blackmailing approaches; **medical researchers** are vulnerable in terms of not being able to analyse biomaterial in biobanks due to a loss of data.

Further hospital staff is both at risk for causing and suffering from cyberattacks; and one participant from the hospital stated that the “hospital as a whole organization” is at risk because cyberattacks could cause high cost to clean up the IT system after an attack.

3.2.2 Insights: Members of pilot organizations on risk awareness, organization policies and experiences with cybersecurity topics

Enriching the finding that patients/residents are considered the most vulnerable group to suffer from the consequences of cyberattacks, it became clear that, in the hospital setting, most of the hospital members that filled in the questionnaire saw medical staff (e.g., physicians and nurses) as not *knowledgeable enough on cybersecurity to prevent dangerous situations*. In addition, the training on cybersecurity for medical staff seems to be not sufficient enough to prevent critical incidents.

Similar findings delivered for administrative staff at the hospital: Although they might be more proficient still both knowledge and training concerning cybersecurity as appears to be insufficient.

For all participants of the internal user requirements analysis, it was found that most participants *self-assessed* an average *knowledge of cybersecurity topics* (n=22). Nevertheless, some stated that they had only below average knowledge (n=3) (Figure 5).

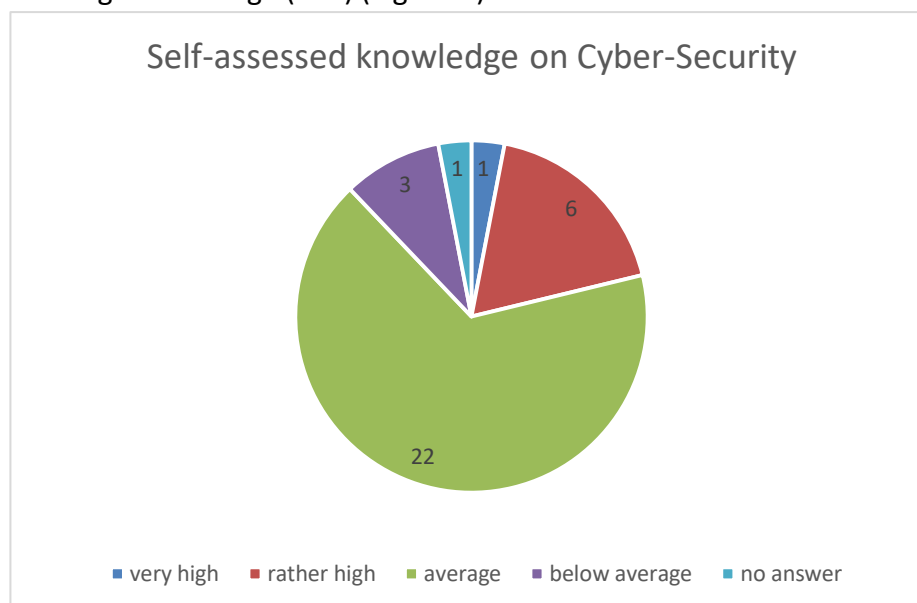


Figure 5: Self-assessed knowledge on cybersecurity topics

Half of the participants (n=17) stated that they have *been trained by their organization on cybersecurity topics*; the other half was not trained (n=17). Four of the 17 participants that have not been trained by their own organization have, nevertheless, been trained by other organizations. Thus, we have the finding that slightly more participants are getting trained on cybersecurity topics.

Is training crucial for preventing cybersecurity breaches? We tried to get more insights into this aspect by asking what measurement concerning cybersecurity was considered as most important: In the hospital setting it appeared that measures concerning training and awareness on cybersecurity were considered as more important than technology-based solutions. Nevertheless, in general, it has been stated that both, technical solutions and the creation of higher awareness, are important to create a more secure cyber environment at the hospital.

When it comes to the *personal experiences with cybersecurity incidents*, it has been found for all pilot sites, that numerous persons have already personally encountered cybersecurity incidents and knew only partially what to do in this situation (Figure 6, Figure 7) .

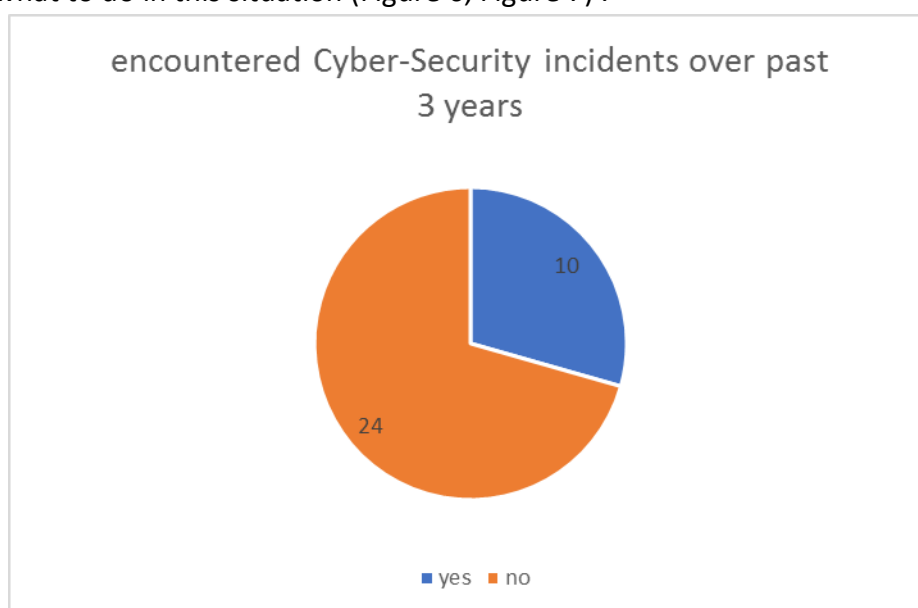


Figure 6: Encountered cybersecurity incidents over the past 3 years



Figure 7: Knowledge what to do in case of cybersecurity incident

The pilot organizations' members are only partially *involved in their company's risk management*: 17 of the participants stated they have been involved, 13 answered that they are not involved yet (Figure 8).



Figure 8: Personal involvement in risk management process

Most participants of all pilot sites are not yet personally *involved in cybersecurity tasks* (n=19); 13 persons have been involved (Figure 9).

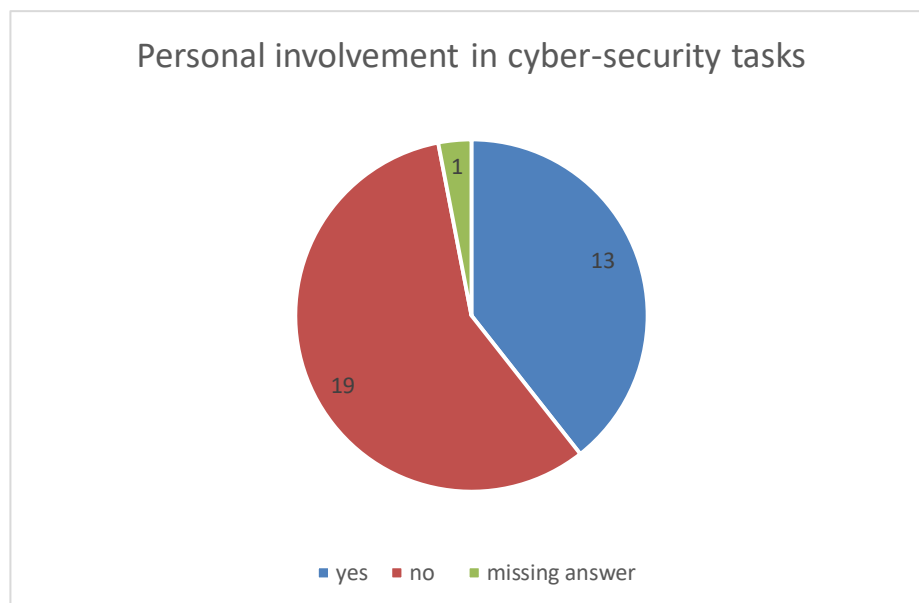


Figure 9: Personal involvement in cyber-security tasks

Those cyber-security related tasks included:

- the management of post sales installations in healthcare IT
- the active protection of the personal IT work environment
- the creation of risk analysis and testing concepts for healthcare IT products
- the specification of a data protection concept in the development of disease management solutions with wearables and biobanks
- the operational monitoring of biobank equipment and infrastructure
- the participation in the development of secured web-based database applications for biobanks which integrates authentication and access-control frameworks
- to take care of possible cyber threats during the development of products
- the development of a complete disaster management concept for the hospital's main IT structures
- the performance or review of policies in reference to cyber-security, e.g., password policy, writing down the technical and organizational measures for processing personal data

Most participants of all pilot sites agreed that their organization's security incident management *policy is able to improve the situational awareness* regarding cybersecurity (n=25 agreed or strongly agreed). Although, a large part (n=6) did not know how to answer this question (Figure 10).

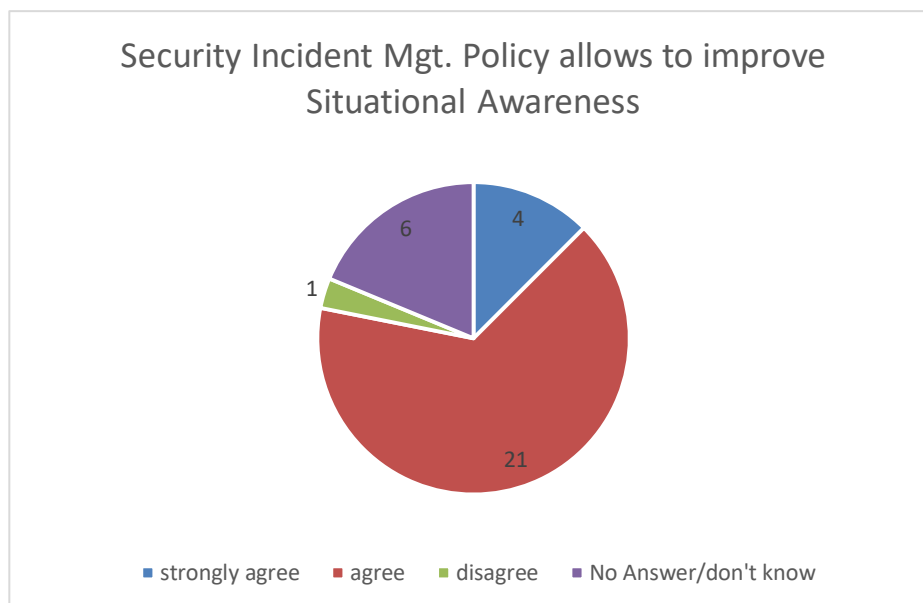


Figure 10: Policy allows to improve situational awareness concerning cybersecurity

Apparently, it is recognized that an organization-wide security-awareness is important. Most participants consider their resp. organization to be *in favour of engaging in a CIP program*; also most participants said that they would find it very useful or useful if their organization participated in a CIP program (several did not answer this question) (Figure 11).

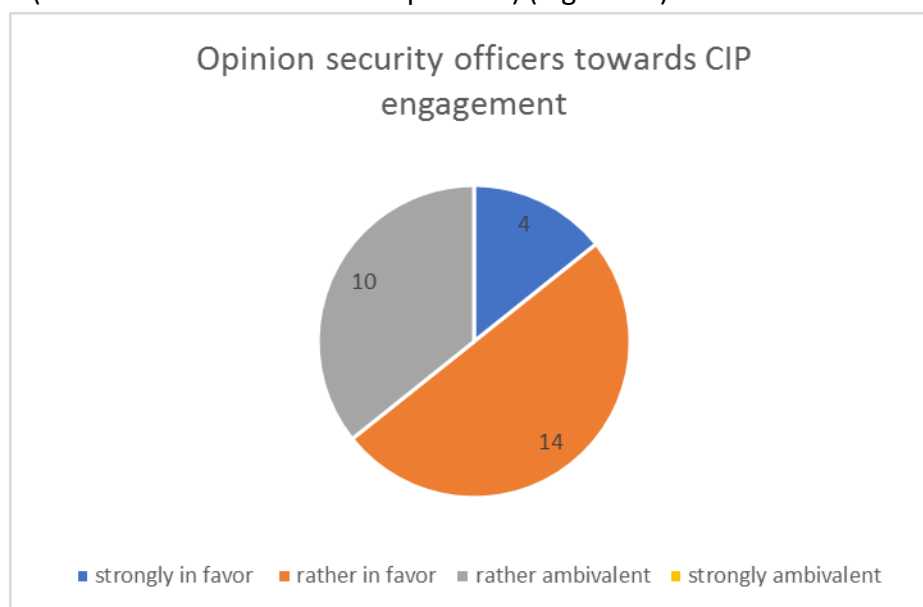


Figure 11: Opinion of organization's security officers towards engagement in CIP program

We got a rather heterogeneous picture regarding the wish for a *visibility of an external cyber-security framework* in the daily work life: 13 participants stated that they would prefer it invisible in the

background, whereas 17 would expect it visible for them, e.g. by providing regular status reports (Figure 12).

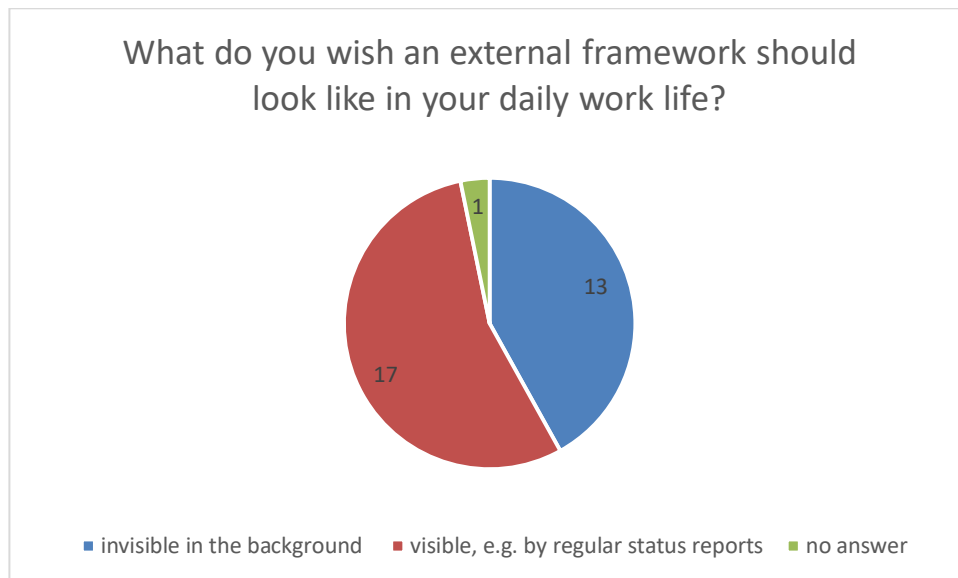


Figure 12: Invisible vs. visible framework

For the question of *how to interact with an external cybersecurity framework* in the daily work life the answers were more clear: 22 participants would like to have a framework that would run completely by itself; 7 would prefer a framework that would need input by the user (Figure 13).

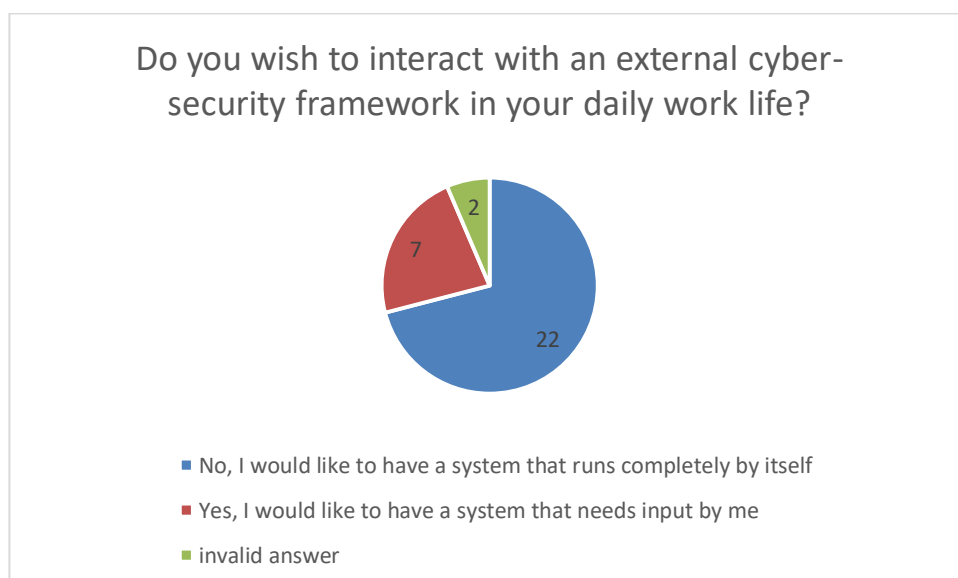


Figure 13: Interaction with external cybersecurity framework

In the following, there is a conclusion of the findings of the internal user requirements analysis mostly regarding the closed questions of the questionnaires:

There is a mixed picture of availability of training on cybersecurity topics and awareness amongst the pilot organizations: Not every staff member has access to training. Nevertheless, most of the participants see themselves as at least average knowledgeable regarding cybersecurity topics.

Especially, medical staff at the hospital might not be trained enough to prevent critical cyber-security situations.

Several participants have yet experienced critical situations themselves and only some of them knew what to do. Notwithstanding, there is a relatively high engagement in cybersecurity tasks among the participants.

In general, the pilot organizations seem to be in favour of engaging in a CIP program as well as to enable a higher situational awareness in the whole organization as well.

An external framework that would help with creating a higher cybersecurity would preferably run by itself but should possibly provide regular status reports.

3.3 External user requirements analysis

From all project partners we collected 30 external user requirements questionnaires. The external organizations that answered the questionnaire originated from the following domains:

- Finance domain
- Health domain
- Logistic domain
- High culture and research on telecommunications and information technologies
- Public administration on digital innovation
- Education (university)
- Archiving and conservation of documents coming from different domains (health, legal, financial)
- Energy
- Non-profit organization
- Insurance
- IT
- Certification body
- Drinking water supply sector

3.3.1 External user requirements analysis: Information on security policies

Inhouse as well as outsourced *security and incident management models* are adopted among the external organizations, several use also mixed approaches where several parts of security management are outsourced, whereas other parts stay inhouse. *Security management standards and protocols* are mainly adopted, including ISO 9001, ISO/IEC 27001, ISO 20000, ISO/IEC 27002, ISO/IEC 27005, NIST SP800-30, NIST SP800-61, the NIST framework for improving critical infrastructure

cybersecurity, the ISO 14001:2015 standard for environmental management systems and other more sector-specific standards.

Policies and procedures concerning such as incident handling response, information security incident management, disaster recovery or security monitoring are mostly common at the external organizations.

Most (n=23/29) organizations have an *incident response team* in case of a security breach; also the most of them (n=22/29) have procedures to cover cyberattacks. With regards the employment of *advanced response capabilities to effectively respond to cybersecurity incidents* the finding was not so clear: 16 of 29 organizations had actually employed such capabilities employ them, whereas 13 did not have such capabilities or the representatives of the organization did not give an answer to this question.

Most organization's representatives answered "yes, the procedures estimate the *cascading effects*". The same finding was true with the question if the organization *cooperated with external entities* to correlate and share incident information to achieve a cross-organizational perspective on incident awareness (n=14 "yes", n=12 "no").

16 organizations representatives stated that they have a *vulnerability management process*. Three participants stated that this process is performed at least yearly, whereas, three reported that it is performed (or planned to be performed) on a daily basis. Some organizations used vulnerability databases such as OWASP, Nessus, CVE, NVD or Secunia.

The majority of the external organizations do not use *tools or suites to run a dynamic (penetration) testing* of their ICT infrastructure (n=13 "no"; n=11 "yes"). Of those who do use such tools most perform the tests yearly (n=3) or ad hoc (n=3) with the help of tools from third parties or, if internal, Nmap, openvas, burp suite, wireshark and other tools.

The largest part of organizations *monitors their infrastructures for malicious activities* (n=22) using mostly antivirus software.

In the most external organization a few, but not all *staff members are skilled and trained on security and incident handling practices* (n=18 "a few"; n=9 "most", n=1 "none") and at the same time most organizations are offering or at least are willing to offer *training programs to its employees* concerning cyber-security awareness (n=16).

3.3.2 External user requirements analysis: Vulnerable groups

Vulnerable groups identified in the questionnaires for the external organizations' representatives included the ones that are presented in Table 2.

Table 2: Vulnerable groups external user requirements analysis

Vulnerable Group	Reason of Vulnerability
Patients/refugees/students/final customers	At risk of suffering consequences from cyber-attacks.
Doctors/Nurses/non-technical staff	At risk causing dangerous cyber-security situations. /

	At risk of suffering consequences from cyber-attacks.
Staff members (e.g., office workers, system admins, crew on board of a ship)	At risk causing dangerous cyber-security situations / At risk of suffering consequences from cyber-attacks.

Similar to the findings from the internal user requirements analysis, we found that external organization's representatives also tend to see person groups as most vulnerable to cause critical situations that are not technical savvy and those groups as suffering from consequences that are linked to sensitive data. It became very clear from the answers that there is a cybersecurity awareness gap amongst staff members which is considered the biggest vulnerability.

3.3.3 Insights: Representative of external organization on risk awareness, organization policies and experiences with cybersecurity topics

The representatives of external organizations appeared to be more *knowledgeable on cyber-security topics* than the members of the internal organizations – 18 reported that they have rather or very high knowledge, 10 stated that they have an average knowledge (Figure 14). This finding might be connected to the fact that external organizations that have been contacted by our project consortium probably provided more, in terms of cybersecurity, experienced staff members to answer the questions than internal organizations did.

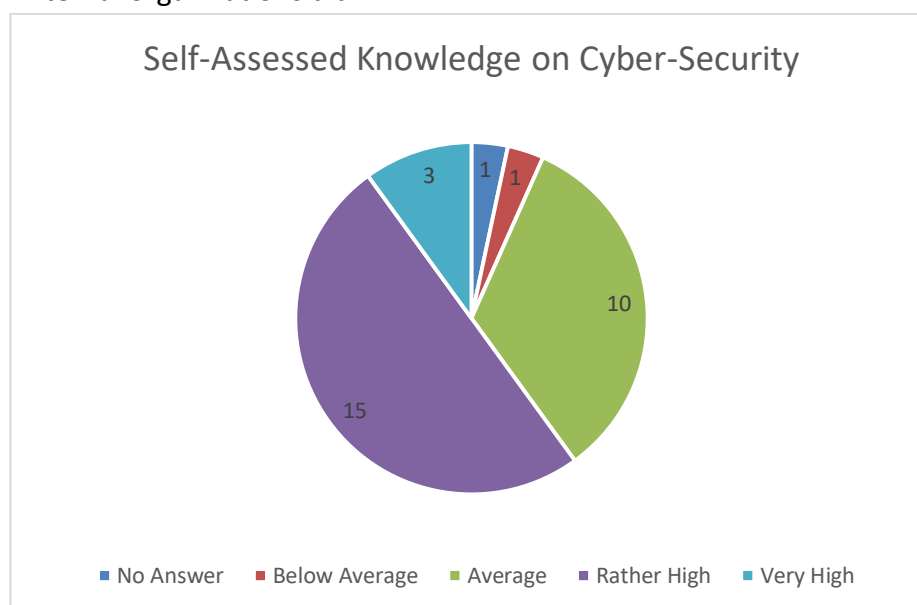


Figure 14: External_self-assessed knowledge on cybersecurity

17 of 29 participants answered that they have not been *trained by their organization on cybersecurity topics*; only 11 have been trained by the respective organization. Nevertheless, eight of the persons that have not been trained by their own stated that they have been trained by another organization

(Figure 15). Similar to the internal organizations' members, the largest part appears trained on cybersecurity topics.



Figure 15: External_trained on cybersecurity by own organization

Regarding *personal experience with cybersecurity incidents*, the largest part indeed did encounter them in the past three years ($n=17$), most of them have also been personally been involved. Of these persons most knew what to do ($n=5$) or got help ($n=1$) (Figure 16, Figure 17).

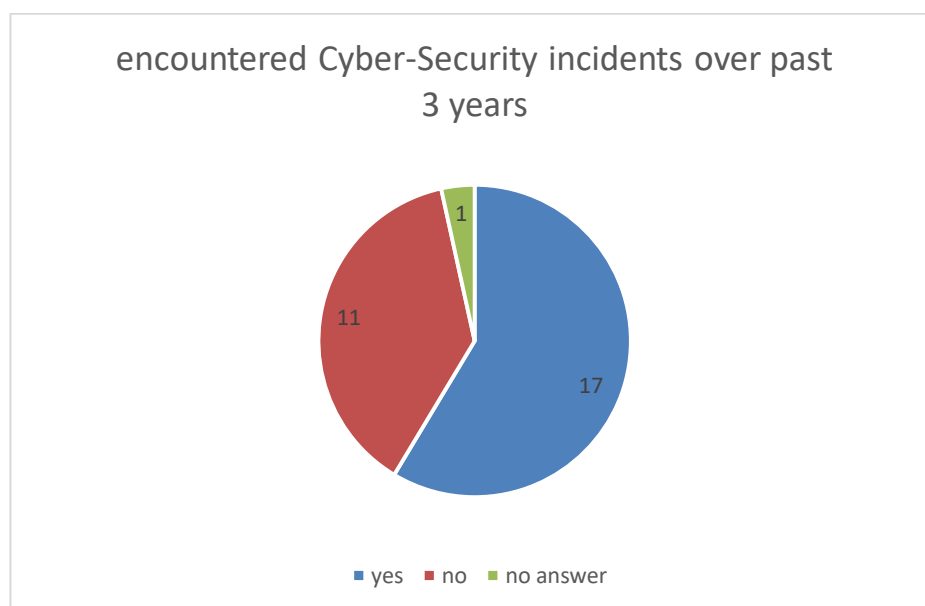


Figure 16: External_encountered cybersecurity incidents

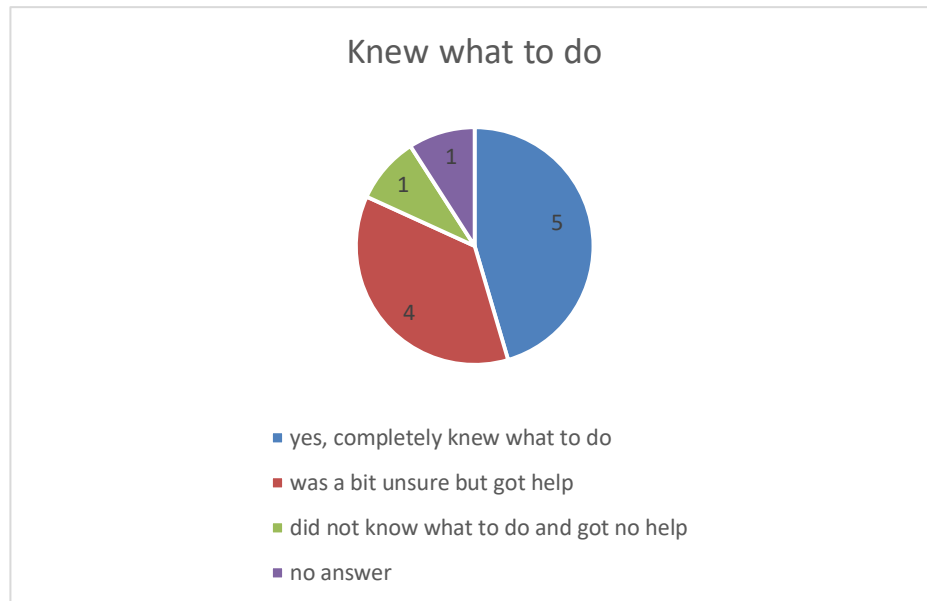


Figure 17: External_knew what to do in case of incident

Similar to the internal organizations, most of the external companies are not yet personally *involved in cybersecurity tasks* (n=16); 13 persons have been involved.

Those tasks in cybersecurity included for the representatives of the external organizations:

- Internal Audits and Vulnerability Assessments
- Team membership of an incident response team
- Logging of user access
- Writing and reviewing of cyber-security policies and procedures
- Access control performance

Most participants of all external organizations agreed or strongly agreed that their organization's security incident management *policy is able to improve the situational awareness* regarding cybersecurity (n=25 agreed or strongly agreed) (Figure 18).

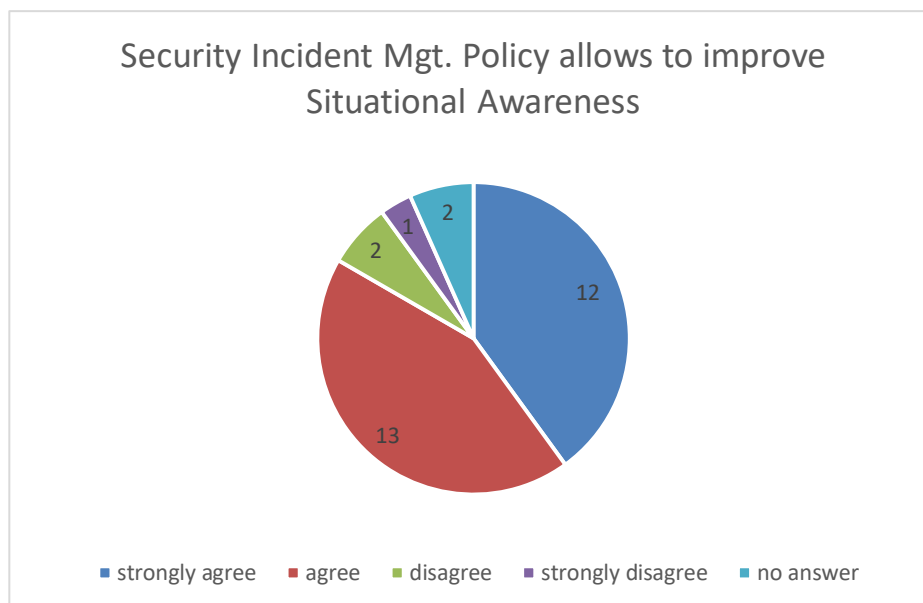


Figure 18: External_Policy allow to improve situational awareness concerning cybersecurity

Most participants saw their resp. organization as being *in favour of engaging in a CIP program* (Figure 19) and, most participants said that they would find it very useful or useful if their organization *participated in a CIP program*.

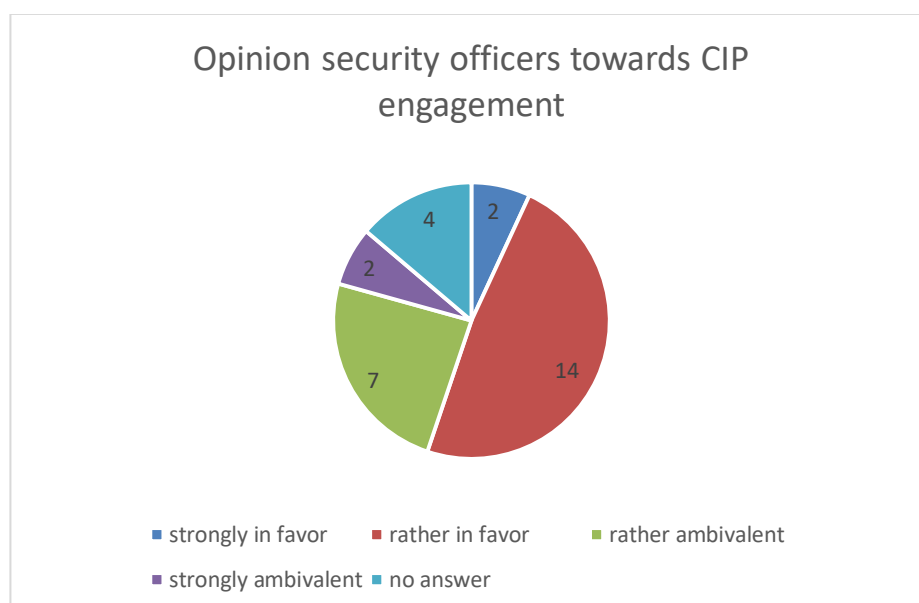


Figure 19: External_Opinion security officers towards CIP engagement

The necessity of that an external cybersecurity framework should be *visible in the daily work life* was clearer than at the internal organizations: 16 representatives of external organizations said that they would prefer a visible framework, e.g. by regular status reports. Ten participants wish to have a framework that runs invisible in the background (Figure 20).

