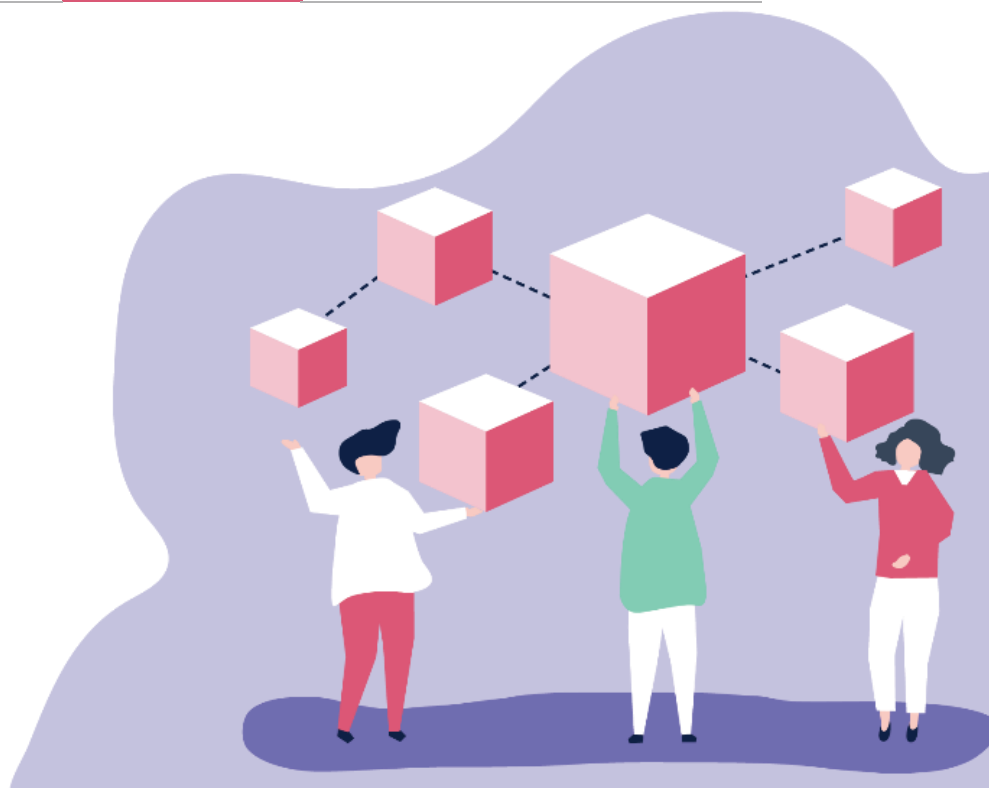




D5.11 Intelligent Connected Care Services and IoT

Deliverable No.	D5.11 (D5.4.2)	Due Date	31/12/2023
Description	This deliverable describes the communication pattern for the integration into the GATEKEEPER platform and their reference implementations.		
Type	Other	Dissemination Level	PU
Work Package No.	WP5	Work Package Title	Integrated Plug & Play GATEKEEPER Dynamic Intervention services
Version	1.0	Status	Final



Authors

Name and surname	Partner name	e-mail
Eugenio Gaeta	UPM	eugenio.gaeta@lst.tfo.upm.es
Manuel Stocker	MEDISANTE	manuel.stocker@medisante-group.com
Mor HersHKovitz	BioBeat	mor@bio-beat.cloud
Albert Pages	SENSE4CARE	albert.pages@sense4care.com
Pilar Sala	MYSphera	psala@mysphera.com

History

Date	Version	Change
25/04/2023	0.1	Table of content
13/07/2023	0.2	MEDISANTE Contribution
20/09/2023	0.3	BioBeat Contribution
22/11/2023	0.4	SENSE4CARE Contribution
15/12/2023	0.5	MYSphera Contribution
19/12/2023	0.6	UPM.
02/01/2024	1.0	Final version Ready for submission

Key data

Keywords	IoT, Intelligent services, Connectors, API, Batch processing
Lead Editor	Eugenio Gaeta (UPM)
Internal Reviewer(s)	Ioanna Drympeta (CERTH), Alessio Antonini (OU)

Abstract

This deliverable reports on the progress of the architecture, software and hardware development of T5.4 with the main goal of describing the communication patterns that have been developed and integrated into the GATEKEEPER platform.

Statement of originality

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Table of contents

TABLE OF CONTENTS.....	4
LIST OF TABLES	5
LIST OF FIGURES	6
1 EXECUTIVE SUMMARY	7
2 INTRODUCTION.....	8
2.1 STRUCTURE OF THE REPORT	8
2.2 CONTEXT OF THE DOCUMENT	9
3 DEVICE CONNECTIVITY WITHIN GATEKEEPER.....	10
3.1 GATEKEEPER CONTAINER PLATFORM, GATEKEEPER SERVICE PLATFORM AND GATEKEEPER BIG DATA PLATFORM	10
3.2 GATEKEEPER SECURE PROCESSING ENVIRONMENTS.....	17
3.2.1 <i>Generic principles for secure processing environments requirements</i>	17
3.2.2 <i>GATEKEEPER secure processing environment principles</i>	19
3.2.3 <i>GATEKEEPER responsible data management findings and experience</i>	22
3.2.4 <i>GATEKEEPER SPE for responsible AI</i>	23
3.3 GATEKEEPER RECOMMENDATIONS FOR EHDS.....	24
3.3.1 <i>Secure image building pipeline for SPEs</i>	24
3.3.2 <i>FHIR based interoperability</i>	27
3.3.3 <i>Secure processing environments for primary and secondary use of data</i>	29
3.3.4 <i>Exchange of AI model and services</i>	32
3.3.5 <i>Proposal for techno-legal specification for generic EHDS secure processing environments</i>	34
4 INTELLIGENT CONNECTORS.....	36
4.1 EXAMPLE USE CASE	36
4.1.1 <i>Update of Intelligent Connected Care Service in Puglia</i>	36
5 IOT WEB CONNECTORS	42
5.1 EXAMPLE USE CASE	42
5.1.1 <i>Update of Web Connector implementation</i>	42
5.1.2 <i>Update of Web Connector implementation</i>	43
5.1.3 <i>Connectors for MAHA and Data Federation</i>	45
6 GATEKEEPER IOT GATEWAY	48
6.1 DEVELOPMENT OF IOT GATEWAY	48
6.2 GATEKEEPER IOT GATEWAY REFERENCE DESIGN	49
6.3 GATEKEEPER IOT GATEWAY PROTOTYPE PRODUCTION	51
6.3.1 <i>Functional description</i>	52
6.3.2 <i>Final product</i>	54
7 CONCLUSIONS	56
APPENDIX A MAHA DATA MODEL	57

List of tables

TABLE 1: GANTT CHART OF THE DEVELOPMENT PLAN OF THE GATEKEEPER IOT GATEWAY	48
TABLE 2: ESTIMATED DELIVERY TIMES OF GATEKEEPER IOT GATEWAY COMPONENTS	49

List of figures

FIGURE 1 – GATEKEEPER PLATFORM	10
FIGURE 2 – SOFTWARE LAYERS OF GATEKEEPER CONTAINER PLATFORM.....	11
FIGURE 3 – GATEKEEPER SERVICE PLATFORM LOGICAL VIEW	14
FIGURE 4 – GATEKEEPER SERVICE PLATFORM COMPONENT DEPLOYMENT MODEL	15
FIGURE 5 – TMS DEPLOYMENT MODEL ON GATEKEEPER CONTAINER PLATFORM	15
FIGURE 6 –GATEKEEPER TENANCY MODEL OF GATEKEEPER CONTAINER PLATFORM	16
FIGURE 7 –OVERALL GATEKEEPER DEPLOYMENT MODEL.....	17
FIGURE 8 – DevSecOps GATEKEEPER APPROACH AS RECOMMENDATIONS FOR EHDS	27
FIGURE 9 – FHIR BASED INTEROPERABILITY APPROACH AS RECOMMENDATIONS FOR EHDS.....	29
FIGURE 10 – GATEKEEPER INTEGRATION FOR PRIMARY AND SECONDARY USE OF DATA.	32
FIGURE 11 – GATEKEEPER AI MODELS AND SERVICES MAPPER TO FHIR OPERATIONS.	33
FIGURE 12 –INTEGRATED HARDWARE CONNECTED TO GATEKEEPER DF	37
FIGURE 13 –ICCS MEASUREMENTS CONNECTED TO GATEKEEPER DF	38
FIGURE 14 – VERIFICATION EXAMPLE PERFORMED IN SWITZERLAND AND GERMANY	39
FIGURE 15 – ARCHITECTURE OF EXTRACTION AND TRANSFORMATION OF QUESTIONNAIRE	43
FIGURE 16 – WEB CONNECTOR FLOW RUC4 PARKINSON BASQUE COUNTRY	44
FIGURE 17 – MAHA CONNECTORS FOR DATA INTEGRATION INTO DATA FEDERATION	45
FIGURE 18: SET OF SCHEMATICS OF THE GATEKEEPER IoT GATEWAY	51
FIGURE 19: GATEKEEPER IoT GATEWAY TOP VIEW	51
FIGURE 20: GATEKEEPER IoT GATEWAY BOTTOM VIEW	52
FIGURE 21: GATEKEEPER IoT GATEWAY FINAL PROTOTYPE TOP AND BOTTOM VIEW.....	54
FIGURE 21: GATEKEEPER IoT GATEWAY FINAL PROTOTYPE WITH 3D PRINTED CASE.....	55

1 Executive summary

The document provides detailed insights into the architecture, implementation, and testing of connectors, emphasizing their role in ensuring interoperability, data integrity, and efficient communication. Connectors are crucial components that contribute to the success of the GATEKEEPER project, enabling the integration of diverse medical devices and fostering a connected healthcare ecosystem. It also highlights the critical role of connectors in facilitating seamless data integration and communication within the healthcare ecosystem. Various connectors, including IoT web connectors and specific device connectors, play a key role in data federation and technical pilot deployments.

Intelligent connectors including CAT-M weight scales and blood pressure monitors, are integrated, showcasing a commitment to providing a comprehensive suite of options for care teams. The transition from Cat-1 to Cat-M technology is emphasized for enhanced accessibility by reducing cellular device costs. Successful integrations and validations in various regions underscore the scalability and adaptability of the developed technologies, with a notable focus on Medisanté's role in adapting architecture and security measures. Collaborations with partners like Beurer and integrations into platforms such as Oviva and EPIC demonstrate the project's global impact.

