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"

D2.3 User Reference Scenarios, evaluation metrics and criteria principles

Due date of deliverable: 31/05/2021
Actual submission date: 31/05/2021

Grant agreement number: 883273
Start date of project: 01/10/2020
Revision 1

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	

D2.3 User Reference Scenarios, evaluation metrics and criteria principles

Editor
Marco Fruscione (EBIT)

Contributors
Lena Griebel (KLINIK)
Haralambos Mouratidis (UOB)
Theo Fotis (UOB)
Vasilis Tountopoulos (AEGIS)
Spyridon Papastergiou (FP)
Stephan Kiefer (Fraunhofer)
Gabriele Weiler (Fraunhofer)
Dmitry Amelin (Fraunhofer)

Reviewers
Andreas Zacharakis (STS)
Djordje Diokic (PN)
Dusan Pavlovic (PN)
Farhan Sahito (PN)

The work described in this document has been conducted within the project AI4HEALTHSEC, started in October 2020. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883273

©Copyright by the AI4HEALTHSEC Consortium.

VERSION	DATE	REVISION AUTHOR	SUMMARY OF CHANGES
0.1	09/03/2021	EBIT	ToC with explanation of sections
0.2	29/03/2021	EBIT FHG-IBMT UOB	Contribution to Chapter 2 in Pilot Description (missiing KN and EBIT completion)
0.3	11/04/2021	EBIT KN	Completed chapter 2 Refinement requested on the definition of Potential Attack Scenario
0.4	26/04/2021	EBIT	General Revision and first proposal on User Requirement Evaluation and KPI
0.5	10/05/2021	FHG-IBMT UOB	Potential Attack Scenario structured with tables
0.6	24/05/2021	KN	Contribution on Chapter 2 Completed KPI and general revision
0.7	27/05/2021	STS PRIVANOVA	Quality Assurance completed
0.8	29.05.2021	EBIT	Addressed reviewers' comments
1.0	31.05.20121	EBIT	Final edits, Updated tables, figures

The work described in this document has been conducted within the project AI4HEALTHSEC, started in October 2020. This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883273

©Copyright by the AI4HEALTHSEC Consortium.

Executive Summary

This document is the report of the AI4HEALTHSEC project that defines the scenarios description associated with the use case pilot of the project. Six Health Care pilot are described that will serve as AI4HEALTHSEC's real-life pilot scenarios, upon which various potential security attack will be deployed and the efficiency of the innovative outcomes of the AI4HEALTHSEC platform measured.

The process of evaluating the effectiveness of the impact in the use of the platform will be pursued, through the implementation and execution of the pilots, in two phases: through the validation of the individual functional requirements and the definition and measurement of KPI, as defined in this document.

The metric needed to assess the impact of the AI4HEALTHSEC system on the pilot use cases has to be defined and specified in the form of Key Performance Indicators (KPIs). The pilot scenarios will be then designed and implemented in WP6 along specific test cases associated with real-life attacks, threats and security incidents pertaining to the pilot sites of the project.

Contents

List of acronyms

AI	**Artificial Intelligence**
AICS	Artificial Intelligence Computer System
CSIRT	Computer Security Incident Response Team
CVSS	Common Vulnerability Scoring System
DCS	Distributed Control System
DDoS	Distributed-Denial-of-Service
DoW	Description of Work
DSAF	Dynamic Situational Awareness Framework
EAB	External Advisory Board
EHR	Electronic Health Record
ENISA	European Union Agency for Cybersecurity
EUDAMED	European Database for Medical Devices and in-vitro Diagnostics
FHG	Fraunhofer Gesellschaft
FHG - IBMT	Fraunhofer IBMT: Fraunhofer Institut Biomedizinische Technik (in English: Fraunhofer Institute for Biomedical Engineering)
FDA	Food and Drug Administration (US)
HCIIs	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
MDR	Medical Device Regulation
PMCF	Post-Market Clinical Follow-up
PMS	Post-Market Surveillance
RAAP	Risk Analysis and Assessment Process
SDLC	Software Development Life Cycles
SI	Swarm Intelligence
VM	Virtual Machine

List of tables

List of figures

1 Introduction

1.1 Scope and link to project objectives

This deliverable includes the description of the reference scenarios of the project as well the criteria and metrics needed to assess the impact of the AI4HEALTHSEC system on the pilot use cases.

The deliverable is the outcome of Task T2.3, devoted to the refinements of the scenario description associated with the pilot operations of the project, specifying and analysing the Health Care pilot which will serve as AI4HEALTHSEC's real-life pilot scenarios, upon which various attack cases will be deployed and the efficiency of the AI4HEALTHSEC approach measured.

Moreover, a methodology and certain metrics for the qualitatively and quantitatively evaluation of the refined requirements of T2.1 to match current and future security and privacy demands mainly in the health sector to be addressed in AI4HEALTHSEC, are developed.

Such metrics and criteria will be specified in the form of Key Performance Indicators (KPIs). The pilot scenarios will be designed in detailed, implemented and executed in WP6 along specific test cases associated with real-life attacks, threats and security incidents pertaining to the pilot sites of the project. Hence, the work in this task has been carried out with the involvement of all pilotsites partners of the consortium.

1.2 Relation to other work packages and tasks

This deliverable D2.3 is the result of Task T2.3 which is part of Work Package (WP) 2 "Refinement of pilot requirements, evaluation metrics and AI4HEALTHSEC Architecture".

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 define the appropriate evaluation methodology and corresponding metrics for the demonstration of the unique characteristics of AI4HealthSec

This WP contributes and it is strictly related to others WPs in the project, such as WP3 and WP4,

Task 2.3 interacts with task 2.4 of WP2.

WP2 is strictly related 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".

Task 2.3 and the outcomes contained in this deliverable are strictly related with WP6 and provides input for WP6 "Pilots development of the AI4HealthSec system", where the activities of pilot design,

implementation and run will be executed, as well the functional requirement and KPI, here defined, will beevaluated.

Task 2.3 and scenario definition is also strictly related with deliverables D9.1 and D9.2 where for each pilot has been described:

1. *Whether humans (patients, health professionals, nurses, etc.) are involved in the pilot*
2. *If 1. is true, how to recruit them (inclusion/exclusion criteria)*
3. *If 1. is true, how to intend to ask for the consent of the humans involved*
4. *If 1. is true, provide a template of the informed consent*
5. *If personal data will be processed, specifying the kind of sensitive information intended to use (describing that they are the minimal data necessary for your research by specifying the purpose of use.*
6. *If personal data will be processed, how it will be protected (e.g. encryption) and if it will be anonymized or pseudonymized.*

1.3 Structure of the document

The document is structured in 3 main sections.

The first section (Chapter 2) describes in detail the reference pilot scenario and specifically for each of the pilots the following main topics are addressed:

- Current description of the real-life scenario
- Description of the architecture of the IT systems involved in the implementation of the pilot and their hardware and software architecture
- Definition of the use cases of the pilot and the type of users of the systems
- Definition of the critical data involved in the scenario
- Definition of potential attack scenarios
- Definition of expectations and challenges that can be expected to be faced and improved with the AI4HEALTHSEC platform.

The second section (Chapter 3) starts from the Business Needs and Requirements defined in D2.1., in order to identify the methodology through which part of those User Requirements will be addressed and evaluated in the pilot execution.

The final section (Chapter 4) identifies the principles and metrics that will be defined in the implementation of the pilots to measure the effectiveness of the AI4HEALTHSEC platform.

2 Pilot Description

2.1 Pilot #1 – Klinik Nurnberg

2.1.1 Description of the pilot

Klinikum Nürnberg is one of the largest, maximum care hospitals in Germany with 2.197 beds at two locations in Nuremberg, encompassing 42 clinics, about 7.000 employees and about 200.000 patients a year (in-patients and out-patients combined). The hospital includes a dedicated IT department responsible for the complete outpatient and inpatient IT infrastructure, systems for clinical R&D and the teaching/training for healthcare professionals with the need for establishing cyber-security measures.

Therefore, the Klinikum offers 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. For example, electronic health records get input by several sources, such as DICOM servers, laboratory software, or surgical software.

Due to the close connection of the IT infrastructure components there evolve potential security risks mainly on data exchange and spreading access permissions. Those risks are even more relevant when being aware that a lot of software tools are used by medical staff in their daily, often stressful daily work routine in parallel to the treatment of patients. If hospital IT is attacked, this easily leads to far-reaching consequences such as the loss or manipulation of sensitive patient data, blocked access to important services and subsequently to the loss of reputation of the hospital and in the worst case to a physical endangerment of patients.

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.

If there is a known security gap within a software product the Klinikum uses, the software manufacturer creates a patch/patches for this gap and offers them to the Klinikum as its customer. The IT department of the Klinikum gets the patch and is not able to bring it into the productive system before testing if the patch might influence the functionalities of other IT systems that are closely linked to the one system with the security gap. For this testing, Klinikum uses a dedicated test environment. If the often time-consuming testing of the security patch shows that there are no problems evolving in another IT system, the patch can be activated in the respective software for operational use.

The time delay in implementing security patches might also arise through a dependency on medical device manufacturers that often have to approve on software updates before the Klinikum is allowed to implement them. For instance, a company-wide malware protection is often not possible due to the fact that this would impair the functional integrity of a medical device based on the MDR.

Another characteristic point of cyber-security in a hospital that needs to be taken into account is that the computers of medical staff mostly are directly connected to sensitive systems such as systems containing patient data. This is because medical staff needs immediate access to patient files when treating a patient.

2.1.2 Current infrastructure and available devices (HW/SW)

For its IT processes, KN uses a hospital information system (HIS) by SAP consisting of SAP IS-H (central patient administration and billing system) and i.s.h.med by Cerner (clinical workstation system). Those systems include numerous functions using further SAP modules and connected IT systems. HIS includes a database server and several application servers. Each patient is administered via an individual IS-H patient number; each medical case gets a separate IS-H case number. AT KN patient documentation at the bedside is paper-based, older patient files are scanned and then stored in the archive system.

The HIS is connected to further hospital applications that can send and receive data using interfaces. For example, there is a system to order drugs, a PACS (Picture Archiving and Communication System), a RIS (Radiology Information System) or special information systems for medical disciplines, e.g. a Cardiology Information System. Using the IS-H staff can start requests for archive documents containing scanned patient files with information on a patient's medical history including important information such as allergies.

Access to IS-H is granted with a personal password. This password is sent to medical staff in a letter when they first start their job at KN. Each password is valid for three months and is changed after those three months. The HIS is accessible at client computers. Those computers also are used to receive and send emails (e.g. for conversation with colleagues inside KN/other hospitals/doctor's offices).

The usage of HIS mostly occurs during everyday medical practice – medical staff needs to access HIS regularly before and after seeing, treating and caring for patients.

2.1.3 Current Use Case Workflow and Target Audience (Final End-User)

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 in parallel with other tools or with the actual treatment of the patient. There is not much time specifically dedicated to the use of the 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 increases the hospital's security claims.

2.1.4 Critical Data Involved

All data that is stored in the HIS is highly critical as it is sensitive and personal data that is crucial for the proper treatment and care of patients. Data contains patient information including diagnoses, procedures, anamnesis including family medical history and personal medical history (e.g. mental illness, genetic dispositions to diseases), names, dates of birth, addresses, risk factors (such as drug abuse or high BMI), and the current ward where the patient is treated. Also treating physicians' and

nurses' names are included in the HIS. It also enables access to hospital systems such as PACS or pharmacy systems including drug orders of staff and important medical images of patients.

2.1.5 Potential Attack Scenario

Possible attack scenarios might look like this:

- **Ransomware attacks**: Physicians work on documents with clinical findings, doctor reports and medical records uses the internal email server by corresponding client software. They use the email account and notices a mail that apparently is sent by KN's internal IT department.

  In this email the "IT department" informs that the staff member has to confirm their HIS password; otherwise, it won't be usable anymore. For this, they should click on an attached link. Now, if the staff member would think thoroughly about this situation, it would potentially be obvious that this mail cannot be correct. But, in stressful work situation, they simply do not have the time to re-think, click on the link and confirm their password. The connected malware will be activated by this user interaction. Now, all text documents based on the staff member's existing access permissions will be encrypted. Based on the role in the hospital (e.g. chief physician, senior physician, ward doctor) the extent of encrypted files grows without having a chance to stop this process or having a key to cancel and decrypt the documents which the staff member needs to treat patients adequately. Affected are not only documents on the respective computer but also documents stored in all other file systems of the hospital.

  An example, a trojan that is inserted via social engineering attacks, is Emotet which encrypts the complete data of the victim and tries to extort ransom. Furthermore, Emotet implements a spy software in the hospital system that collects information such as administrator passwords. If this is successful, the attacker not only has access to patient data but can create new accounts with far-reaching entitlements.

  To make sure that the trojan is removed the whole hospital IT needs to be completely shut down and put on again, every user account needs to be created again, every staff member needs new passwords. The hospital needs to be separated from the Internet for a while. No access to patient data means high risks for patients as their important data cannot be accessed anymore. It also means that no new patients can be admitted and there is potentially an image loss for the hospital.

  This attack scenario is based on a lack of cybersecurity awareness among the medical staff. Through the use of AI4HealthSec framework, it is expected that human errors are prevented by providing information on current threats and active warnings but in a non-intrusive manner that does not interrupt existing work processes.

- **External attack:** A service technician or manufacturer for medical devices had a maintenance assignment for a medical product that is used for diagnostic purposes in the hospital (e.g. MRT). To do that, he uses his service notebook used for all customers of the manufacturer and also for private purposes of the technician (including downloading files and software and surfing on the Internet). During the last use, the notebook was contaminated with malware. When connecting this notebook to a medical device, the malware will be transmitted to this

device which is connected to the LAN of the hospital. The malware now might delete all image files of the medical devices and tries to capture further medical images archived from diagnostic modalities which reside in that network segment of the LAN/WAN.