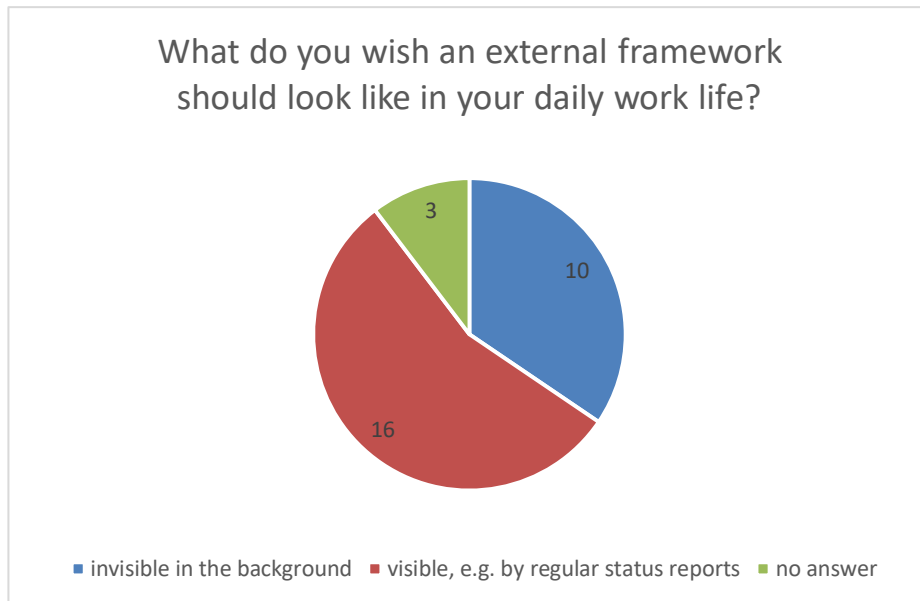


Figure 20: External_Invisible vs. visible framework

For the question of *how to interact with an external cybersecurity framework* in the daily work life the answers were again like the internal organization responses: 14 participants would like to have a framework that would run completely by itself; eleven would like a framework that needs input by the user (Figure 21).

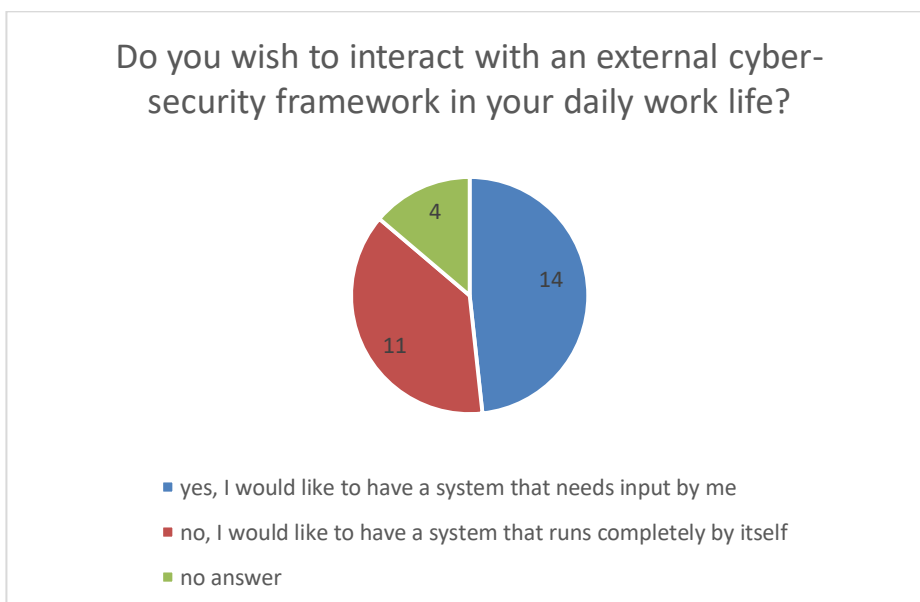


Figure 21: External_Interaction with external cybersecurity framework

In general, the findings of the external user analysis are very similar to the internal analysis, but the participants of the analysis from the external organizations saw themselves as a bit more cybersecurity savvy than the internal organization members.

For both internal and external organizations, it became clear that the organizations were indeed willing to offer the frame for a higher security awareness and that, in fact, most of the persons have been trained on cybersecurity issues either by their own or by another organization. In addition, several people did encounter cybersecurity incidents by person but only partially know what to do in this case.

When it comes to the external organizations' members, an external framework that would help to provide a higher cybersecurity protection would preferably need input by a staff member and provides regular status reports. The internal organization members would also prefer to get regular status reports, but would prefer the system to run completely by itself rather than requiring input by staff members.

3.4 Concerns against AI4HealthSec framework

Concerns regarding the AI4HealthSec framework that we concluded from both the internal and the external questionnaires included:

- Framework vulnerability: New vulnerabilities in the organization due to the framework
- System overloading/Performance reduction: Loss of performance and usability in already existing tools
- Problems with integrating framework into existing IT infrastructure: Compatibility with all type of infrastructure
- Groups/persons needed for the maintenance and support with AI4HealthSec framework and possible support delays: Not enough support if there is a problem with the framework
- Lack of trust in the framework amongst staff
- State-of-the-art of the framework: Framework might not include latest information
- General concerns regarding cloud-based solutions
- High expected effort for initializing and operating the framework
- Missing transparency regarding features, capabilities and limitations of the framework
- Possibility that the existence of a security framework might lead to carelessness
- Ethical concerns of outsourcing the security management
- Fear to take out business secrets to another company, framework might misuse private information
- Staff not trained enough to use such a framework
- Diversity of medical devices, not easy to integrate in one platform
- Framework does not work reliable.
- Some external organizations' members doubted that a framework designed for the healthcare sector would be adaptable to other sectors as well.

Therefore, it should be considered, that an external framework must be compatible to numerous already existing IT structures. Support should be provided (both with maintaining the framework and with initializing it), and it should be guaranteed that the framework constantly is updated regarding the newest cyber-security threats.

The framework might face a trust issues with the organization's staff which is intended to use it because it "is from an external organization". The self-image of the framework should also be promoted as an additional auxiliary source to the existing cybersecurity solutions with the concern that it does not replace them and therefore it does not release individuals from their duty to take care for cybersecurity. Defined solutions for other critical infrastructures should be provided, in case it is planned to extend the framework's focus from the healthcare area to other sectors. In this regard, another in-depth analysis of specifications and users' expectations in each CI area should be probably provided.

3.5 Business Needs/ User Challenges

By analysing both closed and open questions in the internal and external questionnaires, we were able to exclude a list of six business needs (Table 3). Those needs depict challenges that need to be faced when creating an AI4HealthSec framework.

Table 3: Business Needs/ User Challenges

Business Need ID	Title	Description
BN1.	Prediction and Prevention of Attacks	My organization needs to forecast and prevent cyber-attacks.
BN2.	Vulnerability Assessment	My organization needs a framework to assess its cyber-security weaknesses.
BN3.	Awareness Creation and Prevention of Human Errors	My organization needs a better awareness and higher knowledge concerning the staff when it comes to cyber-security topics.
BN4.	Detection of Abnormal Patterns and Creation of Warnings	My organization needs a system to automatically detect abnormal patterns in my IT and create warnings.
BN5.	Simplification of the Process of Risk Assessment	My organization needs a simpler process of risk assessment.
BN6.	Development of Long-Term Strategy of New Protection Solutions.	My organization needs a long-term and comprehensive cyber-security strategy.

At this point of the project, all challenges are rather broad and are to make sure that all the basic needs of potential users are depicted.

4 Results: Domain requirements elicitation and analysis

Enriching the findings from a user perspective, the next pillar includes the elicitation of requirements from the domain perspective.

4.1 Healthcare security management standards and best practices

This section describes a set of international and national standards and best practices and guidelines related to the AI4HealthSec project. In particular, in section 4.1.1 an outline of security management standards, including the ISO/IEC 27000 family of standards, which is the main international standard for information security management systems and the National Institute of Standards and Technology (NIST) SP 800 publication, which provides guidelines for securing IT infrastructure from a

technical perspective. Section 4.1.2 outlines management standards specifically for the health care domain, including the ISO14971, ISO/TR 22696, UEC 80001, ISO13606, the UK National Health Service Data Security Standard and the ISO/IEC 81001-1. Section 4.1.3 concludes with an outline of relevant best practices and guidance from FDA, HIPAA, the EU and ENISA.

4.1.1 Security Management Standards

ISO/IEC 27000:2018. The ISO / IEC 27000 is the family of international standards that define the requirements for setting up and managing the Management System of Information Security. It provides good practice recommendations on Information Security Management Systems (ISMS). The series of ISO / IEC 27000 is broad in scope. It is applicable to all types of organizations (e.g., governmental agencies, large companies, small and medium size enterprises) which intend to manage risks that could compromise the organization's information security. Essentially, the ISO information security risk management process can be applied to the organization as a whole; any discrete part of the organization (e.g., a department, a physical location, a service); any information system; and any existing, planned, or particular aspect of control (e.g., business continuity planning). It includes a family of standards that define requirements for an ISMS and for those certifying such systems, it provides direct support, guidance and interpretation for the overall process to establish, implement, maintain and improve an ISMS, it addresses sector-specific guidelines for ISMS and it addresses conformity assessment for ISMS. The most relevant to AI4HealthSec standards of the 27000 family are outlined below.

ISO/IEC 27001:2013 is a standard that specifies requirements for the establishment, implementation, monitoring and review, maintenance, and improvement of an Information Security Management System. The ISO/IEC 27001 does not mandate specific information security controls but stops at the Management and Operational level. Usually, a group of analysts with high ICT expertise and experience verifies the compliance of the organization with the defined requirements. However, although, the compliance process requires the involvement of multiple users the collaborative abilities of the standard are limited due to its inherent complexity. Practically the standard is mostly used by large scale organizations (e.g., governmental agencies and large companies) since it is considered too heavy for micro, small and medium size businesses. The ISO/IEC 27001 ISMS incorporate continuous improvement processes; such as Plan-Do-Check-Act (PDCA) or Six Sigma's Define, Measure, Analyse, Improve and Control (DMAIC) cycles. For instance, information security controls are not merely specified and implemented as a one-off activity but are continually reviewed and adjusted to take account of changes in the security threats, vulnerabilities and impacts of information security failures, using review and improvement activities specified within the management system.

It should be noted that ISO/IEC 27001 is actually not in effect a method for risk management but rather a compliance standard, reporting a list of controls for good security practices and the requisites that an existing method should have to be standard-compliant. Specifically, it provides generic requirements that a risk analysis and management have to comply to through a recognized method without to provide a specific method.

ISO/IEC 27005:2018. The ISO/IEC 27005 is part of the 27000 family of standards that describes the Risk Management Process and its activities for information security and provides guidelines for Information Security Risk Management and supports the general concepts specified in ISO/IEC 27001:2013 as well as the main principles and rules described in ISO/IEC 27002:2013. The information security risk management process consists of:

- Context Establishment: intends to define the risk management's boundary.
- Risk Assessment (Risk Analysis & Evaluation phases): used to make decisions and consider the objectives of the organization.
- Risk Analysis (Risk Identification & Estimation phases): intends to evaluate the risk level.
- Risk Treatment (Risk Treatment & Risk Acceptance phases): to reduce, retain, avoid or transfer the risks.
- Risk Acceptance: review of the risk treatment, validation of selected solutions, selection of residual risks, accepting a number of risks that can consider itself unable to deal, or are acceptable to the organization
- Risk Communication: to achieve agreement on how to manage risks by exchanging and/or sharing information about risk between the decision makers and other stakeholders.
- Risk Monitoring and Review: to detect any changes in the context of the organization at an early stage, and to maintain an overview of the complete risk snapshot.

However, it should be noted that the objective of this standard is not to constitute a risk management method but rather to fix a minimal framework and to describe requirements for the risk assessment process itself, for the identification of threats and vulnerabilities which are required to estimate risks and their level, and hence be in the position to define an effective treatment plan. ISO 27005 proposes the use of both quantitative and qualitative methods for the calculation of risk levels, however it does not support any specific technique for this purpose or any computational method to analyse and combine the assessment information. The generic nature of the standard does not include aspects that promote the collaboration among the users.

ISO / IEC 27002 provides guidance not related to the protection the information assets of a company, rather it provides recommendations for ensuring information security against risks to the confidentiality, integrity and availability of information. Moreover, the guidelines in ISO / IEC 27002 focus on ensuring the security of all forms of IT systems, networks, including data, and intellectual property. The standard is tailored to the specific information risks and needs of any organisation, irrespective of size or type and offers recommendations on standard security practices that enable an organisation to meet audit, regulatory and legal requirements. Therefore, by adopting ISO / IEC 27002, an organisation can be able to assess its information risks, define control objectives and apply appropriate controls (e.g., asset management, compliance, operations security, communications security etc.)

ISO / IEC 27003 provides guidelines for the implementation of a management system of information security in accordance with ISO 27001. The goal of this standard focuses on the crucial aspects needed for the successful design and implementation of ISMS within an organisation. In particular, it guides the process of obtaining management approval to implement ISMS, defining ISMS project from

planning, inception and design and final implementation phases. Mostly, ISMS comprise a set of activities for the management of information security risks by which an organisation identifies, analyses and addresses risks. ISMS ensure that an organisation's security processes are fine-tuned to address the ever-dynamic security threats and vulnerabilities.

ISO / IEC 27010 primarily focuses on the information exchange and sharing regarding the maintenance and protection of an organisation's CI. It aims at providing general guiding principles for communicating and information sharing about security incidents, threats, vulnerabilities, and controls, between organisations in the same or different sectors to protect CI, meet legal, regulatory or contractual agreements. In addition, it provides the basis and guidance on methods, models, policies, processes, protocols, and controls, for the sharing of information securely with trusted counterparties under all circumstances.

ISO / IEC 27014 provides guidance on principles and processes for the governance of information security, by which organisations can evaluate, direct, and monitor the management of information security. It also provides a structure by which the objectives of an organisation are set, the means of attaining those objectives, and how performance monitoring can be achieved. In general, the standard assists organisations to make informed and timely decisions about information security issues in support of its strategic objectives by aligning security objectives with business strategy, effective investment decisions on information security, ensuring transparency on information security status, as well as achieving compliance with regulatory, contractual, and legal requirements.

ISO / IEC 27032 consists of two focal areas. The first part deals with control measures for addressing cybersecurity issues associated with the Internet, with a particular focus on providing technical guidance for addressing common cybersecurity risks such as social engineering, hacking and malicious software. The standard also provides recommendations with regards to the crucial measures for addressing these risks, including preparing, detecting, and monitoring, and responding to attacks. The second focal area of the standard provides a framework for efficient and effective information sharing, collaboration, coordination, and incident handling amongst organisations. It includes key elements for establishing digital trust and processes for information interchange.

ISO / IEC 27035 is another crucial standard that focuses on information security incident management. It aims to complement other ISO standards that guide the investigation of, and preparation to investigate security incidents. In addition, it provides a basic definition of concepts and phases for information security incident management, including a structured guideline for planning and preparing incident management activities such as detecting, reporting, assessing, and responding to incidents. The guidelines consist of phases for planning and preparing security incident management policies, security policies, establishing incident response team, incident management awareness training, and incident management plan testing.

ISO 27799. It deals with information security management and information security controls in the healthcare industry. The standard provides detailed guidance on how best to protect the confidentiality, integrity, and availability of personal health data for anyone working in the health sector or its unique operating environments. Additionally, it gives guidelines for organisational information security standards and information security management practices including the selection, implementation and management of controls taking into consideration the organisation's information security risk environment.

NIST SP 800. The National Institute of Standards and Technology (NIST) of the United States is responsible for establishing technology, standards, and metrics to be applied to the science and technology industries. The NIST Special Publication (SP) 800 series present information of interest to the computer security community. The series comprises guidelines, recommendations, technical specifications, and annual reports of NIST's cybersecurity activities. Below we present some of the publications, which are of interest to the AI4HealthSec project.

NIST SP 800-30 "Guide for Conducting Risk Assessments". NIST SP 800-30 is a standard developed by NIST, which provides guidelines for securing IT infrastructure from a technical perspective. NIST SP 800-30 was one of the first risk assessment standards, and many other standards are influenced by it. It has been widely used for information security risk assessment globally, and it is relevant to any business with an IT component. Although the standard does not explicitly focus on health care, it provides guidance for critical infrastructures including health care infrastructures. It also guides determining appropriate courses of action in response to identified risks, as well as identifying specific risk factors that are continuously monitored so that an organisation can decide if risks have exceeded organisational risk tolerance and the different courses of actions that should be taken. Generally, the guideline articulates the fundamental concepts associated with assessing security risks within an organisation and an overview of the risk management process.

NIST SP 800-39 "Managing Information Security Risk". The purpose of this publication is to provide guidance for an integrated program for managing information security risks across all levels of organisational operations including reputation, mission, functions, assets, and individuals. It aims to provide complementary enterprise risk management program that supports existing risk-related activities or programs of organisations by providing a structured and flexible approach for managing risks with specific details of assessing, responding to, and monitoring risks continuously.

NIST SP 800-64, Revision 2 "Security Considerations in the SDLC." This publication aims to provide guidelines to assist organisations in incorporating security into the IT systems development process for ensuring a more cost-effective, risk-appropriate security control. It describes the key security roles and responsibilities needed in the development of information systems, as well as the basic understanding of the relationship that exists between information security and SDLC. Overall, the guidance focuses on the security aspects of SDLC.

NIST SP 800-82 "Guide to ICS Security." This publication focuses on providing guidance for ensuring the protection and security of systems that perform control functions such as ICS, SCADA systems, and Distributed Control Systems (DCS). It elaborates the typical overview of ICS, identifies the common threats and vulnerabilities to these systems, and provides different methods, techniques, and recommendations for mitigating the associated risks and security ICS.

NIST SP 800-150 "Guide to Cyber Threat Information Sharing". This publication intends to provide guidance to organisations on gathering, exchange, and sharing information on cyber threats to CI. It addresses the process for sharing of cyber threat information within an organisation, for using cyber threat information received from external sources, as well as for producing threat information that can be shared with other organisations. The publication provides the basic concepts of threat information sharing, the benefits of sharing, challenges associated with sharing capabilities, including important considerations for active participation and sharing relationship between organisations.

NIST SP-184 “Guide for Cybersecurity Event Recovery”. The purpose of this publication is to support organisations in improving their cyber event recovery plans, processes, and procedures to resume normal operations in times of a disaster. The publication aims to extend existing NIST guidelines regarding incident response by providing more detailed and actionable information guidelines on planning, preparing, and recovering from a cyber event, achieving continuous improvement of recovery capabilities as well as integrating these processes into an organisation’s risk management plan.

4.1.2 Health Care domain management standards

ISO14971: Medical devices — Application of risk management to medical devices. The standard ISO 14971 (European version EN ISO 14971) concerns itself with the application of risk management to medical devices including software. The requirements described in the standard provide manufacturers with a framework within which experience, insight and judgement are applied systematically to manage the risks associated with the use of medical devices.¹ The standard covers the whole product lifecycle including post-production.

Certification according to ISO 14971 can be used as step towards certification according to ISO 13485 (Medical devices -- Quality management systems -- Requirements for regulatory purposes)². Which itself can be a step towards fulfilling market-specific regulations, e.g., the Medical Devices Directive 93/42/EEC of the European Union³.

The third edition of ISO 14971 has been published in December 2019. ISO 14971 states the following requirements on the risk management for medical devices⁴:

- a. The manufacturer has to establish, implement, document, and maintain a risk management process. (Chapter 4.1)
- b. The leadership of the manufacturer has to take responsibility for providing enough resources for risk management and to delegate risk management to competent staff. Furthermore, it has to define a policy of risk acceptance criteria, which are based upon relevant regulatory demands. (Chapter 4.2)
- c. Staff planning and implementing risk management has to be qualified accordingly. (Chapter 4.3)
- d. The manufacturer has to establish and document a risk management containing all risk management activities during the product lifecycle. (Chapter 4.4).
- e. The manufacturer has to keep a risk management file, which documents all identified risks or dangers how they were processed to facility traceability of all risk related work. (Chapter 4.5) Especially it must be documented or referred to some documentation, how the following activities were conducted:

¹ <https://www.iso.org/obp/ui/#iso:std:iso:14971:ed-3:v1:en>

² ISO 13485, https://en.wikipedia.org/wiki/ISO_13485

³ ISO 14971, https://en.wikipedia.org/wiki/ISO_14971

⁴ Medizinprodukte - Anwendung_des_Risikomanagements_auf_Medizinprodukte (ISO_14971:2019); Deutsche_Fassung_EN_ISO_14971:2019

- risk analysis
- risk evaluation
- implementation and verification of measures to control risk
- evaluation of remaining risks

In the following, the standard ISO 14971 details the four aforementioned activities. Before a product release, it further requires a validation of the whole risk management process (Chapter 9). Lastly, it mandates activities during and after production (Chapter 10).

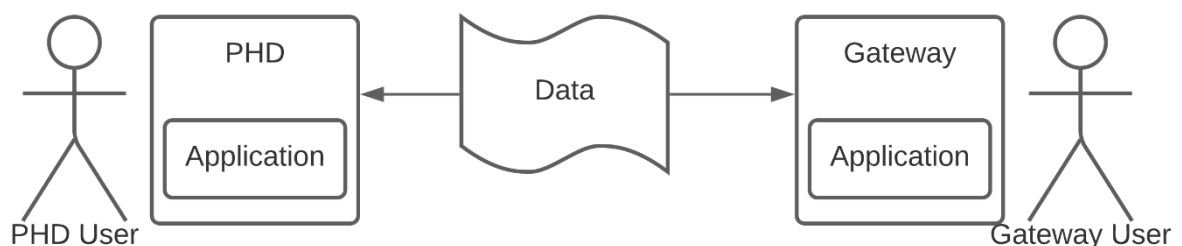
At first, the activity risk analysis has to identify and describe the covered medical device, document the personal and or organization performing the risk analysis, and it has to describe risk analysis itself. Furthermore, it has to define the assigned purpose of the product and reasonably foreseeable misuse. In addition, it has to define safety relevant properties. If applicable safety thresholds have to be defined. Hazards are then to be identified based on the assigned purpose, the reasonably foreseeable misuses and the safety relevant properties. For each of the identified hazards, the manufacturer has assessed the resulting risks. (Chapter 5)

In the evaluation of risks (Chapter 6), ISO 14971 requires evaluation of the assessed risks with regard to the risk acceptance criteria of the risk management plan.

To control risks, the standard mandates (Chapter 7) to analyse the possible options to handle risks and to select options that are to be applied. The selected options have to be implemented and their implementation has to be validated. If the remaining risks after mitigation are still not acceptable, a risk benefit analysis has to be performed. Every mitigation for a risk has to be analysed for new risks arising from its introduction. Lastly, the manufacturer has to ensure that all identified hazards are either mitigated or acceptable. Chapter 8 requires to determine the overall remaining risks⁵.

ISO/TR 22696. ISO-TR 22696 was released in May of 2020 and the main purpose of the document is to provide guidance for managing healthcare service security with connectable personal health devices (PHDs).

The document uses the CIA concept (confidentiality, integrity, and availability) to define cybersecurity focus. In chapter 5, authors state that it is not easy to define which of the three aspects are the most important in the healthcare domain and that all three should be considered equally valuable.



⁵ Medizinprodukte - Anwendung_des_Risikomanagements_auf_Medizinprodukte (ISO_14971:2019); Deutsche_Fassung_EN_ISO_14971:2019

Figure 22 PHD-to-gateway Communication Model

Chapter 6 describes the security vulnerabilities and threads of PHDs. A schematic representation of the bi-directional PHD-to-gateway⁶ communication model is depicted in Figure 22. The document defines 5 possible attack surfaces in that model: physical devices or gateway, users, application, network, and data. The following list contains the attack surfaces which are focused in ISO/TR 22696 with the respective security threats:

1. Physical devices or gateway: jamming scrambling, eavesdropping, exhaustion.
2. Users: device lost or stolen, unskilful device control, malicious intention, social engineering, failure in human resources security.
3. Application: hardcoded password, simple password, malware, reverse engineering, firmware re-flashing, air-gap attack

In chapter 7 it introduces three main objectives on how to prevent the threats.

- “... ensure that the person or entity who has access to devices, PHI, or resources is the legitimate user or entity in accordance with the level(s) access.” (section 7.2.1)
- “... ensure accuracy and consistency of applications for PHDs and gateway.” (section 7.3.1)
- “... allow only authorized people or entities to access devices, PHI, or resources in accordance with level(s) of access...” (section 7.4.1)

Each of them has a subset of the recommendations and the implementation guidance.

For a person or entity, mutual identification, and authentication a procedure on user or entity registration should be established (section 7.2.2). Additionally, all human users (section 7.2.4) and devices (section 7.2.3) should be uniquely identified and authenticated.

To ensure the accuracy and consistency of applications for PHDs and gateway, they should be uniquely identified and authenticated (section 7.3.2). Any unauthorized change in them and information should be detected, recorded, reported, and protected through integrity verification mechanisms (section 7.3.3). To introduce an upgrade to an application and firmware the appropriate security policies and procedures should be established (section 7.3.4). All input data should be verified to prevent malicious tampering attempts (section 7.3.5) and information (stored and transmitted) confidentiality should be protected (section 7.3.6).

To achieve access control secure log-on mechanism should be implemented (section 7.4.2). A special account should be implemented for the emergency cases (section 7.4.3). In a case of inactivity for a defined period, the user should be re-identified and re-authenticated in the system (section 7.4.4). The document also discusses the recommendations in the case of loss or theft (section 7.4.5).

IEC 80001 is a norm to describe the risk management when running IT systems in hospitals and other healthcare providers. It includes requirements for risk management for medical IT networks (MIT), i.e., networks that contain at least one medical device. IEC 80001 is not required by law but describes the state-of-the-art of risk management concerning MIT. It describes the following aspects:

⁶ Gateway - relay mechanism that attaches to two (or more) computer networks that have similar functions but dissimilar implementations and that enables host computers on one network to communicate with hosts on the other

- The hospital/healthcare provider manager should define a risk management strategy, introduce and control risk management processes, name a dedicated risk manager
- The organization should document responsibilities, products and networks
- The risk manager should process the risk management, collect and process relevant information, and conveys between external partners, IT providers, and internal departments
- The product provider should give information on their product (e.g., information flow in the network)

The main objectives of IEC 80001 are to find risks, to assess risks, to control risks, to re-evaluate risks.

ISO 13606 is a standard with the main objective to define a rigorous and stable information architecture for communicating part or all of the Electronic Health Record (EHR) of a single subject of care (i.e. patient). The communication can be between EHR systems, between EHR systems and a centralized EHR data repository. ISO 13606 is also applicable for communication between an EHR system and clinical applications that need to access EHR data. All communication approaches are reached by a Dual Model architecture, which defines a clear separation between information and knowledge. Information is structured through a Reference Model; knowledge is based on archetypes – formal definitions of clinical information models, e.g., discharge reports or glucose measurements. The Reference Model represents data instances and the Archetype Model semantically describes those data.

UK National Health Service (NHS) Data Security Standards. All NHS digital, data and technology services should achieve the Data Security Standards required through the Data Security and Protection Toolkit (DSPT)⁷. DSPT is an online tool that enables relevant organisations to measure their performance against the data security and information governance requirements mandated by the Department of Health and Social Care (DHSC), notably the 10 data security standards set out by the National Data Guardian in the 2016. The self-assessment is accomplished through confirming assertions and providing supporting evidence. Health and social care organisations complete the DSPT as an online self-assessment against the National Data Guardian Standards. They are required to complete the self-assessment every financial year. The self-assessment provides the organisations with a level of Standards Not Met, Standards Met or Standards Exceeded. Once organisations complete their self-assessment, they publish the result. They are required to publish every financial year but can publish more often if the self-assessment have changed.

These Standards along with their relevant mandatory assertions are:

S1. All staff ensure that personal confidential data is handled, stored and transmitted securely, whether in electronic or paper form

- There is senior ownership of data security and protection within the organisation

⁷ <https://www.dsptoolkit.nhs.uk>

- There are clear data security and protection policies in place, and these are understood by staff and available to the public
- Individuals' rights are respected and supported (GDPR Article 12-22)
- Records of processing activities are documented for all uses and flows of personal information (GDPR Article 30 and DPA 18 Schedule 1 Part 4)
- Personal information is used and shared lawfully
- The use of personal information is subject to data protection by design and by default
- Effective data quality controls are in place and records are maintained appropriately
- There is a clear understanding and management of the identified and significant risks to sensitive information and services

S2. All staff must understand their responsibilities under the Data Security Standards, including their obligation to handle information responsibly and their personal accountability for deliberate or avoidable breaches.

- Staff are supported in understanding their obligations under the National Data Guardian's Data Security Standards

S3. All staff complete annual security training that is followed by a test, which can be re-taken unlimited times, but which must ultimately be passed. Staff are supported by their organisation in understanding data security and in passing the test. The training includes a number of realistic and relevant case studies.

- There has been an assessment of data security and protection training needs across the organisation
- Staff pass the data security and protection mandatory test
- Staff with specialist roles receive data security and protection training suitable to their role
- Leaders and board members receive suitable data protection and security training

S4. Personal confidential data is only accessible to staff who need it for their current role and access is removed as soon as it is no longer required. All access to personal confidential data on IT systems can be attributed to individuals.

- The organisation maintains a current record of staff and their roles
- Organisation assures good management and maintenance of identity and access control for its networks and information systems

- All staff understand that their activities on IT systems will be monitored and recorded for security purposes
- You closely manage privileged user access to networks and information systems supporting the essential service
- You ensure your passwords are suitable for the information you are protecting

S5. Processes are reviewed at least annually to identify and improve processes, which have caused breaches or near misses, or which force staff to use workarounds which compromise data security.

- Process reviews are held at least once per year where data security is put at risk and following data security incidents
- A confidential system for reporting data security and protection breaches and near misses is in place and actively used

S6. Cyber-attacks against services are identified and resisted and NHS Digital Data Security Centre security advice is responded to. Action is taken immediately following a data breach or a near miss, with a report made to senior management within 12 hours of detection.

- All user devices are subject to anti-virus protections while email services benefit from spam filtering and protection deployed at the corporate gateway
- Known vulnerabilities are acted on based on advice from CareCERT, and lessons are learned from previous incidents and near misses

S7. A continuity plan is in place to respond to threats to data security, including significant data breaches or near misses, and it is tested once a year as a minimum, with a report to senior management.

- Organisations have a defined, planned and communicated response to Data security incidents that impact sensitive information or key operational services
- There is an effective test of the continuity plan and disaster recovery plan for data security incidents
- You have the capability to enact your incident response plan, including effective limitation of impact on your essential service. During an incident, you have access to timely information on which to base your response decisions

S8. No unsupported operating systems, software or internet browsers are used within the IT estate.