IoT Web Connectors Developed in multiple project tasks, serve as essential components for integrating data within pilot environments. For instance, BioBeat's implementation involves a transformer facilitating the extraction and transformation of questionnaire data from the Samsung Platform into the GATEKEEPER FHIR Data Federation. This structured approach ensures efficient data handling and analysis. BioBeat develops connectors to pull data from its AWS data repository and transforms proprietary measurements data into FHIR patient bundles using LOINC code system. The connectors create secure connections to the Gatekeeper FHIR server, emphasizing the importance of standardized formats for efficient data exchange. Sense4Care utilizes connectors to gather vital parameters from the StatOn Parkinson's device for precise data analysis on the Gatekeeper server. The data flow involves synchronization, transformation into JSON format, and conversion to FHIR specifications, showcasing the adaptability of connectors in handling diverse data sources.

The document also describes the GATEKEEPER IoT Physical Gateway that is a crucial element in the project, enabling seamless integration of diverse medical devices with the GATEKEEPER data federation. Powered by a dual-core ARM Cortex-A7 and Cortex M4 processor, the gateway offers key features like Ethernet connectivity, USB hubs, mSATA support for SSD storage, and interfaces for wireless connections via modems and WiFi/BT. Its adaptability includes support for various SSD models, functioning effectively with or without SSDs. Robust security features such as tamper detection and secure modem connections ensure the protection of healthcare data. The prototype's production, culminating in a 3D printed case, signifies the IoT Physical Gateway's readiness for extensive deployment. Ultimately, the success of the GATEKEEPER project is intricately linked to the gateway's robust design and versatile capabilities, enhancing data flow and integration across the healthcare ecosystem.

The document anticipates a transformative shift in remote patient monitoring over the next five years. GATEKEEPER is positioned as a catalyst for this change, contributing not only to technological innovation but also fostering a shift towards European digital health infrastructure. The success and integration of GATEKEEPER's technologies into existing healthcare ecosystems significantly contribute to EHDS goals and the advancement of patient-centric, digital healthcare models.

2 Introduction

2.1 Structure of the report

The document provides an additional comprehensive overview of the GATEKEEPER platform architecture with a focus on T5.4, detailing various aspects of the project's development and implementation. The structure includes multiple parts covering different themes and it also includes some recommendations for European Health Data Spaces (EHDS).

The initial sections introduce the project, its objectives, and the role of T5.4 in addressing healthcare challenges and data integration.

Subsequent parts delve into the integration of OLTP (Online Transaction Processing) and OLAP (Online Analytical Processing) systems in healthcare data analysis, emphasizing the significance of real-time decision support and historical analysis.

The document then explores challenges associated with resource-intensive OLAP systems and presents solutions through the strategic integration of container-based data science applications, particularly Docker and Kubernetes. This integration is highlighted as crucial for scalability, flexibility, and cost optimization in healthcare analytics.

A section is dedicated to the validation of innovative services, particularly a Bring-Your-Own-Device (BYOD) approach for connectivity in large-scale deployment. The integration of new medical devices, such as CAT-M weight scale and blood pressure monitoring, is discussed, addressing feedback from pilot programs and offering more choices to care teams.

Furthermore, the document details the implementation of web connectors, particularly focusing on BioBeat's approach for updating and implementing web connectors. The process involves token acquisition, data retrieval from the Samsung Platform, transformation into FHIR format, and subsequent submission to the Gatekeeper FHIR Data Federation.

The report also covers implementations in specific regions, such as Basque Country, showcasing the integration of data from devices like StatOn Parkinson's, involving JSON transformation and the use of FHIR connector. Additionally, connectors for MAHA data integration are explained, describing the process of migrating custom data into the GATEKEEPER Data Federation.

The concluding sections detail the development of the GATEKEEPER IoT Gateway, providing a Gantt chart outlining the development plan. The document concludes with insights into the MAHA data model, breaking down key components related to patient/practitioner information, social assessment, habits, clinical activities, prescribed medication, clinical variables, symptoms, forms/questionnaires, and comorbidity.

Overall, the report follows a structured approach, systematically addressing various facets of the GATEKEEPER project, from data processing and integration to the development of key components like the IoT Gateway and the MAHA data model.

2.2 Context of the document

The document is situated within the broader context of the GATEKEEPER project, specifically focusing on Work Package 5 (WP5). This work package is dedicated to addressing crucial aspects of healthcare data processing and integration within the GATEKEEPER initiative. The overall context involves the development of innovative solutions to healthcare challenges, particularly in the realm of remote patient monitoring and data analytics in the GATEKEEPER project's overarching goal of revolutionizing healthcare through innovative data processing and integration solutions. It covers a spectrum of topics ranging from technological implementations to regional case studies, contributing to the broader context of advancements in healthcare technology and patient care.

3 Device connectivity within GATEKEEPER

3.1 GATEKEEPER container platform, GATEKEEPER service platform and GATEKEEPER big data platform

The GATEKEEPER technical innovation revolves around three primary concepts: the GATEKEEPER container platform, the GATEKEEPER service platform, and the Gatekeeper big data platform.

The GATEKEEPER container platform operates as a software layer on HPE-provided hardware within a data center. This platform facilitates the creation and management of virtual nodes utilizing Docker, Kubernetes, and OpenShift technologies. It offers features such as resource management, multi-tenancy, data lake connection, cluster management, and security services.

Built upon the GATEKEEPER container platform, the GATEKEEPER service platform provides a suite of core services emphasizing connectivity (connectors), interoperability (data federation), and scalability (marketplace and GTA). Essentially, the GATEKEEPER service platform consists of services and tools deployed within the GATEKEEPER container platform hosted by HPE.

Similarly, the big data platform, also constructed atop the container platform, encompasses end-user services tailored for the manipulation and processing of vast datasets. It interacts with the service platform through integrated data federation, allowing direct data analysis performed on FHIR.

To ensure segregation, the service and big data platforms are replicated across different tenants, each operating independently of the others. While, under a legally binding agreement, a specific data scientist may have access to multiple tenants, generally, this behavior is restricted, and users belonging to one tenant do not have access to others. Figure 1 shows the tenancy model and how containers, services and big-data tools are related.

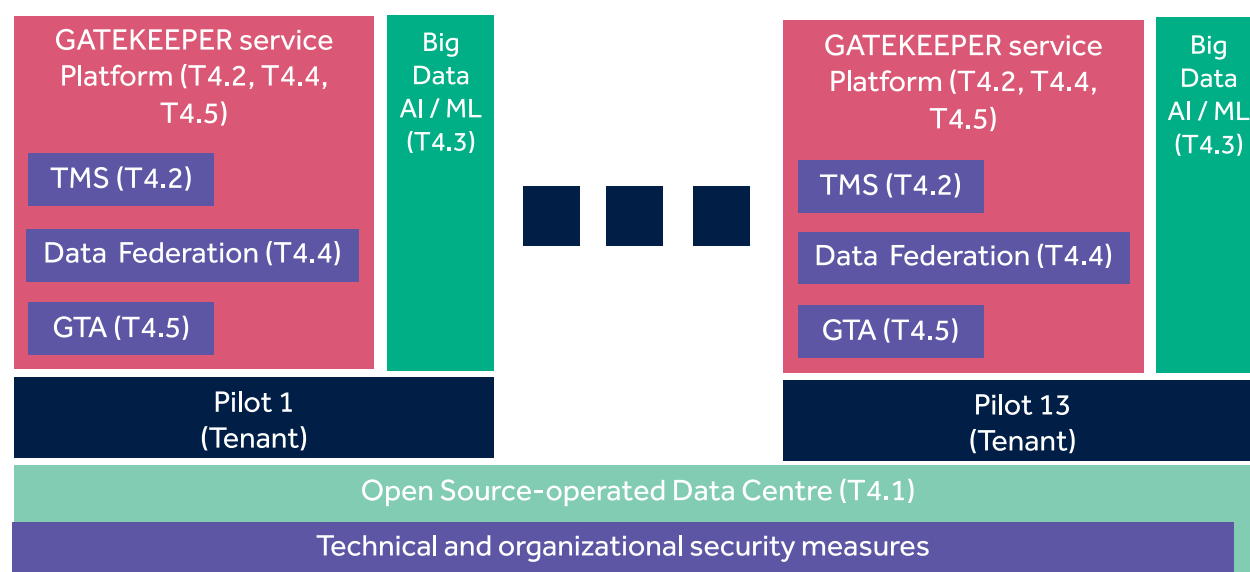


Figure 1 – GATEKEEPER Platform

GATEKEEPER Container Platform can automatically adjust computational resources and server instances according to incoming workloads. This is achieved by relying on open source Container Platform (Figure 2), which enables infrastructure as a code approach. Furthermore, the GATEKEEPER Container platform is infrastructure agnostic. Thanks to its licensing and usage of open-source technologies, it can be instantiated in any hardware compliant environment, including private, public or hybrid clouds. The following figure represents the software layers of the Container Platform:

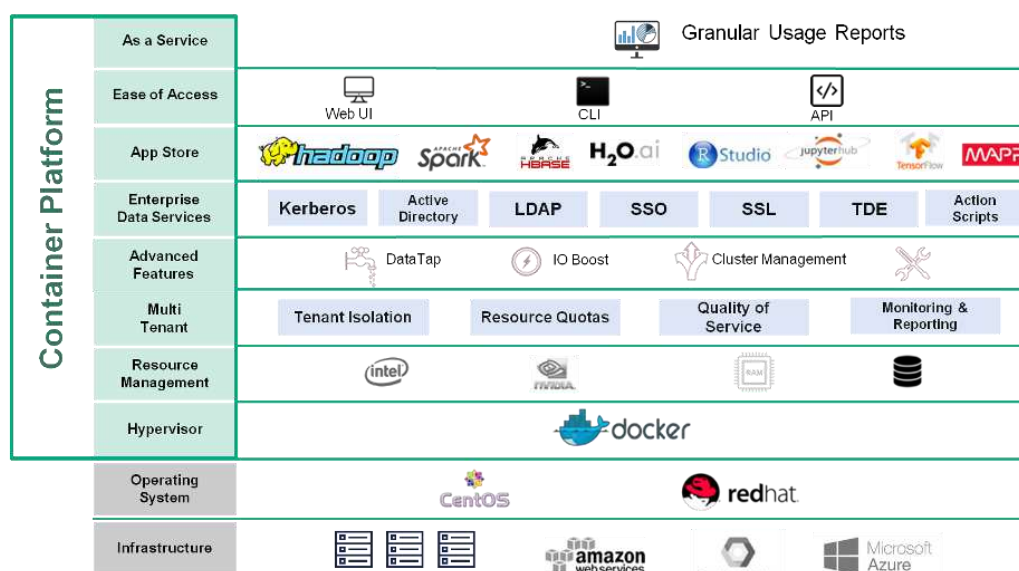


Figure 2 – Software layers of GATEKEEPER Container Platform

The bottom layer represents the infrastructure where the container platform is installed and it is based on HPE-provided physical hardware and data center, it is made of virtual nodes running on OKD, the Community Distribution of Kubernetes (in principle the infrastructure nodes could include bare metal machines, or cloud instances, or both, thus allowing for hybrid deployments). The operating system used is CentOS/Fedora CoreOS (selected since they are open source). Internally, the HPE platform uses Docker as core containerization technology and Kubernetes for orchestration. On top of the containerization layer, there are resource management tools, responsible for administering physical resources of nodes (CPU, RAM, storage) and obtain a complete abstraction of the underlying infrastructure, which is presented as a homogenous resources pool to the overlying layers.