This attack scenario is based on not yet-implemented software patches. Through the use of AI4HealthSec framework, it is expected to provide up-to-date and customized information on current software vulnerabilities and assistance on prioritisation of the implementation of certain patches.

- **Unsecured medical devices**: After a social engineering attack, an attacker succeeds to get physical LAN access by posing as a service technician. Due to openly known admin passwords, he is able to extract patients personal data from the devices. After that, he uses insecure DICOM interfaces and configurations to steal thousands of patient medical records and images. In the evening he leaves the building without any of the internal staff noticing him. After two weeks, all internal log files are overwritten by default because no one had recognized any irregularities. A month later, the hospital gets an extortion letter.

All kind of attacks might be eased due to the fact that there might be an open security gap whose patch has not been tested yet and thus might be open at this exact time point when the attack occurs. In addition what might happen to hospital data in case of an attack is included in the following table.

Scenario ID	Name	Description	Objectives
S1.1	*Encryption of data*	The attacker encrypts sensitive data to disrupt legitimate access.	Data encryption and temporarily unavailability have the effect that the users can not anymore access the encrypted data with potential damage on the patient clinical diagnostic and therapeutical path. Users can not anymore view the data.
S1.2	*Exfiltration of sample and patient data*	Data exfiltration attacks is the act of sensitive data deliberately being moved and	The attacker can steal sensitive data and/or copy or let then available to 3rd party

		extracted from the hospital application to the outside without permission.	for any other type of data management that is not in line with the scope for which the data have been gathered.
S1.3	*Manipulating sample and patient data*	Data manipulation attacks are attacks where an intruder does not take the data, but instead make changes or alter data for some type of gain.	An attacker can make changes to the data.
S1.4	*Data Deletion can e.g. lead to service disruption*	A successful deletion attack can e.g. cause deletion of entire hospital files and disrupt the service	Data Deletion and permanent unavailability has the effect that users can not anymore access the deleted data with potential damage on the patient clinical diagnostic and therapeutical path

Table 1: Pilot#1: Potential Attack Scenario

2.1.6 Security Challenge and Problems

The AI4HEALTHSEC framework raises the awareness of cyber-security topics within the hospital staff. The staff is informed about security gaps, trained in correct behaviour in critical situations and aware of the recognition of dangerous situations. The framework should not only provide mere information via e-mail as the e-mails often get lost in a large number of messages.

Instead, the framework should be able to give instructions in a way that they will indeed be realized.

At the same time, it is important that security notifications for the staff should not be intrusive:

Numerous medical software products provide alerts when it comes to potentially critical situations for patients. Thus, the physician or the nurse permanently has to react to several alerts in electronic form and has to permanently keep important information apart from not so critical information.

Instead of e-mails or pop-up alerts, the AI4HealthSec framework should create innovative ways of informing the staff on cyber-security dangers and of providing training to finally increase the overall awareness of cyber-security with designing a continuous perception that each use of software products includes benefits as well as operational risks.

Furthermore, an innovation of AI4HealthSec should be that a system cushions security gaps that persist due to the time delay of activating security patches.

2.2 Pilot #2 Medical implants – FHG - IBMT

2.2.1 Description of the pilot

Fraunhofer Institute FHG-IBMT has developed a technology platform for programmable active implants with neuro-stimulation and neuro-monitoring functionality in novel clinical applications. Intakt is an application based on the platform, formed by a network of implants and a central external unit that provides a bidirectional real-time radio connection based on the TI CC1350 microcontroller, which can be configured remotely and therefore be susceptible to cyber-attacks.

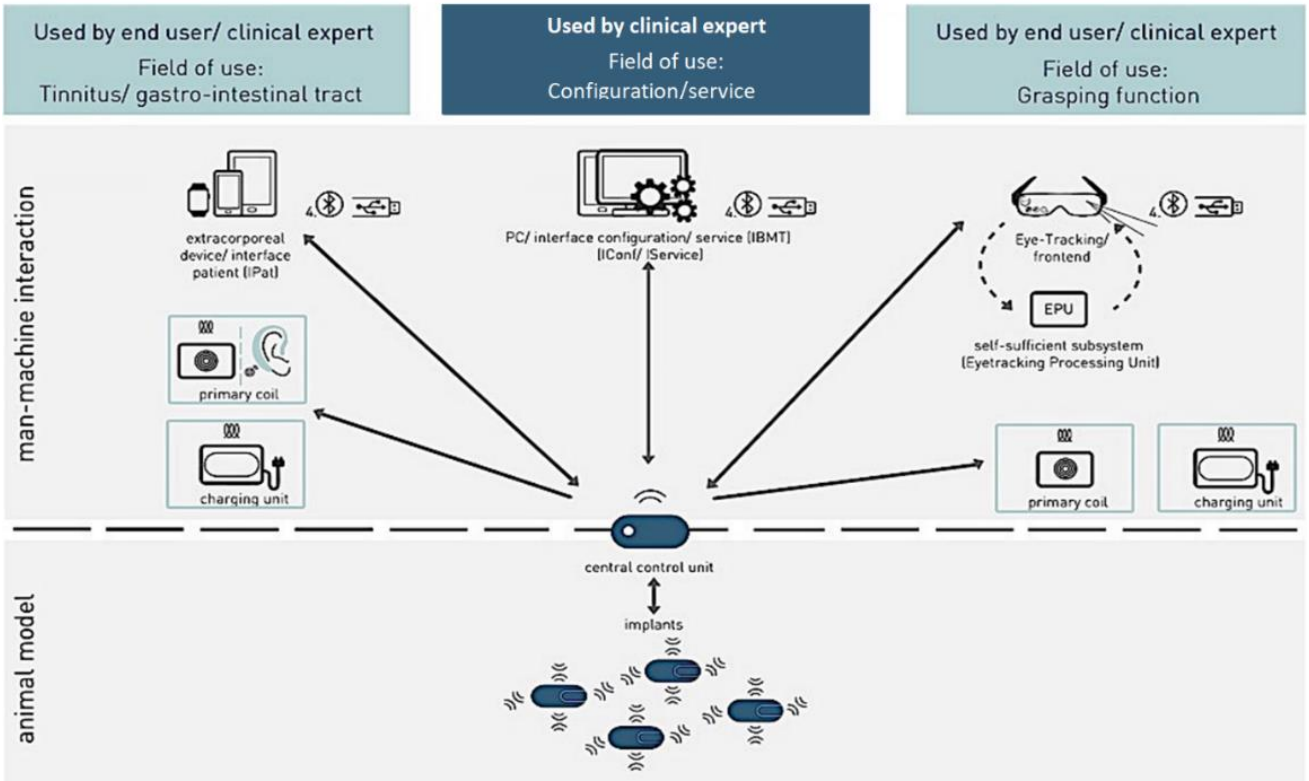

Figure 1. Architecture of Fraunhofer Medical Implants platform

One of the main challenges is to define the ease of data access and implant control, to allow healthcare personnel to take actions on malfunctions or emergencies. Simultaneously, access must be restricted to prevent possible malicious access from endangering the patient's life. The product must be able to get patched or updated, and it is also necessary to identify what types of metadata must be generated by the system to be able to detect if it is being attacked or has been attacked.

Moreover, it is essential to define a standardized way that the system should react in case of malfunction or unauthorized access.

2.2.2 Current infrastructure and available devices (HW/SW)

To simplify the pilot, only one node is planned to be used (more nodes can be used), as well as one central unit and a computer (running Windows 10 OS) with software to interact with implants and the base station. All firmware on the implants and the base station is implemented in the C programming language. The desktop application is implemented in C# with the WPF framework. The communication between the pilot actors happens block-wise, where each block can represent an instruction or data (e.g., sampled EMG signal or electrode impedance). Each block consists of a header and a payload. The header consists of five bytes, and the length of the payload is defined by one byte in the header, limiting the maximum size of the payload to 255 bytes. Moreover, the header contains sender and receiver addresses which are 1 byte long each. One bit in the header defines the type of block (instruction or data). In the case of an instruction, the first byte in the payload defines the instruction type (e.g., new stimulation profile).

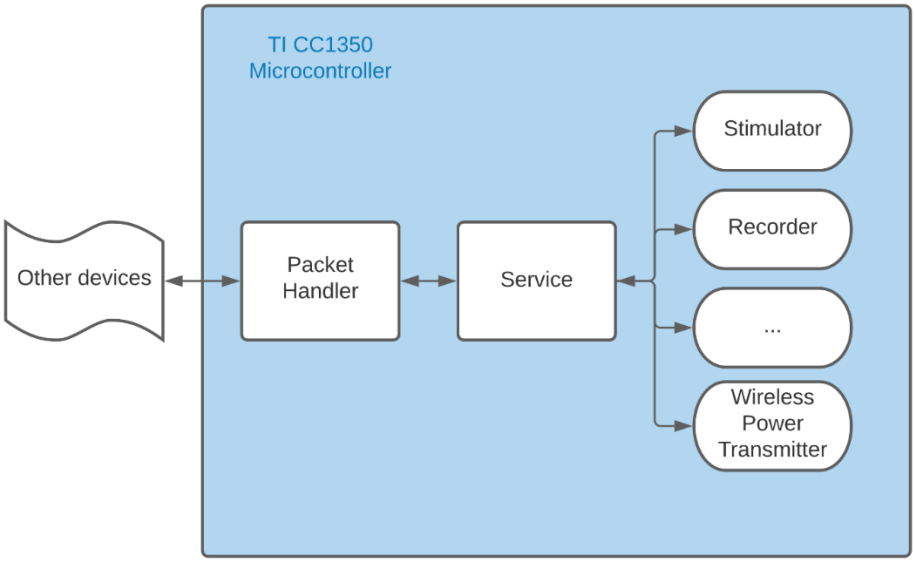

Figure 2. Architecture of Microcontroller

On each of the devices, the following schema to process the blocks applies. Packet handler collects the incoming blocks from all possible sources (e.g., BLE, USB, RF) and decides based on the header's information if the block must be processed locally or forward to another device. If the block must be processed locally, then the payload is sent to the service task, which calls a proper API of the periphery (e.g., stimulator, signal recorder, etc.). The receiving block is acknowledged by the device, and if the

block is an instruction, an answer with results will be sent back to the sender upon the completion of the instruction.

2.2.3 Current Use Case Workflow and Target Audience

The implants' network can be interacted with by a doctor or patient via an external interface, namely the central control unit. In the future, this direct communication should allow physicians to obtain

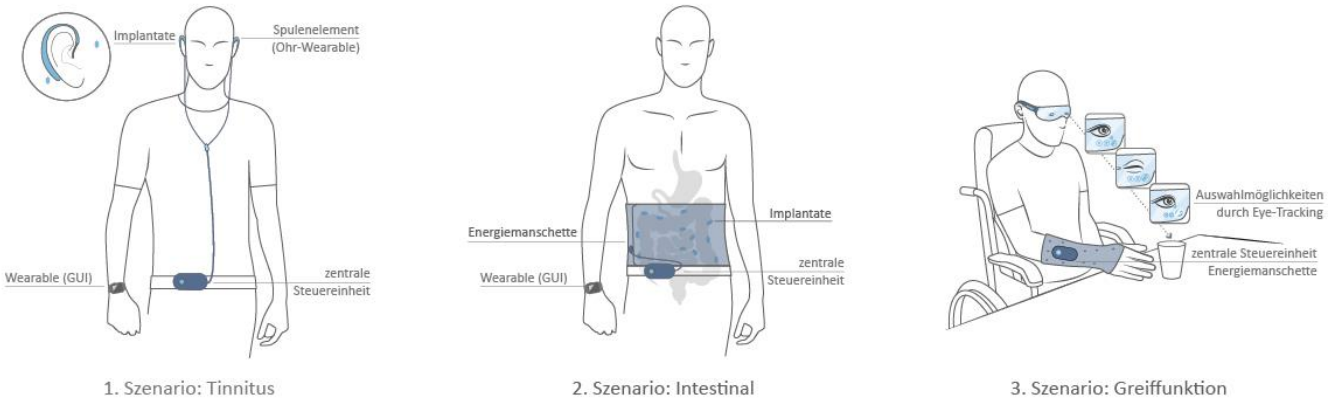

Figure 3. Application Scenario for Medical Implants

diagnostic data for treatment optimization by controlling stimulation parameters and modes. The system has three different applications: treatment of tinnitus, gastrointestinal motility disorders, and partial restoral of gripping function after paralysis.

Dysfunctions such as tinnitus can be suppressed by utilizing a stimulation method known as neuromodulation.

Digestive disorders have various causes; however, a not insignificant proportion can be traced back to a change in motility (movement activity of the digestive tract). Electrical stimulation can be used to restore this activity to normal.

Motor impairments and paralysis are common late effects of strokes or brain tumours. With the help of an implantable and controllable assistance system, intact nerves or muscles that the patient can no longer control could be reactivated by electrostimulation and thus help the patient perform the required movement.

2.2.4 Critical Data Involved

The critical data involved can be divided into two groups:

1. Data that is needed for the normal functioning system, e.g. feedback from wireless power transmission to keep all systems properly supplied and to avoid implants overheat, which can lead to damage of surrounding tissue.
2. Diagnostic data can be used by a doctor to diagnose the progression of the treatment and can be considered private data.

2.2.5 Potential Attack Scenario

There are many possible attack scenarios on implantable medical devices. The following are some potential cyberattacks to consider:

Scenario ID	Name	Description	Objectives
2.2.1	*Battery-depletion*	The attack targets the limited battery power resource of the embedded systems	It may lead to device malfunctioning and cause severe consequence for the user of the devise
2.2.2	*Man-in-the-middle*	Wireless technology solves many problems for implantable medical devices. Still, it provides a new challenge to protect information transmitted over the air which can be easily intercepted.	Intercept data transmitted between the implants and the central control unit.
2.2.3	*Evil-twin (node/base station)*	Attack is a wireless equivalent of a phishing attack, where a fraudulent device tries to appear as a	Receive data from the implants or/and control the implants by sending a command.

		legitimate part of the network	
2.2.4	*Normal vs. Emergency Modes*	As mentioned in the project description, all medical devices have two different modes: normal and emergency, in which medical personal must have more control over the medical device. Therefore, this mode must be protected against malicious intent.	An attacker can get access to functionality that must be available only to medical personal during an emergency.