- All software and hardware has been surveyed to understand if it is supported and up to date
- Unsupported software and hardware is categorised and documented, and data security risks are identified and managed
- Supported systems are kept up-to-date with the latest security patches

- You manage known vulnerabilities in your network and information systems to prevent disruption of the essential service

S9. A strategy is in place for protecting IT systems from cyber threats which is based on a proven cyber security framework. This is reviewed at least annually. NHS Digital Data Security Centre assists risk owners in understanding which national frameworks do what, and which components are intended to achieve which outcomes.

- All networking components have had their default passwords changed
- A penetration test has been scoped and undertaken
- Systems which handle sensitive information or key operational services shall be protected from exploitation of known vulnerabilities
- You securely configure the network and information systems that support the delivery of essential services
- The organisation is protected by a well-managed firewall

S10. IT suppliers are held accountable via contracts for protecting the personal confidential data they process and meeting the Data Security Standards.

- The organisation can name its suppliers, the products and services they deliver and the contract durations
- Basic due diligence has been undertaken against each supplier that handles personal information in accordance with ICO and NHS Digital guidance

ISO/IEC 81001-1 Health software and health IT systems safety, effectiveness and security (current status under publication). This standard focuses on the importance of information transfer as a product moves from manufacturer to implementer & integrator to user, identifying and defining also common terms to harmonize the definitions used across the lifecycle where possible. This information would relate to risk, usability, configuration, and other important information that is necessary for stakeholders to transfer and maintain ownership of the product.

Whilst the version is under final approval before publication, Part 1 'Principles and concepts' states that 'managing safety, effectiveness and security for health software and health IT systems (including medical devices), requires a comprehensive and coordinated approach to optimizing safety, effectiveness, and security.

The health care sector is a very complex one as several different stakeholders with separate roles are involved throughout the life cycle of health software and health IT systems. According to the standard, the lifecycle of a product development can be divided into three phases:

1. 'Design and Development Phase' where the identified accountability is with the manufacturer and includes the following steps:
 - a. Concepts and requirements definitions

- b. Design
 - c. Development
 - d. Testing, Verification, Documentation and
 - e. Production and Release
- 2. 'Implementation Phase' where the accountability sits with the health care delivery organisation and includes the following steps:
 - a. Acquisition
 - b. Installation, Customisation and Configuration
 - c. Integration, data migration, transition, and validation
 - d. Implementation, workflow optimisation and training
- 3. 'Clinical Use Phase' where accountable is again the health care delivery organisation and includes:
 - a. Operations and maintenance
 - b. Decommissioning

The framework identifies two main core themes and includes terms, definitions, and concepts. The two core themes are Governance and Knowledge transfer.

The Governance includes:

- 1. Organisation culture, roles, and competencies
- 2. Quality management
- 3. Information management
- 4. Human factors/ usability

The Knowledge transfer includes:

- 1. Risk management
- 2. Safety management
- 3. Security management
- 4. Privacy management

According the standard's scope, it can be used by all stakeholders involved in the 'health software and health IT systems life cycle' including:

- Organizations, health informatics professionals and clinical leaders (including health software developers)
- medical device manufacturers, system integrators, system administrators
- Healthcare service delivery organizations, healthcare providers and others who use these systems in providing health services
- Governments, commissioners, monitoring agencies, professional organizations and customers seeking confidence in an organization's ability to consistently provide safe, effective and secure health software, health IT systems and services
- Organizations and interested parties seeking to improve communication in managing safety, effectiveness and security risks through a common understanding of the concepts and terminology used in safety, effectiveness and security management
- Providers of training, assessment or advice in safety, effectiveness and security risk management for health software and systems
- Developers of related safety, effectiveness and security standards.

4.1.3 Best Practices and Guidance

US Food & Drug Administration (FDA). In the US, the FDA⁸ has the responsibility to protect public health among others by ensuring the safety, efficacy, and security of drugs, biological products and medical devices. The FDA have regulatory responsibilities and enforce relevant laws and regulations, approve FDA-regulated products, and provide guidance documents as well.

In a medical device security guidance⁹, the FDA instructed manufacturers to include cybersecurity risks assessment during the design and development of their devices. This guidance was not an enforceable regulation but informed manufacturers of established best practices and of cybersecurity issues that should be addressed. The guidance states that “cybersecurity threats to the healthcare sector have become more frequent, more severe, and more clinically impactful” and observed that recent cybersecurity attacks have made medical devices and hospital networks inoperable and led to delays and disruption with the potential to cause patient harm. The FDA regards medical device security as a shared responsibility among health care facilities, patients, health care providers, manufacturers of medical devices, and other relevant stakeholders.

As part of the software validation and risk analysis required by 21 CFR 820.30(g), software manufacturers are advised to include a cybersecurity vulnerability and management approach,

⁸ fda.gov

⁹ <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/content-premarket-submissions-management-cybersecurity-medical-devices>

including cybersecurity controls that maintain safety and effectiveness, where appropriate. Manufacturers are advised to apply a risk-based approach when determining the security-relevant design features and the level of cybersecurity resilience required. A Cybersecurity Bill of Materials (CBOM) is considered a “critical element in identifying assets, threats, and liabilities”.

The FDA advises to decide on the needed security controls based among others on the intended use, the functionality of the data interfaces, the type of cybersecurity vulnerabilities, the exploitability of the vulnerability and the risk of patient harm in the case of a breach.

Several key elements were proposed to be considered when addressing cybersecurity: (1) identification of assets, threats, and vulnerabilities; (2) assessment of the impact of threats and vulnerabilities on functionality and end-users; (3) assessment of the likelihood that a vulnerability is exploited; (4) identification of risks levels and mitigations; and (5) assessment of the residual risk and risk acceptance criteria.

According to their cybersecurity risks based on the above elements, the FDA identified two tiers of devices: (1) Higher Security Risk and (2) Standard Security Risk. For Tier 1 devices, pre-market submitted documentation should demonstrate how the device design and risk assessment incorporate the cybersecurity design controls. For Tier 2 documentation should either demonstrate that the specific design features and cybersecurity design controls are included or provide a risk-based rationale for why the specific cybersecurity design controls are not appropriate.

The key design controls are as follows:

- Identify and protect device assets and functionality
 - Prevent unauthorized use
 - Ensure trusted content by maintaining code, data, execution integrity
 - Maintain integrity of data
- Detect, respond, recover
 - Design the device to detect cybersecurity events in a timely fashion
 - Design the device to respond to and contain the impact of a potential cybersecurity incident
 - Design the device to recover capabilities or services that were impaired due to a cybersecurity incident

This FDA guidance, for which compliance is voluntary, is expected to have a significant role in improving cybersecurity¹⁰.

¹⁰ <https://www.databreachtoday.com/fda-issues-more-medical-device-security-guidance-a-8805>

Health Insurance Portability and Accountability Act (HIPAA). HIPAA¹¹ is a US federal law that required the creation of standards to protect sensitive patient health information from disclosure without patient consent or knowledge. The US Department of Health and Human Services (HHS) issued the HIPAA Privacy Rule to implement the requirements of HIPAA. The Privacy Rule standards address the use and disclosure of individuals' health information ("protected health information") by entities subject to the Privacy Rule ("covered entities"). It specifies permitted uses and disclosures when data can be shared without an individual's authorization (e.g., for public health purposes, when required by law, and in several other well-specified situations). In all other cases, consent needs to be obtained or the data needs to be adequately de-identified.

The HIPAA Privacy Rule standardizes as well data de-identification to protect patients' data and to prevent identity disclosure following the release of patient data for secondary use. HIPAA proposes two de-identification methods, Expert Determination and Safe Harbour. This standard has global relevance as is the most prescriptive standard for data de-identification and can be effectively translated into policies, procedures, and processes. Both methods have been widely implemented, with the Expert Determination method reaching increased adoption recently due to its 4 key characteristics: (1) applies generally-accepted statistical or scientific principles, (2) quantifies the risk for re-identification and limits it to a very small risk deemed acceptable, (3) documents and reports on the process and on the results, (4) is carried out by an expert. This methodology allows for the risk to be quantified and effectively balances risk with data utility¹².

The HIPAA Privacy Rule safeguards Protected Health Information (PHI). The HIPAA Security Rule protects a subset of information covered by the Privacy Rule, i.e., all individually identifiable health information a covered entity creates, receives, maintains, or transmits in electronic form. To comply with the HIPAA security rule, a covered entity must comply with several requirements¹³.

- Ensure confidentiality, integrity, and availability of all electronic PHI
- Detect and safeguard against anticipated threats to the security of the PHI
- Protect against anticipated impermissible uses or disclosures
- Certify compliance by their workforce

In the Security Rule confidentiality means that electronic PHI is not available or disclosed to unauthorized persons. Integrity can be defined as the requirement that e-PHI is not altered or destroyed in an unauthorized way. Availability means that e-PHI is accessible and usable on demand by authorized persons¹³. The covered entities are enabled to decide which security measures to use, but they need to consider (1) their size, complexity, and capabilities, (2) their technical, hardware, and software infrastructure, (3) the costs of the planned security measures, and (4) the likelihood and

¹¹ <https://www.cdc.gov/phlp/publications/topic/hipaa.html>

¹² K. El Emam and L. Arbuckle, "Anonymizing Health Data", O'Reilly, 2013

¹³ <https://www.hhs.gov/hipaa/for-professionals/security/laws-regulations/index.html>

impact of potential risks to e-PHI. The security measures need to be regularly reviewed to ensure continuous protection that deals with changes in the environment.

The covered entities need to carry out risk analysis as an ongoing process, to review records, track access and detect incidents. Both the effectiveness of the security measures and the potential risks need to be regularly assessed. Covered entities are required as well to put in place administrative, physical and organizational safeguards to protect the e-PHI. The rule requires as well that appropriate policies and procedures that are in place, are adequately documented. The documentation is reviewed and updated periodically.

EU Regulation 2017/745 on Medical Devices = Medical Device Regulation (MDR) or European Medical Devices Regulation.

This Regulation applies in all EU member states and it repeals Directive 93/42/EEC, concerning medical devices and Directive 90/385/EEC concerning active implantable devices. Regulation 2017/745 focuses on:

- Unified designation and control of Notified EU Bodies based on concrete requirements
- Creation of a coordination group (Medical Devices Coordination Group, MDCG) consisting of Notified Experts from all EU member states
- Implementation of a mean of control for the conformity assessment of medical devices with high risks by including a panel of experts (scrutiny approach)
- Detailing the requirements for a clinical assessment
- Detailing the process of allowing clinical assessments of medical devices and performance studies of in-vitro-diagnostics
- Stricter regulations on vigilance system
- Rules concerning the re-use of medical one-time products
- Provision of a Unique Device Identification number (UDI)
- Widening the European database for medical devices and in-vitro diagnostics (EUDAMED) and providing partially public access to the EUDAMED
- New classification rules for in-vitro diagnostics so that it gets similar to the four-classes-system of medical devices
- Inclusion of European reference labs to assess in-vitro diagnostics belonging to the highest class of risk
- Introduction of a concept to clinically assess in-vitro-diagnostics

ENISA: The European Union Agency for Cybersecurity (ENISA) is the EU's agency for achieve a high common level of cybersecurity across Europe. For this, ENISA cooperates closely with the EU member states and other stakeholders. It aims to provide advice and solutions and to improve the member states' cyber-security capabilities. Furthermore, ENISA supports the development of cooperative responses to large-scale cyber-security incidents crossing national borders. The agency aims is to provide a centre of expertise for member states and EU institutions (e.g., the Commission) where it is possible to seek advice.

4.1.3.1 Incident handling of Medical Devices

The Article 2 of Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017¹⁴ (Medical Device Regulation (MDR)) defines a “medical device” as *any instrument, apparatus, appliance, software, implant, reagent, material or other article intended by the manufacturer to be used, alone or in combination, for human beings for one or more of the following specific medical purposes:*

- *diagnosis, prevention, monitoring, prediction, prognosis, treatment or alleviation of disease,*
- *diagnosis, monitoring, treatment, alleviation of, or compensation for, an injury or disability,*
- *investigation, replacement or modification of the anatomy or of a physiological or pathological process or state,*
- *providing information by means of in vitro examination of specimens derived from the human body, including organ, blood and tissue donations,*

and which does not achieve its principal intended action by pharmacological, immunological or metabolic means, in or on the human body, but which may be assisted in its function by such means.

The Annex I of the MDR also states that *manufacturers shall set out minimum requirements concerning hardware, IT networks characteristics and IT security measures, including protection against unauthorised access, necessary to run the software as intended.*

With the purpose of supporting healthcare stakeholders in respecting the regulations and the requirements of the MDR, the *Medical Device Coordination Group* (MDCG) (MDR Article 103) has been established by EC. This Group is composed of representatives of all Member States and it is chaired by a representative of the European Commission. Endorsed by the EC, MDG is actually providing a set of guidance documents to assist stakeholders in implementing the medical devices that respect the actual regulations¹⁵. Among these documents, the MDGC released in December 2019 the *Guidance on Cybersecurity for medical device*¹⁶ where a comprehensive highlight of incident handling procedures related to medical devices is reported.

In general, in the case of medical device, an incident can be defined as an event that causes, or has the potential to cause, unexpected or unwanted effects involving the health and safety of patients, users or other persons. General incidents in medical devices may arise due to:

- shortcomings in the design or manufacture of the device itself;
- inadequate instructions for use;
- inadequate servicing and maintenance;
- locally initiated modifications or adjustments;
- inappropriate user practice;
- inappropriate management procedures;
- inappropriate environment in which a device is used or stored;
- selection of the incorrect device for the purpose.

¹⁴ <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32017R0745&from=EN>

¹⁵ https://ec.europa.eu/health/md_sector/new_regulations/guidance_en

¹⁶ https://ec.europa.eu/health/md_sector/new_regulations/guidance_en

Focusing on *security incidents* and according to MDR, a security incident is any malfunction or deterioration in the characteristics or performance of a device made available on the market, including use-error due to ergonomic features, as well as any inadequacy in the information supplied by the manufacturer and any undesirable side-effect. Furthermore, MDR distinguishes *serious incident*, defining them as any incident that directly or indirectly led, might have led or might lead to any of the following:

- the death of a patient, user or other person;
- the temporary or permanent serious deterioration of a patient's, user's or other person's;
- state of health;
- a serious public health threat.

A list of examples provided by ¹⁷ on the distinction between incidents and serious incidents arising from medical devices from the point of view of cybersecurity are reported in the next Table, which also shows the corresponding foreseen control measures, security control/incident handling measures and, finally, safety controls that are needed to be implemented in order to eliminate or mitigate the risk of patient harm (safety harm) caused by incidents. In this way, the Table also provides a representation of the relationship between cybersecurity risk management and patient safety management.

Table 4: Examples of medical devices' incidents and corresponding severity, security harm and control and safety harm and control

Serious Incident (Yes/No)	Risk Relationship	Device	Security Harm	Security Control	Safety Harm	Safety Control
Yes	Security risk with a safety impact	External Programmer for an implantable Deep Brain Stimulator	Custom malware is installed on the External Programmer / Modification of External Programmer function, including stimulation parameters.	Establish message authentication between Programmer and IPG and Programmer prevents installation of third-party applications and limits access to the programmer device OS.	Increased, decreased, and/or an intermittent stimulation not intended in the current programming parameters; or, inability to change programs or control the amplitude using the patient programmer.	N/A
Yes	Security risk with a safety impact	External Programmer for an implantable Pacemaker	External Programmer is used by an unauthorized user to adjust therapy settings without the patient's knowledge.	Implement User Authentication on External Programmer.	Increased, decreased, and/or an intermittent stimulation not intended in the current programming parameters.	Inductive Programming Wand is required to start communication session with the IPG (requires close patient proximity)
Yes	Security risk with a safety impact	Implantable Sensor used to monitor	An attacker modifies or creates patient data in transit to or from	Connection protocol from electronics unit to clinician website	Physician fails to treat based on incorrect low PA	N/A

¹⁷ https://ec.europa.eu/health/md_sector/new_regulations/guidance_en

		Pulmonary Artery pressures in Heart Failure Patients	the external electronics unit, causing misdiagnosis that affects patient care.	uses SSL/TLS encryption.	pressure readings leading to worsening of patient's heart failure condition.	
Yes	Security risk control with a safety impact	Pacemaker	An unauthorized person is able to fatigue the device by overwhelming the device of requests.	Avoid possibility to overwhelming the device.	Avoid possibility to overwhelming the device. A premature battery depletion may occur.	N/A
Yes	Security risk control with a safety impact	A smart infusion pump with its remote control	Patient may reconfigure the device.	User type and access right should well be defined.	The smart infusion pump infuses more or less insulin than what was prescribed by an authorized user.	N/A
Yes	Security risk control with indirect safety impact (device availability)	Any Medical Device with Windows OS	Network-spread malware (worm) encrypts the content of the system hard drive.	Disconnect devices from network.	No direct safety harm. (Indirect: MD not available).	Use of alternative devices.
Yes	Security risk with a safety impact	Anaesthesia device	An unauthorized user with physical access to the device guesses the weak password for the service account and manipulates the configuration settings.	Access control without password complexity enforcement.	The anaesthesia device supplies a wrong anaesthetic concentration	N/A
No	Security risk only	Warming therapy device for premature babies	An unauthorized user with physical access to the device guesses the weak password for the service account and exports therapy and patient data via the USB interface.	Access control without password complexity enforcement.	None	N/A
Yes	Security risk with a safety impact	Warming therapy device for premature babies	An attacker floods the network interface with tons of malformed service requests which causes the system to crash.	N/A	The therapy functionality of the device is not available.	N/A
No	Security risk only	Monitoring System	An attacker eavesdrop the network communication between a local patient monitor and the central monitoring station. Therefore the attacker gains possession of sensitive health information of the patient.	N/A	None	N/A
Yes	Security risk with a safety impact	Monitoring System	An attacker with physical access to the network manipulates a ventilator's alarm messages sent to the central monitoring system.	N/A	Emergency measures are not carried in time	N/A

Yes	Security risk with a safety impact	PACS	An unauthorized user gains access to the local network and manipulates the network traffic between a device and the PACS Software.	Network Access Security.	There is the danger of manipulation of medical image data and thus the danger of false diagnoses.	User checks data display directly on device.
Yes	Health damage caused by unavailability	PACS	An unauthorized user deploys malware (ransomware, scareware).	Security Awareness Training, Firewall, Antivirus Solution, secure infrastructure, Backups.	Health damage caused by unavailability.	User checks data display directly on device.
No	Security risk only	PACS	Employee stealing data with mobile USB storage on a client pc.	Implement a User and Usergroup Permission Environment.	None.	N/A
No	No Impact, annoyance of the patient	MR	Network based infection, leading to contaminated system. System performs its functions, but slows down (at same time notifies the operator)	N/A	None	N/A
Yes	Security risk control with a safety and security impact	X-ray Machine	DICOM objects infected with executable malware imported and exported spreading across PACS and medical device network.	Hardening / Whitelisting blocking execution of DICOM objects.	Delayed diagnosis and treatment due to unavailability of compromised networked systems.	N/A

As seen from the previous Table, the severity of a cybersecurity incident arising from a medical device, as well as the security controls to minimise/handle each one of them, depends not only on the medical device itself (class, purpose, application, use, etc.), but also on the specific incident's type. This is also valid from the cybersecurity's point of view: risk mitigation, countermeasures and incident handling procedures are strictly bounded with the device type, purpose and application, as well as with the security violation. For these reasons, the need of identifying specific incident minimisation/management approaches for each device and cyberattack is a great challenge.

Another crucial aspect for the incident handling of medical device is the *post market surveillance and vigilance*, which is mandatory for medical devices manufacturers to be implemented. The rapid evolution and changes of cybersecurity vulnerabilities could make the controls and incident handling procedures implemented during pre-market activities inadequate to maintain an acceptable benefit-risk level. An effective and successful post-market cybersecurity surveillance program should be defined, including the following aspects:

- operation of the device in the intended environment;
- sharing and dissemination of cybersecurity information and knowledge of cybersecurity;
- vulnerabilities and threats across multiple sectors;
- vulnerability remediation;

- incident response.

The post-market surveillance is implemented by the manufacturer by putting in place a *Post Market Surveillance (PMS) system* and actively keeping the PMS system up to date (in accordance with MDR Art. 83). Cybersecurity considerations for medical devices should be part of this PMS system. The PMS system includes the active and regular collection of user experience from devices on the market (including third party software and hardware components), the review this collection and to timely implement necessary corrective action, taking into account the nature and risks in relation to the device. The manufacturer will involve the distributors of the device and, where applicable, the authorised representative and importers of the device in his system, in order to obtain the relevant information from the market. The PMS activities must be supported by a manufacturer's PMS plan (MDR art. 84), where a set of information described in MDR Annex III are reported. In particular, the PMS plan shall address the collection and utilization of available information:

- information concerning serious incidents, including information from PSURs, and field safety corrective actions;
- records referring to non-serious incidents and data on any undesirable side-effects;
- information from trend reporting;
- relevant specialist or technical literature, databases and/or registers;
- information, including feedbacks and complaints, provided by users, distributors and importers; and
- publicly available information about similar medical devices.

The post-market surveillance plan shall cover at least:

- a proactive and systematic process to collect any information referred to the previous pointed list. The process shall allow a correct characterisation of the performance of the devices and shall also allow a comparison to be made between the device and similar products available on the market;
- effective and appropriate methods and processes to assess the collected data;
- suitable indicators and threshold values that shall be used in the continuous reassessment of the benefit-risk analysis and of the risk management as referred to in Section 3 of Annex I of MDR;
- effective and appropriate methods and tools to investigate complaints and analyse market-related experience collected in the field;
- methods and protocols to manage the events subject to the trend report (MDR Art. 88), including the methods and protocols to be used to establish any statistically significant increase in the frequency or severity of incidents as well as the observation period;
- methods and protocols to communicate effectively with competent authorities, notified bodies, economic operators and users;

- reference to procedures to fulfil the manufacturers obligations laid down in MDR Articles 83, 84 and 86;
- systematic procedures to identify and initiate appropriate measures including corrective actions;
- effective tools to trace and identify devices for which corrective actions might be necessary;
- a *Post-Market Clinical Follow-up* (PMCF) plan, as referred to in MDR Part B of Annex XIV, or a justification as to why a PMCF is not applicable.

PMCF is a continuous process that updates the clinical evaluation and shall be addressed in the manufacturer's PMS plan. When conducting PMCF, the manufacturer shall proactively collect and evaluate clinical data from the use in or on humans of a device which bears the CE marking and is placed on the market. The manufacturer as well shall perform this proactive collection and evaluation if the device is put into service within its intended purpose as referred to in the relevant conformity assessment procedure with the aim of confirming the safety and performance throughout the expected lifetime of the device, of ensuring the continued acceptability of identified risks and of detecting emerging risks on the basis of factual evidence.

The PMS report must be prepared, summarizing the results and conclusions of the analysis of all the data from the market. Data gathered from PMS system must be used to actively update:

- the clinical evaluation;
- the benefit-risk determination and to improve the risk management;
- the design and manufacturing information, the instructions for use and the labelling;

Handling and remediation of cybersecurity incidents and vulnerabilities reported through the PMS and vigilance systems shall be carried out conforming to the Security Risk Management procedures, with regard to:

- Assess the need for reporting serious and non-serious incidents and of carrying-out field safety corrective actions;
- Enhancing security capabilities;
- Update the original Security Risk Assessment;
- Update the Verification and Validation;
- Update the original Security Benefit Risk Analysis
- Update the Technical Documentation.

Risk Management is generally understood as the discipline of identifying and measuring risks towards safety and effectiveness resulting from the intended use and foreseeable misuse of a medical device and reducing them “as far as possible” to an acceptable level. The general approach to risk management for medical devices according to the state-of-the-art can be found in the MDR Annex I, Section 3, as well as in relevant harmonized standards published in the Official Journal. Risks related to data and systems security are specifically mentioned within the scope of the risk management process, to avoid any misunderstanding that a separate process would be needed to manage security

risks related to medical devices. Specific methods and requirements are, however, used for security risks. As an example, ‘blanking’ a screen might be an appropriate security control to mitigate the disclosure of personal data, but when the medical device is used for interventional use or the display of vital signs, then ‘blanking’ the screen is a safety concern and thus, it should not be implemented: the challenge in this case is to satisfy both security and safety requirements, which could have contrasting requisites. Security vulnerabilities may affect the product’s safety or effectiveness. A product risk analysis for safety should therefore consider the effects of security vulnerabilities to the essential functioning of the product. The safety risk assessment might list generic security related hazards identified for the product, such as but not limited to: denial of service, execute code, memory corruption, gain information, gain privilege, etc. This is to avoid detailing every possible security attack vector, which does not result in a different hazard for the product.

A clear requirement for medical devices’ incident handling arises from MDR, which states that any risks associated with the operation of medical devices must be acceptable to enable a high level of protection of health and safety. As mentioned above, this can be only achieved through the establishment of an adequate balance between benefit and risk during all possible operation modes of a medical device. To this end, there is a need to consider the relationship between “safety and security” as they relate to risk.

Finally, we remark that also ENISA published in January 2021 a report focused on Cloud Security for Healthcare Services [3], which aims to provide Cloud security practices for the healthcare sector and to identify security aspects. The report includes relevant data protection aspects, to be considered when procuring Cloud services for the healthcare industry, allowing in this way the identification of the main incident sources and suggesting the correct handling procedures for them. It also identifies in a clear manner a reference Cloud architecture, the factors to be considered during risk assessment, and the risk mitigation measures, applying them to a typical use case that can be used to better describe and introduce the incident handling related to medical device. In this use case scenario, medical device data is made available to different stakeholders using Cloud technologies, for example to enable remote patient monitoring for heart disease or diabetes patients. Medical device manufacturers also provide medical device monitoring using Cloud computing technology. In particular, in the use case framed by ENISA in [3] to highlight the cybersecurity risk and incidents related to medical devices, a manufacturer produces a device to measure certain patient data (e.g. a pacemaker measuring heartbeat). The device itself is not able to communicate over the internet. However, it can transfer measurements via Bluetooth to smartphones with an appropriate app from the device manufacturer. The app can then transfer the aggregated measurements for a month to a Cloud file storage provider and share this information with the treating doctor, following the schema depicted in next Figure 23.

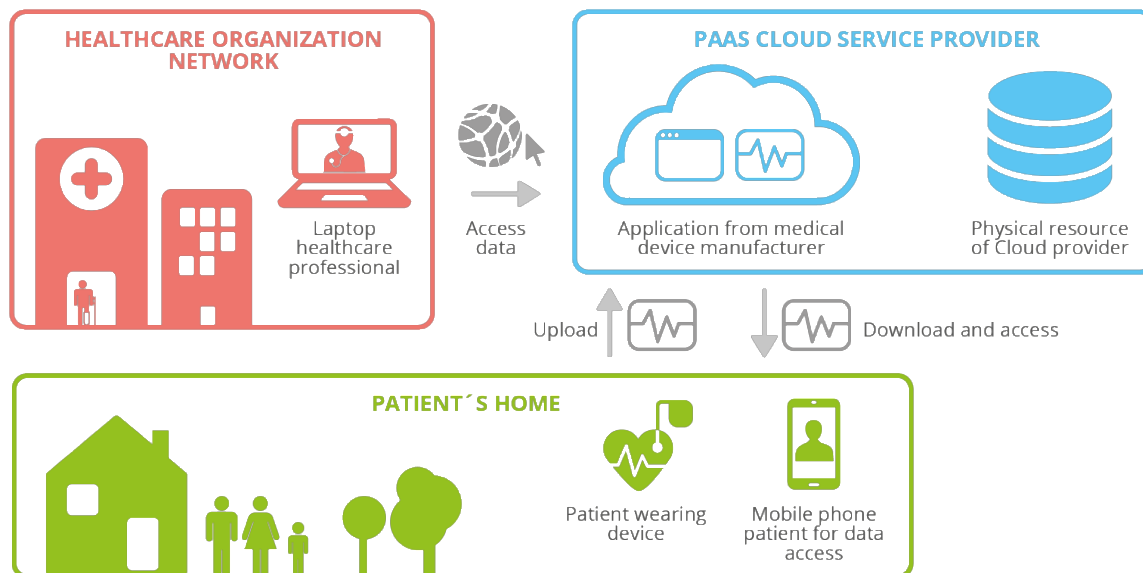


Figure 23: ENISA cloud architecture model for medical device (extracted from [3])

When conducting a risk assessment on use cases such as the one depicted in the Figure, healthcare organisations should be considered the possible impact of a cybersecurity incident on *confidentiality* (e.g., data breach leading to exposed patient data), *integrity* (e.g. alteration of important patient data) and *availability* (e.g., timely access to patient data), concerning the results taken from the literature analysis regarding the incident. This would allow the healthcare organisation to assign an appropriate quantitative or qualitative value to the risk impact depending on the specific risk assessment methodology used. While this specific use case only involves collection of patient data that is then subject to examination by medical staff, other use cases involving medical devices may include the device itself taking actions based on measurements, resulting in a drastically different risk profile.

In summary, the aforementioned risk factors from which a cybersecurity incident related to medical devices can arise, and that allow the identification of the main requirements for incident handling are:

- **Confidentiality:** loss of confidentiality for similar use cases may cause data subjects to encounter significant adverse effects from unauthorised disclosure of their health data. Within the scope of the specific processing operation, the impact from loss of confidentiality is not necessarily considered critical since the disclosure of measurements such as heartbeats is usually not as severe as disclosing other health data. However, if the data is exchanged in its entirety through unsecure means (i.e. email) poses a risk in itself. In a broader context, the impact of loss of confidentiality for use cases involving medical devices depends on the nature of the data involved in the operation.
- **Integrity:** in the case of loss of integrity, data subjects may encounter significant or even irreversible consequences from unauthorized alteration of health data. For instance, doctors may prescribe inappropriate medication. This impact is heavily influenced by the overall

treatment process; for instance, a doctor might notice sudden deviations from regular measurements and doctors usually explain treatment procedures or changes in medication via personal conversations, which might reveal the alteration of data. In the case of more automated processes or even processes where the device can even act based on the data, the impact of loss of integrity may be significantly higher.

- *Availability*: the impact of loss of availability may range from moderate to critical depending on the frequency by which the measurements need to be made available to medical staff or even the nature of the measurements (e.g., when an anomaly in measurements may indicate a life-threatening circumstance). The lack of data may affect the patient's health because unavailability affects intervention options.

As a result of the analysis presented above, the main challenges in preventing and handling medical device security incidents comes from the large diversity in devices (type, class, scope, applications, use, etc.), the need of balancing utility and safety with security and privacy, and their compliance with regulations [3].

4.2 Analysis of healthcare security domain requirements and challenges

4.2.1 Security of AI4HEALTHSEC Circles of consideration

AI4HEALTHSEC defines four Circles of Consideration that enables to structure and deal with the interconnections and complexity of health care infrastructures. The first circle of consideration includes health care components (e.g., implants), the second includes medical devices (e.g., wearables). The third circle encloses the two previous ones and incorporates the individual health care information infrastructures (HCIs). The fourth and outer circle contains all the other circles and represents the interdependent HCIs composing the entire health ecosystem. This ecosystem is a widely distributed, interconnected set of entities (i.e., organizations, individuals or/and CIs), processes and services that relies upon interconnected ICT infrastructures, establishing a dynamic Health Care Supply Chain (HCSC). These HCSCs are characterized by a high degree of complexity and interconnectivity of the ICT systems. The four circles are identified and distinguished based on the homogeneity of characteristics (safety, technical requirements, architectures etc.) identified in each one of them. They are not independent from each other and need to be all secured. Inner circles can be seen as the building blocks of the external ones, meaning that the security of the external circles is directly affected by the inner ones. Thus, the security of the interdependent HCIs and the HCSCs, is directly affected by the security of the individual HCIs that compose it. However, the overall system is not secured by simply securing its “building blocks”, as interdependences between the different layers have their own specificities and require cross layer coordination.