The multi tenancy level is responsible for segregating the resources pool into tenants. Tenants can be defined at whatever level: in GATEKEEPER, tenants are GATEKEEPER platform instances and can be created for each partner, or each pilot, or whatever other entity. The system administrator can assign Resource Quotas to tenants in order to limit their resource consumption (in terms of CPU cores, storage and memory) with respect to the total resources availability.

On top of this, the container platform also offers other high-level functionalities: the capability to connect to existing data lakes without moving data, and cluster management functions allowing users to create new clusters, stop them, expand/contract them depending on their needs, or create auto-sizing rules.

The Enterprise Data Services layer includes security services applied at container level: authentication, authorization, encryption, TDE, TLS, and SSL.

For each GATEKEEPER platform instance, the Applications catalogue of the container platform offers a number of reference Kubernetes applications available to users: they can be used as they are, or they can be customized for specific requirements.

GATEKEEPER service platform, is one of consumer of the Container Platform and can be accessed through a rich web user interface environment that includes a Developer Portal and a Marketplace. The Developer Portal is the fundamental element enabling the community building and ecosystem networking effect around the GATEKEEPER open source project. The Developer Portal allows developers to access GATEKEEPER components and combine them with third party services to create new services and applications which can then be made available in the platform through the marketplace. The marketplace is the entry point for the final user and it provides services, data and devices in a rich user interface.

GATEKEEPER container platform and GATEKEEPER service platform do not share the same user. The GATEKEEPER container platform is a super user for GATEKEEPER service platform and is linked to the HPE (Authorization Authentication Infrastructure) while the GATEKEEPER platform users are managed by the GATEKEEPER Trust Authority (GTA) that provide a high customizable service for user management (Keycloak).

The GATEKEEPER service Platform is managed by the Web of Things Thing Management System (TMS), the central gateway that orchestrates all the interactions to, from and within the Platform. All components that are registered in the TMS have a W3C Thing representation that holds information about its functionalities and security features. The certification of the Things is the responsibility of the GATEKEEPER Trust Authority (GTA), which also offers user management, and manages authentication, authorization and auditing.

Through the GATEKEEPER Web Environment, developers can register new things or obtain access to the existing ones through the Developer Portal, while Business users and administrators can use the Marketplace to browse the available things or manage the platform, respectively. All these users are managed by the GTA and there are not the same users of the GATEKEEPER container platform.

Data sources produce data that enters the platform through connectors, which differ according to the type of devices and protocols they interact with:

- Gateway BLE / FHIR for Bluetooth devices compliant to BLE Continua Standard Health Profiles;
- Web Data Connectors for accessing online data files or Web based APIs from device vendors;
- Intelligent Medical Device Connectors to manage fully connected devices;
- Multi Robot Connectors to access data from robot sensors;
- Personal Health GATEKEEPER App, an extension of Samsung Health App that allows the GATEKEEPER Platform to interact with the Samsung Health ecosystem, and more.

In order to inject the data into GATEKEEPER a connector need the VPN access with a user account of the GATEKEEPER container platform. For security reasons and accountability purposes the GTA users are not allowed to write or delete data into the Data Federation.

All connectors registered to GATEKEEPER Platform send health data to the Data Federation Integration Engine in the original producer format. It is the responsibility of this component to translate it to the common GATEKEEPER data format, that is the GATEKEEPER FHIR IG, and store it in the Data Federation FHIR Server, thus ensuring semantic interoperability.

Data are also forwarded to the Big Data platform (T4.3), which offers Big Data analysis services, and can combine data coming from the Data Federation and potentially external data sources.

Data stored in the Big Data infrastructure are used by the AI Reasoning Framework to realize the intelligence of the platform by means of diagnostic and prognostic algorithms for several diseases, early detection of changes in patient conditions and the risk factor evaluations, as well as personalized monitoring, prevention and intervention.

Both the output of the computations of the AI Reasoning Framework and the integrated data stored in the FHIR Server represent the added value of the GATEKEEPER Platform, and they are presented to the final users through customer's services that can interact directly with the platform, or using the easily customizable dashboards built using the Authoring Tool for Dashboards.

GATEKEEPER service Platform is capable of acquiring data from heterogeneous Data Sources through its Data Connectors. These services allow the platform to receive data decoupling the protocol used for the collection or the details on the interaction from the single vendors and communication standards and allow the rest of the platform to rely on a well-defined REST-HTTP communication. The platform offers a set of connectors that will be available to integrate a large variety of Devices:

- Web Data Connectors, to interact with Device vendors cloud services;
- Gateway BLE / FHIR, that can collect measurements from devices following the BLE Continua standard;
- Personal Health GATEKEEPER App, a customization of Samsung Health mobile app, tailored to collect data for the GATEKEEPER service Platform;

The Platform also supports the input of data from Multi Robot Connector, a connector that allows collecting data coming from Robots.

Data sources that are able to communicate directly with the platform can do so, provided they are certified in the platform – as will be described below – and their requests to the platform secured. An example of such sources is represented by the Intelligent Medical Device Connectors.

The Platform is also able to poll directly data from pre-configured data sources, such as EHRs, to retrieve related historical data.

Having a syntactic integration provided by connectors, the platform can also provide Semantic Data Integration through the Integration Engine of the GATEKEEPER Data Federation. This component is responsible for integrating and federating data coming from the connectors described above. It supports different data acquisition modalities (data can be sent explicitly exploiting the REST interface provided by this component or periodically fetched directly from the external data sources). Using semantic models, data are transformed in a unique format, the GATEKEEPER FHIR Data.

Federated Data will then be stored in the Data Federation FHIR Servers, which will keep a representation in GATEKEEPER FHIR Profile in its FHIR Server, or a RDF representation in its RDF Server.

Data integrated in the Data Federation integration engine are also pushed to the Big Data platform, where they can be further processed and merged with further external data sources.

This infrastructure will provide services to perform Data Analysis and AI/ML software and libraries that can be exploited from the AI Reasoning Framework. In this Framework we can find:

- AI Personalized Risk Detection & Assessment, that provides diagnostic and prognostic algorithms that can help both professionals to support their decisions and elderly with no technical knowledge to improve their independence and ability over the time.;
- Home and Health Activity Monitoring that can combine Personal Health Background and Environmental Measurements, mapping of daily activities and environmental threats at home, to identify and notify abnormal conditions;
- Medical Based AI Algorithms that support the whole AI reasoning framework;

VERSION 4 - LOGICAL VIEW

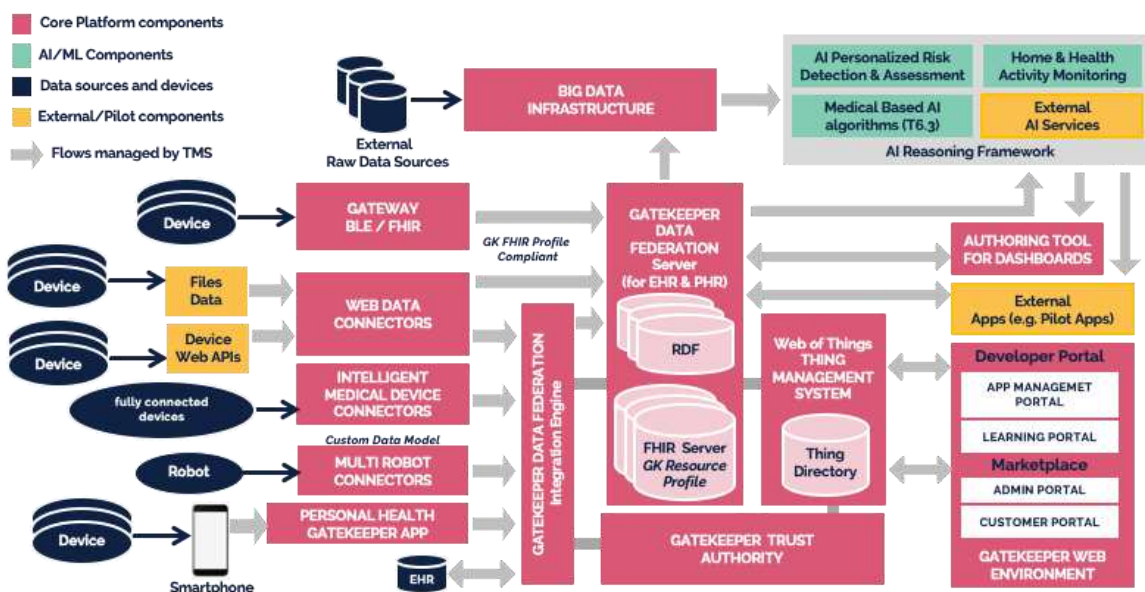


Figure 3 – GATEKEEPER service Platform logical view

In the Gatekeeper logical view are visible some high level interactions among the components. The main interaction that is visible in the figure 3 is related to the Web of Things Thing Management System and the Gatekeeper Trust authority that interact each other in order to provide access to the user of the Gatekeeper Web Environment that are managed by GTA.