Table 2: Pilot#2: Potential Attack Scenario

Assets	**Scenario ID or Referred Step**	**Technologies (of the IT system or Databases)**	**Incident-related information and data (i.e., alarms, alerts, logs) that can be collected, processed, stored …**
Implantable Medical Device (IMD)	All scenarios	Custom firmware and hardware	Communication logs between IMDs and central control unit (incl. status information e.g. temperature, voltage, current and etc)
Central Control Unit	All scenarios	Custom firmware and hardware	
PC	2.2.4	Windows 10 OS	

Table 3: Pilot#2: Potential Attack Scenario – Assett Definition

2.2.6 Security Challenge and Problems

During the research phase of the Intakt project, encryption is disabled, which makes the system more susceptible to a cyberattack. Another challenge is that any software on a medical device must go through certification as a medical product before releasing it to the market, which should be considered in the case of integration part of AI4HEALTHSEC into source code. Additionally, a medical

product must be supported until its recycling, which can apply constraints to the AI4HEALTHSEC as part of the medical device.

2.3 Pilot #3 - Personal Health Systems with on-body-sensors/actors ('Wearables') –FHG IBMT

2.3.1 Description of the pilot

Fraunhofer Institute (FHG-IBMT) developed various innovative Personal Health Systems (PHS) for the management of chronic diseases such as cancer, diabetes, liver and cardiovascular diseases but also specific theranostic solutions such as intelligent shutter glasses for amblyopia therapy in the context of public research projects. Medical information managed via these applications has to be armored against cyber-attacks and data breaches in order to ensure data confidentiality and integrity. The pilot for wearables focus on a patient monitoring and symptom reporting platform of Fraunhofer IBMT in the context of the corona pandemics and addresses relevant cybersecurity issues of this typical scenario for e-health solutions with wearables from the perspective of a non-profit organization for research and technology development.

2.3.2 Current infrastructure and available devices (HW/SW)

The corresponding personal health system consists of the commercial smart watch ScanWatch[1] of the company Withings in combination with the app Corona Diary of Fraunhofer IBMT that is used in a clinical pilot of FhG-IBMT and Saarland University Medical Center to collect self-reported symptom data from COVID-19 patients in home quarantine for research purposes. FhG-IBMT is actually extending its reporting app with the Withings ScanWatch to receive oxygen saturation values measured by this watch. The whole personal health system for this pilot scenario consists of the Withings ScanWatch together with the corresponding app of Withings to receive monitoring data of the watch that the app forwards to a server of Withings, and FhG-IBMT's app Corona Diary together with the health data integration platform XplOit of FhG-IBMT. The app Corona Diary on the smartphone of the patient downloads the patient's monitoring data acquired by the watch from the Withings server using a public API of Withings and the patient's credentials on the Withings platform. The app Corona Diary sends data received from the Withings platform together with data on symptoms that the patient enters in this app to the XplOit server, a data integration and management system of Fraunhofer IBMT where the data is used for research purposes on COVID-19.

[1] Available at https://www.withings.com/eu/en/scanwatch, accessed on 27th May 2021

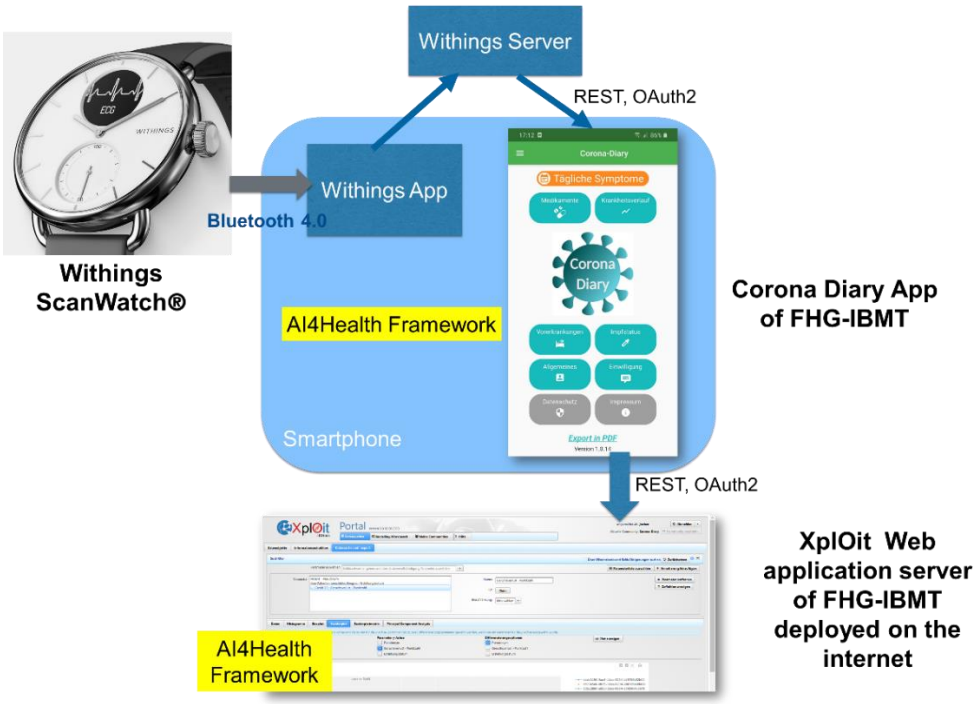

Figure 4: Component Diagram of Corona Diary App

Technologies and Frameworks

In the following the technologies and frameworks used for the components and interfaces of the Corona Diary App (Fig 4) are described:

- The **Corona Diary App** is developed with the Flutter development kit (Google) in the language dart. It is developed with the Android Studio Tool.
- The App communicates with the **XplOit Server** over a REST Interface. The REST server is implemented in Jersey and secured with Spring-Security.
- Withings is called over a REST Interface. The authentication is carried out with OAuth 2.0. For implementation the Withings API is used.
- To deploy the XplOit Server, Docker and Docker-Compose is used.
- The XplOit Server uses the following technologies and frameworks:
    - Spring and Spring-Security,
    - JSF and Primefaces,
    - MongoDB and Openlink-Virtuoso.

2.3.3 Current Use Case Workflow and Target Audience

The Corona Diary App is currently in use for a clinical study to collect patient reported data from COVID-19 cases in domestic quarantine to better understand and predict severe courses of the disease. The device interface for the Withings ScanWatch is currently being added to the app to complement the patient reports with objective information on the patients' vital signs, in particular SPO2.

The pilot for AI4HealthSec will be installed in a lab environment. The XplOit server can be provided on a virtual machine on the internet, the app can be downloaded from a private link to be used in this pilot, the watch must be purchased.

2.3.4 Critical Data Involved

The critical data involved can be divided into three groups (s. Figure 4):

- Oxygen saturation is measured by the **Withings ScanWatch**. This data can be watched from the patient in the **Withings App**, where also personal information of the patient is visible and can be entered. This information is stored on the Withings Server.
- The **Corona Diary App** can pull personal data from the Withings Server. Additionally, the patient can enter his health data into the app.
- The app Corona Diary sends data received from the Withings platform together with data on symptoms that the patient enters in this app to the **XplOit server**, a data integration and management system of Fraunhofer IBMT, where the data is used for research purposes on the health data. The data in the XplOit platform is pseudonymized.

Please note: The pilot will **not** involve patients and their personal data nor physicians, but just test users and test data to evaluate the AI4HealthSec framework.

2.3.5 Potential Attack Scenario

The incorporation of a cyber-security framework to PHS and related applications that manage medical information poses a number of challenges in terms of integration, non-invasive operation, privacy and security. Personal health systems are subject to a large number of possible cyber-security threats and attacks, which correspond to different circles of consideration of the AI4HealthSec framework. The following several possible attacks are listed:

Scenario ID	Name	Description	Objectives
2.3.1	*Attacks on wearables*	An attacker steals the wearable and uses direct access to the hardware to exfiltrate personal data or user credentials.	With such information, the attacker could access the data in the XplOit platform.
2.3.2	*Direct attack on a wearable component over Bluetooth*	An attacker exploits vulnerabilities in the wireless network stack (i.e. the Bluetooth stack).	That could allow the attacker to access the data on the ScanWatch.
2.3.3	*Man in the middle attack between a wearable and a smart phone.*	An attacker performs a man in the middle attack between the wearable and the app on a smart phone. Such an attack leads the watch to believe the attacker is the app and leads the app to believe the attacker is the watch.	That allows the attacker to collect and /or modify all transmitted personal data.
2.3.4	*Attack on the software running on the smart phone*	An attacker attacks the mobile OS of the smart phone, the Corona Diary app or any other app on the smart phone followed to gain access to the smart phone.	The attacker escalate privileges, which could allow him to access the data of the app, modify the app or access the cloud services with the credentials of the app.
2.3.5	*Intrusion*	Intrusion attacks are those in which an attacker enters the XplOit web application	Read, damage or steal data.

		to read, damage, and/or steal the data.	
2.3.6	*Exfiltration of patients' vital signs data*	Data exfiltration attacks is the act of sensitive data deliberately being moved from inside the XplOit web application to the outside without permission.	To move data from the inside of the XplOit application to the outside and misuse it.
2.3.7	*Manipulating of patients' vital signs data*	Data manipulation attacks are attacks where an intruder does not take the data from the XplOit platform, but instead make changes to the data.	Regular users can not anymore view the right data but only the manipulated.
2.3.8	*deletion, can e.g. lead to service disruption*	A successful deletion attack can e.g. cause deletion of entire database tables or other resources in the XplOit web application.	Regular users can not anymore access the deleted data.
2.3.9	*Encryption of data*	The attacker encrypts sensitive data to disrupt legitimate access.	Users can not anymore view the data.
2.3.10	*Denial of Service Attack*	An attacker overloads the XplOit system with specially crafted requests to prohibit	Regular users can not anymore access the application

		normal usage of the system.	

Table 4: Pilot#3: Potential Attack Scenario

Assets	**Scenario ID or Referred Step**	**Technologies (of the IT system or Databases)**	**Incident-related information and data (i.e., alarms, alerts, logs) that can be collected, processed, stored …**
XplOit Application Server	All scenarios	Apache Tomcat Webserver	Log files of Apache Tomcat
XplOit Database	All scenarios	MongoDB	Log files of MongoDB
Withings ScanWatch	S 2.3.1, S2.3.3	Withings API	Log files of the Withings server
App Corona Diary	All scenarios	Flutter	Runs on a smartphone, where alerts can be shown
Withings Server	All scenarios	Withings API	Log files of the Withings server

Table 5: Pilot#3: Potential Attack Scenario – Asset Description

2.3.6 Security Challenge and Problems

Often wearables have no encryption on the data that is stored on them. There are often no credentials that secure the data. Mostly no biometric security and no user authentication is required to access data on a wearable. If it is stolen, sensitive data could be accessed very easily. [2]

The ScanWatch connect to smartphones wirelessly using the protocol Bluetooth. We may have Bluetooth on our smartphone turned on all the time now so they can sync with the ScanWatch, but

[2] https://www.csoonline.com/article/3054584/7-potential-security-concerns-for-wearables.html

what else could be connecting? Many of these wireless communications are insufficiently secure to guard against a brute-force attack. [3]

We have identified several attacks on how the Diary App and the Xploit Server can be attacked (s. Sec 2.2.5). To avoid these attacks, we propose that a component of the AI4HealthSec runs on the Diary App and one component on the XplOit Server to monitor cybersecurity.

2.4 Pilot #4 - Human biobanks and related biobank information systems - FHG IBMT

2.4.1 Description of the pilot

Fraunhofer Institute (FHG-IBMT) collects and maintains important biorepositories and provides human biomaterial for research purposes. One example is represented by the European Bank for induced Pluripotent Stem Cells EBiSC that is a collection of human iPS cells being made available to academic and commercial researchers for use in disease modelling and other forms of preclinical research[4]. FHG-IBMT also collects and stores human samples from specific cohorts of donors to monitor people's exposure to contaminants in the environment on behalf of the German Environment Agency UBA. In this context, the agency conducts the so-called German Environment Surveys that addresses specific research on specific groups, to study their exposure to the environment. For the German Environment Survey on children and adolescents (2014-2017), Fraunhofer IBMT developed the specimen management system UBA-PVS to collect, process, store and manage the specimen and related data of around 2500 participants and more than 70000 samples[5]. The sources of the web application UBA-PVS, which is also used in the next environment survey of UBA, are publicly available and can be downloaded from the web[6]. UBA-PVS represents the information system for the pilot on cybersecurity in biobanks.

This use case will use the AI4HEALTHSEC framework to ensure that i) services are not interrupted and corrupted; ii) privacy rights of donors are respected and re-identification of donors is prevented; iii) falsification of health-related information and sample data is discovered; and iv) illegal copies of the information contained in the biobanks are prevented. A risk assessment that will conduct to a formative plan on privacy by design principles for biobank information systems will be required. The

[3] https://www.csoonline.com/article/3054584/7-potential-security-concerns-for-wearables.html

[4] Available at https://ebisc.org, accessed on 27th May 2021

[5] https://www.umweltbundesamt.de/en/topics/health/assessing-environmentally-related-health-risks/german-environmental-surveys/german-environmental-survey-2014-2017-geres-v

[6] Available at https://sourceforge.net/projects/uba-pvs/, accessed on 27th May 2021

high-level requirements coming from GDPR include: i) Data protection by design; ii) Ensuring of privacy rights of donors; iii) Accountability of the structure entitled of personal data management; iii) Intrusion detection system to minimize data breaches; iv) maintenance of the link between donor information and sample.

In the following sections, the UBA-PVS web application will be described in more detail.