The distinct circles of consideration have a complex and interconnected nature that is characterized by the distribution of services, data sharing, the dynamic nature of collaborations and the significant (inter)dependencies among the involved actors, requiring new approaches for the efficient evaluation and treatment of all internal, external and diffused cyber- threats and risks, the estimation of their cascading effects and the thoroughly investigation of a cybersecurity incident (e.g., collection of evidential data). Novel multi-stage attacks can exploit vulnerabilities of the interconnected ICT systems to cross the organization’s boundaries, enabling the attackers to move within the health ecosystem across multiple critical HCSCs and functions. New approaches are required in order to deal with cascading effects of threats, and propagated vulnerabilities and to react on the security events in their interconnected infrastructures as a sole intelligence. The idea of SI is based on the organizational format observed in natural communities, where individual members perform very simple actions co-operating with one another. These actions gradually accumulate, to form a higher-level intelligence which does not exist in any of the individual members contributing to it.

4.2.2 Analysis of healthcare security domain and challenges

Healthcare organizations are increasingly affected by cybersecurity attacks. These incidents can have numerous devastating effects on healthcare organizations, from the inadvertent release of protected health information to disruptions in clinical care [6]. According to Ponemon institute, “healthcare organizations are in the cross hairs of cyber attackers” that grow increasingly frequent [7]. On average, US healthcare facilities have been victims of one cyber-attack per month, and half of them “have experienced the loss or exposure of patient information during this same period (26% of the other half is unsure)”. Moreover, based on the most recent ENISA report at the end of 2018 [8],

cybersecurity incidents have shown that the healthcare sector is one of the most vulnerable. This phenomenon can be explained by combining two factors: (i) the high value of healthcare facilities' assets and (ii) the ease with which they can be compromised. Medical data is 10-20 times more valuable than financial data for the reason that healthcare records can continue being exploited even after resolving the security breach, which released them. At the same time, the healthcare industry is behind other industries in protecting its infrastructure and data.

Taking to account all the above, IT security in healthcare systems, services and applications are positioned as a major concern due to the high privacy and confidentiality requirements of sensitive healthcare data and faces many security challenges, which are highlighted below [9]:

- **Systems availability:** It is about continuous accessibility of critical health information by authorized professionals in order to ensure the best healthcare services. Systems availability may relate to physical systems function (e.g., networks, storage) and affect significantly the healthcare delivery.
- **Lack of interoperability:** The high-level interoperability aims to guarantee that information of healthcare infrastructures is transmitted safely through individual information systems, health service institutions, healthcare providers and patients. It is important as many diverse systems and applications interconnected at various scales i.e. a medical device collecting clinical data can be linked in the same network that a computer uses to access Internet.
- **Access control and authentication:** Authentication is the initial stage of the users' validation in order to determine their identity, which is necessary to ensure that they are authorized to access the system, which is a key-security feature in healthcare infrastructures [10].
- **Data integrity:** It purposes to ensure the quality and integrity of the data that are stored and exchanged for clinical and administrative purposes; a crucial part of healthcare systems for the reason that errors in personal or clinical data may affect a person's medical treatment, insurance or employability [11].
- **Network Security:** It is a fundamental challenge in securing healthcare infrastructures, especially when the system is network based (e.g., EHR/PHR, cross border eHealth).
- **Security expertise and awareness:** A critical factor, including the adequate and sufficient organisational structure and especially the role of a security officer.
- **Data loss:** It is mentioned to the protection of the data from loss; it is considered a very important part of the healthcare sector as a significant amount of vital, personal, and confidential data is stored in digital format.
- **Incident handling:** As typical security incident handling, it includes the incident response and management, which is the protection of an organisation's information by developing and implementing an incident response process (e.g., plans, defined roles, training, communications, management oversight), in order to quickly discover an attack and then effectively contain the damage, eradicate the attacker's presence, and restore the integrity of the network and systems [12].

4.2.3 Incidents handling of healthcare security domain

Based on ENISA analysis [9] **Errore. L'origine riferimento non è stata trovata.**, incidents handling is one of the major challenges in the healthcare security domain. Although that the majority of organizations implement security policies in their healthcare systems and/or infrastructures, there are incidents that can be neither anticipated nor avoided. In fact, security incidents root causes include, human errors, natural phenomena, malicious actions (DDoS attack, MITM attacks, etc.) and system failures (including third party failure, i.e., hardware failure). It is worthwhile noting that system failures and human errors account equally for most of the incidents reported.

Additionally, deliberate human intervention to disrupt the workflow (i.e. malicious actions) also accounts significantly for security risk. On the other hand, the impact of natural phenomena is responsible for a small only percentage of the reported security incidents. It has to be noted that human factor may also relates to malicious actions, with the prospect of causing system holes through negligence or oversights, which could lead to system failures, hence the infrastructures can be vulnerable to possible attacks. Moreover, the incorrect security practices by personnel are included in human error that can result in security incidents; thus, apart from implementing cyber security measures, awareness raising, and training have a significant role in building a secure system. Therefore, healthcare organisations need to have an incident response capacity, in order to timely identify incidents and restore and reconstitute systems and services in a trusted manner. Indeed, there is a vital need for the development of a healthcare specific incident reporting, classification, and alerting mechanism in pan European level. International good practices could be consulted towards this direction.

In [13], NIST presents a guide to the Cyber-Incident Handling process that is flexible and adaptable, and it can be used by healthcare organizations as a baseline. Specifically, it proposes an incident response Life Cycle, which has four phases, as highlighted in Figure 24.

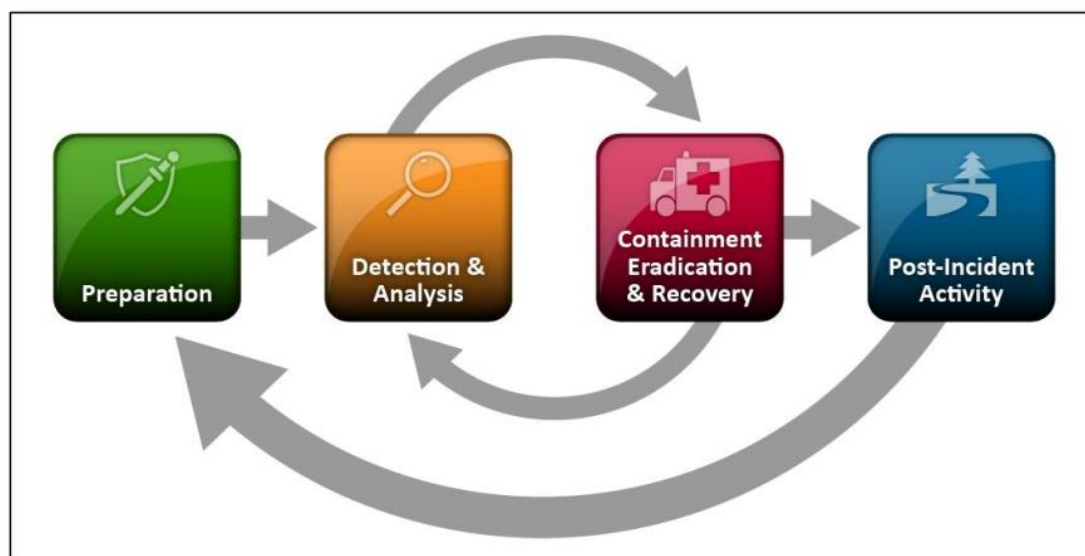


Figure 24: Incident Response Life Cycle [13]

The initial phase involves both the establishment and training of an incident response team, and acquiring the necessary tools and resources. Additionally, during preparation, the organization aims to limit the number of incidents that will take place by selecting and implementing a set of controls based on the results of risk assessments. Although, the residual risk will not be eliminated after controls are implemented. Thus, the phase of detection of security breaches is necessary in order to alert the organization whenever incidents occur. Depending on the severity of the incident, the organization can mitigate the impact of the incident by containing it and eventually recovering from it. During this phase, an activity often returns to detection and analysis - for example, to determine if additional hosts have been infected with malware while a malware incident is being eliminated. After the adequate handling of the incident, the organization issues a report detailing the cause and cost of the incident and the stages that the organization should follow to prevent future incidents. The major phases of the incident response process are analysed below [14]:

- **Preparation:** It contains the steps that are taken before an incident occurs, such as training, writing incident response policies and procedures, and providing tools such as laptops with sniffing software, crossover cables, original OS media, removable drives, etc. In fact, preparation should include anything that may be required to handle an incident or will make incident response faster and more effective.
- **Detection and analysis:** It is the phase in which events are analysed in order to determine whether these events might comprise a security incident.
- **Containment, eradication and recovery:** The containment phase of incident response is the point at which the incident response team attempts to keep further damage from occurring as a result of the incident (i.e. taking a system off the network, isolating traffic, powering off the system). The eradication phase involves the process of understanding the cause of the incident so that the system can be reliably cleaned and ultimately restored to operational status later in the recovery phase. The recovery phase involves cautiously restoring the system or systems to operational status.
- **Post-incident activity:** It includes the creation of a follow-up report and each incident response team should evolve to reflect new threats, improved technology, and lessons learned aiming to reduce the probability of a similar incident happening again and to improve incident handling procedures.

Furthermore, ENISA in [15] suggests, as one of the most effective ways to address cybersecurity threats, the creation of a global ecosystem of Computer Security Incident Response teams (CSIRTs) and security operations centres which should communicate, share information, and respond to cyber-threats effectively. Specifically, it is mentioned that the teams responsible for incident response handling are CSIRT, CERT, and SOC. CSIRT has become a generic name for a team that is involved in a set of services such as information and cybersecurity incident handling (core service), security monitoring, vulnerability management, situational awareness, and cybersecurity knowledge management. SOC provides an incident detection service by observing technical events in networks

and systems and can also be responsible for incident response and handling. In fact, in large enterprises, SOC's most of the times focus only on monitoring and detection services and then hand over incident handling to a separate CSIRT. On the other hand, in smaller organisations, CSIRTs and SOC's are often considered as the same team.

The establishment of CSIRTs includes five different phases that are highlighted in and overviewed as following [15]:

- **Assessment for readiness:** It is the beginning of the establishment of a new CSIRT, that contains a discussion about the reasons and necessity for establishing a CSIRT and the approval of an initial budget and shaping requirements for the design phase.
- **Design:** This phase identifies detailed plans for the next step and its prerequisites are all of the outcomes from the assessment for readiness phase.
- **Implementation:** It covers organisational matters: governance, people, processes, services and technology.
- **Operations:** During this phase, a CSIRT delivers the CSIRT services.
- **Improvement:** It is the phase that a CSIRT formulates requests for improvements, prioritises initiatives and receives an approved budget for following the 'design–implementation–operation–improvement' cycle. It should be noted that the existing CSIRTs can follow the guidelines from this phase rather than from the assessment for readiness phase.



Figure 25: Lifecycle of CSIRT [15]

Both of the above solutions can be adapted in order to cover the AI4HEALTHSEC requirements and needs of the pilots that are associated with security incident management.

4.2.4 Risk management of healthcare security domain

Within Horizontal Layer 1: Risk and Privacy management & Cyber Attack Forecasting Based on the “assumed breach” approach, the HCII’s resilience risk management principles and practices are applied to identify and prioritize current and emerging threats, risk, and potential evidence source and type. We follow an evidence-driven Risk Assessment model to capture and deal with cascading effects, risks, threats, and vulnerabilities, associated with the interdependent HCII's. The layer

incorporates a set of security and privacy processes, including threat assessment, identification, evaluation of all risks, impact analysis and estimation of the propagated effects for risk mitigation, to provide the risk and privacy assessment performance in accordance with existing security standards and regulations (e.g., ISO27001, ISO27005, ISO28000, GDPR). Next, in order to mitigate the identified risks, security measures are implemented.

AI4HEALTHSEC aims to support organizations to leverage security and privacy information for risk assessment and for limiting the risk in an optimal way. The project will develop data mining techniques and models to detect, evaluate and prioritize cybersecurity and privacy risks. Healthcare organisational communication abnormalities can be detected by collecting, analysing, correlating and sharing all individual risks pertaining to the HCII environments. The approaches will provide high detection accuracy in the risk assessment process to better evaluate, model, anticipate, treat and predict risks and security incidents and draw meaningful insights from sophisticated, multidimensional cyber-attacks.

AI4HEALTHSEC will develop as well visualization methods and forecast propagation and cascading effects of attacks in Interdependent HCIs and anticipate how attacks propagate across the HCSCS. The project will develop propagation models for describing spreading of failures in the physical and cyber systems, and to describe and analyse potential threat and attack paths for complex threat scenarios.

Artificial Intelligence/machine learning techniques and propagation models facilitate the analysis and correlation of security and risk-related information in order to achieve a two-fold goal: (i) to optimize the risk estimation, evaluation and strategy mitigation processes; and (ii) to verify the compliance of the healthcare business processes and practises with the European and national privacy and data protection regulations (e.g., GDPR). Specifically, the AI4HEALTHSEC approach will integrate information retrieved from the underlined infrastructures, knowledge and data collected from multiple distributed sources as well as evidence, proofs and findings generated from experiments and simulations models in order to be used for the risk and privacy assessment process. Simulation models will produce timely, accurate, objective and high-quality evidence based on which the multi-dimensional risks and threats associated with the HCIs will be assessed. These models will give the ability to healthcare operators to experiment on security scenarios allowing them to investigate further the risk of cascading effects on their assets. The simulation environment of the AI4HEALTHSEC will encapsulate big data analytics and machine learning techniques to support the attack-path graphs generation giving better insight to the attacking process and increasing the understanding of the healthcare interconnected environment and the relevant cyber security and privacy threats.

Within the last decade, the focus of supply chain security experts regarding the security, privacy and data protection turns on harvesting large data sets and using advanced machine learning techniques to train sophisticated intrusion detection models, analyse data travelling in the networking infrastructures and subsequently detect in near-real-time or real-time breaches and threats. An important technique to detect the presence of an APT is the analysis of file signatures. Nevertheless, due to the surreptitious nature of the APT, the pattern of the attack keeps on changing, and this hardens the process of detecting them on a real-time basis, unless a match is found with the previously known APT repository. Dynamic detection of such attacks is possible by applying unique characteristics of an APT to train the machine learning algorithm. Moreover, machine learning

algorithms can facilitate data automation collection and processing; they can promote integration to the existing HCII, elicit unstructured data from disparate sources to provide context on IoCs and TTPs of threat agents.

5 Results: Healthcare security management solutions relevant to AI4HealthSec

Based on the process that we defined in Section 2.4, in this section, we present a set of existing tools and components, provided by the AI4HealthSec partners, which have been used to address cyber security and attacks related scenarios in various business domains, including e-health, and they can be placed together to eventually develop the envisaged framework. The scope is to highlight technology related challenges, arising from the integration of the different solutions to implement the risk assessment and incident handling processes, and analyse them with respect to the business needs and domain requirements that have been elicited from the work presented in Sections 3 and 4.

5.1 Evidence-driven Maritime Supply Chain Risk Assessment (MITIGATE) System

5.1.1 Short Description

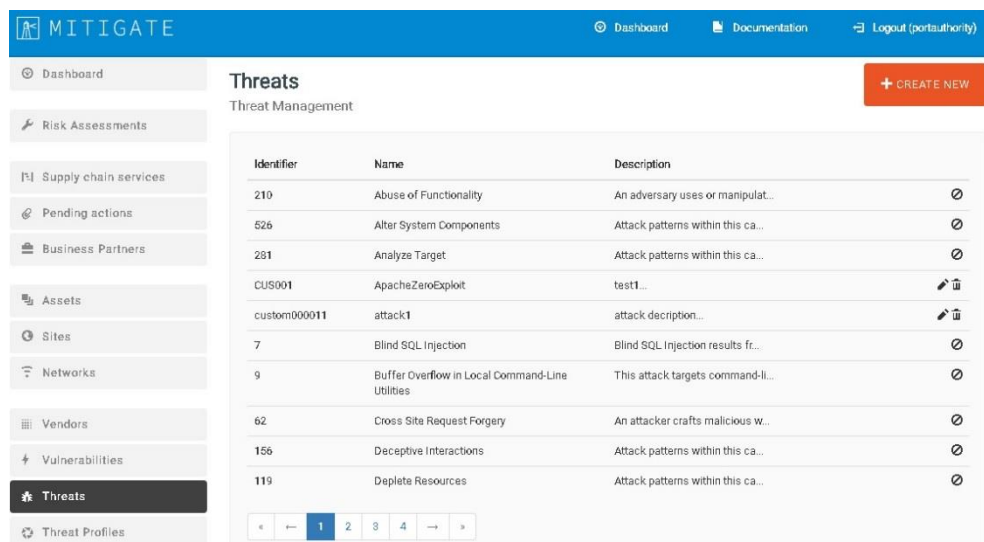
MITIGATE [1][2][3][4] targets to contribute to the effective protection of the Information Communication Technologies (ICT)-based ports SCs that arise from the ICT interconnections and interdependencies of a set of maritime entities (e.g. port authorities, ministries, maritime companies, ship industry, customs agencies, maritime insurance companies other transport CIIs (e.g. airports) and other CIIs (e.g. transport networks, energy networks, telco networks)). This is achieved by treating the resolution of the ICT maritime SC risks as a dynamic experimental environment that can be optimized involving all relevant maritime actors. MITIGATE approach based on simulations facilitates the identification, analysis, assessment and mitigation of the organization-wise and interdependent cyber threats, vulnerabilities and risks.

In the literature, the analysis and evaluation of the cyber risks are based on a straightforward approach that combines a set of parameters and features such as the likelihood of a security event and the consequences of the event itself, the exploitation level of a vulnerability etc. MITIGATE aims to support this approach with rational decision making. The pursuit of MITIGATE is to support risk analysis with security-related information obtained from online repositories strengthening the rational analysis. MITIGATE's objective is to promote a more rigorous, rational approach that gathers, critically appraises and uses high quality research information either produced by well-defined simulation experiments or are available online to enhance the risk assessment process.

In particular, MITIGATE shares the view that process of evaluation and mitigation of the cyber issues is neither objective nor neutral; it should be an inherently rational process that relies on well-defined and widely acceptable security-related data and not only upon highly personalised experience, expertise and judgment of individuals.

5.1.2 Key Features

MITIGATE (see Figure 26) aims at realising a radical shift in risk management for the maritime sector towards a collaborative evidence-driven Maritime Supply Chain Risk Assessment approach. To this end, MITIGATE has integrated an effective, collaborative, standards-based risk management system for port's CIIs, which shall consider all threats arising from the SC, including threats associated with port-CIIs interdependencies and associated cascading effects. The proposed system enables port operators to manage their security in a holistic, integrated and cost-effective manner, while at the same time producing and sharing knowledge associated with the identification, assessment and quantification of cascading effects from the ports' SC. In this way, port operators are able to predict potential security risks, but also to mitigate and minimise the consequences of divergent security threats and their cascading effects in the most cost-effective way i.e. based on information associated with simulation scenarios and data acquired from online sources and repositories (e.g. National Institute of Standards and Technology (NIST) Repositories).



Identifier	Name	Description
210	Abuse of Functionality	An adversary uses or manipulat...
526	Alter System Components	Attack patterns within this ca...
281	Analyze Target	Attack patterns within this ca...
CUS001	ApacheZeroExploit	test1...
custom000011	attack1	attack description...
7	Blind SQL Injection	Blind SQL Injection results fr...
9	Buffer Overflow in Local Command-Line Utilities	This attack targets command-li...
62	Cross Site Request Forgery	An attacker crafts malicious w...
156	Deceptive Interactions	Attack patterns within this ca...
119	Deplete Resources	Attack patterns within this ca...

Figure 26: Evidence-driven Maritime Supply Chain Risk Assessment (MITIGATE) System

In order for the system to meet its objectives has been empowered by: (i) a range of reasoning, data mining, crowd-sourcing and Big Data analytics techniques that incorporate and leverage a variety of data sources and data types, enabling efficient handling of data that are incomplete, uncertain, and probabilistic; (ii) pioneering mathematical techniques for predicting and analysing threats patterns; and innovative visualisation and simulation techniques, which optimise the automatic analysis of diverse data. These ICT solutions/technologies and mathematical instruments provide a basis for implementing a variety of mechanisms and processes that facilitates collaboration between the various maritime agents enabling them to:

- Identify and model assets, processes, risks, stakeholders' relationships/interactions and dependencies.
- Analyse threats, vulnerabilities and countermeasures accumulated in various online sources and repositories.

- Identify, evaluate and classify various ICT-based risks, while at the same time facilitating the risk resolution.
- Design, execute and analyse risks and threat simulation experiments in order to discover viable attack paths in the SCs. These attack paths consist of vulnerability chains that can be exploited by attackers in order to accomplish their malicious goals.
- Exploit the simulation results towards formulating effective evidence-based mitigation plans.
- Support continual Webs' vast reserve of open, distributed data uptake, integration, state assessment, decision analysis, and action assignment based on large-scale high- performance open computing infrastructures so that all agents may access and analyse a plethora of collected data and information.

5.1.3 Component Advantages

The MITIGATE system adopts an integrated framework for identifying, classifying, assessing, simulating and mitigating risks associated with port CII and cybersecurity incidents.

In particular, the MITIGATE system enables the involvement and participation of all stakeholders (e.g., port security operators, port facility operators, and SC participants) in the Cyber-Security management. In order to meet its objective, this system has been empowered by a range of: (i) reasoning, data mining, crowd-sourcing and Big Data analytics techniques that incorporate and leverage a variety of data sources and data types (e.g. vulnerabilities) retrieved from online repositories; (ii) pioneering mathematical techniques for predicting and analysing threats patterns; (iii) innovative visualisation and simulation techniques, which will optimise the automatic analysis of diverse data; and (iv) innovative game theory techniques in order to link optimisation and simulation. All these technologies and techniques have been combined for implementing a variety of services (Collaborative Risk Assessment and Mitigation Services, Open Simulation Environment (ORASE) and Simulation Services, Risk and Vulnerability Visualisation Services and Prediction, Forecasting, Social Engineering and Open Intelligence Services) as part of the project's risk assessment system that enable maritime agents to:

- Design, execute, analyse and optimise risks and threat simulation experiments that will produce the appropriate evidence, information, indicators, factors and parameters.
- Exploit the simulation results towards formulate of effective evidence-based mitigation plans.

However, it should be noted that the complicated nature of the ports' SCs' environment raises a set of additional issues concerning the effective and efficient handling of their security issues. In this context, taking into account the MITIGATE experience, there is a set of research challenges and issues (e.g. usage of machine learning methods such as Naïve Bayes, Random Forests or Neural Networks for the classification of the predicted attack paths, usage of use of distributed computation methods, such as multi-agent systems for more effective cyber-attack path discovery), regarding the distributed and interconnected nature of complex, interrelated SCs' physical and cyber components, network and operating environments, that need to be covered.

5.1.4 Examples Usage Scenario

MITIGATE system aims to provide a holistic solution regarding risk management in the frame of Supply Chain Services (SCS). To this end, specific set of services have been developed and integrated in a seamless manner. Such services include assessment of risk in a collaborative manner and advanced simulation and visualisation of potential attacks.

In this context, the following components of the MITIGATE system can be used and adapted:

- The Asset Modelling component allows the interconnected organizations to declare their assets along with the cyber relationships. The creation of a valid asset cartography within the frame of an organisation is the first step towards the realisation of a collaborative risk assessment. Thus, this component allows the creation of an IT asset inventory of all computing and networking related devices owned, managed, or otherwise used by the organisations involved in the SCS.
- The Vulnerabilities Management component replicates all the vulnerabilities from the CVEDetails portal and associates them with the declared assets. All the vulnerability information are based on the CVE naming standard, and are organized according to severity determined by the Common Vulnerability Scoring System Version 2 (CVSSv2) standard. Therefore, according to the CVE metamodel, a unique ID is declared, the value of the CVSS Score (ranging from 1 to 10) is determined and the access complexity, authentication, exploitability and the various impacts (in confidentiality, integrity and availability) are estimated.
- The Threats and Controls Management component acts as a comprehensive dictionary of known threats as well as the corresponding mitigation controls that can be used to advance organizations understanding and enhance their defences. It should be noted that the MITIGATE system has adopted the MITRE classification; in particular the Threats and Controls Management component synchronizes the MITRE attack identifiers and associates the identified vulnerabilities with one or more weakness identifiers.
- The Simulation component has a twofold goal. On the one hand it is responsible for the discovery of attack paths given a specific asset cartography and a specific SCS and on the other hand it is responsible to propose the best defensive strategy regarding the protection of a specific asset.
- The Collaborative Risk Assessment component is responsible to provide guidance for the conduction of a risk assessment for a specific SCS. More specifically, MITIGATE introduced a detailed multi-step process **Errore. L'origine riferimento non è stata trovata.**, in order to calculate the SCS-related risks. According to the proposed approach, three different qualitative risk levels are evaluated and derived. The individual risk refers to the impact of potential exploitation of several vulnerabilities at the asset-individual level. On the other hand, the cumulative and the propagated risks quantify the effect of an exploitation at a vulnerability chain level, taking under consideration that the assets which participate in a risk assessment are interconnected to each other. Cumulative risk quantifies the effect of incoming attacks to a specific asset while propagated risk quantifies the effect of an exploitation towards the adjacent network.

- The Notification and Reporting component is responsible to provide push notifications to the business partners regarding any type of messages that are published in the pub/sub queue. Since MITIGATE involves many time-consuming operations (e.g. the conduction of a vulnerability assessment, the calculation of risks, the processing of open information sources) every time that such an operation is completed a specific message is placed in a predefined topic of the pub/sub queue. The specific component consumes all messages that relate to notification topics and presents them in a structured way to the user.
- The Visualisation component provides a visualization of the entire infrastructure along with the linked security and risk related information such as threats, vulnerabilities and attack-types that are relevant to the individual assets that have been declared.

5.1.5 Expected extensions and potential new implementations

In AI4HEALTHSEC, the MITIGATE system can be used as a risk assessment tool to assess the assets, threats and vulnerabilities associated with the digital healthcare ecosystems. This system incorporates a bundle of automated processes and routines and integrates a wide range of ICT tools, which enable port operators in structuring, organising and managing assets and threats, as well as in executing simulation scenarios and deriving evidence-based knowledge that will be used for the identification, classification, assessment, simulation and mitigation of risks associated with port CII. Hence, these concepts can also be used for AI4HEALTHSEC to facilitate the analysis and propagation of a threat, risk or incident from in a structured and well-defined way. In particular, MITIGATE is a good candidate tool to be used in order to drive the design and development activities in **Horizontal Layer 1 – Risk and Privacy management & Cyber-Attack Forecasting** of the AI4HEALTHSEC framework. It should be noted that the main services that have been integrated in the MITIGATE system and will be adapted and enhanced in the context of the development of the Horizontal Layer 1 include:

- The Risk Assessment and the Visualisation functionalities aim to quantify the risks that derive from the various vulnerabilities associated to specific assets required for the provision of the Healthcare Supply Chain Service (HSCS).
- The Risk Management functionalities aim the generation of an optimal mitigation strategy given a specific HSCS.
- The Simulation functionalities facilitate the design, execution and analysis of risk and threats simulation experiments in order to generate the chain of sequential vulnerabilities on different assets that arise from consequential multi-steps attacks.
- The prediction and forecasting functionalities provide automated identification of potential vulnerabilities and attacks in the Healthcare ecosystem.

5.2 Security & Privacy Assurance Platform

5.2.1 Short Description

The Security & Privacy Assurance Platform (also referred as Assurance Platform) is a model driven platform that can make hybrid security and privacy assessments that involve threat and vulnerability analysis, static analysis, penetration testing and continuous runtime monitoring, in order to provide a comprehensive analysis of the security and privacy posture of a system.

The platform has interoperability with various platforms and programmatic connectivity to different systems through appropriate probes (e.g., event captors, test tools), enabling it to obtain the monitoring and/or test the evidence required for assurance as well as certification assessments. Since it is a model driven platform, it can be customized to enable the realization of different security standards and risk management requirements.

5.2.2 Key Features

The Assurance Platform forms the basis of this component (see Figure 27) and is used for monitoring, testing, and assessing the protected framework. The platform is responsible for monitoring the execution times and accuracy of each component of the framework, ensuring and assuring the proper functioning of the whole framework. The Assurance Platform is an integrated framework of models, processes, and tools to enable the certification of security properties of services. It uses different types of evidence to demonstrate the support for the required properties and award the corresponding certificate. The types of evidence foreseen include monitoring data and testing data, while additional sources may be added as the development progresses.

The Security & Privacy Assurance Platform:

- combines runtime monitoring and dynamic runtime testing to ensure correct and effective operation of security controls;
- can be hooked to different systems programmatically through appropriate probes (e.g., event captors, test tools) in order to obtain the monitoring and/or test evidence required for assurance and/or certification assessments;
- operates based on models that determine the operational evidence that should be captured from systems and how it should be assessed (e.g., what conditions it should satisfy) in order to assess the correctness and effectiveness of implemented system security controls;
- enables the runtime assessment of temporal event patterns and rules that can express signature or anomaly-based patterns.

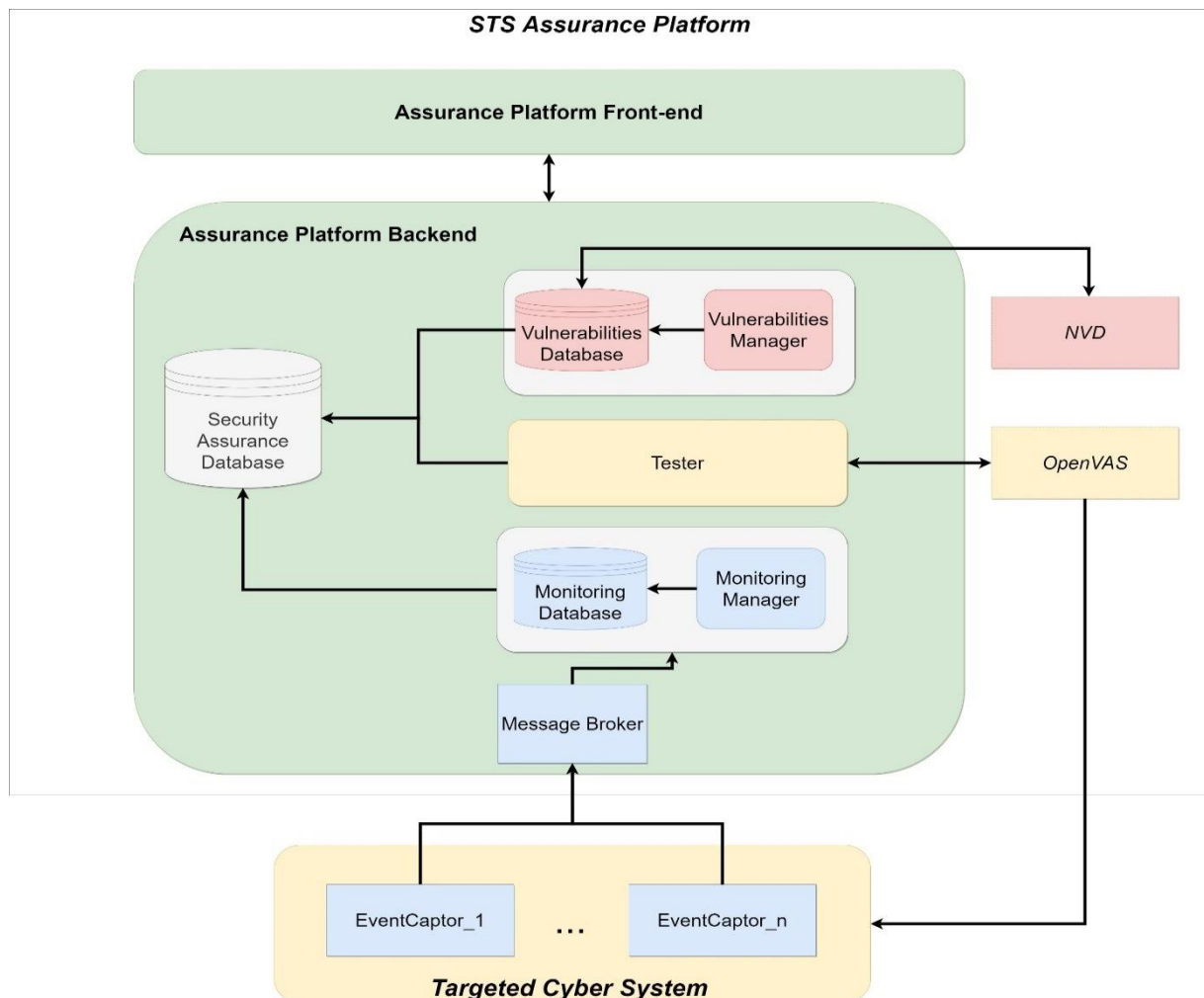


Figure 27: The high-level architecture of the Security & Privacy Assurance Platform.