The following figure 4 is presenting the component that are deployed within the GATEKEEPER infrastructure hosted on the HPE data center.

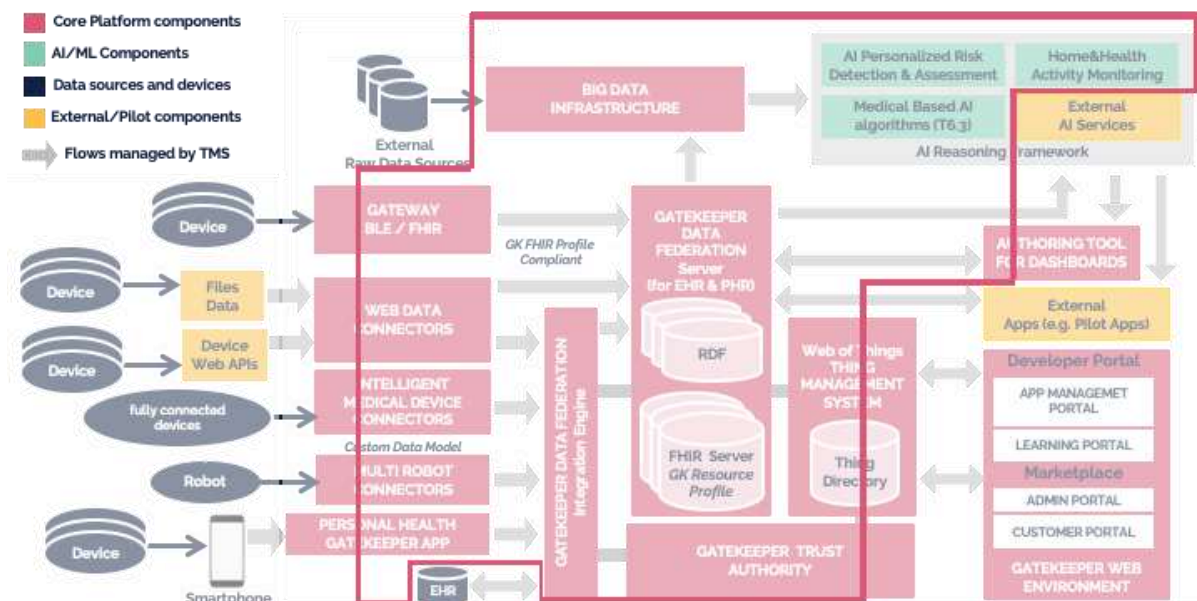


Figure 4 – GATEKEEPER service Platform component deployment model

There are some components such as TMS, GTA and connectors that are not entirely deployed in the Gatekeeper infrastructure. This is the case of all those components that need a public access without VPN connection.

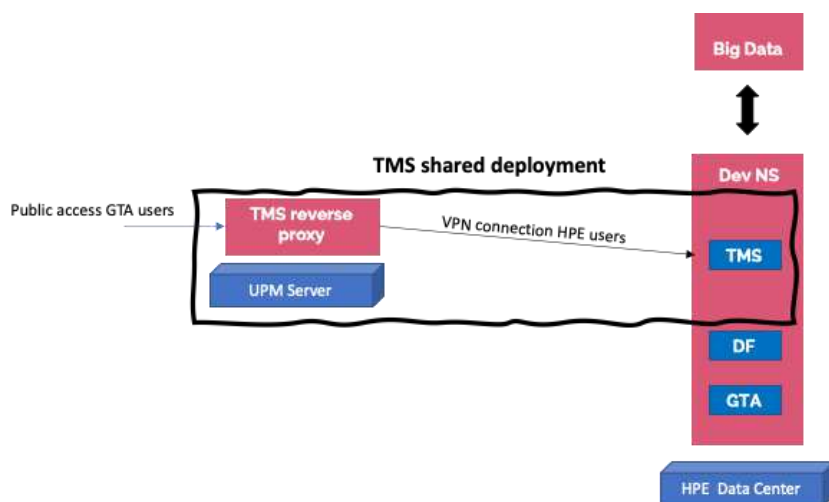


Figure 5 – TMS deployment model on GATEKEEPER container platform

In the following figure is presented the real deployment of the TMS that is shared across the HPE infrastructure and the UPM server that is hosting a reverse proxy with a VPN site to site connection to the HPE infrastructure. The same deployment model also applies to GTA, marketplace connectors and pilot apps that are used by final users.

Tenancy model for environments isolation

As already mentioned, GATEKEEPER service platform is segregated into tenants. It uses a tenancy model based on Kubernetes namespaces. A multi-tenant model based on Kubernetes namespaces is a way to share a Kubernetes cluster among multiple tenants by using namespaces as virtual isolation units. Each tenant has its own namespace, which contains most of the Kubernetes objects such as pods, services, deployments and so on. Namespaces also allow for role-based access control (RBAC), which defines who can do what on the Kubernetes API.

However, namespaces do not provide complete workload or user isolation, as some resources are shared across namespaces such as nodes, persistent volumes and cluster-level objects. Therefore, additional techniques such as network policies, resource quotas and pod security policies may be needed to enforce data plane isolation and resource management.

In GATEKEEPER a tenant is identified by the Kubernetes namespace associated to it. In each tenant is deployed an instance of the service platform and an instance of big data platform based on KubeFlow and Ezmeral technologies. The following figure?? shows the detailed tenancy model of GATEKEEPER.

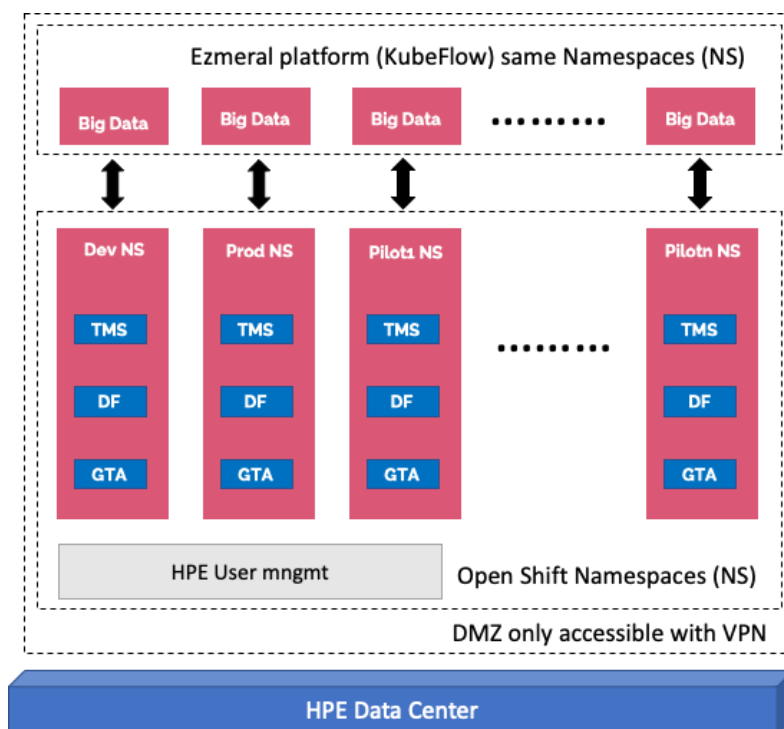


Figure 6 –GATEKEEPER tenancy model of GATEKEEPER container platform

Anyway when we work with final user additional public infrastructure is needed in order to provide reverse proxy engines to gather request inside the VPN of the Data Center.

The following figure shows a more detailed deployment where HPE Data Center and other public cloud or private partner infrastructure are involved and related to provide access to final user.

For instance, UPM infrastructure is used to provide a reverse proxy engine to publicly connect the dev, test and prod tenants. It also provides the access to the developer portal connected with all these environments.

Some GTA services and Marketplace are also deployed outside the HPE data center infrastructure, as well as some application used by pilots.

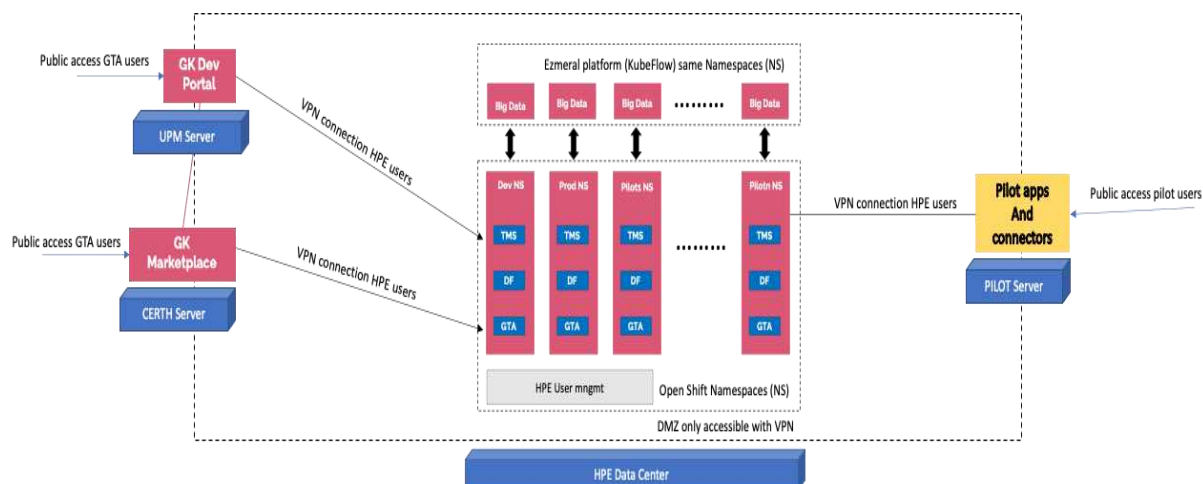


Figure 7 –Overall GATEKEEPER deployment model

3.2 GATEKEEPER secure processing environments

GATEKEEPER has a similar concept of secure processing environments that is envisioned within the European Health Data Spaces (EHDS) framework. It plays a crucial role in ensuring the privacy and security of sensitive health data. The Gatekeeper pilot platform acts as a protective gateway, controlling access to the pilot data and monitoring activities within it. It authenticates users, manages their access privileges, and enforces strict security policies with multi factor authentication. By securely managing permissions and facilitating secure data exchanges, the Gatekeeper contributes to the establishment of a trusted and secure processing environment safeguarding the integrity and confidentiality of health information. Such environments are endorsed through legal agreements signed among pilot and technical partner of the project establishing a techno-legal framework to interoperable data sharing of health sensible data for common research purpose that could be used for data holders in the context of EHDS.