2.4.2 Current infrastructure and available devices (HW/SW)

In Figure 5 the Architecture diagram for the web application UBA-PVS is depicted. A classical three-tier-architecture is used for the design of the application:

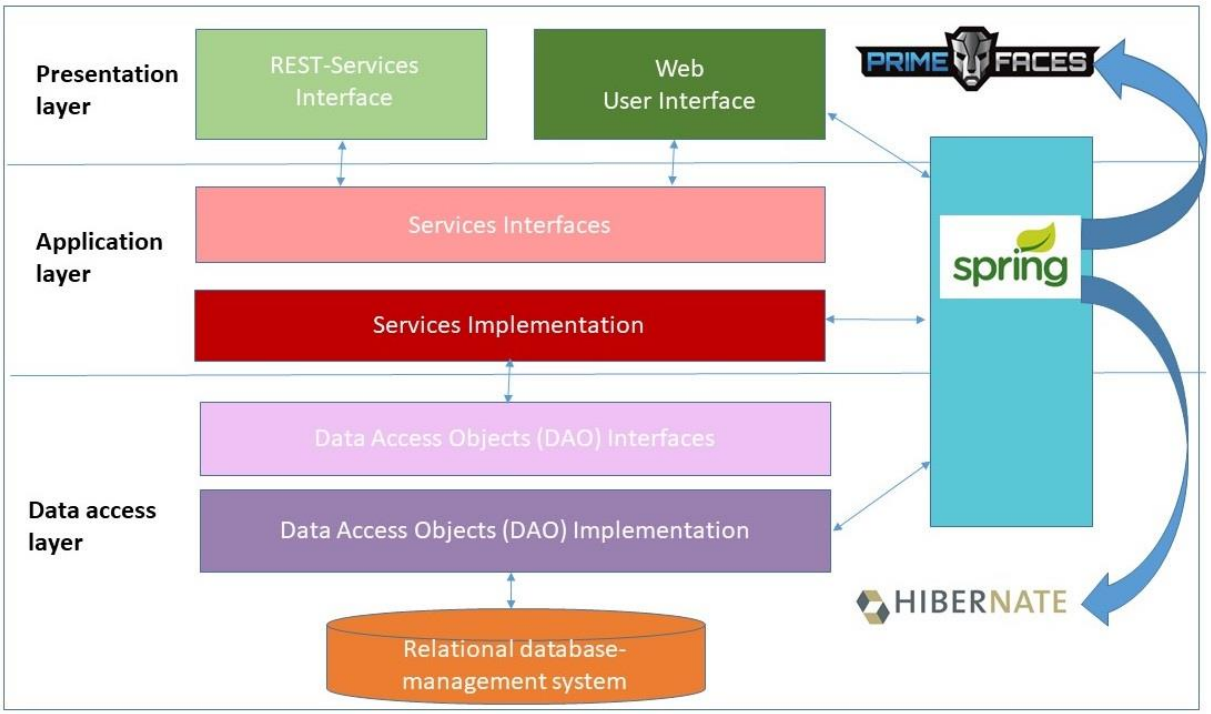

Figure 5. Architecture diagram of UBA-PVS web application

- The presentation layer is responsible for the representation of the data. In this layer, the graphical user interface and the REST-services for the UBA-PVS are settled.
- The application layer implements the complete business logic of UBA-PVS. The layer comprises the application logic, which can be called through service methods.
- The data access layer comprises the relational database management system and the implementation of the data model. The layer encapsulates access to persistent data.

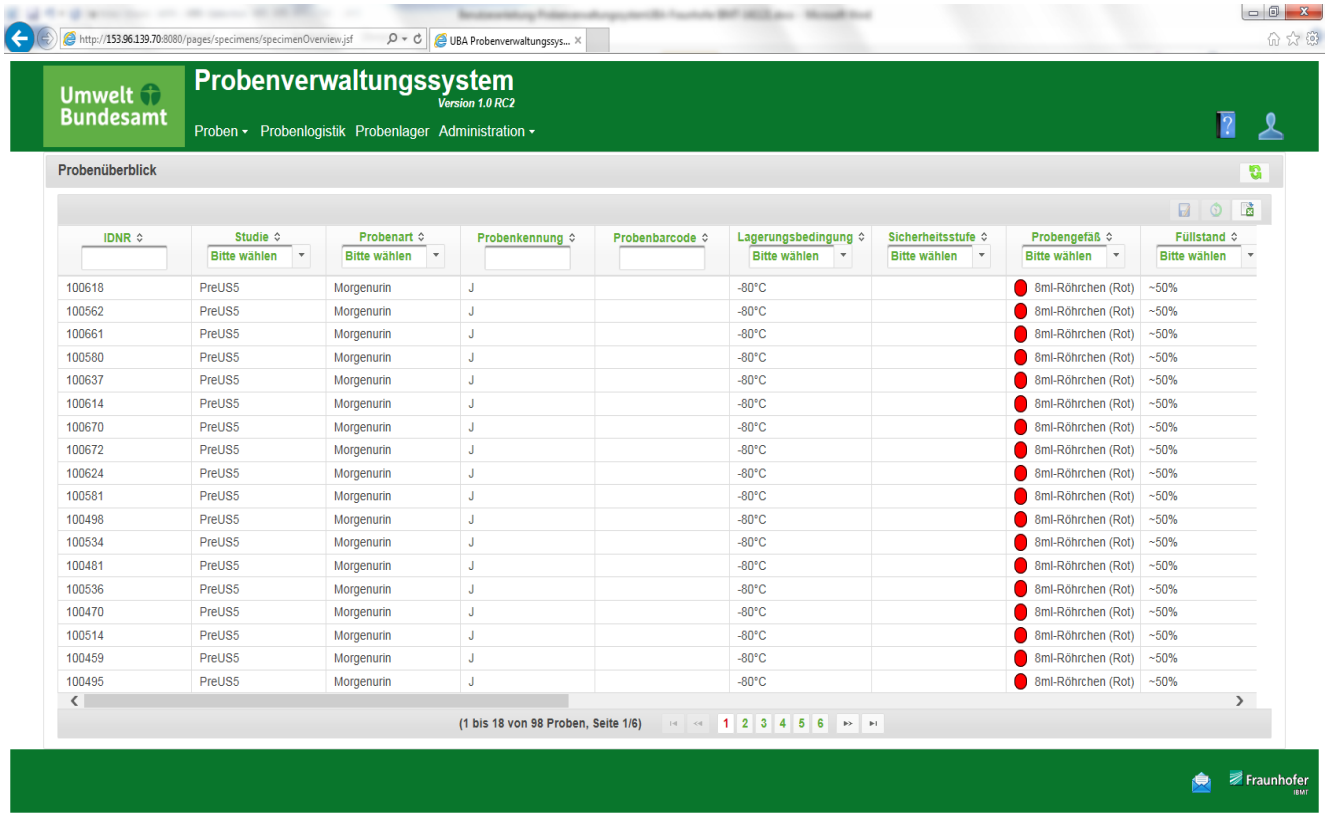

Figure 6: GUI of UBA-PVS

Technologies and Frameworks

UBA-PVS is implemented using the following technologies and frameworks for the single layers:

- *PrimeFaces* is a component-library for web development with Java Server Faces (JSF), a framework-standard for developing graphical user interfaces for web applications.
- *Hibernate* is an open-source-framework for the persistent storage of data in Java.
- The *Spring Application Framework* is an open-source framework for the Java platform. PrimeFaces and Hibernate can be integrated very well with Spring in the tier architecture and together they realize a robust framework for the development of professional web applications. With Spring Security a user authenticaiton with role based access rights on the application resources is realized.

The following Hardware- and Software-Resources are required for installing UBA-PVS:

- Server-Hardware with at least 1 GHz CPU und 1 GB RAM.
- Linux or Windows operating system.
- Java Runtime Environment (JRE) 7.
- Apache Tomcat Version 7 Webserver.

- PostgreSQL Version 9.2 or 9.3 (including pgAdmin Tool).
- Client-Browser: IE 11 or FireFox 32 (or higher).

2.4.3 Current Use Case Workflow and Target Audience

In the following, the interaction of users with the UBA-PVS system is described. In the system users can be created and assigned with specific roles and assigned with an institution. The following roles are implemented for the system:

- Standard-User: Only read (View data and export, execute queries).
- Sample Agent: Write on selected data sets (e.g. samples)
- Sample Manager: Write on sample-specific basis data (e.g. Sample repository, sample kind), Deletion of samples.
- Application administrator: Write on the Basis Data, User Administration.

In the following, the features of the system are shown:

Features

- Management, search and retrieve of specimens data (including historical data).
- Registration of new specimen data by editing forms or importing via XLS.
- Documentation of specimen's storage and other specimens processes (removing, relocating, pooling and aliquoting of specimens).
- Management of the repository (organisation of locations and containers).
- Management of logistical transactions (deliveries, shipments).
- Administration of application users (including access restrictions by authority assignment applying Spring Security).
- Maintenance of application base data (like surveys, institutions, specimen and container types).
- Dynamic extension of specimen's parameters.
- RESTful services for accessing specimens data (applying Apache CXF).
- Java Web-Application in a three-tier-architecture applying Hibernate, Spring and JSF/PrimeFaces.
- Data persistence in a RDBMS (including audit trail) applying PostgreSQL as pilot system.

2.4.4 Critical Data Involved

The application will be deployed in a lab environment on the internet. Simulated data on samples, their location in a storage system and information on donors will be generated that allow studyingthe above cybersecurity risks. The critical data involved is described in the following:

- **Samples data** can be registered, stored, outsourced and transferred. A sample contains e.g. the following values: Id, Trial, Sample type, Sample id, Sample bar code, etc.

- **Sample logistics** describes the logistic procedures (deliveries, transmissions and retour), which over the sample processes for samples, attached to trials, are registered.
- **Data for sample storage** describes the sample storage with tanks (freezer, boxes etc.). A user can create, view and manage these data.
- **User data**: Users of the application can be viewed, created, edited, activated and deleted. A user account can be used for registration and can be linked with trials and as a person in charge record sample processes.

The application will be deployed in a lab environment on the internet. The pilot will **not** involve individuals and their personal data nor users outside our organization (FHG IBMT).

The pilot will **not** involve individuals and their personal data nor users outside our organization.

2.4.5 Potential Attack Scenario

In the following, we will describe some attack scenarios that we would expect to be detected by the AI4HealthSec framework.

Scenario ID	Name	Description	Objectives
2.4.1	*Intrusion*	Intrusion attacks are those in which an attacker enters the UBA-PVS web application to read, damage, and/or steal the data.	The attacker can read, damage and/or steal the data.
2.4.2	*Exfiltration of sample and patient data*	Data exfiltration attacks is the act of sensitive data deliberately being moved from inside a web application to the outside without permission.	The attacker can steal the data.
2.4.3	*Manipulating sample and patient data*	Data manipulation attacks are attacks where an attacker does not take the data but instead make changes on the data for some type of gain.	An attacker can make changes to the data.

2.4.4	*Deletion can e.g. lead to service disruption*	A successful deletion attack can e.g. cause deletion of entire database tables or other resources.	Regular users can not anymore access the deleted data.
2.4.5	*Encryption of data*	The attacker encrypts sensitive data to disrupt legitimate access.	Users can not anymore view the data.
2.4.6	*Denial of Service Attack*	An attacker overloads the system with specially crafted requests to prohibit normal usage of the system.	Regular users can not anymore access the application
2.4.7	*Disclosure of sensitive data of donors; re-identification of donors*	The attacker finds a way to disclose sensitive data of donors and reidentify donors.	Donors are reidentified and have possibly damages.
2.4.8	*Illegal copies of the information contained in biobank information systems*	The attacker extracts illegal copies of the information contained in biobank information systems.	Possibly, illegal information is extracted from the system.

Table 6: Pilot#4: Potential Attack Scenario

Assets	**Scenario ID or Referred Step**	**Technologies (of the IT system or Databases)**	**Incident-related information and data (i.e., alarms, alerts, logs) that can be collected, processed, stored …**
UBA-PVS Application Server	All scenarios	Apache Tomcat Version 7 Webserver	Log files of Apache Tomcat

UBA-PVS Database	All scenarios	PostgreSQL Version 9.2 or 9.3	Log files of PostreSQL
UBA-PVS Web Application and Data	All scenarios	PrimeFaces, Hibernate, Spring Application Framework	Log files of the Web Application

Table 7: Pilot#4: Potential Attack Scenario – Assett Description

2.4.6 Security Challenge and Problems

The web application UBA-PVS is from 2014 and wasn't maintained since then. Some libraries used are outdated and can be a risk for security problems. We expect that the AI4HealthSec infrastructure can find and identify these security problems.

2.5 Pilot #5– Secure Access and Sharing of Clinical Data via VNA and Portal systems - EBIT

2.5.1 Description of the pilot

The large amounts of digital clinical, biomedical and health data, are 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 document types 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.

Clinical Information, acquired with diagnostic equipment (personal, wearable, small as well sophisticated medical equipment) are archived, reported, distributed to different health care providers (General Practitioners, Administrative Personnel, First Care HC Providers, Reporting Physicians, Second Opinion Physician, Healthcare Specialist, Insurance Company, Clinical Researchers etc…), who needs to use the information in a different phase of the process and for a different purpose.

Last but not least, or better first, the owner of the information is the patient itself that need to be empowered on this big amount of clinical data.

Latest technologies (cloud, high network bandwidth, processing power) enables new clinical use cases, which distribute the data much more. For example, performing a diagnosis by the best available specialist in the country for a specific clinical question from a remote hospital or clinic as wellviewing medical image data at any location by a specialist, when he is not in the hospital. Additional services in cloud, like Artificial Intelligence/machine learning, poses new access, privacy and security challenges.

This pilot use case will be based on the use of the EBIT solution SUITESTENSA VNA and Portal integrated with the AI4HEALTHSEC security framework to tackle the security challenges raised at any level of the solution.

EBIT SUITESTENSA Vendor Neutral Archive (VNA) and Portal solution (see next figure) provides interconnectivity with the latest technologies and infrastructures, such as cloud, Machine Learning SaaS and high-net-work bandwidth.