5.2.3 Component Advantages

The Security & Privacy Assurance Platform is model-driven and, thus, by design supports the ability to be adapted to every usage scenario, generating the needed evidence for its assessment. Following the definition of the model of the organisation that needs to be protected, the platform can find the known vulnerabilities affecting each asset modelled, as well as additional vulnerabilities and misconfigurations found through dynamic testing of the system. Furthermore, based on the infrastructure asset relations and models provided, it can calculate and highlight the most pertinent risks stemming from said findings. In terms of real-time protection, the Platform supports monitoring the system for erroneous behaviour using probes that can be easily integrated with the infrastructure and does not require any implementation changes to the current system infrastructure. The only dependency that is needed is that the probes should be able to communicate with the assurance platform.

5.2.4 Examples Usage Scenario

The Security & Privacy Assurance Platform is comprised by different components that work seamlessly to create security and privacy assessments. In the next lines, we present 3 different use cases and how the different components can be used.

- Vulnerability testing

To create a vulnerability assessment, initially the user uses the asset loader component, which is responsible for receiving the security assurance model for the target organization. This model includes the assets of the organisation, security properties for these assets and the assets relations. Afterwards, the vulnerability loader component is invoked. The component responsible to find the vulnerabilities of the assets that comprise the security assurance model provided by the Asset Loader.

- Penetration (dynamic) testing

As aforementioned, there needs to be a security assurance model created for the organisation. Then when the penetration tool is initiated it performs penetration testing assessments. This process includes discovering new vulnerabilities and determining if present vulnerabilities are exploitable. To achieve that it utilizes a combination of various open-source penetration testing tools to passively and actively interact with a target system. As an additional functionality the module can discover and report assets that were not defined in the current asset model of the system.

- Monitor-based assessments

Commonly, also the initial step for this assessment is to have a security assurance model. Additionally, for monitor to function, there must be probes installed at the assets that we need to monitor for security or privacy violations. After the setup of the probes and the security assurance model creation the user initiates the monitor process using predefined monitor assessment profiles. Those profiles are used to monitor different security properties of the infrastructure, such as the availability of a component.

5.2.5 Expected extensions and potential new implementations

In the context of AI4HEALTHSEC, the main functions of the Assurance Platform will be further expanded to meet the security requirements of AI4HEALTHSEC framework, as they are introduced in this deliverable and will be instantiated in the overall system architecture specifications.

5.3 Cloud - based Intrusion Detection System

5.3.1 Short Description

The cloud-based solution is capable of detecting possible attacks that take place within a host running many VMs. Virtual hosts hosted under the same Hypervisor are able to produce orders of magnitude more network throughput than conventional communication over the internet. This happens as VMs are using the internal CPU BUS to communicate, which can lead to throughput over 30 GB/s, and thus an infected VM could produce massive DDoS or other attacks against other co-hosted VMs. The

detection system that we have devised, is based on a well-known IDS (SNORT) which is deployed within the host OS of the server hosting the VMs. The system includes a database, a log-processing engine and a web-based interface to visually present the results. The hypervisor of the system is configured to centrally monitor and log all the “malicious” activity related to the VMs of the specific machine and provide results through a web interface or in the form of raw data. Thus, the solution is able to identify intra-VM attacks and Inter-VM attacks, as well as attacks originating from wherever on the internet, that cannot be identified by an IDS monitoring the uplink of a cloud infrastructure. The solutions need the installation of a special hypervisor and an intrusion detection system on top of it. The solution can be deployed either to the Cloud or locally. Currently, the feasibility of our solution has been tested using the XEN and KVM hypervisor which we have used and tested for our current implementation. Additionally, events and alerts generated by our system are reported to our dashboard and to the XL-SIEM through syslog.

5.3.2 Key Features

The solution provides various means of alerting the user. It can either provide periodical reports via email (pdf format) or Simple Message Service (SMS) alerts in case there is an SMS gateway available.

5.3.3 Component Advantages

The cloud-based IDS tool has the advantage of providing threat detection, intelligence analysis and correlation capabilities for: (i) near real-time identification of anomalies, threats, risks and faults and the appropriate reactions; (ii) proactive reaction to threats and attacks based on simulation and pattern matching processes of the upper layers; and (iii) dynamic decision-making according to the end user's needs and the identified incidents/threats.

5.3.4 Examples Usage Scenario

The cloud-based solution can be one of the main security assets within DDoS detection core in AI4HEALTHSEC framework. Specifically, each pilot needs to satisfy a series of security features, such as alerting, availability, confidentiality, ease of control/administration, real-time response, reliability and scalability, the said tool could be used. Additionally, it does not require external data just monitors the inbound and outgoing network traffic of the VMs.

5.3.5 Expected extensions and potential new implementations

The cloud-based IDS is linking to Horizontal Layer 2 – “Incident Identification” of the AI4HEALTHSEC framework, which aims to detect and assess possible security incidents at existing assets of the HCIs based on the dynamic, distributed management of the auto-exposed/revealed data. It is already able to provide information about potential DDoS attacks that are active in the Internet. The results produced can be used by system and network administrators. The accuracy of the produced results is proportional to the amount of the dark IP address space monitored. Within the AI4HEALTHSEC project, we plan to further increase the level of accuracy of the results and consequently increase the trustworthiness of the service as a whole. Towards that target, we could correlate and combine information provided by other security related online sources (i.e. blacklists) or data produced by other partners' solutions.

5.4 Data Harmonization & Pattern Recognition

5.4.1 Short Description

These tools allow for formally expressing data using the terms of a common Reference Model and include:

- **Metadata Extraction Tool:** This tool automatically detects the event parameters recorded in a file and captures their names, their data type (i.e., Number, Date, String) and the range of values or the terminology used. The data should follow a predefined format (in the current version, the input should be an Excel file with the name of the fields and data recorded for each event in a separate row) so that it can be automatically processed by the developed software tools.
- **Metadata Mapping Tool:** This tool allows for the semi-automatic specification of the mappings between the extracted metadata of a dataset (in our case being the Event Parameters) and the entities (Classes/Properties) of a Reference Model as part of the data harmonization process. This process encompasses the software-based analysis of the metadata for detecting common patterns used in the source files as well as the development of the appropriate data transformation services that could be instantiated by the mapping tool.
- **Data Transformation Tool:** This tool automatically expresses the data recorded in the source file using the terms of the Reference Model, based on the Mapping Rules specified, producing thus the Harmonized Data. The output of this tool is an OWL ontology with the harmonized data. The latter are also provided in JSON format.

5.4.2 Key Features

The main features of this set of tools, include:

- To allow for dealing with a plethora of semantic and structural heterogeneity issues between datasets for data harmonization purposes.
- To establish a semi-automatic data harmonization process, including in sequential order a fully automatic metadata extraction step, a semi-automatic mappings specification step and a fully automatic data harmonization step.
- To introduce a data-blind harmonization process and is GDPR compliant.

5.4.3 Component Advantages

Key advantages are:

- The data harmonization process is constituted of two automatic and one semi-automatic step.
- It enables users to automatically express their data using the terms of a common Reference Model, and hence facilitate the integration of existing systems and resources with a new one.
- Constructive interaction among users from different domains of expertise (from the Data Provider's part and from the Data Integrator's part) is enabled during data harmonization.

- GDPR compliant data harmonization.

5.4.4 Examples Usage Scenario

The files with security-related information produced by software systems (e.g., log files) are often heterogeneous in terms of semantics and structure. Processing of these data by tools and/or systems handling cybersecurity events requires addressing the different cases of heterogeneity met, ideally with a flexible and expandable way. Within the context of AI4HealthSec, the data harmonization process encapsulating the tools presented in the previous sections requires a Reference Model which specifies the parameters of particular interest for the events captured and accordingly the customization, adaptation and use of the tools for linking the log files with the elements of the model and translating their data into the AI4HealthSec semantics.

For this purpose, the Metadata Extraction tool is applied for detecting the elements recorded for each Log file, such as the date of occurrence; the component having produced this log, the debugging level, etc, along with the format of the data and the value range or vocabulary used for each one of them (e.g., list of software components). The automatic extraction of the metadata is followed by a round of communication between the AI4HealthSec tech expert and the Data Provider in order to ensure the correct and complete capturing of the semantics. For specifying the correspondence among such terms with the elements introduced in the common Reference Model, the Metadata Mapping tool is used. Through the latter a series of mapping rules is described followed by the development of the respective transformation services. The last step of the process lies in the incorporation of the mapping rules to the Data Transformation Tool which then automatically applies them to the Log data and produces the harmonized log data (i.e., log data expressed with the AI4HealthSec semantics).

5.4.5 Expected extensions and potential new implementations

This module can be used for the integration of the data produced by existing software systems and security mechanisms with the core components of the AI4HealthSec platform. In particular, the data captured by those systems can be formally expressed using the terms of a common Reference Model so that they can be accordingly mined for detecting potential threats by the AICS nodes. This component will be part of the Horizontal Layer 2.

For the successful integration of this component with the AI4HealthSec platform, the data should follow a predefined format. Depending on the complexity and variability of the data files, the metadata extraction tool could be potentially enriched in order to support different formats. Also, a Reference Model needs to be specified, which will encapsulate the key parameters of particular importance, along with their data type and the terminology being used (in case they come from a controlled set of terms) and will be formally expressed in an OWL ontology. The analysis of the parameters of each file to be integrated as well as the specification of the mapping rules for each dataset with the Reference Model will be on the basis of close collaboration between AI4HealthSec partners and the security expert(s) of each entity.

Moreover, in the AI4HealthSec we are planning to improve the functionality provided by the existing internal components of the Data Harmonization system using AI techniques that accelerate the

design of the Reference Model, the detection of the patterns used for the expression of user data as well as the mapping of them with the Reference Model terms.

5.5 Reasoning Engine

5.5.1 Short Description

This component exploits semantic relations and expands queries in order to mine all the semantically correct results across datasets. For this purpose, it applies formal semantic reasoning based on the explicitly defined relations of concepts in an ontology and detects additional terms with the same or narrower meaning. Missing data are also taken into consideration during the search process.

5.5.2 Key Features

Key features of the Reasoning Engine include:

- The search process focuses on the meaning of terms rather than the sequence of characters being used.
- There is clear separation of the knowledge base from the inference process.
- Missing data are taken into consideration during the evaluation of the user-defined queries.

5.5.3 Component Advantages

The Engine requires an ontological Representation of the knowledge base (e.g., Reference Model vocabularies) using Semantic Web technologies. It, also, focuses on the classification of the ontology terms (i.e., a subset of axioms that are often specified in an OWL ontology). Based on these assumptions and prerequisites, the advantages of the Reasoning Engine are:

- It deals with the semantic distance between user-defined queries and data.
- It allows domain experts to independently develop and update the knowledge base, with the changes made being promptly available by the system.
- The complexity of the user-defined queries does not change. As a result, the response time of the query system will not be significantly affected.

5.5.4 Examples Usage Scenario

The data produced by the software systems deployed in each pilot is of particular importance for detecting potential security issues. Of particular interest is to query and process data from different datasets, which could potentially reveal the threat significance, the impact of an attack and the extend of its consequences.

The Reasoning Mechanism can be applied for mining all the events that belong to a particular category (e.g., DoS attack), despite the fact that the particular events detected across pilots may be captured with different, yet semantically linked, concepts. As a precondition, the ontological representation of the security events needs to be specified along with the relations among them. Through the application of this mechanism the additional facts about the events can be inferred or unveiled.

5.5.5 Expected extensions and potential new implementations

This module can be used for inferring additional information about the data produced by the software components and systems of each pilot or particular events detected (either by the existing security mechanisms or the AI4HealthSec components themselves) and hence facilitate the analysis of the processes taking place. The functionalities provided by this component can be part of the Horizontal Layers 2 and 3.

In order to meet the needs of AI4HealthSec, an ontological representation of the security relevant concepts (depending on the queries of interest they could include risks, threats, assets, etc) of particular interest is necessary so that they can be accordingly mined through a reasoner. Also, the data collected so far can be filtered by the component on condition that they are maintained within a relational database and the terminology being used is already specified in the aforementioned ontology. The search process is currently based on the classification of the terms existing in a knowledge base as well as the data recoded so far. In many cases, part of the data produced by existing systems may not comply with a data structure and hence the application of alternative approaches stemming from the fields of data mining and machine learning in combination with the existing well-defined knowledge bases can significantly improve the outcome of the search process.

5.6 *Advanced Visualization Toolkit*

5.6.1 Short Description

The Advanced Visualisation Toolkit (AVT) is a set of data collection, processing, and presentation components and tools, which can assist users in examining and analysing digital information collected from the monitored sources (i.e., computers, network devices, switches, structured datasets, etc.), in order to investigate abnormal system operation and further explore related data insights. The tools combine the data collected from many different sources and provide a multi-level and user-specific scenario driven overview of the system operation and/or the dataset under inspection. Through different view, timeline control and graph-based visualisations, an investigator may move back in time, narrow down the time frame of data exploration and inspection, and compare two different snapshots and timeframes of the system operations, in order to get insight of how the system is functioning either normally or beyond the detection of a breach or other anomaly.

5.6.2 Key Features

Timeline and graph-based inspection of data and scenario-driven preconfigured views are the key innovative features of AVT, which in conjunction with effective visualisations and data exploration mechanisms allow fast discovery, creation, and presentation of correlations between data.

5.6.3 Component Advantages

AVT goes beyond existing approaches by introducing innovative aspects in data inspection, exploration, and visualisation, through the following mechanisms:

- Timeline analysis, which provides the ability to “travel back in time” and compare the current situation with similar events that occurred in the past. This allows new data to be compared

against patterns encountered before. In this way, the intended users of AVT (such as forensic investigators, security officers, decision makers, etc.) are able to identify abnormal behaviours in the current data-driven process execution operations, and eventually develop response strategies to be deployed, based on past knowledge from both successful and unsuccessful outcomes.

- Graph-based visualisations, which provide an efficient way to advance the existing knowledge in the exploration of data correlations and identify new information paths that present the conceptual and/or semantic relationship between different types of labelled data properties.
- Preconfigured views, which provides the ability to adapt the display of information based on previously encountered situations. For example, if the investigator/Data analyst has created a specific “view” consisting of multiple data sources and presentation modes (e.g. specific relationships to group and associate data items) to deal with a specific incident in the past, this view can be saved and reused, either manually or automatically, to present data associated with a new case, or an incident that is currently playing out.

These mechanisms enable the AVT end users to quickly establish a solid understanding of an event and benefit from existing knowledge gained from past interactions, so as to identify the root cause of incidents and speed up the initiation of proper incident response actions. Information may be automatically presented in a way that enhances situational awareness, which allows the user to concentrate on the analysis and exploration rather than the configuration of the visualization system. In addition, the need for technical support at the end-user level is minimized which can be particularly beneficial in forwarding deployments.

5.6.4 Examples Usage Scenario

AVT has been used as the visualization framework for a variety of EU-funded projects and commercial activities, including digital forensics analysis and investigations, and big data advanced visualizations. In the manufacturing sector, AVT has been used in order to provide visualisations towards situation awareness, reporting of normal operations and anomaly detection and predictive maintenance, based on multisource and IoT-enabled data. Using almost real-time connectors, and automatically updated interactive visualisations, users can gain situational awareness in short time, being capable of quickly managing emerging maintenance events. The toolkit can combine different timescale datasets, like the combination of real-time streams with the analysis of historical data to enhance predictive maintenance and manufacturing planning and optimisation.

A specific instantiation of the AVT toolkit has been developed for the forensics domain and the digital forensics investigation sector (the Forensics Visualisation Toolkit – FVT), in which relevant investigators exercise their intelligence in developing data exploration scenarios that exploit the integration of AI-driven and analytics-based cyber forensics services. Through FVT, such users are empowered with correlation algorithms and innovative capabilities on the forensic investigation of digital and physical assets, including post-mortem analysis and robust processing capabilities of physical media, advanced reporting features and almost real-time mitigation actions.

5.6.5 Expected extensions and potential new implementations

The knowledge of AVT implementation and its customisation to forensics will be exploited in the AI4HealthSec Framework, with the aim to facilitate the expected functionalities of Horizontal Layer 3 and 4 and the capabilities of Vertical Layer 3. The expected implementation will focus on the presentation of different data streams produced in these layers to facilitate the development of user scenarios for: i) the assessment of security events, based on the collected data and their patterns, ii) the realisation of the impact of such events to the development of attack pathways along the candidate affected assets, based on the risk assessment methodology and threat hunting mechanisms, and iii) the enactment of mitigation actions, based on the support for decision management tactics that the project will investigate.

5.7 Sharing Platform

5.7.1 Short Description

Sharing platform provides functionalities related to cybersecurity information exchange, potential integration with MeliCERTes¹⁸ and propose cybersecurity solutions. Specifically, this tool includes:

- the Early Warning Intrusion Detection System (EWIS), which is a honeypot-based solution where the so-called sensors VMs can be deployed in an infrastructure and attract potential attacks;
- the Sharing Platform, which is a Malware Information Sharing Platform (MISP) instance that gathers information from EWIS. It is connected to a central MISP instance managed by the Greek National Computer Emergency Response Team (CERT).

In fact, honeypots are used as a proactive measure to assist security teams to capture and analyse attacks, in a safe sandbox environment; MISP gathers, shares, stores and correlates Indicators of Compromise of targeted attacks, threat intelligence, financial fraud and vulnerability information. Finally, from the architectural perspective, honeypots are connected to ISP through a middleware.

5.7.2 Key Features

This solution is able to detect a variety of attacks running in parallel with the production network/systems. A variety of detection sensors can be deployed, according to the needs of the specific installation. EWIS alerts combined with gathered intelligence from the Dionaea honeypot and the KIPPO SSH specific honeypot are aggregated and provided via the MISP interconnection.

5.7.3 Component Advantages

AI4HEALTHSEC will exploit the integration of Sharing Platform with MeliCERTes CSP in order to create to interchange cybersecurity information, to receive or produce alerts for uprising threats.

¹⁸ <https://github.com/melicertes/csp>

5.7.4 Examples Usage Scenario

The Sharing platform can be used to enable the dissemination and sharing of critical information, between individual nodes of HCIs, related to vulnerabilities, risks, threats, and incidents. In this context, it can be involved in the aim of AI4HEALTHSEC to build trust chains among the interdependent HCIs and the interactive stakeholders within the HCSC, to ensure the privacy and protection of the threat intelligence and incident-related shared information and extinguish the possibility of revealing sensitive data. For example, FORTH, as a CSIRT, will explore best practices for sharing AI4HEALTHSEC information with relevant parties at all required levels. Finally, this tool can manipulate Indicators of Compromise (IoC) and security event descriptions following the MISP format. Thus, specific converters can be employed in order to transform other kind of data to the MISP format.

5.7.5 Expected extensions and potential new implementations

The Sharing platform is linking to Vertical Layer 1 – “Individualised Autonomous Networking”, which has the responsibility to disseminate and share information among the individual AICS nodes of the Interdependent HCIs. It is already able to provide functionalities related to cybersecurity information exchange, however, within the AI4HEALTHSEC project, we plan to further improve the visual capabilities of the Dashboard producing more useful graphs and alerts based on the feedback we receive from the system administrators and security officers that participate in the project.

5.8 CHIMERA

5.8.1 Short Description

Chimera is a dataflow application, integrated in a Web User Interface that can communicate with the Orchestration-Frameworks APIs allowing a user to manipulate knowledge and data generated by other tools. Chimera can safeguard access to data by performing attribute-based anonymizing on dynamically defined semi-structured data. The data to be anonymized is collected, processed, transformed and filtered in order to discard what is not relevant. It can support auto detection of personal data through scanning of existing files (e.g. documents, pdf, spreadsheets, txt, SQL databases, etc).

The anonymisation module has high throughput for processing large text datasets in unstructured formats and perform user-defined transformations to clean, bake, structure, anonymise and or encrypt. Since formats change greatly and often, the tool needs to be customisable and support a dynamic language to define which fields should be transformed. The Chimera Anonymisation Language (CAL) is leveraged by domain specific problems, like the ones addressed in the AI4HEALTHSEC Project, as a way to formally define how personal data at rest should be handled, transformed, anonymised, encrypted or decrypted, to keep relevant data protected from prying eyes. Chimera aims to be a swiss army knife to deal with unstructured information, allowing for operators to use schema on read approaches, and define which fields should be extracted dynamically from the information they need to process.

On the structured information side, a few well known formats are supported as inputs, such as Excel, Word, PDF, CSV. Out of the box Chimera parses these formats using third party libraries, converts

then into the memory data structure representation of their textual values and from there they can be dealt with as if they were extracted from plain text files.

Common operations on these types are provided to ease the setup of new workflows, such as converting PDF to text, performing redaction on PDFs image layers, detecting PII and redacting PII.

5.8.2 Key Features

Chimera provides a backend written in a language that compiles to binary format (elf / exe) for performance reasons. This backend has a hand-written parser and lexer that creates an Abstract Syntax Tree (AST) for the CAL language. Further, a tree walker interpreter provides the runtime for CAL. Message passing between AST nodes is performed through shared memory and all nodes follow the same specifications for accessing data and creating new data. To ease the usage of the domain specific language, a web frontend with a single page application is developed to help operators to design the workflows visually and then export the rules in CAL to a standard format that can be passed into the backend runtime.

Anonymization Framework

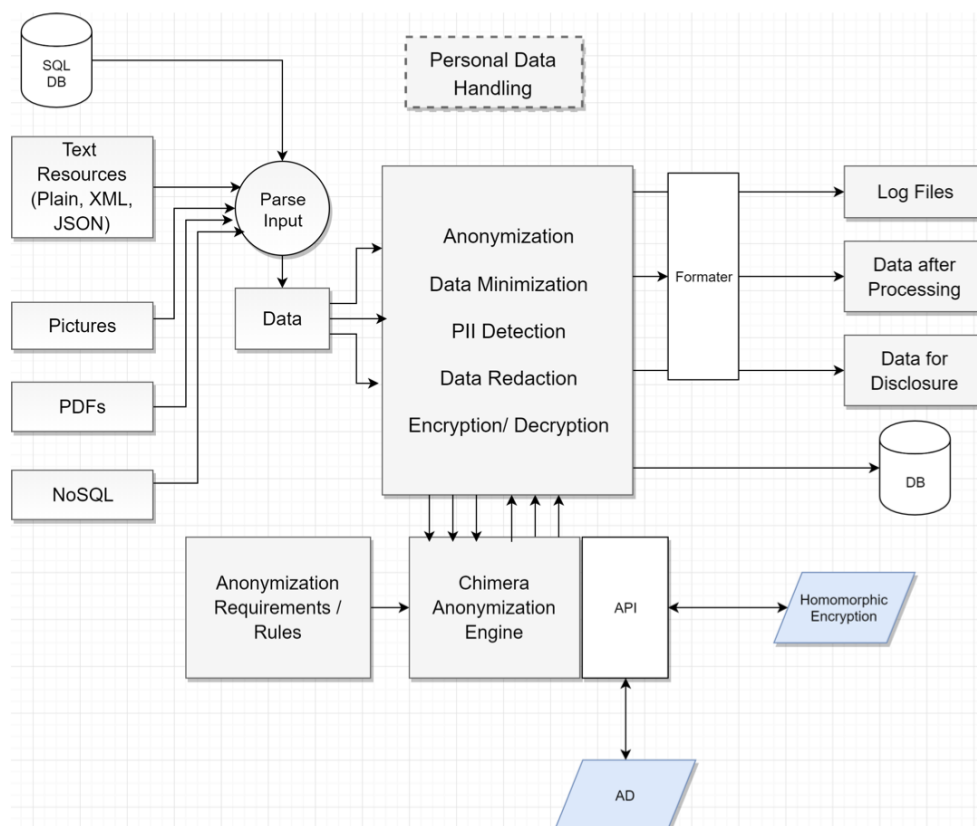


Figure 28 - Chimera Component Diagram.

The anonymisation module provides an API that is able to remotely transform data, by either performing encryption/decryption or by applying anonymisation techniques such as one-way hashing combined with k-anonymity/t-closeness/p-sensitive. These techniques are realized through the implementation of generalization, masking and tokenization algorithms, combined with statistical combinatory analysis.

The detection API is designed and implemented for sensitive data detection, allowing the searching of structured and unstructured data and reporting found evidence of privacy data, it supports crawling through SQL databases, unstructured text files, and PDF files.

This tool provides the current forms of interaction:

- Standalone as GUI with CHIMERA_STUDIO for data exploration & data workflow design
- Webservice Integrated with OF for CHIMERA queries & data workflows. Feature wise we can split the tool into 5 different topics, Data Collection, Data Encryption, Data Anonymization, PII detection, and Reporting.
- Data Collection: On the data collection front, the tool supports direct SQL integration with (Oracle, SqlServer, Sybase, Mysql, Postgresql and Sqlite3) for structured formats. On the semi-structured end JSON and XML are both supported as well as CSV and any formats that can be parsed through regular expressions or key-values mapping. Data parsing features are available to process and convert data into an in-memory format that is used for all inputs, ensuring any transformation tools that are later invoked operate always under the same format structure.
- Data Encryption: Chimera leverages openssl for the encryption and hashing primitives, plus the ability to perform AES-FF1 for format preserving encryption and attribute based encryption, as well as full file encryption. The secret keys can be stored locally in insecure configuration files, that can be obfuscated with internal existing PubPriv crypto. Or can be more securely obtained from third party vaults such as Hashcorp Vault¹⁹.
- Data Anonymization: Regarding anonymization we support the following techniques, like IP masking, location generalization (Local -> Region -> Country -> Continent), GeoLocation generalization (Reducing the decimal precision), tokenization (replacement with pre-defined list values), masking (replacing part of the content), suppression, and PII detection.
- Reporting: Currently reporting can be performed into Excel files, HTTP API calls and SQL databases and plain old text files.

5.8.3 Component Advantages

Chimera brings benefits to the intended users, as it presents the following strengths:

- High Performance: A single CPU core is able to process around 200k record per second, and rules and processing pipelines scale near linearly if more CPUs are used.

¹⁹ <https://www.vaultproject.io/>

- Zero dependencies: Deployment can be made with a single statically compiled binary that doesn't depend on third party libraries. This simplifies deployment and reduces install friction.
- Convention over Configuration: Even though the tool is extensible and hugely configurable it comes with sane defaults that make it work out of the box for most use cases.
- Random Sampling: At small scale everything is possible, at medium scale it's a matter of resources, at large scale is all about smart reasoning so for identifying PII in large datasets Chimera supports random sampling, big enough to be accurate & fast enough for quick turn around. Custom sampling frequencies 1/100, 1/1000, 1/10000 are available to tune for performance vs accuracy.

5.8.4 Examples Usage Scenario

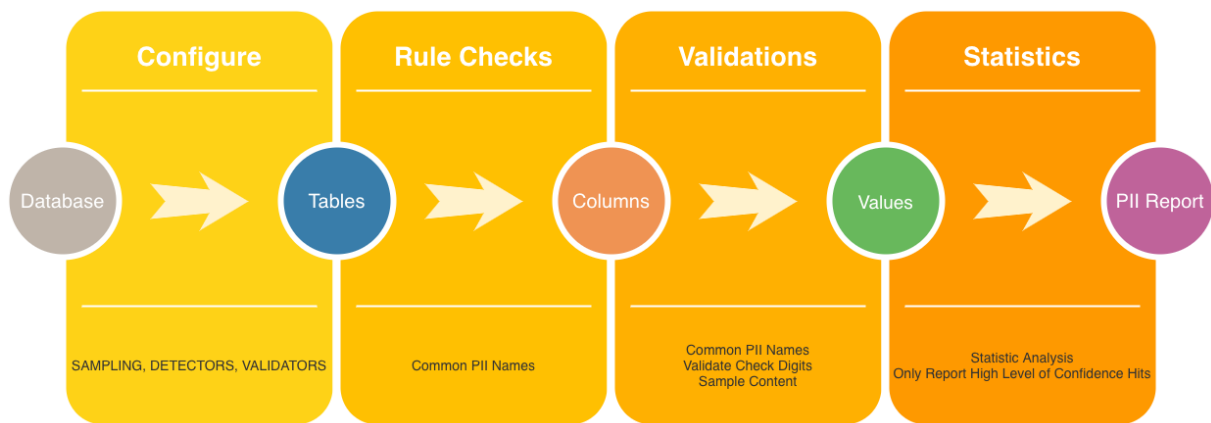


Figure 29: Normal workflow of example usage scenarios

A normal workflow (Figure 29) includes the following steps:

1. Read input data (either from files, URLs, or databases)
2. Define rules to extract dynamic fields
 - a. Either through regular expressions or custom-made parsers (key=value, csv, xml, json, etc)
3. Define the rules to transform fields (either add, remove or mutate in place)
 - a. Can anonymise, encrypt, decrypt, refactor data
4. Write output data (either to files, URLs or databases)

5.8.5 Expected extensions and potential new implementations

The tool is intended for use in Vertical Layer 2 of the AI4HealthSec Framework. In this context, Chimera can be extended to support data structures that are stored in no relational databases (e.g. MongoDB, Redis and ElasticSearch), and are described in commonly used formats (e.g. STIX 2.x, Apache Parquet, Apache Arrow). Since the framework will implement a Swarm Intelligence approach,

Chimera will be extended to be able to keep referential integrity between multiple independent runs, which can be scheduled to occur at different places in time, support automation, through the inclusion of cron like scheduling agent, and be capable of operating distributed workflows between multiple instances for increased performance.

5.9 Asset Explorer tool

5.9.1 Short Description

The Asset Explorer tool facilitates accurate, semantically-rich and user-friendly definition of population characteristics in order to capture population features in full detail leveraging knowledge encoded in the terminologies, ontologies and standards used. It provided efficient selection of cohorts matching those characteristics out of vast repositories. While it was developed for large archives of healthcare data, it is suitable for the management and analysis of security- and privacy-specific information. The tool incorporates support for basic analytics, such as visualization of distributions of variables in the data.