3.2.1 Generic principles for secure processing environments requirements

GATEKEEPER implements the concept of Secure Processing Environments that provides support for both EHDS and GDPR regulations.

EHDS considerations

Following the final goal in European Health Data Spaces that is to provide a secure and trustworthy platform for the exchange of health data between healthcare providers, researchers, and other relevant stakeholders. We can summarize the EHDS requirements of the data and the application focused on ensuring privacy, security, trust, interoperability and data quality of health data, supporting the following features:

- **Privacy:** The EHDS is required to comply with relevant privacy regulations, such as the General Data Protection Regulation (GDPR), to ensure that patients' personal and health data is protected and kept confidential.
- **Isolation:** The EHDS must protect sensitive health information from falling into the wrong hands and to maintain the privacy and confidentiality of patients. Isolation can be achieved through various security measures, such as firewalls, encryption,

and access controls, which limit the flow of information between different systems and networks.

- **Interoperability:** The EHDS must support the exchange of health data between different systems, technologies, and organizations, to ensure that data is accessible and usable for all stakeholders. This requires the implementation of common data standards and protocols to ensure that data can be seamlessly shared and processed.
- **Security:** The EHDS must implement robust security measures to protect health data from unauthorized access and cyberattacks. This includes implementing encryption, authentication, and access controls to secure data in transit and at rest.
- **Trust:** The EHDS must establish trust between stakeholders, including patients, healthcare providers, and researchers, to ensure that data is used appropriately and for the benefit of patients. This requires transparent processes for data access and usage, as well as clear communication and engagement with stakeholders.
- **Data quality:** The EHDS must ensure the quality of health data, including accuracy, completeness, and consistency, to ensure that data is usable and relevant for healthcare purposes.

These features are essential for ensuring that the GATEKEEPER (and EHDS as well) supports the development of innovative health services and solutions, while protecting the privacy and security of health data and fostering trust between stakeholders.

Even if the identified principles are the ones promoted by EHDS, at the moment when EHDS is still a proposal also the most restrictive related to the GDPR have to be taken into account and technical measures have to deal with them.

Organizations that process personal data of EU residents must comply with the GDPR or risk significant fines and reputational damage. These requirements of the GDPR have to be supported by each components and components and can be summarized with the following features:

- **Lawful basis for processing:** Organizations must have a lawful basis for processing personal data, such as consent from the data subject or a legitimate interest.
- **Transparency:** Organizations must provide clear and transparent information about their data processing activities to data subjects. This includes the purpose of the processing, the types of data being processed, and who the data is shared with.
- **Data minimization:** Organizations must only collect and process the minimum amount of personal data necessary to fulfill the purpose of the processing.
- **Data security:** Organizations must implement appropriate technical and organizational measures to protect personal data from unauthorized access, alteration, or loss.
- **Data access and rectification:** Data subjects have the right to access their personal data and request that any inaccuracies be corrected.
- **Data erasure ("right to be forgotten"):** Data subjects have the right to have their personal data erased in certain circumstances, such as when the data is no longer necessary for the purpose for which it was collected.
- **Data portability:** Data subjects have the right to receive their personal data in a commonly used and machine-readable format and to have it transmitted to another controller.

- **Data breach notification:** Organizations must report personal data breaches to the relevant supervisory authority without undue delay and, where feasible, within 72 hours.
- **Privacy by design:** Organizations must implement privacy considerations into the design and development of products and services that process personal data.
- **Data protection officer:** Certain organizations must appoint a data protection officer to oversee their data protection activities.

3.2.2 GATEKEEPER secure processing environment principles

By analysing the considerations related to EHDS and GDPR we have defined a set of principles by design that need to be supported and verified at technical level by any component of the GATEKEEPER overall platform to ensure that every DPO of every pilot is confident and trust in the solution that the project is proposing.

These principles are:

- **Privacy by design:** incorporate data privacy safeguards into the design of systems, products, and services based on seven principles: proactive not reactive, privacy as default setting, privacy embedded into design, full functionality, end-to-end security, visibility and transparency, and respect for user privacy.
- **Trust by design:** establish trust between stakeholders, including patients, healthcare providers, and researchers, to ensure that data is used appropriately and for the benefit of patients. This requires transparent processes for data access and usage, as well as clear communication and engagement with stakeholders.
- **Security by design:** approach to software and hardware development that implements robust security measures to protect health data from unauthorized access and cyberattacks. This includes implementing encryption, authentication, and access controls to secure data in transit and at rest. The objective is to make systems as free of vulnerabilities and impervious to attack as possible through measures such as continuous testing, authentication safeguards, best programming practices, and automated security controls. It follows some principles such as minimizing attack surface area, establishing secure defaults, and failing securely.
- **Ethics by design:** ethical considerations are part of the design process for products, services, and systems. This approach seeks to create ethical products, services, and systems that are mindful of how their design, use, and impact will affect individuals, the environment, and society.
- **Data Isolation by design:** sensitive health information must be protected from falling into the wrong hands and to maintain the privacy and confidentiality of patients considering data protection and privacy issues upfront in everything can be done on data and implementing appropriate technical and organisational measures to minimise the attack surface area and establish secure defaults.
- **Data Interoperability:** the exchange of health data between different systems, technologies, and organizations must be supported from the early phased of design of the system to ensure that data is accessible and usable for all stakeholders. This requires the implementation of common data standards and protocols to ensure that data can be seamlessly shared and processed.

- **Data quality by design:** the quality of health data must be ensured and published, including accuracy, completeness, and consistency, to ensure that data is usable and relevant for healthcare purposes. Data quality by design is a method used in application development to address data quality up front, that is, to design features and functions that ensure a high quality of data is captured. It can also involve deriving an ordinary conceptual schema from application-specific goals and incorporating quality goals into the design process.
- **Lawful basis for data processing by design:** Organizations must provide system such as consent management from the data subject aligned with a lawful basis for processing personal data. Lawful basis for data processing by design is a concept that involves choosing and applying one of the six available lawful bases¹ for processing personal data according to the purpose and relationship with the individual. It also involves documenting and communicating the lawful basis to the data subjects.
- **Data transparency by design:** Organizations must provide clear and transparent information about their data processing activities to data subjects. This includes the purpose of the processing, the types of data being processed, and who the data is shared with. It involves embedding transparency measures into the design of data processing and artificial intelligence systems, instead of adding them as an afterthought. It aims at providing clear and understandable information about the data sources, methods, purposes, outcomes and impacts of these systems to the relevant stakeholders.
- **Data minimization by design:** Organizations must only collect and process the minimum amount of personal data necessary to fulfill the purpose of the processing. It is a principle that requires processing only personal data that is necessary for a specific purpose. It is related to privacy by design, which aims to protect personal data automatically in any system or practice.
- **Data access and rectification by design:** Data subjects have the right to access their personal data and request that any inaccuracies be corrected. Data access and rectification by design is a concept that requires enabling data subjects to access and correct their personal data easily and transparently. It is based on the

¹ According to the GDPR, the six available lawful bases for processing personal data are (privacypolicies.com):

- Consent: the individual has given clear consent for you to process their personal data for a specific purpose.
- Legal obligation: the processing is necessary for you to comply with the law (not including contractual obligations).
- Contractual obligation: the processing is necessary for a contract you have with the individual, or because they have asked you to take specific steps before entering into a contract.
- Legitimate interest: the processing is necessary for your legitimate interests or the legitimate interests of a third party unless there is a good reason to protect the individual's personal data which overrides those legitimate interests. (This cannot apply if you are a public authority processing data to perform your official tasks.)
- Vital interest: the processing is necessary to protect someone's life.
- Public task: the processing is necessary for you to perform a task in the public interest or for your official functions, and the task or function has a clear basis in law.

rights of access and rectification under the General Data Protection Regulation (GDPR).

- **Data erasure (“right to be forgotten”) by design:** Data subjects have the right to have their personal data erased in certain circumstances, such as when the data is no longer necessary for the purpose for which it was collected. It is a concept that requires deleting personal data of individuals on their request, unless there is a valid reason to keep it. It is based on the right to erasure under the General Data Protection Regulation (GDPR).
- **Data portability by design:** Data subjects have the right to receive their personal data in a commonly used and machine-readable format and to have it transmitted to another controller. Data portability by design is a concept that requires enabling data subjects to receive their personal data in a structured, commonly used and machine-readable format, and to transmit those data to another controller without hindrance. It is based on the right to data portability under the General Data Protection Regulation (GDPR).
- **Data breach notification by design:** Organizations must report personal data breaches to the relevant supervisory authority without undue delay and, where feasible, within 72 hours. Data breach notification by design is a concept that requires informing data subjects and supervisory authorities about any personal data breach that poses a risk to their rights and freedoms without undue delay. It is based on the obligation to report personal data breaches under the General Data Protection Regulation (GDPR).

Some technical measures to address the principles are:

- **Privacy by design:** Technical measures to support privacy by design could include incorporating privacy-preserving technologies like differential privacy or homomorphic encryption, implementing data minimization techniques like pseudonymization or anonymization, and conducting privacy impact assessments (PIAs) to identify and mitigate potential privacy risks.
- **Trust by design:** To establish trust between stakeholders, technical measures could include implementing transparent processes for data access and usage, such as access logs and audit trails, and providing clear and concise documentation on data usage policies and procedures.
- **Security by design:** To ensure robust security measures are in place, technical measures could include implementing multi-factor authentication, using strong encryption algorithms, ensuring proper access control mechanisms are in place, conducting regular vulnerability assessments and penetration testing, and using security-focused development methodologies like DevSecOps.
- **Ethics by design:** Technical measures that could be implemented to support ethical considerations might include the use of explainable AI to ensure that decision-making algorithms are transparent and understandable, conducting ethics reviews during the design phase, and implementing measures to mitigate the impact of biases on algorithms.
- **Data Isolation by design:** Technical measures to support data isolation could include implementing access controls, network segmentation, and secure data storage and transfer protocols.
- **Data Interoperability:** Technical measures to support data interoperability could include using standardized data exchange formats like HL7 or FHIR, implementing

APIs and web services for data exchange, and using common data models to facilitate interoperability.