The use-case led by EBIT pilot is related to the security issues raised in the information integration workflow behind the distribution of clinical imaging and report information within healthcare providers and third-party applications through a VNA and Portal Solution.

Healthcare providers may be:

- Second Opinion Specialist
- Pre-Operative Healthcare Professional
- General Practitioners

All the previously mentioned different users may access information on a variety of devices starting from sophisticated report workstation, to smartphone or tablet devices, accessing the webportal of the NHS System or Regional HC system and so on.

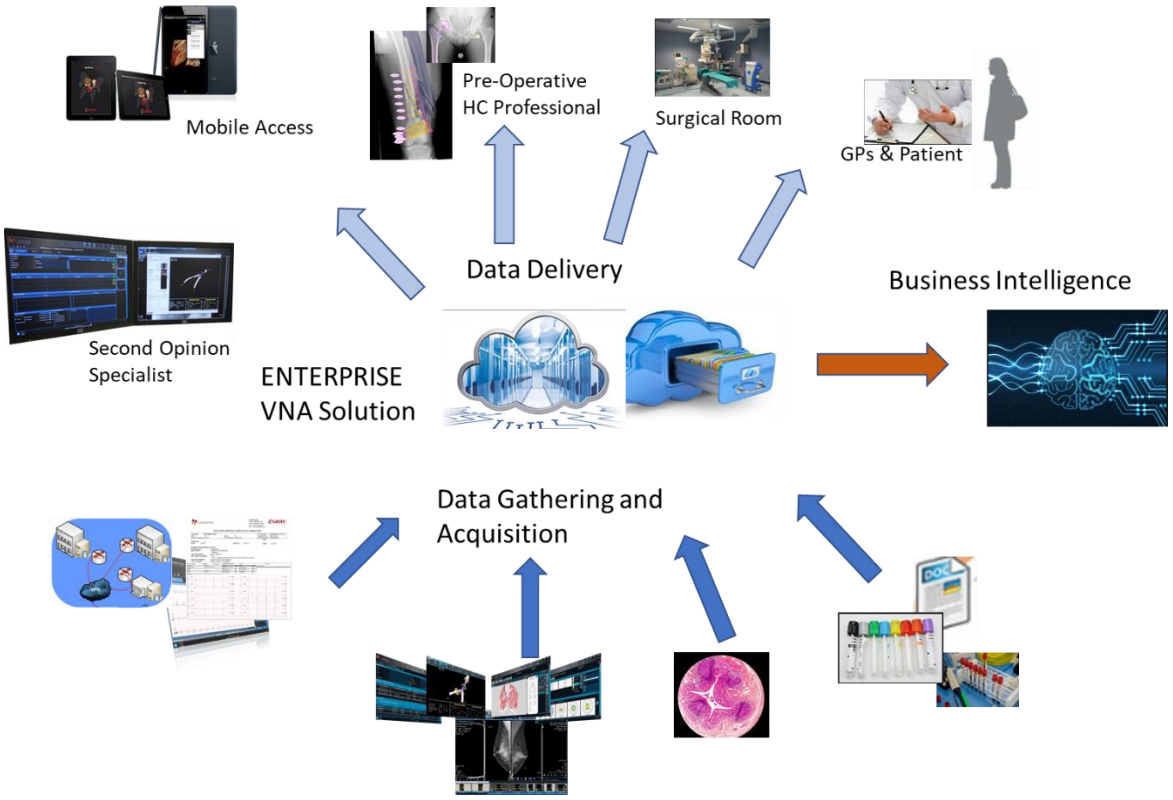

Figure 7: Enterprise VNA solution within an Healthcare Organization

This use-case aims to identify the security and privacy challenges related to a fine-grained access to clinical data by various types of users and from various devices in a VNA and Portal solution. The goal is to support and enhance the current software platform and the new ones with the DSAF, in order to ensure a secure, reliable and trustworthy delivery of personal clinical data to address the legitimate concerns of security, scalability and privacy of electronic medical records.

2.5.2 Current infrastructure and available devices (HW/SW)

The following diagram shows the current SUITESTENSA system software architecture with respect to the different types of client supported: web based, desktop thick client and WADO services.

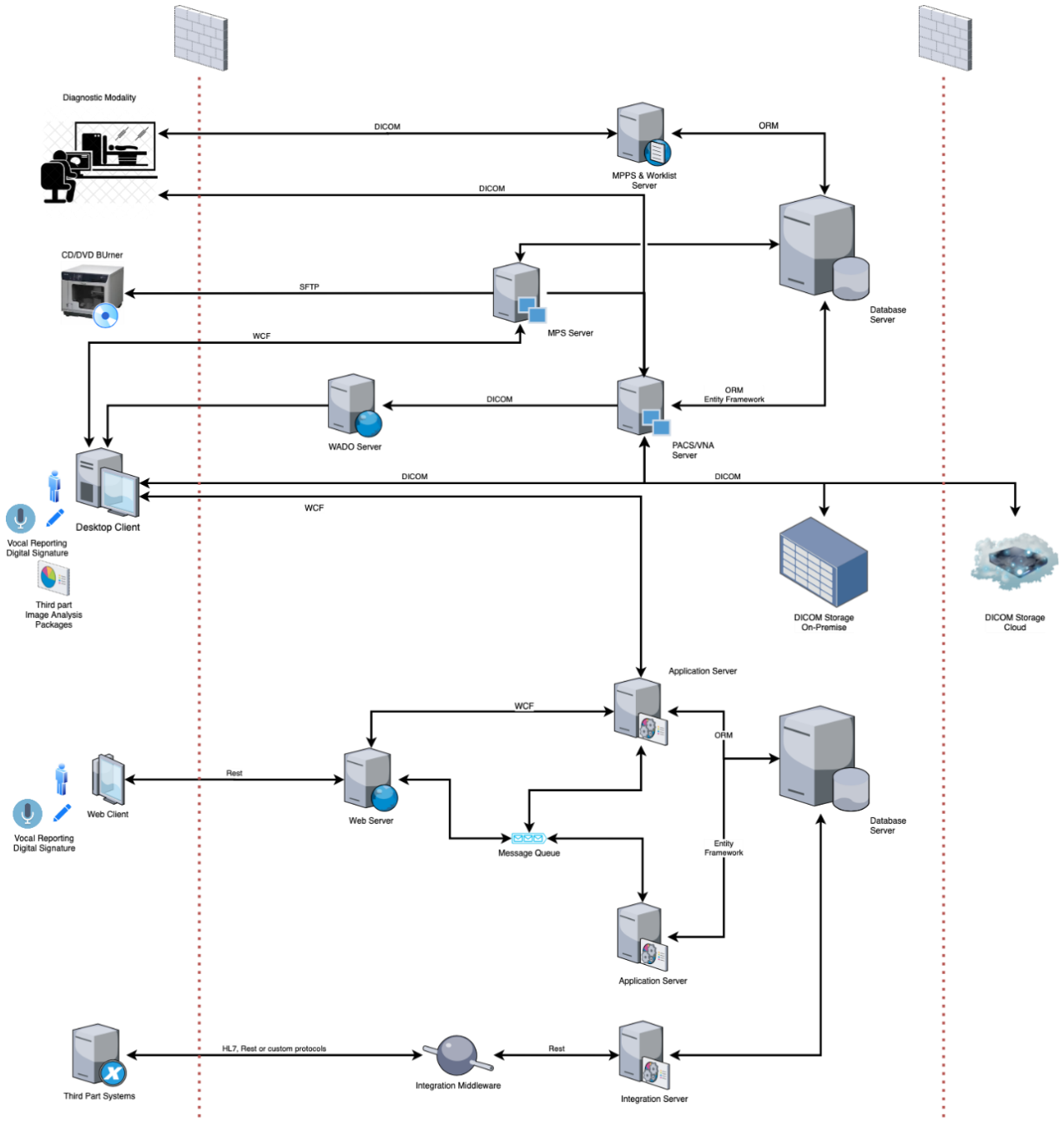

Figure 8: EBIT SUITESTENSA Software Architecture

In the following schema it is also highlited in the red box the potential access attack surface and in the green box the safe environment.

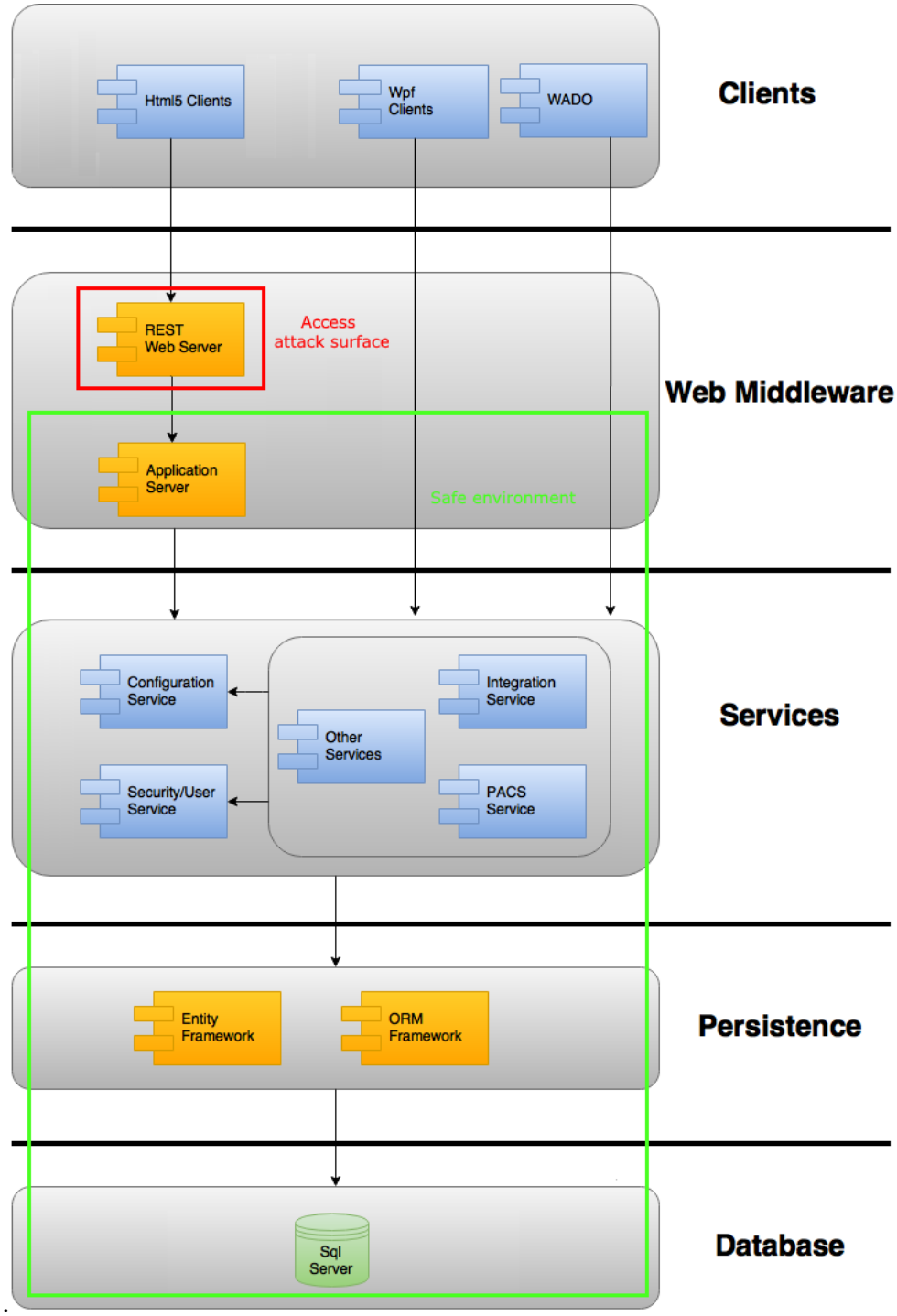

Figure 9: EBIT SUITESTENSA Logical Software Architecture

In the following schema it is also highlighted in red and blue the secure protocol transfer to be implemented and monitored in SUITESTENSA and in the integration with other application:

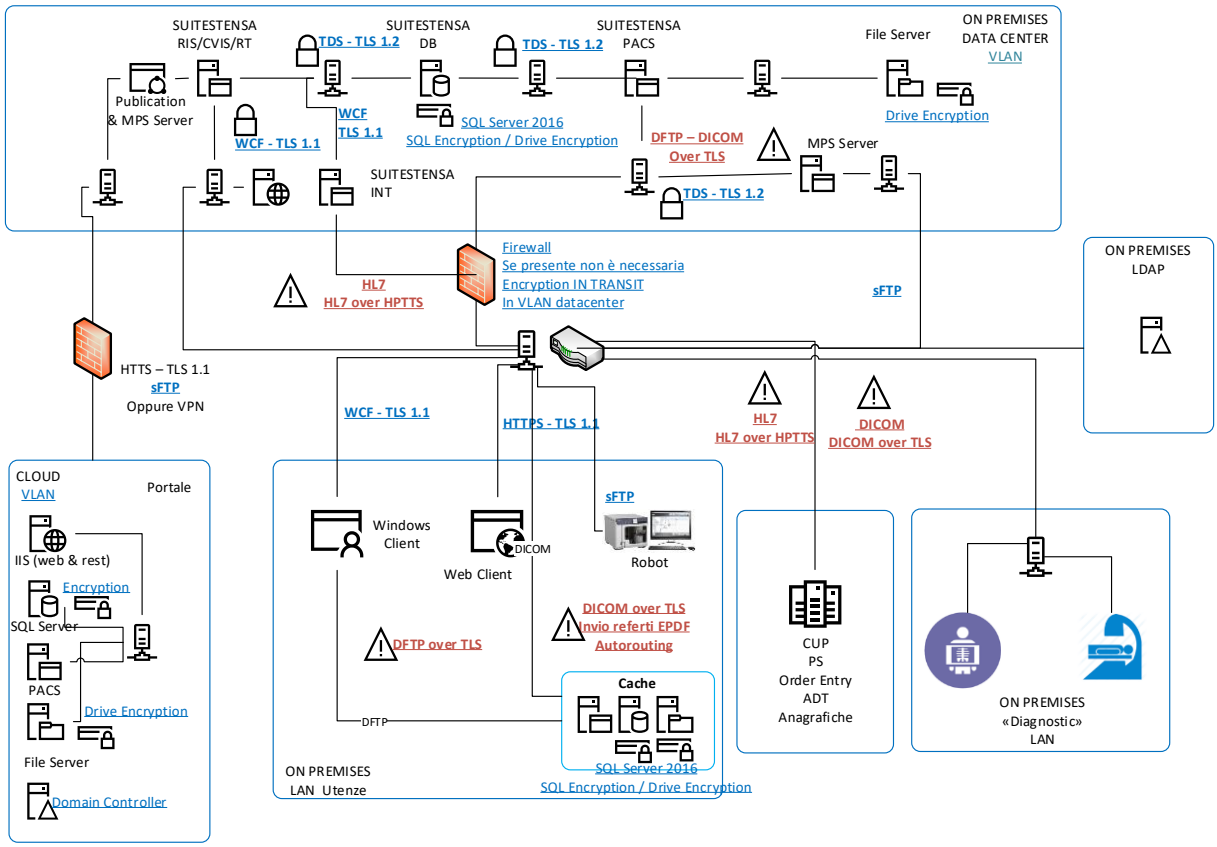

Figure 10: SUITESTENSA – Secure transfer Protocol applied

2.5.3 Current Use Case Workflow and Target Audience (Final End-User)

The use-case led by EBIT pilot is related to the issues raised in the information integration workflow behind the distribution of clinical imaging and report information within healthcare providers and third-party applications through a VNA and Portal Solution.