5.9.2 Key Features

The tool builds on a harmonized repository (with rich semantics and standard-based representations) and provides a flexible UI that enables filtering to extract relevant subsets of the data for a particular analytics task. For the power user, the tool provides a Domain Specific Language that is highly readable and enabled the definition of expressions of desired complexity. The tool enables saving filters and sharing of filter stacks among users in a single deployment and across deployments, enabling users to carry out the same analyses without sharing data. The users are supported with reasoning on their data (making use of the underlying ontologies and standard representation of data).

5.9.3 Component Advantages

The tool enables efficient and highly expressive exploration of diverse data out of large repositories, supporting the definition of the filtering criteria with ontology-based semantics, standardized representations, and complex expressions that include temporal statements (both absolute and relative ordering of events), computations and comparisons. It will add value for the management of the large amounts of heterogeneous information collected and managed in the scenarios of the project. The tool has been developed for the management of cohorts out of large repositories of heterogeneous healthcare data, which means that for the use of the tool in the context of the project, extensions are needed with respect to the standards and ontologies used, metadata extracted and persisted, information model and provenance model.

5.9.4 Examples Usage Scenario

Deployed on harmonized, semantically-rich data, the component will be a first step to analytics and model development, enabling the efficient selection of the datasets relevant for a particular analytics task. It can as well be used to automatically detect relevant anomalies and risks in the data.

5.9.5 Expected extensions and potential new implementations

The tool will be extended in the context of Vertical Layer 3 of the Framework to enable the management and efficient exploration of security and privacy data and other relevant assets (logs, processes, etc.). As the tool was developed focusing on medical data, this will require the extension of the information model (including semantics and selected metadata), and of the provenance model. The new implementations will also enable selecting, filtering and exploring cohorts of data, logs and other relevant assets in a federated environment and flexibly enabling the desired level of sharing across deployments (e.g. only counts, metadata, processes, risks, propagation graphs, etc.). Leveraging information from real-life medical environments and domain knowledge specific to security and privacy management, AI techniques will be used to validate and extend the existing data transformation and semantic annotation approaches.

5.10 Data Sharing Management

5.10.1 Short Description

This component mainly enables a handshaking process among two different type parties, i.e. (a) a data processor and (b) one or more data provider(s), in order to facilitate the sharing and usage of the dataset(s) in a GDPR-compliant manner.

5.10.2 Key Features

Key features of the Data Sharing Management platform include:

- It addresses the need to offer **GDPR-compliant access to data**.
- It enables **data processors** to request **access** to a number of selected **datasets** for usage of the data with a selected service.
- It enables **data providers** to **respond** (accept or reject) to the **data access requests** that pertain to their dataset(s) and provide more information to the data processor.
- Through easy-to-use dashboards, both **data processors and data providers** can **assess** all **relevant** information and details of requests, responses and their statuses.

5.10.3 Component Advantages

The Handshaking module undertakes the task to mediate between a data processor, who wishes to access the data of one or more dataset and process it in a certain manner via one of the available data processing services, and the individual data providers of the selected datasets, who are responsible for responding to data access requests made by data processors regarding the usage of their data. The module's primary role is to allow each data provider maintain control over the data of their dataset. The module allows a data processor to:

- Create and submit a new data access request (DAR). In order to do so, the user selects one of the available data processing services and one or more of the available datasets, includes the reason for requesting access to the data and their intended usage and submits the data access request.

- View the details of all data access requests that they submitted in the past, along with the response status of the data providers.
- For each of their data access requests, the data processor can further view the detailed responses per dataset as provided by the respective providers, along with the data usage terms for approved requests or the rejection justification for disapproved ones.

It further allows a data provider to:

- View a list of the new and existing data access requests for their dataset(s).
- Respond or edit their existing response to a request for each dataset individually. If they have provided a response in the past, they can change any of the already provided response details, which are displayed on screen. If they are providing a response for the first time, the initially displayed response details are blank. Each request that is approved or rejected can be accompanied with a text message (the terms of usage of the data or the justification of rejection, correspondingly), which the data provider includes (types or pastes) in an appropriate field or uploads (e.g., in the case of data usage terms) and the Requester can subsequently view.

The Handshaking service can further be included within data processing workflows which consist of one or more data processing service. Depending on its configuration and its positioning in the workflow, the module can further store the exact configuration specified by the data processor for the selected data processing service to be performed on the dataset(s) for which approval has been obtained. This way, a) the execution of the data processing service can be automated for all approved requests and b) only the exact usage defined by the data processor and reviewed and approved by the data provider can be performed by the data processing service. The following relevant functionality is enabled by the service:

- Store service-specific configuration or parameters associated with a specific data access request and data processing service.
- Retrieve the service-specific configuration before executing a service in accordance to that configuration, i.e. after approval has been granted by one or more data provider.

Finally, it should be noted that the data provider and data processor roles could be replaced by computer agents with the specific tasks and with predefined policies.

5.10.4 Examples Usage Scenario

The following usage scenario is the simplest one, which presents the stand-alone functionality of the handshaking module (i.e., not as part of a data processing service workflow). It assumes the existence of a simple user interface that enables the interaction between the various end user roles and the respective module services.

- The data processor enters the *Requests from you to others* page. Some previously made requests are already there. She clicks “*Create New Data Access Request*” and is redirected to the *New Data Access Request* page. She fills in the details of the new DAR and clicks “*Submit*”.

- She is automatically redirected back to the *Requests from you to others* page where the new request also appears.
- The data provider enters the *List of Requests from others to you* page, where he sees some older requests along with the request just submitted by the data processor. He clicks "*Edit response*" and is redirected to the *Edit Response to Data Access Request* page. He fills in the required data and clicks "*Approve*" or "*i*". He is then automatically redirected back to the *List of Requests from others to you* page, where the status of the data processor's request appears updated.
- The data processor enters again the *Requests from you to others* page, where she sees that the status of her latest request has been updated. She clicks on "*view detailed responses*" and then "*view usage instructions*" or "*view rejection justification*" to see if her DAR has been accepted or rejected and view the respective instructions or justification message.

5.10.5 Expected extensions and potential new implementations

For covering the requirements and scopes of AI4HealthSec project the Data Sharing Management component can be extended and used as part of Vertical Layer 2, in order to manage the sharing of thread-incident and intelligence information among the various AICS nodes deployed either (a) within a large HCII or (b) among different entities.

Since this information could potentially contain sensitive data, its sharing could be handled in different manners, based on its characteristics/ type, e.g.: (i) In most cases permanent sharing permissions could be granted for sharing trivial information among certain AICS nodes, while (ii) in case of exceptional circumstances the sharing of higher-importance or more sensitive information could be manually granted by an assigned security officer.

The component could be adapted to smartly and effectively handle the incoming handshaking requests for sharing of information in such cases. Given the availability of highly detailed information description and trust models describing its desired level of protection, predefined policies regarding "what can be shared and with whom" could be utilised so that the user's intervention could be only requested in cases that this is deemed necessary. By enabling such advanced distributed data management capabilities, administrative overhead can be reduced, trust can be enhanced and the collaboration and interaction among various entities and HCIIs can be optimised while retaining secure and privacy-aware sharing of information.

5.11 Contribution to the AI4HealthSec Requirements Process

In this section, we analysed a set of existing tools that are provided by the AI4HealthSec partners and they could be adopted in the project to develop the expected capabilities of the respective Framework. All these tools have been used and adopted in similar research and innovation activities in past and running projects and they have been validated in the security domain for a variety of business sectors.

The technical angle that we introduced in this section poses for additional requirements that the Framework should implement. More specifically, the adoption of technical means to support the implementation of the two main processes of the AI4HealthSec Framework, namely risk management

and assessment and incident handling, introduces a set of challenges for the users of the Framework, which should be considered in the project requirements elicitation process, along with the business challenges that the Framework should address, as they were analysed in the previous sections. These technical-oriented challenges relate to:

- The specification of a risk management and assessment methodology, which is based on the collection of evidence from the interconnected IT systems and devices in the healthcare sector (and beyond). This methodology is driven by a swarm intelligence approach and must be able to address issues with respect to the context of the risk methodology, the compliance of the risk management approach, the process for the identification and predication of risks, the specification of the risk modelling and the associated control mechanisms, the definition of the risk assessment process, and the extensibility and applicability of the whole risk management and assessment practices to a variety of critical domains, other than healthcare.
- The cyber-security and risk-based incident handling methodology and practices, which must address challenges with respect to the specification of a multi-level evidence collection environment from disperse vulnerable sources exposed to cyber-attacks, and the introduction of mechanisms for the correlation of attack related information so that incidents are detected in a more efficient manner. These challenges drive the requirements for an in-depth analysis of the runtime operations across all the layers of an IT healthcare supply chain ecosystem, by expediting the analysis of security events and supporting risk-based decisions towards the management of the detected incidents and the implementation of mechanisms in response to cyber-attacks.

In section 7 of this deliverable, we will present the details of the user requirements for each of these two main families of technical challenges.

6 Results: Input by External Advisory Board

The video call with members of the project's EAB was held in March 2021. Before the call each EAB member received a presentation with a short report via e-mail containing the main information on the AI4HealthSec project and framework and on the results from the user requirements analysis. This report can be found in the appendix of this deliverable.

During the call the three EAB members were present accompanied by project consortium members of the pilot sites and of project technical partner organizations. The main ideas of AI4HealthSec as well as the pilot scenarios were presented focusing the requirements elicited during this task T2.1. After the short presentations there was an open discussion with the EAB members to find their opinion and ask for potential further ideas about the requirements elicitation process. It was explicitly not the objective of this call to validate the obtained requirements, but to get more feedback on the previously performed actions and hints on how to proceed.

In the call with experts from the Advisory Board and project members the project's main objectives have been presented as well as requirements elicited from the three pillars (user challenges, domain requirements, technical requirements).

Following the presentation there was an open discussion. All EAB members agreed on the basic methods applied to identify the requirements. It has been reported that hospitals offer special challenges when creating an external cybersecurity framework due to their interconnectivity of

systems that needs to be considered as well as the sensitive nature of their IT structure. It was discussed to use simulated data to build test scenarios at KLINIK – nevertheless, KLINIK’s representatives at the AI4HealthSec project already planned to rework their pilot scenario so that it takes more aspects of cyber-security awareness into account (e.g., preventing social engineering attacks). This pilot re-definition will be part of Task T2.3.

Moreover, the presentation of the requirements to the AI4HealthSec framework and the subsequent discussion lead to the conclusion that the project consortium should not focus on prioritising requirements at this early stage of the project but should consider all identified requirements as the basis for further developments. Therefore, the wording of the requirements was changed from including the word “should” to the word “shall” – this is to emphasize that the requirements are a basis for the project and might be further refined.

An EAB member proposed to use structured models to track down the process of the fulfilment of the requirements during the ongoing project, such as IBM Doors. In addition, it was proposed to use tools for modelling the hospital’s IT infrastructure, for example 3LGM2 Model by Winter and colleagues²⁰. Furthermore, the project AI4HealthSec should re-check literature on Artificial Intelligence and resulting vulnerabilities.

7 AI4HealthSec Requirements

The elicitation of requirements was performed in perspective of three pillars:

- a. User’s Wishes/Challenges for the development of the AI4HealthSec framework from user perspective
- b. Technical Requirements
- c. Domain Requirements

To elicit users’ wishes and therefore to get a basic understanding of the challenges the framework will face, we created questionnaires to be fulfilled both by internal project partners and external organizations from further critical infrastructures (besides healthcare, e.g. financial sector, transportation sector).

In parallel technical requirements were elicited by intense discussions with the technical project partners.

Additionally, the healthcare domain concerning relevant policies and standards have been analysed by reviewing existing literature.

The analysis of the existing set of tools that will be adopted from the AI4HealthSec project to develop the respective framework, concludes the results of this deliverable with the presentation of technical challenges that this framework should address, through the research and innovation activities that will be performed in the remaining of the project. These challenges relate to the six business needs that were presented in Section 3 of this deliverable. Table 5 includes those main two technical challenges that are linked with the six business needs.

²⁰ <https://www.3lgm2.de>

Table 5: Technical Challenges linked to Business Needs

Technical Challenge ID	Description	Relevance to Business Needs
TC1.	Evidence-based, Swarm-driven Risk Management and Assessment Methodology <i>To address issues for the context and compliance of the management approach, identification and predication of risks, the approach for risk assessment management, modelling and control and the applicability to other domains</i>	<ul style="list-style-type: none"> • BN1: Prediction and Prevention of Attacks • BN2: Vulnerability Assessment • BN3: Awareness Creation and Prevention of Human Errors • BN5: Simplification of the Process of Risk Assessment • BN6: Development of Long-Term Strategy of New Protection Solutions.
TC2.	Cyber-security Risk-based Incident Handling Methodology <i>To address issues for multi-level evidence collection, correlation of information to detect incidents and analyse security events and support for incident management and response</i>	<ul style="list-style-type: none"> • BN1: Prediction and Prevention of Attacks • BN3: Awareness Creation and Prevention of Human Errors • BN4: Detection of Abnormal Patterns and Creation of Warnings • BN6: Development of Long-Term Strategy of New Protection Solutions.

We summarise in the next lines the respective requirements extracted from a technical perspective.

TC1: Evidence-based, Swarm-driven Risk Management and Assessment Methodology

Requirements for Risk Management Context and Compliance

- REQ1:** The risk assessment /management models and process shall be considered from a holistic view of internal (i.e., organisational, technical, medical devices) and external context of the complex health care system.
- REQ2:** The introduction of risk assessment/management models and processes in the AI4HEALTHSEC methodology shall adequately consider the complexity of the ICT infrastructure and technical evolution of medical devices that underpin security processes of health care complex adaptive system.
- REQ3:** The risk management approach shall provide an informed real time decision making for managing cyber security risks and ensuring overall business continuity.

- REQ4:** The methodology shall define the organisation cyber security needs, risk appetite, and risk tolerance for the key healthcare ICT infrastructure areas.
- REQ5:** The risk assessment /management approach shall alleviate the limitations of existing risk management methodologies in terms of their ability to deal with ICT systems in the critical infrastructures.
- REQ6:** The methodology shall leverage, use and implement existing cyber security, information security risk management, information security incident management standards including ISO 31000, ISO27001, ISO 27005, ISO 27031, and ISO 27032 associated with the protection of the complex ICT infrastructure.
- REQ7:** The methodology shall offer compliance with the relevant regulation necessary to compliance with the health care information system sector.

Requirements for Risk Identification and Predication

- REQ8:** The methodology shall automatically detect potential cyber-attack and adversary actions using autonomous intelligence swarm agents and reporting to the supervisor agents, so that evidences are combined and correlated with the existing data for the attack predication and new attack vector discovery.
- REQ9:** The methodology shall include a real time communication, interaction, and feedback among hierarchy-based multiple agents including supervisor and swarm agents and create an overall dynamic cyber security situational awareness.
- REQ10:** The methodology and associated risk management framework shall consider organisation-wide vulnerabilities detection using collective behaviour of swarm intelligence taking into account the underlying complexity of the ICT infrastructure and interoperability and interconnectivity among various sub components including medical devices.
- REQ11:** The methodology shall consider depth of access by measuring how far threat actors reach within the ICT infrastructure by collective swarm intelligence data for the risk identification and predication.
- REQ12:** The methodology shall introduce a risk management system, which will consider the nature and interdependencies of cybersecurity and medical assets and as well as their implications on overall business continuity

Requirements for Risk Assessment and Modelling

- REQ13:** The methodology shall adopt an evidence-driven Cyber Security Risk Assessment model in order to capture and deal with cascading effects of risks, threats and vulnerabilities, associated with the health care ICT infrastructure
- REQ14:** The methodology shall help elicit, understand and analyse risk management requirements for the health care ICT infrastructure, with particular emphasis on requirements associated with the overall complex system and its supply chain context.

- REQ15:** The methodology shall consider all organisation wide vulnerabilities by correlating data from the swarm agents and its impact for the net risk calculation.
- REQ16:** The risk assessment approach shall follow quantitative assessment methods to determine the risk level, based on existing consistent cyber security threat data
- REQ17:** The risk assessment approach shall consider Cyber Threat Intelligence (CTI) information including relevant threat actors, their capabilities, skills, motivations, and underlying TTP and IoC.
- REQ18:** The methodology shall consider cyber risk modelling considering assets and their dependencies, vulnerabilities within the assets, possible attack paths, threat intelligence properties, and risks.
- REQ19:** The methodology shall leverage simulation models combined with a multi-criteria decision making approach in order to produce timely, accurate, relevant and high quality evidence, information, indicators, factors and parameters associated based on which the multi-dimensional risks will be assessed.
- REQ20:** The methodology shall use graphs to discover and represent possible attacks plans and patterns and will adopt a general approach to integrate several aspects of both vulnerabilities and threat agents.
- REQ21:** The methodology shall identify and model assets, processes, risks, stakeholders' relationships/interactions and dependencies.
- REQ22:** The methodology shall create a range of metrics covering reliability, credibility, acceptance, timeliness, realism of risk management goals and the level of integration of the risk management approach in decision making structures. These metrics should be able to be measured across all cyber-security assets, medical device, and ICT systems available within health care infrastructure.

Requirements for Risk Management and Control

- REQ23:** The methodology shall determine the level of assurance based on the evidence of existing controls and their effectiveness and recommend alternative courses of action for responding to risks.
- REQ24:** The methodology shall explore new techniques/methods for the credible calculation of insurance premiums.
- REQ25:** The risk management approach shall ensure the constant vigilance of existing risks, by offering mechanisms to understand status of residual value of risk and identifying any new risk using intelligence swarm agents.

Requirements for Incident Management

- REQ26:** The risk analysis methodology shall provide real-time decision making support for incident response and post incident review activities.
- REQ27:** The risk identification, forecasting and analyse shall provide a better understanding of the cyber security incident related information.
- REQ28:** The risk management methodology shall align with the incident response and post-incident activities to ensure eradication of the threats and risks and overall business continuity.
- REQ29:** The risk assessment methodology should support updating threat intelligence information and incident response planning, through lessons learn from the evolving threats, risks and related incidents.

Requirements for Contribution to other Domains

- REQ30:** The risk management methodology shall consider publishing best practices that include blueprints and guidelines for adapting the approach to other critical infrastructures sector, such as smart grid cyber physical systems.
- REQ31:** The AI4HEALTHSEC project shall contribute best practices associated with the deployment and operation of its framework for risk management in health care sector of any type and size.

TC2: Cyber-security Risk-based Incident Handling Methodology

Requirements for Multi-Source Evidence Collection and Preparation

- REQ32:** The incident handling methodology shall support evidence collection on both real time and historic data from the various evidence collection sources to assist incident detection.
- REQ33:** The evidence collection process shall include batch data (i.e., collection of raw data over a specific period), including, but not limited to, log files from vulnerable systems and network traffic.
- REQ34:** The evidence collection process shall include configurable steps, allowing for the specification of the type, format and location of the incoming data sources such as log files.
- REQ35:** The evidence collection process shall consider anonymization of raw data collected by various sources.
- REQ36:** The evidence preparation process shall consider the semi-structured nature of different datasets.
- REQ37:** The data collected shall include records about network usage and bandwidth, and should allow the identification of network traffic anomalies and excessive bandwidth usage.

- REQ38:** The data collection process shall take into consideration and be at least partially aligned with existing industry proprietary or non-proprietary data exchange protocols, with particular interest in understanding to some extent the messages exchanged, including network packages and messages from the interaction among systems.
- REQ39:** The incident handling process should be able to monitor the availability of signals and system web sources or services and calculate their response time for further analysis.
- REQ40:** The incident handling approach shall support normalization and transformation of raw data coming from semantically relevant sources to perform system independent data processing and sharing across the AI4HEALTHSEC Framework.
- REQ41:** The incident handling approach shall consider for managing structural and semantic mismatches across the different datasets collected.
- REQ42:** The incident handling approach shall support normalization and transformation for the unified representation of cyber security threats detected by internal or external components of this platform.
- REQ43:** The evidence preparation process shall support preliminary filtering of raw data, using predefined criteria over the parameters collected from raw data, so that irrelevant one can be removed and/or not taken into consideration in the incident handling process.

Requirements for Evidence Chain Generation and Security Incident Detection

- REQ44:** The incident detection and event analysis approach shall be able to process streaming, batch and historical data.
- REQ45:** The incident detection and event analysis approach shall consider data uncertainty and incompleteness, so that the processing of the provided raw data can be feasible even in the absence of some elements.
- REQ46:** The organization and filtering of the incoming raw data (across all the available data sources) is essential for the further analysis of the current status of the systems. During this process, the evidence chains would be generated and the relevant data would be collected and stored for latter usage.
- REQ47:** The incident detection and event analysis approach shall support the preliminary analysis of relevant raw data (e.g., deviation from normal patterns) to identify potential security incidents.
- REQ48:** The security event analysis approach shall support semantic and structural decisions regarding the description of the different type of incidents so that further processing of the information generated can be feasible and meaningful.
- REQ49:** The incident detection and event analysis approach shall utilize existing knowledge sources with security data (including either external knowledge used for training purposes or other security related knowledge acquired by other modules of the system) for correlating evidence to incidents and security events.

- REQ50:** The incident detection and event analysis approach shall be customizable to further domains, other than health ICT infrastructures.
- REQ51:** The incident handling methodology shall maintain a knowledge base with information about actual successful attack scenarios.
- REQ52:** The incident detection and event analysis approach shall support decision making, towards developing more efficient and effective defence strategies, based on evidence from past detected incidents, extracted from the knowledge base.
- REQ53:** The incident handling methodology shall provide cyber-attacks related information that can be shared with other organizations in a secure and privacy preserving way.

Requirements for Incident Management and Response

- REQ54:** The incident handling methodology shall identify the on-going attacks and related information at all times.
- REQ55:** The incident handling methodology shall be able to predict possible scenarios of future attacks.
- REQ56:** The incident handling methodology shall provide a visual representation of the cyber-attack path.
- REQ57:** The incident handling methodology shall assure an acceptable risk level for the cooperating stakeholders.
- REQ58:** The incident handling methodology shall promote the necessary defensive capabilities and provide a rational decision-making to help stakeholders in determining which security controls shall be implemented to encounter the identified security issues and cyber-risks.
- REQ59:** The incident handling methodology shall support matching evidence collected in real time with archived information for cyber-attack scenarios.
- REQ60:** The incident handling methodology shall be able to provide comparison among the patterns of data collected at the infrastructure nodes and the normal state of operations.
- REQ61:** The incident handling methodology shall allow decision makers in predicting the assets that are exposed to risks when a security event is detected.
- REQ62:** The incident handling methodology shall support decision makers in exploring different attack scenarios on potential harmfulness of a detected anomaly to the infrastructure.
- REQ63:** The incident handling methodology shall present the attack path of a detected incident across all impacted assets.
- REQ64:** The incident handling methodology shall present sufficient information to decision makers to enable them to understand the risk of cyber-attacks detected in real time on the infrastructure.

- REQ65:** The incident handling methodology shall provide decision makers with access to the results of the risk assessment process at any time, to understand the consequences of a detected cyber-attack.
- REQ66:** The incident handling methodology shall provide recommendations to decision makers on the most suitable security controls to mitigate the risks from detected security events and cyber risks.
- REQ67:** The incident handling methodology shall allow decision makers understand the impact from the implementation of a defensive mechanism to support informed decisions when selecting the appropriate security controls.

8 Conclusions

By using our approach that took all three relevant pillars into account we were able to provide a broad picture of challenges from a users' perspective, technical requirements, and domain requirements. We found that especially patients and other clients of organizations are most vulnerable to cyber-threats, whereas physicians, nurses and other non-technical staff is the main weak point at which cyber-attacks might occur most easily. This is due to a missing cybersecurity awareness of numerous organizational members which should be trained more to ensure a higher cyber-security level in the future.

Most members, both of project organizations and external companies, are rather knowledgeable when it comes to cybersecurity but at the same time, training in this field seems sometimes to be lacking efficacy in regards of its ability to prevent dangerous situations – especially in the hospital setting. Most organizations are in favour of engaging in CIP programs and on improving the overall situational awareness. An external framework on cyber-security might offer tools that fit into the existing IT infrastructure and include support for users to enable organizations to act more cyber-secure aware. The findings, if a framework should be more running in the background or not and if it should provide interaction with the user are rather heterogeneous. For both external and internal organizations, a little bit more persons stated they wanted it visible with input needed by the user.

In the analysis of the user's perspective, we elicited six business needs that can be seen as user-based challenges that should all be considered when designing the future AI4HealthSec framework. This set the basis for the subsequent domain and technical requirements analysis pillars.

At this state of the project, the very beginning, all requirements are to be seen as the foundation for the further tasks. The MoSCoW method²¹ will be considered in further stages of the project to enable a further prioritization of requirements.

²¹ <https://www.projectsmart.co.uk/moscow-method.php>

9 References

- [1] Spyridon Papastergiou and Nineta Polemi. Mitigate: A dynamic supply chain cyber risk assessment methodology. In *Smart Trends in Systems, Security and Sustainability*, pages 1–9. Springer, 2018.
- [2] Eleni Maria Kalogeraki, Despina Polemi, Spyros Papastergiou, and Themis Panayiotopoulos. Modeling SCADA Attacks. In: Yang XS., Nagar A., Joshi A. (eds) *Smart Trends in Systems, Security and Sustainability. Lecture Notes in Networks and Systems*, vol 18. Springer, Singapore, pages 47–55. 01 2018.
- [3] Spyridon Papastergiou and Despina Polemi. Securing maritime logistics and supply chain: The medusa and mitigate approaches. *Marit. Interdiction Oper. J.*, 14:42–48, 2017.
- [4] Eleni Maria Kalogeraki, Spyros Papastergiou, Haris Mouratidis and Nineta Polemi (2018) “A novel risk assessment methodology for SCADA maritime logistics environments”, *Applied Sciences*, MDPI AG, Switzerland, 8(9): 1477, ISSN: 2076-3417, <https://doi.org/10.3390/app8091477>.
- [5] Eleni Maria Kalogeraki., Dimitrios Apostolou, Nineta Polemi, Spyros Papastergiou S. (2018) "Knowledge Management Methodology for Identifying Threats in Maritime/Logistics Supply Chains" in S. Durtst, P. Evangelista (Eds) (SI) “Logistics knowledge management: state of the art and future perspectives”, *Knowledge Management Research and Practice Journal*, Taylor and Francis, ISSN: 1477-8238 (Print) 1477-8246, DOI: 10.1080/14778238.2018.14867What do external organizations recommend?
- [6] Mohammad S Jalali, Bethany Russell, Sabina Razak, William J Gordon, “Journal of the American Medical Informatics Association”, Volume 26, Issue 1, January 2019, Pages 81–90, <https://doi.org/10.1093/jamia/ocy148>
- [7] <https://ponemonsullivanreport.com/2016/04/>
- [8] ENISA, “ENISA programming document 2019-2021,” [Online]. Available: <https://www.enisa.europa.eu/publications/corporate-documents/enisa-programming-document-2019-2021>.
- [9] <https://www.enisa.europa.eu/publications/security-and-resilience-in-ehealth-infrastructures-and-services>
- [10] Health care and cyber security: Increasing Threats Require Increased Capabilities, KPMG, 2015.
- [11] Ensuring Data Integrity in Health Information Exchange, AHIMA Thought Leadership Series, American Health Information Management Association, 2012 p.2.
- [12] ENISA “Strategies for Incident Response and Cyber Crisis Cooperation”, version 1.1, August 2016.
- [13] NIST SP 800-61 Rev. 2 "Computer Security Incident Handling Guide" (2012) Supersedes: SP 800-61 Rev. 1 (03/07/2008). National Institute of Standards and Technology.
- [14] Computer Security Incident Handling Guide - an overview | ScienceDirect Topics.” <https://www.sciencedirect.com/topics/computer-science/computer-security-incident-handling-guide>.
- [15] ENISA “How to set up CSIRT and SOC”, December 2020 <https://www.enisa.europa.eu/publications/how-to-set-up-csirt-and-soc>.

10 Appendix

10.1 Introduction Text Questionnaire

With your participation to this questionnaire, you agree to the processing of your given information, that is the job title and opinion on cyber-security issues, by the AI4HealthSec consortium to study and extract anonymous information regarding user requirements. The objective is the extraction of user requirements for the AI4HealthSec framework. For this you will represent your organization and will not be asked to answer any questions that are directly linked to you as a private person. The AI4HealthSec consortium—particularly all those organizations involved in gathering, processing and analysing end-user needs— is aware of the possible sensitive nature of the subject and will not include any information that is not suitable for the public domain. At the same time, all participant's personal details (job title, opinion on cyber-security issues) will remain anonymous and 'firewalled'. The data will be anonymized and aggregated to keep it confidential.

It is possible that you in your job function will be contacted again after fulfilling this questionnaire. This will be the case when project partners need more detailed information on how to link the developed AI4HealthSec framework with the technical cyber-security details of your organization. Again, no personal information will be asked.

Research results will be used for extracting user requirements in the project AI4HealthSec and to link the AI4HealthSec framework to those wishes and expectations.

The data will be stored until the end of the AI4HealthSec project, 30/09/2023, and then permanently deleted.

Participation in this study is voluntary. If you choose to participate, you can nonetheless leave this questionnaire at any time without being required to provide any explanation. Should you wish to withdraw your consent regarding the processing of your personal data, you can contact: lena.griebel@klinikum-nuernberg.de.

10.1Part A: Questionnaire Internal Organization

Questionnaire only for 1 representative of pilot organization (i.e. direct project consortium member)

1. Please select the type of your organisation from the list below:

- ☐ Public
- ☐ Private
- ☐ Other, which?

2. What is the size of your organisation?

- ☐ Very Large (> 250 employees)
- ☐ Large (100 - 250)
- ☐ Medium (50 - 100)
- ☐ Small (10 - 50)
- ☐ Micro - Very Small (<10)

3. Which type of security and incident management model does your organization adopt?

- ☐ Outsource (supported by external organization)
- ☐ Inhouse (internal support)
- ☐ Other, which?

4. Which Security Management standards, security protocols and proved guidelines have your organization adopted? (more than one answer possible)

- ☐ ISO 9001
- ☐ ISO/IEC 27035
- ☐ ISO/IEC 27001
- ☐ ISO/IEC 27002
- ☐ ISO 20000
- ☐ ISO/IEC 27005
- ☐ NIST SP800-30
- ☐ NIST SP800-61
- ☐ NIST Framework for Improving Critical Infrastructure Cybersecurity
- ☐ Other, which?

5. In case of a security breach:

- a. Do you have an Incident Response Team?
 - ☐ Yes

☐ No

b. Do your procedures cover cyber-attacks/incidents?

☐ Yes

☐ No

c. Does your organization employ advanced response capabilities to effectively respond to security incidents?

☐ Yes

☐ No

d. Do your procedures estimate the cascading effects of a security events?

☐ Yes

☐ No

e. Does your organization cooperate with external entities to correlate and share incident information to achieve a cross-organization perspective on incident awareness and more effective incident responses?

☐ Yes

☐ No

6. Does your organization employ automated mechanisms to support the incident handling process?

☐ Yes

☐ No

7. Are you collecting security related data (e.g., logs, attacks, etc.) from your applications and how do you store them?

- a. We collect data and store them in files.
- b. We collect data and store them in databases.
- c. We do not collect such data.
- d. We outsource such tasks to external companies.

Other, please specify:

8. Are you using any tool/software/methodology to evaluate, monitor, detect and manage possible organization wise and supply chain threats and risks?