- **Data quality by design:** Technical measures to ensure data quality could include implementing data validation and cleansing routines, using standard data definitions and formats, and using automated data quality checks.
- **Lawful basis for data processing by design:** Technical measures to support lawful data processing could include implementing consent management systems, maintaining detailed records of data processing activities, and ensuring that data processing activities are aligned with the lawful basis for processing.
- **Data transparency by design:** Technical measures to support data transparency could include implementing data dashboards and visualizations, providing clear and concise explanations of data processing activities, and using open data formats and APIs to enable data sharing.
- **Data minimization by design:** Technical measures to support data minimization could include implementing data retention policies, using data masking techniques to hide unnecessary data, and ensuring that data collection activities are aligned with the principle of data minimization.
- **Data access and rectification by design:** Technical measures to support data access and rectification could include implementing user self-service portals for data access and correction, using role-based access controls to ensure appropriate access to data, and implementing secure data transfer protocols to enable data sharing.
- **Data erasure ("right to be forgotten") by design:** Technical measures to support data erasure could include implementing data retention policies, maintaining detailed records of data processing activities, and ensuring that data erasure requests are processed in a timely and secure manner.
- **Data portability by design:** Technical measures to support data portability could include implementing APIs and web services for data transfer, using standard data formats, and ensuring that data portability requests are processed in a timely and secure manner.
- **Data breach notification by design:** Technical measures to support data breach notification could include implementing intrusion detection and prevention systems, conducting regular vulnerability assessments and penetration testing, and maintaining detailed incident response plans to facilitate timely notification of relevant stakeholders.

Within Gatekeeper we have gone a step behind, providing standard legal clauses reused across the different pilot that local lawyers have accepted, and Gatekeeper platform has been considered "trustable" for all pilot sites.

3.2.3 GATEKEEPER responsible data management findings and experience

Within Gatekeeper we have provided a standard means of verification of the above principles, by using standard legal clauses reused across the different pilots that local lawyers have accepted, and Gatekeeper platform has been considered "trustable" for all pilot sites.

Some of the aspects mentioned earlier, such as security by design, data isolation by design, and lawful basis for data processing by design, are indeed related to the concept of a Secure Processing Environment. Other aspects, such as data transparency by design and ethics by design, are more broadly related to the concept of responsible and ethical data management.

To summarize, a Secure Processing Environment is an important component of responsible data management, but it is not the only consideration. Other aspects such as privacy by design, trust by design, data quality by design, data access and rectification by design, data erasure by design, data portability by design, and data breach notification by design are also important components of responsible data management.

Using a Secure Processing Environment is an important step towards responsible data management. Here are some steps that we found can be taken to utilize a Secure Processing Environment for responsible data management:

- **Conduct a privacy impact assessment:** Before setting up the Secure Processing Environment, conduct a privacy impact assessment (PIA) to identify any potential privacy risks associated with the data being processed. This will help you identify the necessary measures to implement in the Secure Processing Environment to mitigate any risks.
- **Implement technical and organizational measures:** Once you have identified the risks associated with the data being processed, implement the necessary technical and organizational measures to protect the data. This can include measures such as access controls, encryption, and data isolation to ensure that only authorized individuals can access the data and that the data is kept secure.
- **Establish clear policies and procedures:** Establish clear policies and procedures for the use of the Secure Processing Environment. This includes guidelines for access control, data retention, and data disposal. Ensure that all employees and third-party contractors who have access to the Secure Processing Environment are trained on these policies and procedures.
- **Monitor and audit the Secure Processing Environment:** Regularly monitor and audit the Secure Processing Environment to ensure that the established policies and procedures are being followed and that the technical measures are working effectively. This includes monitoring for unauthorized access, data breaches, and other potential security incidents.
- **Continuously improve the Secure Processing Environment:** Continuously assess the Secure Processing Environment to identify areas for improvement. This can include updating policies and procedures, implementing new technical measures, and providing ongoing training to employees and contractors.

By following these steps, a Secure Processing Environment can be used for responsible data management, protecting the privacy and security of the data being processed.

3.2.4 GATEKEEPER SPE for responsible AI

GATEKEEPER secure processing environment serves as a crucial measure to oversee and hold accountable any unethical utilization of AI models. To promote ethical practices in this realm, organizations are advised to adopt a series of strategic steps.

Firstly, establishing clear ethical guidelines is paramount. These guidelines should comprehensively cover the spectrum of AI model usage, encompassing aspects such as

ethical data collection, model development, and deployment. The foundation of these guidelines should rest on industry best practices and align with prevailing legal requirements, thereby offering a robust framework for ethical AI utilization.

Regular audits of AI models constitute a key practice to ensure ongoing adherence to ethical standards and compliance with established guidelines. These audits involve a comprehensive examination of data sources, model training data, and the overall performance of the AI models.

Continuous monitoring of AI model performance is imperative to verify that these models align with expectations. Key performance metrics, including accuracy, precision, recall, and others, are scrutinized to promptly identify and rectify any biases or discriminatory outcomes.

Implementing stringent access controls is another critical measure. By adopting mechanisms like role-based access control and multi-factor authentication, organizations can ensure that only authorized individuals have the requisite access to interact with the AI models and associated data.

Maintaining detailed logs of all interactions with AI models is essential. These logs, capturing inputs, outputs, and decisions made based on model outputs, not only facilitate audits for ethical usage but also play a pivotal role in identifying and resolving any potential ethical concerns that may arise.

Incorporating alerting mechanisms into the system is the final recommendation. This involves the implementation of alerts that promptly notify relevant personnel when AI models exhibit unexpected behaviour or when potential ethical concerns are detected. These alerts, accompanied by detailed logs, empower organizations to swiftly investigate and address any arising issues.

By meticulously following these steps, organizations can establish a robust framework for the ethical use of AI models, ensuring compliance with legal requirements. The GATEKEEPER secure processing environment technical measure are able to support each one of these steps, furthermore it enhances data protection efforts, reducing the risk of unauthorized access and fortifying defences against potential data breaches and security incidents.

3.3 GATEKEEPER recommendations for EHDS

3.3.1 Secure image building pipeline for SPEs

GATEKEEPER has implemented a secure image building pipeline that is a comprehensive and proactive approach to constructing and deploying container images with a strong emphasis on security. It aligns with the principles of DevSecOps, seamlessly integrating security into the entire development and deployment lifecycle. More in details, a secure image building pipeline is a structured and controlled process within software development that prioritizes security considerations throughout the construction and deployment of containerized applications. In the landscape of modern DevOps practices, where containers are widely utilized, this pipeline serves as a crucial element to ensure that the images used for deploying applications are not only functional but also devoid of vulnerabilities, compliant with security policies, and aligned with best practices.

The process initiates with the source code residing in a version control system, typically Git. As developers make changes and commit code, the pipeline kicks into action.

Automated build tools, such as Docker, take charge of compiling, configuring, and packaging the application code into containerized images.

One of the key decisions in the pipeline involves the selection of a secure and up-to-date base image. This choice is critical for minimizing security risks, and using official and well-maintained base images is a common practice.

Security scans and risk analysis are conducted at various stages of the pipeline. Dependency scanning examines libraries and third-party components for known vulnerabilities, while static code analysis tools inspect the source code for security issues and adherence to coding best practices.

Configuration management is another vital aspect, ensuring that security configurations, environment variables, and sensitive information are handled in a secure manner. The pipeline incorporates measures to securely manage secrets and prevent inadvertent exposure within the container image.

To safeguard the integrity and authenticity of images, digital signatures are applied during image signing, and verification processes are implemented to ensure that only legitimate and unaltered images are deployed.

The pipeline also includes compliance checks to ensure that the constructed images comply with security policies, regulatory requirements, and organizational standards. Furthermore, the concept of immutable image builds is embraced, emphasizing that once an image is created, it remains unchanged to reduce the risk of introducing vulnerabilities through manual modifications.

Images are stored in a secure container registry, which is configured with robust access controls, authentication mechanisms, and encryption. This registry acts as a centralized repository for securely storing versioned container images.

Continuous monitoring is an integral part of the process, extending beyond the image building phase. The pipeline incorporates tools and processes for ongoing assessment of container images in production, identifying and addressing vulnerabilities and security incidents.

GATEKEEPER has adopted this DevSecOps approach to proactively respond to eventual data breach and accomplish with the GDPR recommendations on data breach notifications.

A secure image building pipeline stands as the cornerstone also for Secure Processing Environments in European Health Data Spaces. It establishes a robust foundation for deploying containerized applications, ensuring data integrity, and compliance with regulatory standards while actively mitigating potential security risks.

In the European Health Data Space (EHDS), an EHDS secure processing environment, as defined in Article 50, is a controlled setting where electronic health data is accessed and processed. It is characterized by rigorous technical and organizational measures to ensure data security and interoperability. Key features include restricted access, state-of-the-art data protection, limited operations, individualized access for data users, logging and auditing, continuous compliance monitoring, and support for data uploading and downloading. Regular audits are mandated to assess security effectiveness. At the moment the European Commission defines technical and security standards at glance, they will be detailed through implementing acts. The GATEKEEPER secure image building pipeline aligns with high level EHDS security principles by incorporating secure practices during software construction, contributing to overall security and compliance in health data processing environments. The pipeline supports the secure development and deployment of any applications within the EHDS.

At its core, the image building pipeline is designed to guarantee the reliability of the container images employed in health data processing. By incorporating secure practices throughout the image construction process, it certifies that the software components and configurations within the containers are devoid of vulnerabilities and adhere to stringent security requirements.

One of the primary objectives is the mitigation of vulnerabilities. Through continuous security scanning and strict adherence to best practices, the pipeline actively identifies and addresses potential weaknesses in the software stack. This proactive approach is essential to thwart potential exploits and fortify the security posture, safeguarding the confidentiality, integrity, and availability of health data.

Compliance with security standards is a paramount consideration. The image building pipeline is crafted to align seamlessly with the security requirements (article 50) set by European Health Data Spaces. This alignment ensures that the constructed container images adhere meticulously to the prescribed security measures, encompassing access restrictions, encryption, and comprehensive audit logging.

The concept of immutable image builds, where the integrity of an image is maintained post-creation, adds another layer of security. By limiting the potential for unauthorized modifications to containerized applications, this approach enhances the overall resilience of the processing environment.