Healthcare providers may be:

- Second Opinion Specialist
- Pre-Operative Healthcare Professional
- General Pratictioneers

All the previous mentioned different users may access information on a variety of devices starting from sophisticated report workstation , to smartphone or tablet devices, accessing web Portal of the NHS System or Regional HC system and so on.

Clinical information acquired with diagnostic equipment (personal, wearable, etc.) is archived, reported, distributed and managed also by AI/ML Business Intelligence to different healthcare-related end-users (Healthcare Government, Physicians, Administrative Personnel, Insurance Companies, etc.), who need to access the information in different phases and for different purposes. Last but not least, the owner of the information is the patient itself who should be empowered on accessing his/her own clinical data (a requirement also mandated by GDPR).

Further use case workflow is related to advancement in imaging equipment, for example with portable tablet-based UltraSound equipment or X-ray equipment which makes it suitable for ambulatory care in the ambulance or at home. Professional reading is required of the UltraSound and X-ray images, by the specialist in the hospital or clinic. Secure and safe transfer of image data is required by the clinician.

2.5.4 Critical Data Involved

The critical data involved in the pilot are imaging and clinical data used to create a diagnosis and therapeutical plan. These data involve sensitive patient data at all levels. Imaging studies and clinical data must be stored, viewed and exchanged. The data must be complete in order to create a diagnosis.

Data must be confidential:

- The data must be encrypted at rest and in transit
- The data must only be accessible by authorized persons or systems

Data must be accurate: complete, correct and correlated to the right patient identity:

- Incomplete data can lead to misdiagnosis
- Data correlated to the wrong patient will lead to the wrong diagnosis, which may have dramatic impact due to follow-up treatment
- Data may not be deleted, typically for 7 years after the creation of the imaging study

Data must be available and accessible:

- Unavailability of image study may lead to the new examination, which may include contrast agents or additional x-ray.
- Unavailability may lead to delayed diagnosis with potential consequences. This is especially applicable for trauma patients where diagnosis must be instantly available.

Compliance with regulation (subset, to be extended):

- Data must be encrypted at rest and in transit
- Data is only accessible for authorized and identified users

- Every access to data and CRUD operations must be audited

Security requirements at high level and information asset are:

- Secure medical devices: no tampering, safe operation and data privacy
- Secure access, exchange and storage within the hospital
- Secure exchange from hospital to cloud and vice versa
- Secure access and storage in the cloud
- Compliant with regulatory requirements (HIPAA, GDPR, etc)

2.5.5 Potential Attack Scenario

In the following some attack scenarios that we would expect to being detected by the AI4HealthSec framework in the curret scenario description:

Scenario ID	Name	Description	Objectives
S5.1	*Intrusion*	Intrusion attacks are those in which an attacker enters the SUITESTENSA application.	The attacker can read, damage and/or steal the data.
S5.2	*Exfiltration of sample and patient data*	Data exfiltration attacks is the act of sensitive data deliberately being moved and extracted from SUITESTENSA application to the outside without permission.	The attacker can steal sensitive data and/or copy or let then available to 3rd party for any other type of data management not in line with the scope for which the data have been gathered.
S5.3	*Manipulating sample and patient data*	Data manipulation attacks are attacks where an intruder does not take the data, but instead make changes or alter data for some type of gain.	An attacker can make changes on the data.
S5.4	*Data Deletion, can e.g. lead to service disruption*	A successful deletion attack can e.g. cause deletion of entire SUITESTENSA database	Data deletion, and permanent unavailability has the effect that users can not anymore access the

		tables or other resources and disrupt the service	deleted data with potential damage on the patient clinical diagnostic and therapeutical path
S5.5	*Encryption of data*	The attacker encrypts sensitive data to disrupt legitimate access.	Data encryption and temporaryunavailability have the effect the users can not anymore access the encrypted data with potential damage on the patient clinical diagnostic and therapeutical path Users can not anymore view the data.
S5.6	*Denial of Service Attack*	An attacker overloads the system with specially crafted requests to prohibit normal usage of the system.	Regular users can not anymore access the application

Table 8: Pilot#5: Potential Attack Scenario

Assets	**Scenario ID or Referred Step**	**Technologies (of the IT system or Databases)**	**Incident-related information and data (i.e., alarms, alerts, logs) that can be collected, processed, stored …**
SUITESTENSA Image Storage Server and Image Database	All scenario	Microsoft SQL Server 2016 Database supporting Dynamic Data Masking and TLS Criptography Multiple Level of Image Data Storage Area Network on premises managed by state of	data extracted fromSUITESTENSA Sherlogic Monitoring System and related to CyberSecurity attack and not on system Performance

		the art Storage System Hw and Sw solution - High Performance SSD for Short Term Archiving System (eg: 10 TBytes) - HDD Disk for long term (eg hundred of Terabytes) Disaster Recovery on Cloud (eg Azure)	
SUITESTENSA Database Configuration	All scenario	Microsoft SQL Server 2016 Database supporting Dynamic Data Masking and TLS Criptography	Not Applicable
SUITESTENSA Integration with 3rd party System	S5.2, S5.3, S5.5, S5.6	Netwrok Communication and DICOM and HL7 network protocol even protected through TLS and HTTPS security	SUITESTENSA Sherlogic Monitoring System detect no communication or communication interrupt with 3rd party
SUITESTENSA Portal Image & Data Database System	All scenario	Microsoft SQL Server 2016 Database supporting Dynamic Data Masking and TLS Criptography Multiple Level of Image Data Storage Area Network on premises managed by state of the art Storage System Hw and Sw solution - High Performance SSD for Short Term Archiving System (eg: 10 TBytes) HDD Disk for long term (eg hundred of Terabytes	data extracted from SUITESTENSA Sherlogic Monitoring System that could be related to CyberSecurity attack and not on system Performance

Table 9: Pilot#5: Potential Attack Scenario- Asset Definition

2.5.6 Security Challenge and Problems

Security and privacy challenges related to this use case include the fine-grained authorization of data access during normal operations, as well as cross-institutional access control, i.e. from individuals who can access their records across institutions to see a lifetime history of their health records and decide which physicians can see which records. With patient consent, anonymized and aggregated data could be made available to researchers and other organizations that benefit from access to total population health data. In addition, other security challenges that are going to be analysed in this pilot involve certification of data acquired by different sources, and common challenges with Pilot 1 such as secure data transmission, storage and privacy in the cloud. The study of the use case with the assistance of an HC provider will help us to better identify and consider the actual functional and non-functional constraints of the relevant medical services behind a VNA and Portal Solution.

Security and privacy have become a hot topic over the years. Regulatory and hospitals raise high demands on security and privacy. For example, Europe has defined GDPR and United States has defined HIPAA requirements. Compliance with regulation is a prerequisiteto be able to be in business. Next to regulatory also privacy has become critical. Security and privacy breaches get high attention in media today. Breaches result in reduced or lost trust in products/companies. Therefore, appropriate security and privacy are critical for staying in business.

2.6 Pilot #6 –Digital Health Living Lab – University of Brighton

2.6.1 Description of the pilot

The UoB Digital Health Living lab is a user-centred, open innovation ecosystem based on systematic user co-creation/ co-production approach, integrating research and innovation processes in a real life setting and reflects the European Network of Living Labs (ENoLL), Living Labs (LLs) definition. According to the UK's Department of Health's Personalisation Communications Toolkit, co-production is when individuals influence the support and services they receive, or when groups of people get together to influence the way that services are designed, commissioned and delivered. Integration of research and innovation processes is achieved through use of the lab as a tool, an engine, where a range of stakeholders can get tailored value depending on their needs and ambitions, including:

The residents - contribute to health innovation in a new way. They get the opportunity to help other residents and can be key partners in inspiring health innovation for the greater good. They get access to new technology and will see new possibilities before others – they can get early access to new services before they go to market. **The local council** - have the opportunity to use the Lab for inspiration in innovating new services for the future. The aim is to create sustainable welfare solutions for the general population across Brighton and Hove. The staff can test new services and approaches in delivering welfare services to the residents living in Leach Court and nearby. The staff will also gain knowledge and in delivering healthcare with technology embedded. **Academia** - are able to utilise the Lab for research and training purposes. It can act as a clinical placement opportunity for both

undergraduate and postgraduate students and also a field for research projects on digital health. **Digital Health Technology developers** – can engage with citizens as end users to provide feedback throughout all the phases of a product development including in ideation, building and evaluation.

The Living Lab development was built up on work that started back in 2017 as part of the Leading Places initiative (a national initiative aiming to set up and develop meaningful relationships between universities and regional ecosystems) and a collaboration between the University of Brighton, Sussex, Kent Surrey and Sussex Academic Health Science Network, Brighton and Hove City Council and Brighton and Hove Clinical Commissioning Group. The aims were to aid the development of strategies in self-managed care for older people by focusing a range of interventions at a group of people living in supported housing development Leach Court, in order to identify ways to prevent or delay them entering into more intensive and expensive care programmes. The initial pilot project was chosen because projected levels of demand for adult social care services outstrip the city council's available resources and measures are being explored to support the most vulnerable residents in the city, to help them to remain as independent as possible.

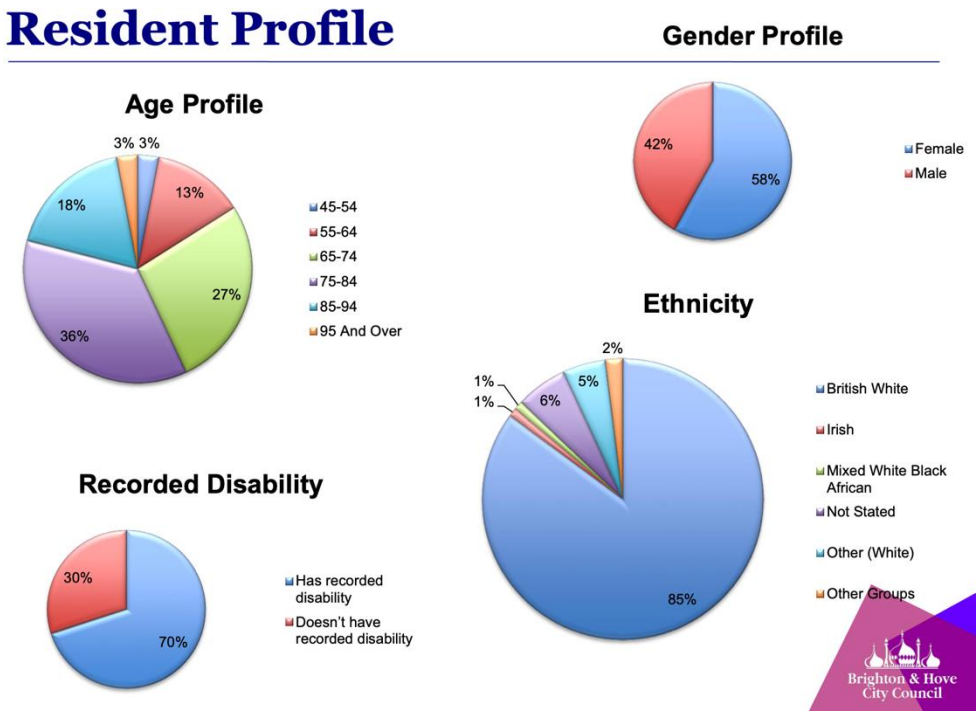

Figure 11: Living Lab resident profile.

The Living Lab is currently been utilized for two other European projects:

I**nnovateDignity**, where citizens will provide their views on developing dignified digital health technologies (https://innovatedignity.eu)

Empowercare, where citizens will test technologies aiming to tackle loneliness and isolation and they will provide their views on efficient ways to implement similar solution in the community and empower them (https://www.interreg2seas.eu/en/EMPOWERCARE)

2.6.2 Current infrastructure and available devices (hw/sw)

As an open innovation ecosystem, the living lab act a unique test bed for developing and testing prototypes or more mature digital health solutions.

Figure 12 shows the current infrastructure of the Living Lab, highlighting the interconnections between the stakeholders.