- a. Yes.
- b. No.
- c. I do not know.
- d. We outsource such tasks to external companies.

If yes, please specify which tool/software:

9. Are you exchanging security related practices and information with other relevant organizations in the field?
- Yes, we both share data and consume information.
 - Yes, we simply share (part of) our data.
 - Yes, we just consume attacks related data from other organizations.
 - No, we do not use or share such kind of data.
 - No, we outsource such tasks to external companies.

If yes, please provide example organizations (or types, such as healthcare, ministry, NGO):

10. Have you ever performed a cyber-risk assessment?

- ☐ Yes
☐ No
☐ Does not apply

11. Does your Organization provide an effective IT Security management plan?

- ☐ Yes
☐ No
☐ I do not know/ Does not apply

12. Security Policies and Procedures that are in place within your organization/company (more than one answer possible)

- ☐ Incident Handling and Response Policy or/and Procedure
☐ Information Security Incident Management or/and Procedure
☐ Disaster Recovery and Data Backup Policy or/and Procedure
☐ Business Continuity Policy or/and Procedure
☐ Security Monitoring Policy or/and Procedure
☐ Access Control Policy or/and Procedure
☐ Security Monitoring Policy or/and Procedure
☐ Malicious Software Policy or/and Procedure
☐ Network Access Policy or/and Procedure
☐ Identification and Authentication Policy or/and Procedure
☐ Third Party Connectivity Policy or/and Procedure
☐ Other, which?

13. Are there skilled and trained personnel on security and incident handling practices?

- ☐ Yes, most of the personnel is skilled and trained
- ☐ A few of the personnel are skilled and trained
- ☐ None of the personnel is skilled and trained

14. Does your organization offers / is willing to offer training programs on its employees about cyber-security awareness?

- ☐ Yes
- ☐ No

15. If yes, how frequently are drills provided?

- ☐ Occasionally
- ☐ Annually
- ☐ 2-3 times a year
- ☐ More often

16. If 9=yes, in what form are drills provided? (more than one answer possible)

- ☐ Regularly offered programs with fix agenda
- ☐ Information given when needed to employee (on the job)
- ☐ Introduction to a new employee about cyber-security awareness when they start their job at your organization
- ☐ Other, how?

17. Does your organization employ a centralized solution to correlate incident information and individual incident responses in order to achieve an organization-wide perspective on incident awareness and response?

- ☐ Yes
- ☐ No

10.2Part B Fraunhofer: Questionnaires Internal Organizations

Questionnaire for scenario “Implantable Medical Devices”

1. What are you responsible for in the development of implantable medical devices (IMD)?
 - ☐ Software development
 - ☐ Hardware development
2. How long have you been in your current position?
 - ☐ Under 1 year
 - ☐ 1-5 years
 - ☐ 5-10 years
 - ☐ Longer
3. Have you been trained on cyber security issues by your organization?
 - ☐ Yes
 - ☐ No
4. If no: Have you been trained on cyber security issues by other organizations?
 - ☐ Yes
 - ☐ No
5. How high would you say is your knowledge regarding cyber-security topics?
 - ☐ Very high
 - ☐ Rather high
 - ☐ Average
 - ☐ Below average
 - ☐ Very low
6. What computer software do you typically use in your daily work life? (also email programs, office etc.)
7. What computer hardware do you typically use in your daily work life?
8. Have you been professionally trained on the computer hardware and software that you use in your daily work life?
 - ☐ Yes, on hardware

- ☐ Yes, on software
- ☐ Yes, on both
- ☐ No

9. Do you personally perform any particular tasks in the field of cyber security?

- ☐ Yes
- ☐ No

If yes, which?

10. Have you personally been involved in the risk management process?

- ☐ Yes
- ☐ No

11. Have you encountered cyber security incidents in your professional area over the past 3 years?

- ☐ Yes
- ☐ No

If yes: Have you personally been involved in a cyber-security incident (i.e. incident in your direct work environment)?

- ☐ Yes
- ☐ No

If you have been personally involved in a cyber-security incident:

Did you feel secure with handling the incident? Did you know what to do?

- ☐ Yes, completely knew what to do
- ☐ Was a bit unsure but got help
- ☐ Did not know what to do and got no help

12. Where do you personally see the biggest threats by cyber-attacks regarding IMDs? Where could criminals attack most easily?

13. Which person groups do you think are most vulnerable to negative consequences of cyber-attacks regarding medical implants? (e.g. patients, physicians) and in what way?
14. What do you think are the worst possible consequences of cyber-attacks regarding IMDs?
15. Do you agree: Your organization well-formed security incident management policies allow you to improve the company's situational awareness
- ☐ Strongly Agree
 - ☐ Agree
 - ☐ Disagree
 - ☐ Strongly disagree
16. In recent years: What do you think was the overall opinion of security officers and IT employees in your organization towards the engagement in a CIP (critical infrastructure protection) program?
- ☐ Strongly ambivalent
 - ☐ Rather ambivalent
 - ☐ Rather in favor of engaging in a CIP program
 - ☐ Strongly in favor of engaging in a CIP program
17. Do you think it would be useful if your organization had a CIP program?
- ☐ Not useful at all
 - ☐ Rather not useful
 - ☐ Somewhat useful
 - ☐ Very useful
18. In your opinion: What are the main benefits AI4HealthSec components could offer to you and your work?
19. Which features should an external dynamic and self-organized cyber security framework as AI4HealthSec aims to provide include that it could reduce the risks of cyber-attacks in your work environment?

20. What are your main concerns regarding an external cyber security framework like AI4HealthSec?
21. What do you wish an external cyber security framework should look like in your daily work life?
- ☐ Visible, e.g. by regular status reports
 - ☐ Invisible in the background
22. Do you wish to interact with an external cyber security framework in your daily work life?
- ☐ Yes, I would like to have a system that needs input by me
 - ☐ No, I would like to have a system that runs completely by itself

Questionnaire for scenario “Wearables”

1. What are you responsible for in the development of medical wearables?
☐ App development
☐ Backend development
2. How long have you been in your current position?
☐ Under 1 year
☐ 1-5 years
☐ 5-10 years
☐ Longer
3. Have you been trained on cyber security issues by your organization?
☐ Yes
☐ No
4. If no: Have you been trained on cyber security issues by other organizations?
☐ Yes
☐ No
5. How high would you say is your knowledge regarding cyber-security topics?
☐ Very high
☐ Rather high
☐ Average
☐ Below average
☐ Very low
6. What computer software do you typically use in your daily work life? (also email programs, office etc.)
7. What computer hardware do you typically use in your daily work life?
8. Have you been trained on the computer hardware and software that you use in your daily work life?
☐ Yes, on hardware
☐ Yes, on software

☐ Yes, on both

☐ No

9. Do you personally perform any particular tasks in the field of cyber security?

☐ Yes

☐ No

If yes, which?

10. Have you personally been involved in the risk management process?

☐ Yes

☐ No

11. Have you encountered cyber security incidents in your professional area over the past 3 years?

☐ Yes

☐ No

If yes: Have you personally been involved in a cyber-security incident (i.e. incident in your direct work environment)?

☐ Yes

☐ No

If you have been personally involved in a cyber-security incident: Did you feel secure with handling the incident? Did you know what to do?

☐ Yes, completely knew what to do

☐ Was a bit unsure but got help

☐ Did not know what to do and got no help

12. Where do you personally see the biggest threats by cyber-attacks regarding medical wearables? Where could criminals attack most easily?

13. Which person groups do you think are most vulnerable to negative consequences of cyber-attacks regarding wearables? (e.g. patients, physicians) and in what way?
14. What do you think are the worst possible consequences of cyber-attacks regarding medical wearables?
15. Do you agree: Your organization well-formed security incident management policies allow you to improve the company's situational awareness
- ☐ Strongly Agree
- ☐ Agree
- ☐ Disagree
- ☐ Strongly disagree
16. In recent years: What do you think was the overall opinion of security officers and IT employees in your organization towards the engagement in a CIP (critical infrastructure protection) program?
- ☐ Strongly ambivalent
- ☐ Rather ambivalent
- ☐ Rather in favor of engaging in a CIP program
- ☐ Strongly in favor of engaging in a CIP program
17. Do you think it would be useful if your organization had a CIP program?
- ☐ Not useful at all
- ☐ Rather not useful
- ☐ Somewhat useful
- ☐ Very useful
18. In your opinion: What are the main benefits AI4HealthSec components could offer to you and your work?

19. Which features should an external dynamic and self-organized cyber security framework as AI4HealthSec aims to provide include that it could reduce the risks of cyber-attacks in your work environment?
20. What are your main concerns regarding an external cyber security framework like AI4HealthSec?
21. What do you wish an external cyber security framework should look like in your daily work life?
- ☐ Visible, e.g. by regular status reports
 - ☐ Invisible in the background
22. Do you wish to interact with an external cyber security framework in your daily work life?
- ☐ Yes, I would like to have a system that needs input by me
 - ☐ No, I would like to have a system that runs completely by itself

Questionnaire for scenario “Biobanks” for developers, IT managers, and biobank operators

This questionnaire is for the following groups

Developers of biobank application

IT managers

Biobank operators

For the questionnaires for biologist see B.5

1. What is your job position?
 - ☐ Developer of biobank application
 - ☐ IT manager
 - ☐ Biobank operator

2. How long have you been in your current position?
 - ☐ Under 1 year
 - ☐ 1-5 years
 - ☐ 5-10 years
 - ☐ Longer

3. Have you been trained on cyber security issues by your organization?
 - ☐ Yes
 - ☐ No

4. If no: Have you been trained on cyber security issues by other organizations?
 - ☐ Yes
 - ☐ No

5. How high would you say is your knowledge regarding cyber-security topics?
 - ☐ Very high
 - ☐ Rather high
 - ☐ Average
 - ☐ Below average
 - ☐ Very low

6. What computer software do you typically use in your daily work life? (also email programs, office etc.)
7. What computer hardware do you typically use in your daily work life?
8. Have you been trained on the computer hardware and software that you use in your daily work life?

- ☐ Yes, on hardware
- ☐ Yes, on software
- ☐ Yes, on both
- ☐ No

9. Do you personally perform any particular tasks in the field of cyber-security?

- ☐ Yes
- ☐ No

If yes, which?

10. Have you personally been involved in the risk management process?

- ☐ Yes
- ☐ No

11. Have you encountered cyber security incidents in your professional area over the past 3 years?

- ☐ Yes
- ☐ No

If yes: Have you personally been involved in a cyber-security incident (i.e. incident in your direct work environment)?

- ☐ Yes
- ☐ No

If you have been personally involved in a cyber-security incident: Did you feel secure with handling the incident? Did you know what to do?

- ☐ Yes, completely knew what to do
- ☐ Was a bit unsure but got help

☐ Did not know what to do and got no help

12. Where do you personally see the biggest threats by cyber-attacks regarding biobanks? Where could criminals attack most easily?

13. Which person groups do you think are most vulnerable to negative consequences of cyber-attacks regarding biobanks? (e.g. patients, physicians) and in what way?

14. What do you think are the worst possible consequences of cyber-attacks regarding biobanks?

15. Do you agree: Your organization well-formed security incident management policies allow you to improve the company's situational awareness

☐ Strongly Agree

☐ Agree

☐ Disagree

☐ Strongly disagree

16. In recent years: What do you think was the overall opinion of security officers and IT employees in your organization towards the engagement in a CIP (critical infrastructure protection) program?

☐ Strongly ambivalent

☐ Rather ambivalent

☐ Rather in favor of engaging in a CIP program

☐ Strongly in favor of engaging in a CIP program

17. Do you think it would be useful if your organization had a CIP program?

☐ Not useful at all

☐ Rather not useful

☐ Somewhat useful

☐ Very useful

18. In your opinion: What are the main benefits AI4HealthSec components could offer to you and your work?
19. Which features should an external dynamic and self-organized cyber security framework as AI4HealthSec aims to provide include that it could reduce the risks of cyber-attacks in your work environment?
20. What are your main concerns regarding an external cyber security framework like AI4HealthSec?
21. What do you wish an external cyber security framework should look like in your daily work life?
- ☐ Visible, e.g. by regular status reports
 - ☐ Invisible in the background
22. Do you wish to interact with an external cyber security framework in your daily work life?
- ☐ Yes, I would like to have a system that needs input by me
 - ☐ No, I would like to have a system that runs completely by itself

Questionnaire for scenario “Biobanks” for Biologists

1. How long have you been in your current position?
☐ Under 1 year
☐ 1-5 years
☐ 5-10 years
☐ Longer
2. Have you been trained on cyber security issues by your organization?
☐ Yes
☐ No
3. If no: Have you been trained on cyber security issues by other organizations?
☐ Yes
☐ No
4. How high would you say is your knowledge regarding cyber-security topics?
☐ Very high
☐ Rather high
☐ Average
☐ Below average
☐ Very low
5. What computer software do you typically use in your daily work life? (also email programs, office etc.)
6. What computer hardware do you typically use in your daily work life?
7. Have you been trained on the computer hardware and software that you use in your daily work life?
☐ Yes, on hardware
☐ Yes, on software
☐ Yes, on both
☐ No
8. Do you personally perform any particular tasks in the field of cyber-security?

☐ Yes

☐ No

If yes, which?

9. Have you personally been involved in the risk management process?

☐ Yes

☐ No

10. Have you encountered cyber security incidents in your professional area over the past 3 years?

☐ Yes

☐ No

If yes: Have you personally been involved in a cyber-security incident (i.e. incident in your direct work environment)?

☐ Yes

☐ No

If you have been personally involved in a cyber-security incident: Did you feel secure with handling the incident? Did you know what to do?

☐ Yes, completely knew what to do

☐ Was a bit unsure but got help

☐ Did not know what to do and got no help

11. Where do you personally see the biggest threats by cyber-attacks regarding biobanks? Where could criminals attack most easily?

12. In your opinion: What are the main benefits AI4HealthSec components could offer to you and your work?

13. Which features should an external dynamic and self-organized cyber security framework as AI4HealthSec aims to provide include that it could reduce the risks of cyber-attacks in your work environment?

14. What are your main concerns regarding an external cyber security framework like AI4HealthSec?
15. What do you wish an external cyber security framework should look like in your daily work life?
- ☐ Visible, e.g. by regular status reports
 - ☐ Invisible in the background
16. Do you wish to interact with an external cyber security framework in your daily work life?
- ☐ Yes, I would like to have a system that needs input by me
 - ☐ No, I would like to have a system that runs completely by itself

10.3Part B UoB: Questionnaire Internal Organization

1. What is your role in UoB's Living Lab?
☐ Researcher
☐ IT specialist
2. How long have you been in your current position?
☐ Under 1 year
☐ 1-5 years
☐ 5-10 years
☐ Longer
3. Have you been trained on cyber security issues by your organization?
☐ Yes
☐ No
4. If no: Have you been trained on cyber security issues by other organizations?
☐ Yes
☐ No
5. How high would you say is your knowledge regarding cyber-security topics?
☐ Very high
☐ Rather high
☐ Average
☐ Below average
☐ Very low
6. What computer software do you typically use in your daily work life? (also email programs, office etc.)
7. What computer hardware do you typically use in your daily work life?
8. Have you been professionally trained on the computer hardware and software that you use in your daily work life?
☐ Yes, on hardware
☐ Yes, on software
☐ Yes, on both

☐ No

9. Do you personally perform any particular tasks in the field of cyber security?

☐ Yes

☐ No

If yes, which?

10. Have you personally been involved in the risk management process?

☐ Yes

☐ No

11. Have you encountered cyber security incidents in your professional area over the past 3 years?

☐ Yes

☐ No

If yes: Have you personally been involved in a cyber-security incident (i.e. incident in your direct work environment)?

☐ Yes

☐ No

If you have been personally involved in a cyber-security incident:

Did you feel secure with handling the incident? Did you know what to do?

☐ Yes, completely knew what to do

☐ Was a bit unsure but got help

☐ Did not know what to do and got no help

12. Where do you personally see the biggest threats by cyber-attacks regarding University of Brighton's (UoB) Living Lab? Where could criminals attack most easily?

13. Which person groups do you think are most vulnerable to negative consequences of cyber-attacks regarding medical UoB's Living Lab? (e.g. patients, physicians) and in what way?

14. What do you think are the worst possible consequences of cyber-attacks regarding UoB's Living Lab?

15. Do you agree: UoB's Living Lab's well-formed security incident management policies allow you to improve the company's situational awareness
- ☐ Strongly Agree
 - ☐ Agree
 - ☐ Disagree
 - ☐ Strongly disagree
16. In recent years: What do you think was the overall opinion of security officers and IT employees in UoB's Living Lab towards the engagement in a CIP (critical infrastructure protection) program?
- ☐ Strongly ambivalent
 - ☐ Rather ambivalent
 - ☐ Rather in favor of engaging in a CIP program
 - ☐ Strongly in favor of engaging in a CIP program
17. Do you think it would be useful if UoB's Living Lab had a CIP program?
- ☐ Not useful at all
 - ☐ Rather not useful
 - ☐ Somewhat useful
 - ☐ Very useful
18. In your opinion: What are the main benefits AI4HealthSec components could offer to you and your work?
19. Which features should an external dynamic and self-organized cyber security framework as AI4HealthSec aims to provide include that it could reduce the risks of cyber-attacks in your work environment?
20. What are your main concerns regarding an external cyber-security framework like AI4HealthSec?
21. What do you wish an external cyber-security framework should look like in your daily work life?

- ☐ Visible, e.g. by regular status reports
- ☐ Invisible in the background

22. Do you wish to interact with an external cyber security framework in your daily work life?

- ☐ Yes, I would like to have a system that needs input by me
- ☐ No, I would like to have a system that runs completely by itself

10.4Part B EBIT: Questionnaire Internal Organization

1. What is your role/your job at the hospital you are currently working at (e.g. IT security officer)?
2. How long have you been in your current position?
 - ☐ Under 1 year
 - ☐ 1-5 years
 - ☐ 5-10 years
 - ☐ Longer
3. Have you been trained on cyber security issues by your hospital?
 - ☐ Yes
 - ☐ No
4. If no: Have you been trained on cyber security issues by other organizations?
 - ☐ Yes
 - ☐ No
5. How high would you say is your knowledge regarding cyber-security topics?
 - ☐ Very high
 - ☐ Rather high
 - ☐ Average
 - ☐ Below average
 - ☐ Very low
6. Do you agree?

I think that **medical staff** like physicians and nurses at my hospital is proficient enough when it comes to cyber-security to prevent dangerous situations.

 - ☐ I fully agree
 - ☐ I partially agree
 - ☐ I mostly disagree
 - ☐ I completely disagree

7. Do you agree?

I think that the training in cyber-security for the **medical staff** at my hospital is sufficient to prevent dangerous situations?

- ☐ I fully agree
- ☐ I partially agree
- ☐ I mostly disagree
- ☐ I completely disagree

8. Do you agree?

I think that **administrative staff** at my hospital is proficient enough when it comes to cyber-security to prevent dangerous situations.

- ☐ I fully agree
- ☐ I partially agree
- ☐ I mostly disagree
- ☐ I completely disagree

9. Do you agree?

I think that the training in cyber-security for the **administrative staff** at my hospital is sufficient to prevent dangerous situations.

- ☐ I fully agree
- ☐ I partially agree
- ☐ I mostly disagree
- ☐ I completely disagree

10. What computer software do you typically use in your daily work life? (also email programs, office etc.)

11. What computer hardware do you typically use in your daily work life?

12. Have you been professionally trained on the computer hardware and software that you use in your daily work life?

- ☐ Yes, on hardware
- ☐ Yes, on software
- ☐ Yes, on both
- ☐ No

13. Do you personally perform any particular tasks in the field of cyber security?

- ☐ Yes
☐ No

If yes, which?

14. Have you personally been involved in the risk management process?

- ☐ Yes
☐ No

15. Have you encountered cyber security incidents in your professional area over the past 3 years?

- ☐ Yes
☐ No

If yes: Have you personally been involved in a cyber-security incident (i.e. incident in your direct work environment)?

- ☐ Yes
☐ No

If you have been personally involved in a cyber-security incident:
Did you feel secure with handling the incident? Did you know what to do?

- ☐ Yes, completely knew what to do
☐ Was a bit unsure but got help
☐ Did not know what to do and got no help

16. Where do you personally see the biggest threats by cyber-attacks regarding your hospital?
Where could criminals attack most easily?

17. Which person groups do you think are most vulnerable to negative consequences of cyber-attacks regarding your hospital? (e.g. patients, physicians) and in what way?

18. What do you think are the worst possible consequences of cyber-attacks regarding your hospital?

19. Do you agree: My hospital's well-formed security incident management policies allow my hospital to improve the overall's situational awareness
- ☐ Strongly Agree
 - ☐ Agree
 - ☐ Disagree
 - ☐ Strongly disagree
20. In recent years: What do you think was the overall opinion of security officers and IT employees in your hospital towards the engagement in a CIP (critical infrastructure protection) program?
- ☐ Strongly ambivalent
 - ☐ Rather ambivalent
 - ☐ Rather in favor of engaging in a CIP program
 - ☐ Strongly in favor of engaging in a CIP program
21. Do you think it would be useful if your hospital had a CIP program?
- ☐ Not useful at all
 - ☐ Rather not useful
 - ☐ Somewhat useful
 - ☐ Very useful
22. In your opinion: What are the main benefits AI4HealthSec components could offer to you and your work?
23. Which features should an external dynamic and self-organized cyber security framework as AI4HealthSec aims to provide include that it could reduce the risks of cyber-attacks in your work environment?
24. What are your main concerns regarding an external cyber-security framework like AI4HealthSec?
25. What do you wish an external cyber-security framework should look like in your daily work life?

- ☐ Visible, e.g. by regular status reports
- ☐ Invisible in the background

26. Do you wish to interact with an external cyber security framework in your daily work life?

- ☐ Yes, I would like to have a system that needs input by me
- ☐ No, I would like to have a system that runs completely by itself

27. Do you agree on the following statements?

- a. I would find the AI4HealthSec framework useful in my job.
 - ☐ Fully agree
 - ☐ Partially agree
 - ☐ Mostly disagree
 - ☐ Completely disagree
- b. I think that the AI4HealthSec framework is a good concept.
 - ☐ Fully agree
 - ☐ Partially agree
 - ☐ Mostly disagree
 - ☐ Completely disagree
- c. A specific group or person would be available for assistance with difficulties with the AI4HealthSec framework.
 - ☐ Fully agree
 - ☐ Partially agree
 - ☐ Mostly disagree
 - ☐ Completely disagree

10.5Part B Klinik: Questionnaire Internal Organization

1. What is your role/your job at the Klinikum Nürnberg (e.g. IT security officer)?
2. How long have you been in your current position?
 - ☐ Under 1 year
 - ☐ 1-5 years
 - ☐ 5-10 years
 - ☐ Longer
3. Have you been trained on cyber security issues by Klinikum Nürnberg?
 - ☐ Yes
 - ☐ No
4. If no: Have you been trained on cyber security issues by other organizations?
 - ☐ Yes
 - ☐ No
5. How high would you say is your knowledge regarding cyber-security topics?
 - ☐ Very high
 - ☐ Rather high
 - ☐ Average
 - ☐ Below average
 - ☐ Very low
6. Do you agree?

I think that **medical staff** like physicians and nurses at Klinikum Nürnberg is proficient enough when it comes to cyber-security to prevent dangerous situations.

 - ☐ I fully agree
 - ☐ I partially agree
 - ☐ I mostly disagree
 - ☐ I completely disagree
7. Do you agree?

I think that the training in cyber-security for the **medical staff** here at Klinikum Nürnberg is sufficient to prevent dangerous situations?

 - ☐ I fully agree

- ☐ I partially agree
- ☐ I mostly disagree
- ☐ I completely disagree

8. Do you agree?

I think that **administrative staff** here at Klinikum Nürnberg is proficient enough when it comes to cyber-security to prevent dangerous situations.

- ☐ I fully agree
- ☐ I partially agree
- ☐ I mostly disagree
- ☐ I completely disagree

9. Do you agree?

I think that the training in cyber-security for the **administrative staff** here at Klinikum Nürnberg is sufficient to prevent dangerous situations.

- ☐ I fully agree
- ☐ I partially agree
- ☐ I mostly disagree
- ☐ I completely disagree

10. Which factors do you think offers the largest cyber-security risks at Klinikum Nürnberg? (e.g. mistakes by staff, old software...)

11. What computer software, also including software for medical technology, do you typically use in your daily work life? (also email programs, office etc.)

12. What computer hardware, also including medical technology, do you typically use in your daily work life?

13. Have you been professionally trained on the computer hardware and software that you use in your daily work life?

- ☐ Yes, on hardware
- ☐ Yes, on software
- ☐ Yes, on both
- ☐ No

14. Do you personally perform any particular tasks in the field of cyber security?

- ☐ Yes
- ☐ No

If yes, which?

15. Have you personally been involved in the risk management process?

☐ Yes

☐ No

16. Have you encountered cyber security incidents in your professional area over the past 3 years?

☐ Yes

☐ No

If yes: Have you personally been involved in a cyber-security incident (i.e. incident in your direct work environment)?

☐ Yes

☐ No

If you have been personally involved in a cyber-security incident:

Did you feel secure with handling the incident? Did you know what to do?

☐ Yes, completely knew what to do

☐ Was a bit unsure but got help

☐ Did not know what to do and got no help

17. Where do you personally see the biggest threats by cyber-attacks regarding Klinikum Nürnberg? Where could criminals attack most easily?

18. Which person groups do you think are most vulnerable to negative consequences of cyber-attacks regarding Klinikum Nürnberg? (e.g. patients, physicians) and in what way?

19. What do you think are the worst possible consequences of cyber-attacks regarding Klinikum Nürnberg?

20. Do you agree: Klinikum Nürnberg's well-formed security incident management policies allow you to improve the company's situational awareness

☐ Strongly Agree

☐ Agree

☐ Disagree

☐ Strongly disagree

21. In recent years: What do you think was the overall opinion of security officers and IT employees in Klinikum Nürnberg towards the engagement in a CIP (critical infrastructure protection) program?
- ☐ Strongly ambivalent
 - ☐ Rather ambivalent
 - ☐ Rather in favor of engaging in a CIP program
 - ☐ Strongly in favor of engaging in a CIP program
22. Do you think it would be useful if Klinikum Nürnberg had a CIP program?
- ☐ Not useful at all
 - ☐ Rather not useful
 - ☐ Somewhat useful
 - ☐ Very useful
23. What do you think would be the most important measure Klinikum Nürnberg should take that you and your colleagues would feel more secure when it comes to cyber-security? *(even if you think that several alternatives are important please select the one that you think is the most important)*
- ☐ More training on cyber-security
 - ☐ Organization-wide awareness campaigns on cyber-security
 - ☐ Technology-based solutions such as firewalls
 - ☐ Others, which
24. In your opinion: What are the main benefits AI4HealthSec components could offer to you and your work?
25. Which features should an external dynamic and self-organized cyber security framework as AI4HealthSec aims to provide include that it could reduce the risks of cyber-attacks in your work environment?
26. What are your main concerns regarding an external cyber-security framework like AI4HealthSec?

27. What do you wish an external cyber-security framework should look like in your daily work life?

- ☐ Visible, e.g. by regular status reports
- ☐ Invisible in the background

28. Do you wish to interact with an external cyber security framework in your daily work life?

- ☐ Yes, I would like to have a system that needs input by me
- ☐ No, I would like to have a system that runs completely by itself

29. Do you agree on the following statements?

- a. I would find the AI4HealthSec framework useful in my job.
 - ☐ Fully agree
 - ☐ Partially agree
 - ☐ Mostly disagree
 - ☐ Completely disagree
- b. I think that the AI4HealthSec framework is a good concept.
 - ☐ Fully agree
 - ☐ Partially agree
 - ☐ Mostly disagree
 - ☐ Completely disagree
- c. A specific group or person would be available for assistance with difficulties with the AI4HealthSec framework.
 - ☐ Fully agree
 - ☐ Partially agree
 - ☐ Mostly disagree
 - ☐ Completely disagree

10.6 Questionnaire External Organizations

AI4HealthSec is a EU-funded project with 14 project partners from several European nations. Its main objective is the development of a solution that improves the detection and analysis of cyberattacks and threats on healthcare information infrastructures.

For this AI4HealthSec will create the Artificial Intelligence Dynamic Situational Awareness Framework (DSAF). The DSAF will be built upon a new type of Swarm Intelligence, self-organizing and dynamic collaboration approach. This will be implemented through an individualized autonomous networking protocol that provides an autonomic deployment, cluster formulation and hierarchical communication in healthcare information infrastructures.

AI4HealthSec wants to help healthcare information infrastructures in several ways:

1. Assess the vulnerabilities of cyber assets
2. Forecast and evaluate the probability of cyber-attacks
3. Access and receive warning for upcoming attacks and vulnerabilities
4. See the continuum between indicators of compromise, advanced persistent threats, cyber alerts and adversaries
5. Recreate, visualize and forecast propagation and cascading effects of attacks
6. Providing timely technical assistance and guidance on investigating and handling complex, interrelated cyber-security incidents and data breaches
7. Combine and analyze all security-related information and proofs in an effective and accurate manner
8. Receive guidelines and share information and warnings with other healthcare information infrastructures.

Healthcare information infrastructures will be protected by this solution in terms of a higher situational awareness among stakeholders and by a better incident handling and risk assessment in vulnerable healthcare information infrastructures.

1. Please select the type of your organization from the list below:
 - ☐ Public
 - ☐ Private
 - ☐ Other, which?
2. What domain does your organization come from (e.g. healthcare, logistics, energy supply)?
3. What is the size of your organization?
 - ☐ Very Large (> 250 employees)
 - ☐ Large (100 - 250)
 - ☐ Medium (50 - 100)
 - ☐ Small (10 - 50)
 - ☐ Micro - Very Small (<10)
4. Which type of security and incident management model does your organization adopt?
 - ☐ Outsource (supported by external organization)
 - ☐ Inhouse (internal support)
 - ☐ Other, which?
5. Which Security Management standards, security protocols and proved guidelines have your organization adopted? (more than one answer possible)
 - ☐ ISO 9001
 - ☐ ISO/IEC 27035
 - ☐ ISO/IEC 27001
 - ☐ ISO/IEC 27002
 - ☐ ISO 20000
 - ☐ ISO/IEC 27005
 - ☐ NIST SP800-30
 - ☐ NIST SP800-61
 - ☐ NIST Framework for Improving Critical Infrastructure Cybersecurity
 - ☐ Other, which?
6. In case of a security breach:
 - f. Do you have an Incident Response Team?
 - ☐ Yes
 - ☐ No
 - g. Do your procedures cover cyber-attacks/incidents?
 - ☐ Yes
 - ☐ No

h. Does your organization employ advanced response capabilities to effectively respond to security incidents?

☐ Yes

☐ No

i. Do your procedures estimate the cascading effects of a security events?

☐ Yes

☐ No

j. Does your organization cooperate with external entities to correlate and share incident information to achieve a cross-organization perspective on incident awareness and more effective incident responses?

☐ Yes

☐ No

7. Does your organization employ automated mechanisms to support the incident handling process?

☐ Yes

☐ No

8. Security Policies and Procedures that are in place within your organization/company (more than one answer possible)

☐ Incident Handling and Response Policy or/and Procedure

☐ Information Security Incident Management or/and Procedure

☐ Disaster Recovery and Data Backup Policy or/and Procedure

☐ Business Continuity Policy or/and Procedure

☐ Security Monitoring Policy or/and Procedure

☐ Access Control Policy or/and Procedure

☐ Security Monitoring Policy or/and Procedure

☐ Malicious Software Policy or/and Procedure

☐ Network Access Policy or/and Procedure

☐ Identification and Authentication Policy or/and Procedure

☐ Third Party Connectivity Policy or/and Procedure

☐ Other, which?