An integral aspect of the pipeline is the enforcement of access controls and adherence to the principle of least privilege. It ensures that only individuals with explicit authorization have access to the pipeline and, consequently, to the constructed container images. This practice minimizes the attack surface and elevates the overall security posture.

Secure handling of secrets and sensitive configurations is of paramount importance in a health data processing environment. The pipeline incorporates robust mechanisms to manage and inject secrets securely, mitigating the risk of inadvertent exposure of sensitive information within container images.

Auditability and compliance verification are facilitated through detailed logs maintained throughout the image-building process. These logs serve as crucial artifacts during regulatory audits, providing transparency into the adherence to security measures and overall compliance.

The pipeline's role extends to regular audits and continuous monitoring, aligning with the requirement for ongoing scrutiny of secure processing environments. This ensures that security measures are consistently applied and any emerging threats are promptly addressed.

In the context of European Health Data Spaces, a secure image building pipeline plays a pivotal role in ensuring the integrity, trustworthiness, and compliance of the containerized applications used within Secure Processing Environments.

Gatekeeper integrates an automated, secure image building pipeline to ensure that every artifact deployed within it undergoes rigorous vulnerability risk analysis, ensuring compliance, or is automatically rejected.

Gatekeeper DEVSECOPS is a proactive measure that aligns with best practices in cybersecurity. It not only protects against potential threats but also fosters a culture of security and continuous improvement within the development process.

EHDS recommendations:
Ensuring **the security of every EHDS processing environment is imperative**, and it necessitates the incorporation of essential functionality. This critical feature is particularly crucial during the deployment of artifacts, underscoring the **indispensable need for an automated, secure image building pipeline**.



Figure 8 – DevSecOps GATEKEEPER approach as recommendations for EHDS

3.3.2 FHIR based interoperability

The design philosophy of GATEKEEPER revolves around the "interoperability by design" concept, emphasizing seamless collaboration and integration across diverse pilot sites. The anticipated interoperability format serves as the foundation for the platform's collaborative approach, aiming to deliver several key functionalities:

- **Common Data Model Mapping:** The interoperability format is designed to map the specific needs and requirements of individual pilot sites onto a unified and standardized common data model. This ensures that diverse data structures and formats across sites can be harmonized, promoting a cohesive and interoperable data environment.
- **Health Data Exchange Suitability:** Recognizing the critical nature of health data, the interoperability format is tailored to support secure and efficient health data exchange. This includes considerations for data privacy, security, and compliance with healthcare standards, fostering a reliable foundation for sharing sensitive health information.
- **Semantic Interoperability:** GATEKEEPER places a strong emphasis on semantic interoperability, aiming to facilitate a shared understanding of data meaning and context. By adopting standardized ontologies and ensuring consistent data semantics, the platform enhances the meaningful exchange of information, fostering collaboration and interoperability.
- **Standard Service Definition Support:** The interoperability format is crafted to support standard service definitions. This involves establishing clear and standardized specifications for the services offered by GATEKEEPER, promoting consistency and clarity in the way services are defined and utilized across the federated infrastructure.
- **Standard Extensibility:** GATEKEEPER anticipates the need for adaptability and evolution over time. The interoperability format is designed to support standard extensibility, allowing for the seamless integration of new features, functionalities, and data elements. This ensures that the platform can evolve in response to emerging requirements and technological advancements.
- **User-Friendly and Developer-Friendly:** Acknowledging the importance of usability, the interoperability format is engineered to be user-friendly and developer-friendly. This includes intuitive interfaces for users and developers, comprehensive documentation, and support for standard development practices.

The goal is to foster a collaborative environment that encourages ease of use and encourages developers to contribute effectively to the platform.

In essence, GATEKEEPER's commitment to interoperability by design is reflected in the careful consideration of these functionalities within the interoperability format. By addressing the diverse needs of pilot sites, supporting health data exchange, ensuring semantic interoperability, defining standard services, enabling extensibility, and prioritizing user and developer friendliness, GATEKEEPER aims to create a robust and collaborative ecosystem for innovative healthcare solutions.

GATEKEEPER is one of early adopter of FHIR for the project interoperability approach.

HL7 FHIR is a standard for exchanging healthcare information electronically. It is designed to address the interoperability challenges in healthcare by providing a modern, web-based framework for representing and exchanging clinical data. Let's explore how HL7 FHIR supports the functionalities outlined for GATEKEEPER:

- **Common Data Model Mapping:** FHIR employs a resource-based approach, where resources represent discrete pieces of information such as patients, observations, and medications. The standard defines a set of resource types and attributes, providing a common data model. FHIR's flexibility allows for the mapping of diverse data structures to a standardized format.
- **Health Data Exchange Suitability:** FHIR is specifically designed for health data exchange. It supports secure and efficient data exchange through RESTful APIs, allowing systems to query, retrieve, and update healthcare information. FHIR also incorporates security features, including authentication and authorization mechanisms, to ensure the confidentiality and integrity of exchanged data.
- **Semantic Interoperability:** FHIR promotes semantic interoperability by using standardized and widely adopted code systems, vocabularies, and ontologies. It provides support for coding systems such as SNOMED CT, LOINC, and RxNorm, facilitating a shared understanding of clinical concepts across different systems.
- **Standard Service Definition Support:** FHIR defines a set of standard operations and interactions through its RESTful API. These operations are well-defined and include capabilities for searching, reading, updating, and deleting resources. This standardization supports consistent service definitions, enabling interoperability across different implementations.
- **Standard Extensibility:** FHIR is designed with extensibility in mind. It allows for the addition of custom data elements and extensions to accommodate domain-specific requirements. The extension mechanism uses special resources such as extensions and conformance resources that provide a standardized and machine readable way to enhance both the data model and the behaviour of the RESTful APIs without compromising interoperability.
- **User-Friendly and Developer-Friendly:** FHIR's RESTful API and use of modern web standards make it user-friendly for developers. The standard provides clear and concise documentation, making it easy for developers to understand and implement. FHIR's resource-oriented approach and use of JSON or XML for data representation contribute to its developer-friendly design.

FHIR aligns with the functionalities outlined for GATEKEEPER by providing a standardized and flexible framework for healthcare data exchange. Its common data model, support for semantic interoperability, standardized service definitions, extensibility features, and user-friendly design make it a suitable candidate for interoperable health information exchange. However, it's important to note that the

specific implementation and success of interoperability also depend on how FHIR is implemented and utilized within the context of specific healthcare systems and applications.

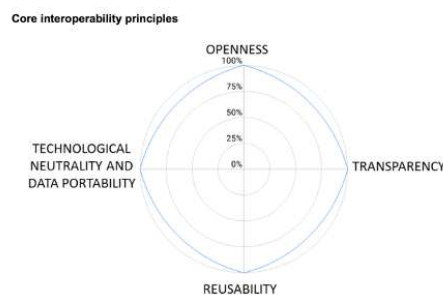
Gatekeeper rigorously assessed various interoperability standards, ultimately designating FHIR as the indisputable leader in providing a standardized approach for health data management.



EHDS recommendations:

Gatekeeper has been an early adopter of FHIR. Gatekeeper has a lot of lessons learnt to share on usage of FHIR within its pilots.

On the other hand THEDAS project has already provided a Common Assessment Method for Standards and Specifications (CAMSS) and suggested to adopt FHIR for EHDS.



Ref: Recommendations to enhance interoperability within [HealthData@EU](#)

Figure 9 – FHIR based interoperability approach as recommendations for EHDS.

In conclusion, as an early adopter, from GATEKEEKEER we strongly recommend the usage of FHIR in EHDS. This advice was also raised in the end of 2022 from the THEDAS project with the evaluation they made with the Common Assessment Method for Standards and Specifications (CAMSS) and suggested the usage of FHIR combined with other standards.

3.3.3 Secure processing environments for primary and secondary use of data

On Line Data Processing (OLTP) and On Line Analytical System (OLAP) environments in healthcare are integral components that contribute to both the immediate operational needs and long-term strategic decision-making within the data processing and analysis landscape.

Primary Use of Data (OLTP):

OLTP systems are specifically designed for real-time transactional processing in healthcare. They play a crucial role in the day-to-day operations related to patient care, encompassing tasks such as recording individual patient encounters, updating medical records, and managing administrative functions. The primary focus of OLTP environments is on ensuring the quick and efficient processing of individual transactions, addressing immediate data needs with a priority on accuracy, reliability, and consistency in real-time. Healthcare professionals heavily rely on OLTP systems for operational decision-making, gaining timely access to accurate patient information, prescribing medications, and managing daily workflows.

Secondary Use of Data (OLAP):

In contrast, OLAP systems are tailored for complex analytical processing and reporting, emphasizing historical analysis and strategic decision support. These systems facilitate in-depth exploration of historical and aggregated data, allowing healthcare organizations to derive insights, identify patterns, and make informed decisions about treatments, disease patterns, and long-term healthcare strategies. The secondary use of data involves extracting comprehensive insights from aggregated and transformed data, providing a holistic view of patient outcomes, resource utilization, and overall healthcare

performance. Healthcare executives and administrators leverage OLAP systems for strategic decision support, enabling activities such as planning resource allocation, optimizing healthcare delivery processes, and adapting to evolving trends in the healthcare landscape.

OLTP and OLAP environments in healthcare are interconnected components of a comprehensive data processing ecosystem. OLTP focuses on the primary use of data for immediate operational needs, while OLAP facilitates the secondary use of data for in-depth analysis, historical insights, and strategic decision support. The integration of these environments ensures a cohesive approach to leveraging healthcare data for both immediate actions and long-term planning.

The integration of OLTP and OLAP environments in healthcare data analysis is essential for obtaining a comprehensive understanding of patient health and operational efficiency. The challenges associated with resource-intensive OLAP systems can be addressed through the adoption of container-based data science applications, providing scalability, flexibility, and cost optimization in the healthcare analytics landscape.

The need for integrating OLTP and OLAP systems in healthcare are rooted in the pursuit of comprehensive insights, real-time decision support, historical analysis, and operational efficiency.

Healthcare organizations deal with vast amounts of data, including patient records and real-time transactions, managed by OLTP systems. Concurrently, OLAP systems specialize in intricate analytics and reporting. The integration of both environments is pivotal, as it allows for a holistic understanding of patient health, treatment outcomes, and operational efficiency by combining real-time transactional data with advanced analytical processing.