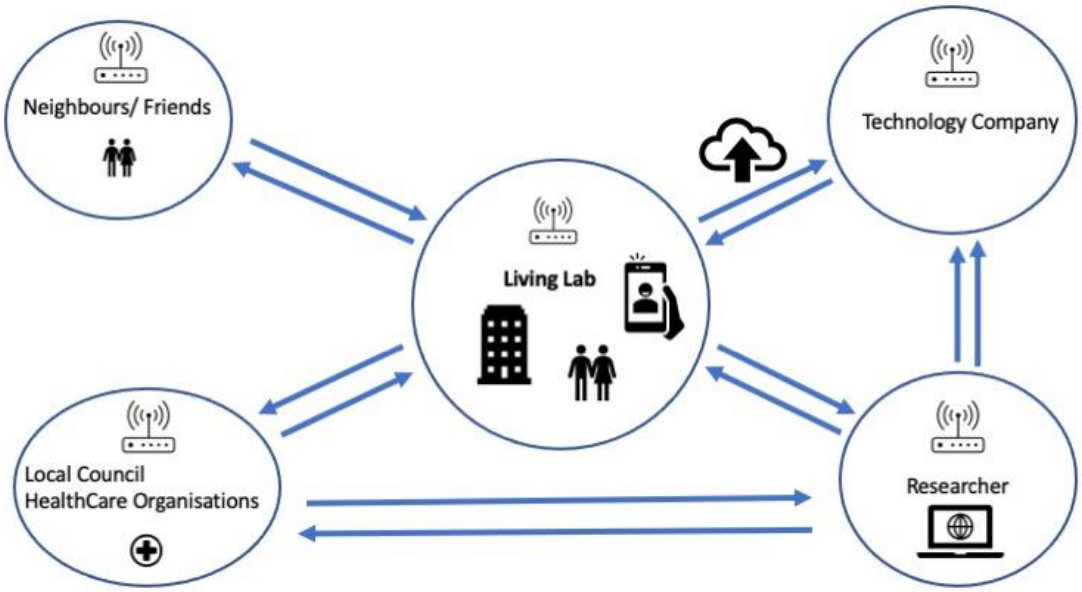

Figure 12: Living Lab Infrastructure

Every stakeholder (Local Council, Researchers, tech companies) engaging with the Living Lab work within their own infrastructures and network connections. As such they connect to the internet through their own Wi-Fi (routers) and communicate through emails (PCs) or their mobile devices (mobile phones, tablets, laptops)

The citizens may test and trial different categories (both for purpose of use and maturity) of Tier 2 and above, Digital Health Technologies (DHTs) as these have been classified by the UKs National Institute for Health and Care Excellence (NICE):

Tier	Functional Classification	Description
Tier 2: DHTs which help users to understand healthy living and illnesses but are unlikely to have measurable user outcomes.	Inform	Provides information and resources to patients or the public.
	Communicate.	Allows 2-way communication between users and professionals, carers, third- party organisations or peers. Clinical advice is provided by a professional using the DHT, not by the DHT itself.
	Preventative behaviour change.	Designed to change user behaviour related to health issues with, for example, smoking, eating, alcohol, sexual health, sleeping and exercise.
Tier 3a: DHTs for preventing and managing diseases. They may be used alongside treatment and will likely have measurable user benefits.	Self-manage.	Aims to help people with a diagnosed condition to manage their health. May include symptom tracking function that connects with a healthcare professional.
	Treat	Provides treatment for a diagnosed condition (such as CBT for anxiety), or guides treatment decisions.
Tier 3b: DHTs with measurable user benefits, including tools used for treatment and diagnosis, as well as those influencing clinical management through active monitoring or calculation. It is possible DHTs in this tier will qualify as medical devices.	Active monitoring.	Automatically records information and transmits the data to a professional, carer or third-party organisation, without any input from the user, to inform clinical management decisions.

The **blue arrows** represent the sharing of data which can have the form of messages, images, videos, videocalls etc. These data, depending on the user, are been stored in different devices: i.e for the researcher at the University of Brighton's OneDrive, for the company in their cloud storage and for the citizen in their mobile device hard drive.

2.6.3 Current Use Case Workflow and Target Audience (Final End-User)

Citizens living independently in their own homes contribute as end users to the trial of digital health solutions. In a hypothetical scenario, the end user (citizen) will trial a communication device to provide feedback the tech company for the usability and ease of use. At the same time a researcher will collect data online by the end users regarding the usefulness of the device. During the trial the end user will use the device to communicate with neighbours and friends through their mobile devises (i.e smartphones)

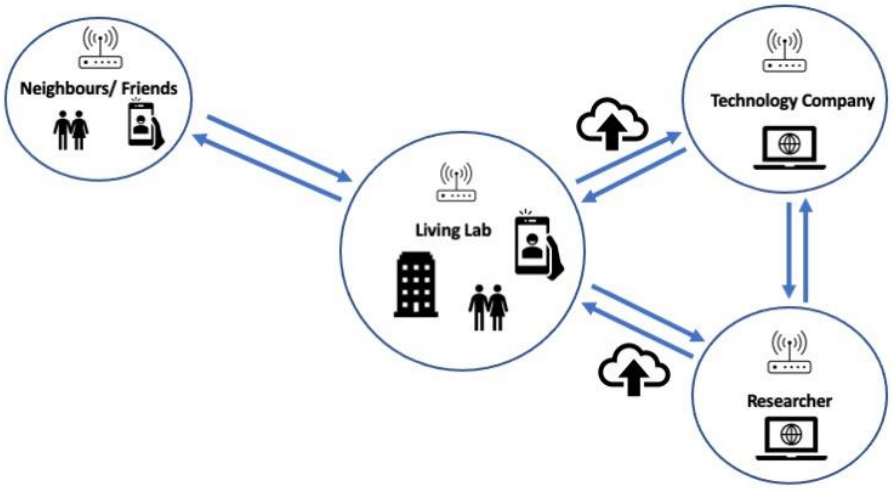

Figure 13: A scenario of testing a telehealth/ communication device

In Figure 13 above we can see an representation of a possible testing scenario. A tech company installs the device in the citizen's home. The device connects to the TV and through the citizens router to the Wi-Fi. The citizen uses the device to communicate (videocall, share of photos) through their TV with their relatives (granted access by the users themselves). The researcher communicates with the citizen through emails and videocalls (i.e through Microsoft Teams) from their own Wi-Fi connection, to conduct interviews and online surveys, collecting feedback on the experience of using the device and stores these data in the institutions OneDrive. The company collects usage data through the internet and stores these in their cloud storage. The researcher and the company share bth of their collected data through emails.

2.6.4 Critical Data Involved

The type of critical data involved in the above scenario, include personal information shared through online discussions, data of device usage (on/ off connection, duration of connection, type of activity during connection i.e call with another device). A researcher will collect data regarding the usability and the interaction of end user with the device and from a business perspective data related to the usability of the device are critical. Another type of critical data involved are healthcare data, as the device can be used as a communication tool between the end user and a healthcare professional to discuss health related issues.

2.6.5 Potential Attack Scenario

The Digital Health Living Lab's unique structure as an open ecosystem, is the main challenge of incorporating a cybersecurity framework. The exchange of sensitive personal data through communication channels (blue arrows) of stakeholders (citizens, researchers, IT product companies, local councils) with different levels of security and privacy awareness as well as different levels of protective mechanisms (hardware, software) raise the cyber security and privacy risk.

A potential attack could target any of the blue arrows where information is exchanged. Reflecting on the described scenario above most potential attack scenarios would include malware, phishing and eavesdropping. As the Living Lab is utilised as a test bed for health tech devices by end users, the potential attack scenarios can be applicable regardless the device. Below the potential attack scenarios relate to a common health care telemedicine device as only a representation of the variety of potential other technologies that can be trialled in a Living Lab.

Scenario ID	**Name**	**Description**	**Objectives**
S6.1	*Attacks on telemedicine device*	An attacker steals the device and uses direct access to the hardware to exfiltrate user credentials.	With such information the attacker could access the data in the manufacturer's platform.
S6.2	*Direct attack on the wireless infrastructure of the living lab*	An attacker exploits vulnerability in the wireless network stack.	That could allow the attacker to access the data on the manufacturer's server.
S6.3	*Indirect Attack: Man in the middle attack between the telemedicine device and a smart phone.*	An attacker performs a man in the middle attack between the device and the app on a smart phone.	That allows the attacker to collect and or modify all transmitted personal data.
S6.4	*Attack on the software running on the telemedicine device*	An attacker attacks the devices operating system, to gain access to the it.	The attacker escalates privileges, which could allow them to access the data of the device, modify it or access the cloud services

			with the credentials of the device.
S6.5	*Deletion, can e.g. lead to service disruption*	A successful deletion attack can e.g. cause deletion of entire database tables or other resources in the company's server.	Regular users can not anymore use the services.
S6.6	*Encryption of data*	The attacker encrypts sensitive data to disrupt legitimate access.	Users cannot anymore use the services.
S6.7	*Denial of Service Attack*	An attacker overloads the telemedicine company's system with specially crafted requests to prohibit normal usage of the system.	Regular users can not anymore access the services affecting the quality of healthcare provision and/ or their health directly
S6.8	*Social Engineering Attack*	An attacker through social engineering techniques tricks the authorised owner of the device to share information with them	The attacker gains access to the personal information of the living lab resident stored on the manufacturer's platform.
S6.9	*Modification*	The attacker modifies the data on the company's server to create data that best suit their needs or to prepare for further attacks	Information provided to legitimate users is incorrect and can lead to wrong decisions

Table 10: Pilot#6: Potential Attack Scenario

2.6.6 Security Challenge and Problems

Reflecting on the infrastructure of the Living Lab which includes the sharing of data between many stakeholders with each one of them managing and storing these data with different approaches and different security standards, the challenge of securing these is clear. Questions and problems:

1. What are the vulnerability of such diverse networks/ ecosystems?
2. Who owns the data and has overall responsibility of their security?

3. How does the level of cybersecurity awareness impact the overall security of the interconnections (**blue arrows)** and affects the vulnerabilities?
4. What type of training/ raising awareness should be provided and to whom?
5. Where should (which domain/ stakeholder) an AI4HealthSec solution be applied?

2.7 Pilot Scenario versus AI4HEALTHSEC Architecture

The definition of the components and services of AI4HEALTHSEC Architecture is going on in parallel and is being defined within the Task 2.4 and will be finalized in deliverable D2.4, due at the same date of this D2.3 deliverable. For this reason at this time if it not possible to depict-in the current pilot scenario definition- the level of the integration of the HC systems of the pilots and the AI4HEALTHSEC platform. This further scenario integration refinements will be part of the WP6 in the Task 6.1 "Pilot implementation strategy and evaluation plan" .

In this paragraph - starting from the description of the architectural levels of the AI4HEALTHSEC platform defined in the DoA and from the pilot definition in chapter 2 - it is possible to characterize the possible levels of integration expected with the AI4HEALTHSEC platform.This paragraph is just addressing the level of interaction and interoperability expected and that could be foreseen between the Healthcare IT System (HCIIs) part of the pilot and the component of the AI4Healthsec level. The integration and interoperability level can be both technological and organizative.

AI4HEALTHSEC Architecture is based on the concept of the following four **horizional layers** – shortly described in the following - that namely address the Risk Assessment Process (RAP) and the Incident Handling Process (IHP).

The real level of integration of the pilot Healthcare IT systems and AI4HEALTHSEC platform has to be defined, confirmed and/or modified after the technical development planned in WP3, WP4 and WP5.

AI4HEALTHSEC Horizional Layer	**Pilot HC Systems expected integration**
AI4HEALTHSEC Horizontal Layer 1 (HL1) – Risk and Privacy management & Cyber-Attack Forecasting The HL1 is the "**Risk and Privacy management & Cyber-Attack Forecasting**". It is planned to have "Assets Management" and "Infrastructure Mapping". This allows the creation of an IT asset inventory of all computing and networking related devices owned, managed, or otherwise used by the Healthcare operators. Additionally, it includes the the specification of the main interrelations and interconnections that exist between the cyber-assets and provides a visual representation of the entire infrastructure.	Organizative Integration in the pilot, by using specific and dedicated tools that will be part of the AI4HEALTHSEC platform, and can be accessibile to the pilot users

The HL1 also comprises the **Threat Assessment service** which consists mainly of three main capabilities; namely: **Vulnerability Management, Threat Management and Control Management.** The Vulnerability Management aims to feed organizations with information concerning the identified vulnerabilities of their underlined cyber assets. The Threat and Control Management provides a dictionary of known threats as well as the corresponding mitigation used to respond to these threats. It incorporates threat intelligence information providing meaningful information to operators.	
AI4HEALTHSEC - Horizontal Layer 2 (HL2) – Incident Identification The second horizontal layer is the "Incident Identification". This layer aims to detect and assess possible security incidents at existing assets of the HCIIs and has two main services which are **Data Sensing and Data Fusion**. Data Sensing service includes three different capabilities namely Source Definition, Low-level Data Monitoring and Data Pre-processing. Similarly, the Data Fusion service includes also three capabilities which are the Evidence Management, the Evidence Chain Generation and the Incident Modelling. On the Data Sensing AI4HEALTHSEC platform will collect raw data (which is a combination of passive and active means) from various sources (such as active vulnerabilities in the infrastructure; misuse detection; availability signals; network usage and bandwidth monitoring; industry proprietary protocol anomalies etc.).	The level of the integration of the HL2 AI4HEALTHSEC layers related to Data Sensing could be through dedicated API. After the definition and implementation of the architecture, in the Pilot evaluation plan and implementation it will be analysed the possibility to integrate the HL2 Data Sensing component with log monitoring systems exisiting in the pilot HC IT systems (HCIIs).
AI4HEALTHSEC Horizontal Layer 3 (HL3) – Security Events Evaluation The third horizontal layer "**Security Events Evaluation**" consisting of two main services which are Anomaly Detection and Anomaly Analysis. The Anomaly detection includes Incident Analysis and Knowledge Investigation while the Anomaly Analysis consisting of Knowledge Sharing and Attack Analysis.	Organizative Integration in the pilot, by using specific and dedicated tools that will be part of the AI4HEALTHSEC platform, and can be accessibile to the pilot users

This layer aims to provides anomalies identification functions that provide effective and efficient identification of possible security incidents such as threats, risks and faults at existing assets of the HCIIs, in order to then support the incident analysis, which includes the scrutiny of the attacker's actions and identification of his employed means.	
AI4HEALTHSEC Horizontal Layer 4 – Analysis and Decision-Making The fourth horizontal layer "Analysis and Decision-Making" has three main services namely Data Evaluation, Simulation and Response. The main responsibility of this layer is to handle incidents and specific rules of engagement and guide the HCIIs stakeholders to further investigate and analyse their occurred security events.	The level of the integration of the HL4 AI4HEALTHSEC layers expected in the pilot is both of organizative type -by using specific and dedicated tools that will be part of the AI4HEALTHSEC platform - as well it could be technological in definition of interoperability with existing business analysis tools existing in the organization

Table 11: Integration of AI4HEALTHSEC Layers

3 Business Needs and User Requirement Validation in the Pilots

As results of the AI4HEALTHSEC task T2.1 , described in the project document deliverable D2.1 "D2.1 – AI4HealthSec Requirements and Research Directives " the elicitation of requirements of the platform has been 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, questionnaires have been created to be fulfilled both by internal project partners and external organizations from further critical infrastructures (besides healthcare, e.g. financial sector, transportation sector).