9. Does your organization have a vulnerability management process?

☐ Yes

☐ No

If yes:

How often do you scan for vulnerabilities?

If yes: What vulnerability databases do you use?

10. Does your organization apply dynamic (penetration) testing of its ICT infrastructure?

☐ Yes

☐ No

If yes:

How often does your organization run the dynamic (penetration) testing of its ICT infrastructure?

If yes:

What tools/suites does your organization use to run the dynamic (penetration) testing of its ICT infrastructure?

11. Does your organization monitor its infrastructure for malicious activities?

☐ Yes

☐ No

If yes:

What tools does your organization use to monitor its endpoints?

If yes:

What tools does your organization use to monitor its network?

12. Are there skilled and trained personnel on security and incident handling practices?

☐ Yes, most of the personnel is skilled and trained

☐ A few of the personnel are skilled and trained

☐ None of the personnel is skilled and trained

13. Does your organization offers / is willing to offer training programs on its employees about cyber-security awareness?

☐ Yes

☐ No

If yes:

How frequently are drills provided?

- ☐ Occasionally
- ☐ Annually
- ☐ 2-3 times a year
- ☐ More often

If yes:

In what form are drills provided? (more than one answer possible)

- ☐ Regularly offered programs with fix agenda
- ☐ Information given when needed to employee (on the job)
- ☐ Introduction to a new employee about cyber-security awareness when they start their job at your organization
- ☐ Other, how?

14. Does your organization employ a centralized solution to correlate incident information and individual incident responses in order to achieve an organization-wide perspective on incident awareness and response?

- ☐ Yes
- ☐ No

15. What is your current job title?

16. How long have you been in your current position?

- ☐ Under 1 year
- ☐ 1-5 years
- ☐ 5-10 years
- ☐ Longer

17. Have you been trained on cyber security issues by your organization?

- ☐ Yes
- ☐ No

18. If no: Have you been trained on cyber security issues by other organizations?

- ☐ Yes
- ☐ No

19. How high would you say is your knowledge regarding cyber-security topics?

- ☐ Very high
- ☐ Rather high
- ☐ Average
- ☐ Below average
- ☐ Very low

20. What computer software do you typically use in your daily work life? (also email programs, office etc.)

21. What computer hardware do you typically use in your daily work life?

22. Have you been professionally trained on the computer hardware and software that you use in your daily work life?

- ☐ Yes, on hardware
- ☐ Yes, on software
- ☐ Yes, on both
- ☐ No

23. Do you personally perform any particular tasks in the field of cyber security?

- ☐ Yes
- ☐ No

If yes, which?

24. Have you encountered cyber security incidents in your professional area over the past 3 years?

- ☐ Yes
- ☐ No

If yes: Have you personally been involved in a cyber-security incident (i.e. incident in your direct work environment)?

- ☐ Yes
- ☐ No

If you have been personally involved in a cyber-security incident:
Did you feel secure with handling the incident? Did you know what to do?

- ☐ Yes, completely knew what to do
- ☐ Was a bit unsure but got help
- ☐ Did not know what to do and got no help

25. Where do you personally see the biggest threats by cyber-attacks in your branch? Where could criminals attack most easily?

26. Which person groups do you think are most vulnerable to negative consequences of cyber-attacks regarding your organization and in what way?

27. What do you think are the worst possible consequences of cyber-attacks regarding your organization?

28. Do you agree: Your organization well-formed security incident management policies allow you to improve your organization's situational awareness

- ☐ Strongly Agree
- ☐ Agree
- ☐ Disagree
- ☐ Strongly disagree

29. In recent years: What do you think was the overall opinion of security officers and IT employees in your organization towards the engagement in a CIP (critical infrastructure protection) program?

- ☐ Strongly ambivalent
- ☐ Rather ambivalent
- ☐ Rather in favor of engaging in a CIP program
- ☐ Strongly in favor of engaging in a CIP program

30. Do you think it would be useful if your organization had a CIP program?

- ☐ Not useful at all
- ☐ Rather not useful
- ☐ Somewhat useful
- ☐ Very useful

31. In your opinion: What are the main benefits AI4HealthSec components could offer to you and your work?
32. Which features should an external dynamic and self-organized cyber-security framework as AI4HealthSec aims to provide include that it could reduce the risks of cyber-attacks in your work environment?
33. What are your main concerns regarding an external cyber-security framework like AI4HealthSec?
34. What do you wish an external cyber security framework should look like in your daily work life?
- ☐ Visible, e.g. by regular status reports
 - ☐ Invisible in the background
35. Do you wish to interact with an external cyber security framework in your daily work life?
- ☐ Yes, I would like to have a system that needs input by me
 - ☐ No, I would like to have a system that runs completely by itself

10.7 Report for EAB

CALL H2020-SU-DS-2018-2019-2020

Digital Security

TOPIC SU-DS05-2018-2019

Digital security, privacy, data protection and accountability in critical sectors

AI4HEALTHSEC

"A Dynamic and Self-Organized Artificial Swarm Intelligence Solution for Security and Privacy Threats in Healthcare ICT Infrastructures"

Report to the Advisory Board

18/03/2021

Grant agreement number: 883273

Start date of project: 01/10/2020

Revision

Lead contractor: CNR

Duration: 36 months

Project funded by the European Commission within the EU Framework Programme for Research and Innovation HORIZON 2020	
Dissemination Level	
PU = Public, fully open, e.g. web	
CO = Confidential, restricted under conditions set out in Model Grant Agreement	
CI = Classified, information as referred to in Commission Decision 2001/844/EC.	
Int = Internal Working Document	✓

Contents

1	Overall concept	149
2	Overview of AI4HEALTHSEC Outcomes	150
3	Pilot scenarios	151
4	Requirements for Healthcare ICT Infrastructure	153

Overall concept

In the digital era the health care ecosystem in Europe has turned into a complex mosaic, composed by large health systems and institutes, single physician practices, device developers etc. This ecosystem can be defined as a widely distributed, interconnected set of entities (i.e., organizations, individuals or/and CIs), processes and services that relies upon interconnected ICT infrastructures, establishing a dynamic Health Care Supply Chain (HCSC). The established interconnections reflect the relationships that exist between the involved entities.

In this context, these HCSCs are characterized by a high degree of complexity and interconnectivity of the ICT systems. As depicted in Figure 1 the health care ecosystem can be represented as being composed by four **circles of consideration** that puts the patient in the centre of attention. The **first inner circle**, our starting point, includes health components that are very close to the user (e.g. implants, sensors). The **second circle** encapsulates the previous one as well as all the medical equipment and devices (e.g. pathology scanners and servers) used in health institutes. The **third circle** encloses the two previous ones and incorporates the **individual Health Care Information Infrastructures (HCIIIs)**. Finally, the **fourth and outer circle** contains all the above circles and represents the **interdependent HCIIIs** composing the whole health ecosystem including the supporting **Health Care Supply Chain Services (HCSCS)**.

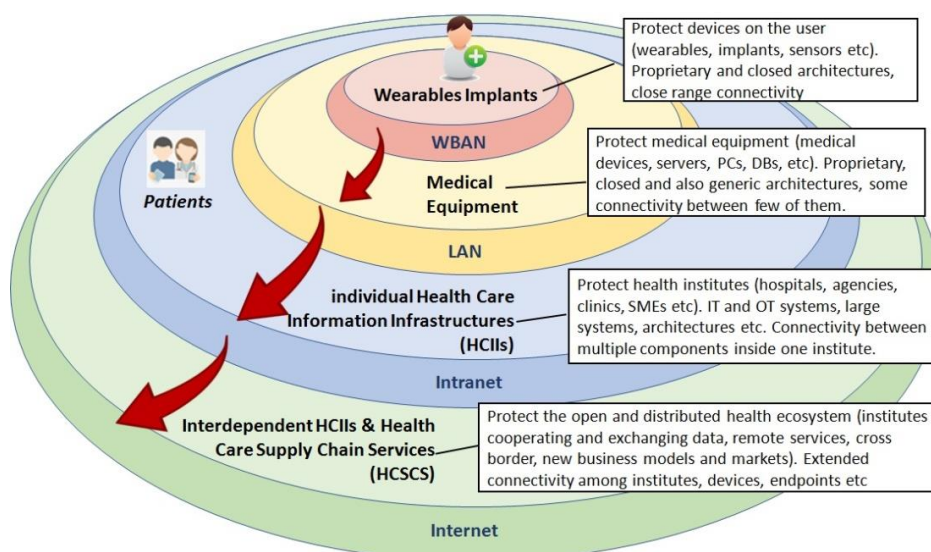


Figure 1. AI4HEALTHSEC Circles of Consideration

However, the evolving digital interconnectivity of medical ICT systems has also changed the threat landscape, as the digitalization of patient data is attracting more attention from cybercriminals, producing a wide range of security and privacy challenges and increasing the danger of potential cybersecurity attacks in Healthcare Infrastructures. Thus, there is an urgent need to ensure that these identified four distinct areas of consideration are all properly secured. However, despite the fact that these areas have their own unique characteristics, they are not independent from each other. Inner

circles can be seen as the building blocks of the external ones, meaning that the security of the external circles is directly affected by the inner ones. Thus, the security of the **interdependent HCIs** and the **HCSCS**, is directly affected by the security of the **individual HCIs** that compose it. However, it should be noted that the overall system is not secured by simply securing its “building blocks”. There are interdependences between the different layers that have their own specificities and require cross layer coordination.

Overview of AI4HEALTHSEC Outcomes

AI4HEALTHSEC’s aim is to enhance the security and resilience of the modern digital healthcare ecosystems and the provided medical supply chain services through the provision of a novel **Artificial Intelligence Dynamic Situational Awareness Framework (DSAF)**. The main goal of the proposed approach is to improve, intensify and coordinate the overall security efforts for the effective and efficient identification, evaluation, investigation and mitigation of realistic risks, threats and multi-dimensional attacks within the cyber assets in the four distinct **areas of consideration** (Figure 1). The proposed approach seeks to support, prepare and help the **Interdependent HCIs** participating in different types of **HCSCS** to: (i) thoroughly assess the vulnerabilities of all cyber assets; (ii) continuously forecast and evaluate the probability of cyber-attacks; (iii) access/receive warnings for upcoming attacks and vulnerabilities; (iv) see the continuum between indicators of compromise, advanced persistent threats, cyber alerts and adversaries (v) easily recreate, visualize and forecast propagation and cascading effects of attacks in their **Interdependent HCIs** and anticipate how these attacks propagate across the **HCSCS**; (vi) follow a targeted step-by-step framework providing timely technical assistance and guidance on investigating and handling complex, interrelated cyber security incidents and data breaches and extracting all relevant information; (vii) combine and analyse all security incident-related information and proofs in an effective and accurate manner; and (viii) receive guidelines and, share information and warnings with all **HCIs**.

In order for **DSAF** to meet its objectives, it consists of consists of 7 main conceptual layers, 4 horizontals (“**Risk and Privacy management & Cyber-Attack Forecasting**”, “**Incident Identification**”, “**Security Events Evaluation**” and “**Analysis and Decision-Making**”) dealing with the situational awareness process and three vertical, the “**Information Sharing & Individualised Autonomous Networking**” responsible to distribute, disseminate, self-publish, broadcast or circulate the security-related information, the “**Security & Privacy**” incorporating a set of security, privacy and data protection features and the “**Context-Rich/ Analytical Exploration**” providing environment that allows the HCIs’ operators to have a better understanding of the cyber environment.

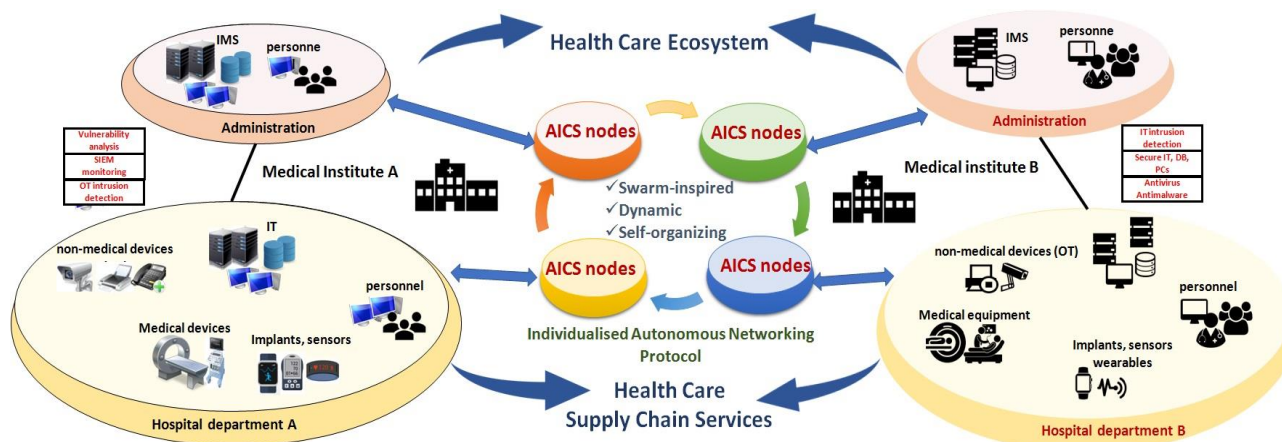


Figure 2. Main Aspects and Principles of the AI4HEALTHSEC framework

In addition the proposed framework will be built upon a new type of Swarm Intelligence (SI), self-organizing and dynamic collaboration approach implemented through an **individualised Autonomous Networking protocol** (Figure 2) that provides autonomic deployment, cluster formulation and hierarchical communication in HCII. This protocol, will connect the four circles of the health ecosystem grouping individual ICT elements, systems and components into a population of simple or group of nodes, named **AICS nodes** (group of ICT assets or individual HCII), allowing them to interact locally with one another and with their Interdependent Health Care environment. In this way, the proposed protocol will build networking infrastructures that manage the effective coordination of the AICS nodes of **Interdependent HCII** by defining and leveraging the actions that should be performed by them. These agents are linked together and cooperate with each other through local interactions to achieve distributed optimization of the risk analysis and incident handling in real time. The continuous diffusion of security-related information across the network enables the agents to optimize the evaluation and mitigation of the interdependent threats and risks as well the investigation of complex security events and data breaches.

Pilot scenarios

The AI4HEALTHSEC proposed solution will be tested, verified and validated in pilots setups, that will be coordinated and executed by consortium partners. In the following a short description of the reference pilot frameworks is provided.

Cybersecurity in Medical Implants, Wearables and Biobanks – executed by Fraunhofer Institute

Medical devices are increasingly becoming targets of hacker attacks. This applies for active implants that often incorporate wireless programming and even remote programming interfaces. For therapeutic implants like pacemakers, malignant change of their programming or deactivating of features could seriously harm the patient to death. Moreover, Personal Health Systems (PHS) in combination with mobile technologies, wearable medical sensors/actors, and related health services which communicate to each other via internet result in an increasing attack surface. Intruders may

interrupt health services, re-configure systems, and get full access to personalized patient data. The latter can be also achieved via attacking central points of patient data collection, like Biobank Information Systems. Human biobanks store biological samples of humans such as body fluids or tissue samples, with associated data on sample collection, analyses, and donor information. Accountability and privacy preservation are keys to ensure their operation.

Secure Access and Sharing of Clinical Data via VNA systems - executed by EBIT (ESAOTE Group) company

The large amounts of digital clinical, biomedical and health data, are a crucial and central source of information to improve the provision of clinical, diagnostic and therapeutic services. Vendor Neutral Archiving (VNA) systems consist a new paradigm of Health Care (HC) IT solution used to manage data types used in the case of PACS and also other type of document and imaging data (Radiology, Cardiology, etc.). A VNA system must comply with enterprise workflows standards by storing information in non-proprietary, interchangeable formats that enable rapid data migration without clinical disruption. Health and clinical governance organizations are interested in such solutions for cost reduction, improved care and real time quantitative analysis of all available data, reducing and optimizing the total cost of treatment wherever possible. On the other hand, cloud connectivity raises high privacy and security challenges for a connected VNA, whereas people's expectation for understanding when and where their health information is shared increases the necessity to ensure trustworthiness.

Creating higher situational awareness of cyber-security amongst hospital staff –executed by Klinikum Nürnberg

Klinikum Nürnberg, partner of AI4HEALTHSEC consortium, offers, as for any healthcare organization a huge amount of different IT systems, such as software for large medical equipment (e.g. CTs, MRTs), electronic health records of patients, special clinical systems like RIS, PACS or LIS, email communication systems or other corporate or administrative software.

One special issue about hospitals in general is the close linkage and deep technical integration of all software solutions including intensive data and information exchange. That implies and establishes dependencies between nearly all parts of the IT.

Another characteristic point of cyber-security in a hospital is that the computers of medical staff mostly are directly connected to sensitive systems such as systems containing patient data.

The users of the hospital IT consists in large part of medical staff, such as physicians and nurses who are responsible for the treatment of patients. The software supports the treatment and enables the documentation of the performed procedures. All staff members use several different software products daily and also several in parallel with other tools or with the actual treatment of the patient. There is not much time specifically dedicated to the use of software. The staff might not be aware enough of cyber-security risks to prevent dangerous situations even though the IT department informs them via internal communication paths such as e-mails about possible security risks. In

consequence errors and deficiencies in the perception of cyber-risks by staff members systematically decreases hospital's security claims.

Security and Privacy in a Digital Health Living Lab – executed by University of Brighton

According to the European Network of Living Labs (ENoLL), Living Labs (LLs) are defined as user-centred, open innovation ecosystems based on systematic user co-creation approach, integrating research and innovation processes in real life communities and settings. LLs are both practice-driven organisations that facilitate and foster open, collaborative innovation, as well as real-life environments or arenas where both open innovation and user innovation processes can be studied and subject to experiments and where new solutions are developed. LLs operate as intermediaries among citizens, research organisations, companies, cities and regions for joint value co-creation, rapid prototyping or validation to scale up innovation and businesses. The concept and methodology of the Living Lab can be implemented in different environments and healthcare settings, ranging from in hospital wards to community services and general practices, providing a variety of pilot settings.

Requirements for Healthcare ICT Infrastructure

The elicitation of requirements was performed in perspective of three pillars:

- d. User's Wishes/Challenges for the development of the AI4HealthSec framework from user perspective
- e. Technical Requirements
- f. Domain Requirements

To elicit users' wishes and therefore to get a basic understanding of the challenges the framework will face, we created questionnaires to be fulfilled both by internal project partners and external organizations from further critical infrastructures (besides healthcare, e.g. financial sector, transportation sector).

In parallel technical requirements were elicited by intense discussions with the technical project partners.

Additionally, the healthcare domain concerning relevant policies and standards have been analyzed by reviewing existing literature.

The analysis of the questionnaires that included numerous closed and open questions on wishes, expectations, fears and on participants' background (e.g. level of proficiency concerning cybersecurity) was condensed in the following list of challenges that evolve from the user's perspective (business needs):

Business Need ID	Title	Description
BN1.	Prediction and Prevention of Attacks	My organization needs to forecast and prevent cyber-attacks.
BN2.	Vulnerability Assessment	My organization needs a framework to assess its cyber-security weaknesses.
BN3.	Awareness Creation & Prevention of Human Errors	My organization needs a better awareness and higher knowledge concerning the staff when it comes to cyber-security topics.
BN4.	Detection of Abnormal Patterns & Creation of Warnings	My organization needs a system to automatically detect abnormal patterns in its IT structures and to create warnings.
BN5.	Simplification of the Process of Risk Assessment	My organization needs an easier process of risk assessment.
BN6.	Development of a Long-Term Strategy of New Protection Solutions	My organization needs a long-term and comprehensive cyber-security strategy.

The numbering of these challenges does not represent any weighting of them, since these business needs are equally important to be taken into account.

Those challenges are the basis for the further elicitation of requirements. From a domain-oriented perspective our approach identified and analysed the health care market based on the AI4HealthSec circles of consideration, i.e. health components (first circle), medical equipment (second circle), individual HCIs (third circle) and interconnected HCIs (fourth circle).

By this approach we found numerous standards that need to be taken into account including for example ISO 17971 on the application of risk management to medical devices. Another example is the standard ISO/TR 22969 which provides guidance for managing healthcare service security with connectable personal health devices. Further domain specifications were found in the incident handling of medical devices.

The big frame for creating an AI4HealthSec framework given by domain characteristics to meet the user challenges coming from the questionnaires was narrowed down to technical challenges and requirements - thus more specific descriptions on how a framework should possibly look like. The categories of technical requirements are:

Technical Challenge ID	Description	Relevance to Business Needs
TC1.	<p>Evidence-based, Swarm-driven Risk Management and Assessment Methodology</p> <p><i>To address issues for the context and compliance of the management approach, identification and predication of risks, the approach for risk assessment management, modelling and control and the applicability to other domains</i></p>	<ul style="list-style-type: none"> • BN1: Prediction and Prevention of Attacks • BN2: Vulnerability Assessment • BN3: Awareness Creation and Prevention of Human Errors • BN5: Simplification of the Process of Risk Assessment • BN6: Development of Long-Term Strategy of New Protection Solutions.
TC2.	<p>Cyber-security Risk-based Incident Handling Methodology</p> <p><i>To address issues for multi-level evidence collection, correlation of information to detect incidents and analyse security events and support for incident management and response</i></p>	<ul style="list-style-type: none"> • BN1: Prediction and Prevention of Attacks • BN3: Awareness Creation and Prevention of Human Errors • BN4: Detection of Abnormal Patterns and Creation of Warnings • BN6: Development of Long-Term Strategy of New Protection Solutions.

We summarise in the next lines the respective requirements that were extracted from a technical perspective.

TC1: Evidence-based, Swarm-driven Risk Management and Assessment Methodology

Requirements for Risk Management Context and Compliance

REQ1: The risk assessment /management models and process should be considered from a holistic view of internal (i.e., organisational, technical, medical devices) and external context of the complex health care system.

REQ2: The introduction of risk assessment/management models and processes in the AI4HEALTHSEC methodology should adequately take into account the complexity of the ICT infrastructure and technical evolution of medical devices that underpin security processes of health care complex adaptive system.

REQ3: The risk management approach should provide an informed real time decision making for managing cyber security risks and ensuring overall business continuity.

REQ4: The methodology should define the organisation cyber security needs, risk appetite, and risk tolerance for the key healthcare ICT infrastructure areas.

REQ5: The risk assessment /management approach should alleviate the limitations of existing risk management methodologies in terms of their ability to deal with ICT systems in the critical infrastructures.

REQ6: The methodology should leverage, use and implement existing cyber security, information security risk management, information security incident management standards including ISO 31000, ISO27001, ISO 27005, ISO 27031, and ISO 27032 associated with the protection of the complex ICT infrastructure.

REQ7: The methodology should offer compliance with the relevant regulation necessary to compliance with the health care information system sector.

Requirements for Risk Identification and Predication

REQ8: The methodology should automatically detect potential cyber-attack and adversary actions using autonomous intelligence swarm agents and reporting to the supervisor agents so that evidences are combined and correlated with the existing data for the attack predication and new attack vector discovery.

REQ9: The methodology should include a real time communication, interaction, and feedback among hierarchy-based multiple agents including supervisor and swarm agents and create an overall dynamic cyber security situational awareness.

REQ10: The methodology and associated risk management framework should consider organisation-wide vulnerabilities detection using collective behaviour of swarm intelligence taking into account the underlying complexity of the ICT infrastructure and interoperability and interconnectivity among various sub components including medical devices.

REQ11: The methodology should consider depth of access by measuring how far threat actors reach within the ICT infrastructure by collective swarm intelligence data for the risk identification and predication.

REQ12: The methodology should introduce a risk management system, which will consider the nature and interdependencies of cybersecurity and medical assets and as well as their implications on overall business continuity

Requirements for Risk Assessment and Modelling

REQ13: The methodology should adopt an evidence-driven Cyber Security Risk Assessment model in order to capture and deal with cascading effects of risks, threats and vulnerabilities, associated with the health care ICT infrastructure

REQ14: The methodology should help elicit, understand and analyse risk management requirements for the health care ICT infrastructure, with particular emphasis on requirements associated with the overall complex system and its supply chain context.

REQ15: The methodology should consider all organisation wide vulnerabilities by correlating data from the swarm agents and its impact for the net risk calculation .

REQ16: The risk assessment approach should follow quantitative assessment methods to determine the risk level, based on existing consistent cyber security threat data

REQ17: The risk assessment approach should consider Cyber Threat Intelligence (CTI) information including relevant threat actors, their capabilities, skills, motivations, and underlying TTP and IoC.

REQ18: The methodology should consider cyber risk modelling considering assets and their dependencies, vulnerabilities within the assets, possible attack paths, threat intelligence properties, and risks.

REQ19: The methodology should leverage simulation models combined with a multi-criteria decision making approach in order to produce timely, accurate, relevant and high quality evidence, information, indicators, factors and parameters associated based on which the multi-dimensional risks will be assessed.

REQ20: The methodology should use graphs to discover and represent possible attacks plans and patterns and will adopt a general approach to integrate several aspects of both vulnerabilities and threat agents.

REQ21: The methodology should identify and model assets, processes, risks, stakeholders' relationships/interactions and dependencies.

REQ22: The methodology should create a range of metrics covering reliability, credibility, acceptance, timeliness, realism of risk management goals and the level of integration of the risk management approach in decision making structures. These metrics should be able to be measured across all cyber-security assets, medical device, and ICT systems available within health care infrastructure.

Requirements for Risk Management and Control

REQ23: The methodology should determine the level of assurance based on the evidence of existing controls and their effectiveness, and recommend alternative courses of action for responding to risks.

REQ24: The methodology should explore new techniques/methods for the credible calculation of insurance premiums.

REQ25: The risk management approach should ensure the constant vigilance of existing risks, by offering mechanisms to understand status of residual value of risk and identifying any new risk using intelligence swarm agents.

Requirements for Incident Management

REQ26: The risk analysis methodology must provide real time decision making support for incident response and post incident review activities.

REQ27: The risk identification, forecasting and analyse should provide a better understanding of the cyber security incident related information.

REQ28: The risk management methodology should align with the incident response and post-incident activities to ensure eradication of the threats and risks and overall business continuity.

REQ29: The risk assessment methodology should support updating threat intelligence information and incident response planning, through lessons learn from the evolving threats, risks and related incidents.

Requirements for Contribution to other Domains

REQ30: The risk management methodology should consider publishing best practices that include blueprints and guidelines for adapting the approach to other critical infrastructures sector, such as smart grid cyber physical systems.

REQ31: The AI4HEALTHSEC project should contribute best practices associated with the deployment and operation of its framework for risk management in health care sector of any type and size.

TC2: Cyber-security Risk-based Incident Handling Methodology

Requirements for Multi-source Evidence Collection and Preparation

REQ32: The incident handling methodology should support evidence collection on both real time and historic data from the various evidence collection sources to assist incident detection.

REQ33: The evidence collection process should include batch data (i.e., collection of raw data over a specific period of time), including, but not limited to, log files from vulnerable systems and network traffic.

REQ34: The evidence collection process should include configurable steps, allowing for the specification of the type, format and location of the incoming data sources such as log files.

REQ35: The evidence collection process must consider anonymization of raw data collected by various sources.

REQ36: The evidence preparation process must consider the semi-structured nature of different datasets.

REQ37: The data collected should include records about network usage and bandwidth, and should allow for the identification of network traffic anomalies and excessive bandwidth usage.

REQ38: The data collection process should take into consideration and be at least partially aligned with existing industry proprietary or non-proprietary data exchange protocols, with particular interest in understanding to some extent the messages exchanged, including network packages and messages from the interaction among systems.

REQ39: The incident handling process should be able to monitor the availability of signals and system web sources or services and calculate their response time for further analysis.

REQ40: The incident handling approach should support normalization and transformation of raw data coming from semantically relevant sources to perform system independent data processing and sharing across the AI4HEALTHSEC Framework.

REQ41: The incident handling approach should consider for managing structural and semantic mismatches across the different datasets collected.

REQ42: The incident handling approach should support normalization and transformation for the unified representation of cyber security threats detected by internal or external components of this platform.

REQ43: The evidence preparation process should support preliminary filtering of raw data, using predefined criteria over the parameters collected from raw data, so that irrelevant one can be removed and/or not taken into consideration in the incident handling process.

Requirements for Evidence chain Generation and Security Incident Detection

REQ44: The incident detection and event analysis approach should be able to process streaming, batch and historic data

REQ45: The incident detection and event analysis approach should consider data uncertainty and incompleteness, so that the processing of the provided raw data can be feasible even in the absence of some elements.

REQ46: The organization and filtering of the incoming raw data (across all the available data sources) is essential for the further analysis of the current status of the systems. During this process the evidence chains would be generated and the relevant data would be collected and stored for latter usage.

REQ47: The incident detection and event analysis approach should support the preliminary analysis of relevant raw data (e.g., deviation from normal patterns) to identify potential security incidents.

REQ48: The security event analysis approach should support semantic and structural decisions regarding the description of the different type of incidents so that further processing of the information generated can be feasible and meaningful.

REQ49: The incident detection and event analysis approach should utilize existing knowledge sources with security data (including either external knowledge used for training purposes or other security related knowledge acquired by other modules of the system) for correlating evidence to incidents and security events.

REQ50: The incident detection and event analysis approach should be customizable to further domains, other than health ICT infrastructures.

REQ51: The incident handling methodology should maintain a knowledge base with information about actual successful attack scenarios.

REQ52: The incident detection and event analysis approach should support decision making, towards developing more efficient and effective defence strategies, based on evidence from past detected incidents, extracted from the knowledge base.

REQ53: The incident handling methodology must provide cyber-attacks related information that can be shared with other organizations in a secure and privacy preserving way.

Requirements for Incident Management and Response

REQ54: The incident handling methodology must identify the on-going attacks and related information at all times.

REQ55: The incident handling methodology should be able to predict possible scenarios of future attacks.

REQ56: The incident handling methodology should provide a visual representation of the cyber-attack path.

REQ57: The incident handling methodology should assure an acceptable risk level for the cooperating stakeholders.

REQ58: The incident handling methodology should promote the necessary defensive capabilities and provide a rational decision-making to help stakeholders in determining which security controls must be implemented to encounter the identified security issues and cyber-risks.

REQ59: The incident handling methodology should support matching evidence collected in real time with archived information for cyber-attack scenarios.

REQ60: The incident handling methodology should be able to provide comparison among the patterns of data collected at the infrastructure nodes and the normal state of operations.

REQ61: The incident handling methodology should allow decision makers in predicting the assets that are exposed to risks when a security event is detected.

REQ62: The incident handling methodology should support decision makers in exploring different attack scenarios on potential harmfulness of a detected anomaly to the infrastructure.

REQ63: The incident handling methodology should present the attack path of a detected incident across all impacted assets.

REQ64: The incident handling methodology should present sufficient information to decision makers to enable them understand the risk of cyber attacks detected in real time on the infrastructure.

REQ65: The incident handling methodology should provide decision makers with access to the results of the risk assessment process at all times to understand the consequences of a detected cyber attack.

REQ66: The incident handling methodology should provide recommendations to decision makers on the most suitable security controls to mitigate the risks from detected security events and cyber risks.

REQ67: The incident handling methodology should allow decision makers understand the impact from the implementation of a defensive mechanism to support informed decisions when selecting the appropriate security controls.