The analysis of the questionnaires (see D2.1 for the questionnaire definition and contents) was condensed in the following lists of business needs, technical challenges and functional requirements, that evolve from the user's perspective (business needs).

D2.1 extracted a list of six main business needs (**Errore. L'origine riferimento non è stata trovata.**). Those needs depict challenges that need to be faced when creating an AI4HealthSec framework. The numbering of these business needs challenges does not represent any weighting of them, since these business needs are equally important to be taken into account.

Business Need ID	Title	Description
BN1.	Prediction and Prevention of Attacks	The organization needs to forecast and prevent cyber-attacks.
BN2.	Vulnerability Assessment	The organization needs a framework to assess its cyber-security weaknesses.
BN3.	Awareness Creation and Prevention of Human Errors	The 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	The organization needs a system to automatically detect abnormal patterns in my IT and create warnings.
BN5.	Simplification of the Process of Risk Assessment	The organization needs a simpler process of risk assessment.
BN6.	Development of Long-Term Strategy of New Protection Solutions.	The organization needs a long-term and comprehensive cyber-security strategy.

Table 12: AI4HEALTHSEC – Business Needs

To meet the user challenges and business needs coming from the questionnaires, a set of narrowed technical challenges and requirements –have been defined in D2.1 as input for the AI4HEALTHSEC platform design and implementation. The following table identifies the different classes of User Requirements that have been defined.

The detailed list of URs is defined in D2.1 and it is not repeated in this deliverable. Some of those User Requirements have also been defined by end-user and stakeholders of the different partners organizations that support the execution of the pilots.

Technical Challenge ID	Description	Relevance to Business Needs
TC1.	**Evidence-based, Swarm-driven Risk Management and Assessment Methodology**	• BN1: Prediction and Prevention of Attacks • BN2: Vulnerability Assessment

	Include requirements of: ***Requirements for Risk Management Context and Compliance*** ***Requirements for Risk Identification and Predication*** ***Requirements for Risk Assessment and Modelling*** ***Requirements for Risk Management and Control*** ***Requirements for Incident Management*** ***Requirements for Contribution to other Domains***	• 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** *Include requirements of:* ***Requirements for Multi-source Evidence Collection and Preparation*** ***Requirements for Evidence chain Generation and Security Incident Detection*** ***Requirements for Incident Management and Response***	• 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.

Table 13: AI4HEALTHSEC – Technical challenges

In the execution and implementation plan of the pilots that will be defined in 6.1, the User Requirements that will be under validation within each specific pilot will be identified within the list of specific User Requirements of D2.1.

The validation and application of the individual requirements will be defined in terms of:
- Functionality (technical and organizative) expected in the requirement
- Presence or not of functionalities
- Levels of usability of the implementation of the requirement (if applicable)
- Performance of the requirement

- Possible Note and Remarks

In addition to the analysis and validation of the individual requirements, the benefits and the overall impact that the adoption of the platform and the compliance with the individual requirements bring to the HC organization and in particular to the different stakeholders involved in the execution of the pilots, are then defined through the definition of specific KPIs and evaluation metrics defined in the next paragraph.

4 Qualitative and quantitative KPIs

To assess the impact of the AI4HEALTHSEC system on the pilot use cases, a specific metric has to be defined and specified in the form of Key Performance Indicators (KPIs).

Analysis of KPIs, key risk indicators (KRIs) and security postures provides a snapshot of how a security framework and a security organizative team and policies are functioning over time and helps the stakeholders and a CISO (Chief Information Security Officier) to better understand what is working and what is worsening, improving decision making about future security projects.

In this paragraph, a set of quantitative and qualitative KPIs are listed and will be the object of definition and evaluation during the execution of each single pilot to evaluate the effectiveness of the application and the results provided by the adoption of the AI4HEALTHSEC platform.

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 (HCIIs)
- 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

The project will develop innovative ways to leverage collected security and privacy information, enabling stakeholders to evaluate the risk and invest to limit that risk in an optimal way.

The adoption of the platform shoud guarantee the following objectives to be evaluated:

(a) evaluation of risk and privacy risk;

(b) identification of propagated vulnerabilities located in interconnected infrastructures;

(c) estimation of the cascading effects of threats or detected events;

(d) detection of security incidents;

(e) uncovering evidence of malicious activities;

(f) extraction and collection of data of particular interest;

(g) analysing and correlating relationships between all recovered forensic artefacts;

(h) anticipation of where an attack is heading;

(i) provision of recommendations, advices and directions on the further investigation of the security incident; (j) proposal of a mitigation/containment strategy.

Metrics provide quantitative information that can be used to show management and board members that the HC Organization is managing the protection and integrity of sensitive information and information technology assets seriously.

4.1 KPI Definition vs Business Needs

With reference to the previous list of Business Needs that are expected to be satisfied by the platform, for each individual pilot the positive impact assessment is defined for each BN through one or more KPIs, as defined in the following table.

Business Need ID	Title	KPI
BN1.	Prediction and Prevention of Attacks	N° of threat considered and managed , as for example: - Intrusion attempts - Unauthorized access - Data Breach access
BN2.	Vulnerability Assessment	No. of critical assets identified in the HC Organization N° of Qualified Risk for each Threat
BN3.	Awareness Creation and Prevention of Human Errors	No of end-users participation in training session No of end-user participation in security awareness focused workshop
BN4.	Detection of Abnormal Patterns and Creation of Warnings *The organization needs a system to automatically detect abnormal patterns in my IT and create warnings.*	No of Security Control considered and implemented MTTD (Meantime to Detect) MTTR – MTTC (Meantime to Resolve and Contain) Identity and Access Management: - N° di tentative di Unauthorized access

BN6.	Development of Long-Term Strategy of New Protection Solutions. *The organization needs a long-term and comprehensive cyber-security strategy.*	No of differents attack defined for the scenario No of system downtime due to Long Term Prevention Strategy No of devices in the network that are fully patched and up to date Identity and Access Management: • Time taken to deactivate credentials • Number of users with excessive entitlements • Third party (Suppliers/Vendors/contractors) access review: # third party unnecessary access removed • Percentage of employees with Privileged access who are monitored • Frequency of review of third party accesses

Table 14: KPI vs Business Needs

In the design and definition of the pilot implementation plan, activity to be done in WP6 and Task 6.1, for each specific pilot the KPIs will be specified in detail and defined in terms of:

- Minimum threshold value expected to reach the KPI
- Expected target value

4.2 Quantitative KPI Description

Intrusion attempts and Security Incidents

Number of times have bad actors attempted to gain unauthorized access.

Number of intrusion that have been detected and number of that have been blocked

Security incidents

Number of time an attacker breached the information assets or networks **(**SIEM/AV/Malware etc.)

Mean Time to Detect (MTTD)

How long do security threats go unnoticed? MTTD measures how long it takes your team to become aware of indicators of compromise and other security threats.

Mean Time To Identify (MTTI), also known as Mean Time To Detect (MTTD) measures how long it takes to detect a breach. To calculate this KPI, count the days, or fraction of days between the beginning of a system outage, service malfunction or other security issues and when the someone identifies the issue. The breach can be detected by the I.T. team, DevOps team or by an external source, in fact, 53% of breaches are reported discovered by an external source.

At large scale, this can be calculated by taking the sum of all the time incident detection times for a given technician or team and divide that by the number of incidents. In this calculation, it may be wise to remove outliers, such as catastrophic errors, to show a true average.

Mean Time to Resolve (MTTR) – Dwell time

The mean response time for your team to respond to a cyber attack once they are aware of it. A great measure of the quality of your incident response plan implementation.

Mean Time to Contain (MTTC)

The time taken to close identified attack vectors. Poor performance in MTTI and MTTC is a huge contributor to breach costs. It's also a good KPI for CISOs to measure and show their Board for long-term improvement. Everyone on the security team should prioritize improving these two KPIs.

Systems Uptime/Downtime

Uptime and downtime simply refer to how often a site and/or an IT System or a service within an IT System is working (uptime) or not (downtime). This is traditionally reflected in a percentage and the two should equal 100%. For example, 97% uptime means a website was working for 97% of the given time (usually a month) and a software had a 3% downtime for either software updates or because of an attack.

Every time the organization has to take something offline to patch security, the HC organization is taking away an important tool for your organization or for your client. Tracking downtime due to security concerns can also help make the case for additional measures when it comes to budget time. In addition to hard costs, there's lost productivity and potentially lost revenue.

Identity and Access management

How many users have administrative privileges? Access control and the principle of least privilege are simple, cost effective methods of reducing privilege escalation attacks.

Best practices in information security management include full control of users' level of access to company resources, it is necessary for an employee to only access data, systems, and assets that are necessary to their work. Identifying the access levels of all network users allows the HC Organization to adjust them as needed by blocking any super user or administrator that does not make sense.

The following KPI and metrics could be used to evaluate the security in identity and access management:

- Time taken to deactivate credentials
- Entitlements review: Number of users with excessive entitlements
- Third party (Suppliers/Vendors/contractors) access review: # third party unnecessary access removed
- Percentage of employees with Privileged access who are monitored: monitoring users that have 'keys to the kingdom' (super-users) provides insight to determine if too many individuals have unlimited network access and restrict access to those who absolutely need it.
- Frequency of review of third party accesses

  Often, IT managers grant access to third parties in their networks to complete a project or activity. It is important to monitor whether the access is canceled at the end of service provisioning. Failure to do so endangers your environment if the third party decides to come back and extract data or carry out other malicious activity – for instance, they may come under the employ of a competitor. Possibly worse, if the 3rd party's network is breached, you could expose your network to the same threat.

Configuration Management:

- % Servers and devices compliant to hardening standards —configuration drift is a risk as IT environments undergo changes, with the widespread adoption of Dev-ops, changes could occur many times daily.
- Firewall/switch audit results

Vulnerability and Patching:

- Number of systems with known critical and high vulnerabilities: While reporting on all systems is the norm, it is preferred to have management reports that focus on the high risk systems and applications (crown jewels)
- Patch levels of High risk systems with known critical and high vulnerabilities: This gives an indication how effective the patching cadence is

- Number of systems with critical and high vulnerabilities that vendors have not released patches yet — alternate mitigation measures applied or if no mitigation is possible, accept risk
- Time taken for vendors to release patches
- Days to roll out patches from vendor release

4.3 Qualitative and organizative KPI Description

Cost Per Incident

Costs go well beyond the technical aspects. Lost revenue, lost company reputation, public notices, employee time, and indirect costs add up quickly.

To truly track the Cost Per Incident of an HC organization, it should be necessarily to correctly bring in all resources, both human and technical, that were required to find the thread and fix it. This should also include missed revenue in terms of actual loss and potential loss. To calculate this KPI, add in three specific categories. Direct costs, such as actual forensic and investigative costs should be added to indirect costs like recovery time and costs to communicate the breach out. Finally, add in the lost opportunity.

Level of preparedness

How many devices on your network are fully patched and up to date?

Unidentified devices on internal networks

Employees can introduce malware and other cyber risks when they bring in their own devices, as can poorly configured Internet of Things (IoT) devices, which is why network intrusion detection systems are an important part of the organization's security.

Security Awareness:

- Training compliance levels- % Completed
- Results of phishing and other social engineering tests on staff: % Failed

Patching cadence

How long does it take the organization to implement security patches or mitigate high risk CVE-listed vulnerabilities. Cybercriminals often use threat intelligence tools and exploit the lag between patch releases and implementation. A great example of this is the widespread success of WannaCry, a ransomware computer worm. While WannaCry exploited a zero-day vulnerability called EternalBlue, it was quickly patched but many organization fell victim anyway due to poor patching cadence.

5 Conclusions

The scenario description of each use case pilots of the AI4HEALTHSEC project has been defined addressing for each of them the following topics: a) description of the current real-life scenario with the architecture of the HCIIs involved b) the use case and the type of end-users that are involved in the scenario c) the potential attack scenario, describing the actions and the potential outcomes of an attack as well the asset and the IT systems involved b) the critical data managed in each scenario and for which integrity, confidentiality and availability must be guaranteed.

A first draft idea of the integration level of the AI4HEALTHSEC platform within the pilot systems has been described and envisaged

In order to assess the effectiveness of the platform, a set of User Requirement, part of the definition of D2.1, have to be selected and evaluated during the pilot execution. Anevaluation methodology has been defined as well a metric and a list of KPIs that will be used to evaluate how the AI4HEALTHSEC platform will improve the performance of a security system within the organization.

The outcomes contained in this deliverable provide input for WP6 “Pilots development of the AI4HealthSec system”, where the activities of pilot design, implementation and run will be executed, as well the functional requirement and KPI defined and evaluated.