In the context of real-time decision support, healthcare professionals rely on timely information for critical decision-making. The integration of OLTP and OLAP facilitates the generation of real-time analytical reports, providing immediate insights into patient conditions, treatment efficacy, and optimal resource utilization.

Historical analysis and trend identification are strengths of OLAP systems. Through integration with OLTP, healthcare organizations gain the capability to analyze historical patient data, identify patterns, and make informed decisions about treatment plans, resource allocation, and overarching healthcare strategies. This historical perspective contributes to more effective long-term planning and management.

The integration of OLTP and OLAP systems brings about operational efficiency by streamlining data processes. Managing separate transactional and analytical systems can be complex and resource-intensive. The integration simplifies this complexity, ensuring quick access to accurate information. In healthcare settings, where time is of the essence for patient care and organizational management, operational efficiency is crucial.

In GATEKEEPER, the integration of OLTP and OLAP serves as a strategic imperative, offering a unified approach to data processing that delivers comprehensive insights, real-time decision support, historical analysis, and operational efficiency. This integration is foundational for healthcare organizations striving to enhance patient care, optimize resource allocation, and make informed decisions based on a holistic view of their data landscape.

OLAP systems are inherently resource-intensive, primarily due to the intricacy of queries they handle and the necessity to process vast amounts of data. This resource intensiveness becomes particularly challenging in healthcare settings, where datasets can be both massive and diverse. Effectively allocating and managing resources in the

face of this complexity becomes a critical challenge, requiring thoughtful strategies to ensure optimal performance.

Additionally, cost considerations play a significant role in the realm of OLAP systems. Traditional OLAP solutions often demand substantial investments in both hardware and software infrastructure. The challenge lies in the associated costs of scaling up this infrastructure to accommodate the ever-growing volumes of data and the increasing demands for analytical processing. For organizations seeking to optimize resource utilization while simultaneously managing costs, striking a balance between performance and financial efficiency becomes a complex yet essential objective.

In GATEKEEPER, we faced the challenges of resource intensiveness and the inherent costs of scaling infrastructure. Overcoming these challenges necessitates careful resource management strategies and a judicious approach to balancing the performance requirements of OLAP with the imperative of cost-effectiveness, especially in dynamic environments like healthcare where data volumes and analytical demands continue to expand.

These challenges in GATEKEEPER have been effectively addressed through the strategic integration of container-based data science applications.

Containerization technologies like Docker and Kubernetes play a pivotal role in enhancing resource efficiency within OLAP environments. By encapsulating data science applications and their dependencies, containers abstract away the intricacies of the underlying infrastructure. This abstraction facilitates efficient resource utilization, allowing for streamlined scaling and allocation of resources based on dynamic workload demands.

The scalability and flexibility offered by container orchestration platforms, exemplified by Kubernetes, prove instrumental in addressing the challenges of resource intensiveness. These platforms enable dynamic scaling, allowing OLAP applications to scale both horizontally and vertically in response to varying demands. This ensures that resources are optimally utilized during peak analytical workloads while providing the flexibility to scale down during periods of lower demand.

Containers also contribute to overcoming challenges related to isolation and portability. By providing a layer of isolation between applications and environments, containers reduce conflicts and ensure consistent execution of OLAP applications across diverse settings. This isolation enhances portability, making it seamless to deploy and manage analytical workloads across different infrastructure setups.

Cost optimization emerges as a significant advantage of containerization. Healthcare organizations can leverage containers to efficiently utilize cloud resources, embrace microservices architectures, and capitalize on cloud-native services. This flexibility enables organizations to adapt OLAP environments to changing demands without incurring unnecessary infrastructure costs, thereby addressing the inherent financial challenges associated with scaling infrastructure.

In conclusion, GATEKEEPER offer a scalable and cost-effectiveness integration of OLTP and OLAP environments in healthcare data analysis, crucial for obtaining a comprehensive understanding of patient health and operational efficiency. The challenges linked to resource intensiveness and scaling costs within OLAP systems find effective solutions through the adoption of container-based data science applications. These applications provide scalability, flexibility, and cost optimization, thereby enhancing the overall landscape of healthcare analytics.

It is a multi-tenant microservice based platform that integrates **secure interoperable transactional systems (OLTP)** with **analytics systems (OLAP)** for **massive data management** and **building AI services**

- The **OLTP** enables **primary use of DATA**
- The **OLAP** enables **secondary use of DATA**

EHDS recommendations

Integrating OLTP and OLAP systems within a shared Secure Processing environment facilitates both primary and secondary data use, ensuring privacy by enabling local data processing.

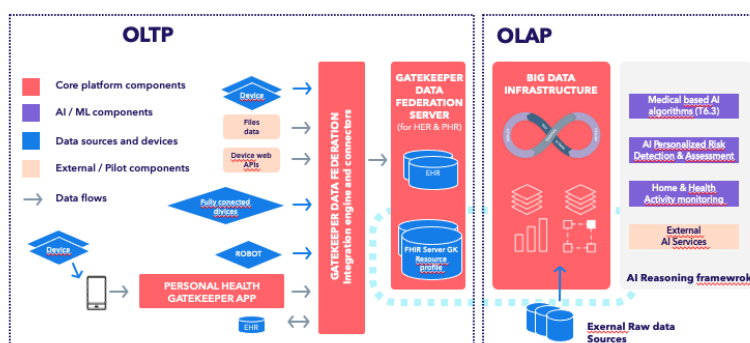


Figure 10 – GATEKEEPER integration for primary and secondary use of data.

The integration of primary and secondary use of health data is pivotal within the European Health Data Spaces (EHDS), serving multifaceted purposes. In real-time transactional processing (OLTP), the focus is on immediate patient care, capturing day-to-day operations, and supporting timely decision-making for healthcare professionals. Conversely, in the realm of online analytical processing (OLAP), historical and aggregated data analysis takes center stage, providing in-depth insights into patient outcomes, treatment effectiveness, and long-term trends.

This integration ensures a comprehensive understanding of healthcare data by combining the immediacy of operational decisions with the strategic depth derived from historical analysis. Primary use facilitates real-time interoperability and meets immediate patient care needs, while secondary use enhances semantic interoperability, allowing a seamless flow of data across diverse healthcare settings and systems.

The synergy between primary and secondary data usage optimizes patient care, supports evidence-based policymaking, and fosters research and innovation. It addresses ethical and privacy considerations through stringent measures in both real-time and historical data handling. Additionally, the integration aids in efficient resource allocation by balancing immediate needs with long-term insights, thus preventing unnecessary costs and promoting cost-effectiveness.

In the context of EHDS, this integration plays a crucial role in fostering cross-border collaboration, providing standardized and interoperable data for regional and European-level healthcare initiatives. Moreover, it establishes adaptability to emerging healthcare challenges, addressing immediate crises while laying the groundwork for proactive planning and resilience.

Ultimately, the harmonious integration of primary and secondary health data usage in EHDS creates a holistic healthcare ecosystem, promoting innovation, research, and data-driven decision-making for the benefit of patients, healthcare providers, and society at large.

3.3.4 Exchange of AI model and services

FHIR allows the definition of custom operations using the OperationDefinition resource. This enables the creation of domain-specific operations that go beyond the standard CRUD operations. Custom operations can encapsulate complex workflows, computations, or interactions tailored to specific healthcare use cases.

GATEKEEPER platform is designed with interoperability in mind, fostering the exchange of AI models and services through a unified approach. This involves the utilization of a Common Data Model (CDM) and leveraging FHIR operations.

The Common Data Model serves as a standardized representation of healthcare data, ensuring consistency and interoperability across diverse sources. Meanwhile, FHIR operations provide a standardized and efficient means of interacting with healthcare data electronically. The platform encapsulates AI models within containers, using technologies like Docker or Kubernetes, for streamlined deployment and portability.

Gatekeeper take advantage of flexibility of FHIR standard for describing services for exchange AI models and services across different pilots that share the same FHIR implementation guide.

EHDS recommendations

Reproducibility of AI training and AI validation is fundamental for healthcare data spaces, Gatekeeper provides and easy and standard based solution for AI models and services.

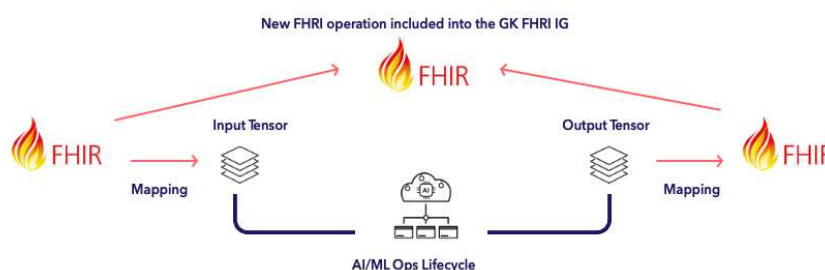


Figure 11 – GATEKEEPER AI models and services mapper to FHIR operations.

By customizing these FHIR operations (figure 11), healthcare organizations and developers can refine the capabilities of FHIR servers to align with particular needs, accommodate workflows specific to healthcare domains, and seamlessly incorporate with varied healthcare ecosystems. This adaptability supports the standardization of AI models and services, fostering interoperability and flexibility. It significantly contributes to the efficient exchange and utilization of healthcare data while adhering to standardized approaches tailored to AI integration within the healthcare landscape.

This lesson learnt in GATEKEEPER provide a strong recommendation that should be taken into account in the EHDS regulation. The standardization and exchange of AI models and services within EHDS unleash a powerful network effect, fostering a transformative impact on the entire healthcare ecosystem. This positive feedback loop is characterized by several interconnected facets.

Firstly, the adoption of standardized practices ensures interoperability, enabling seamless communication and data exchange among diverse healthcare entities within the EHDS. This interoperability forms the foundation for collaborative endeavors and innovative solutions, as stakeholders converge to share expertise and contribute to a collective pool of knowledge.

The efficiency gains from standardization attract an increasing number of participants, driving scalability within the network. As more stakeholders join, the EHDS becomes a dynamic environment capable of accommodating a growing array of standardized AI solutions. This scalability enhances the reach and effectiveness of the network.

Moreover, standardization contributes to improved data quality and consistency. Consistent formats and processing methods elevate the reliability of insights derived from AI applications. This reliability builds trust among participants, fostering a sense of security and encouraging broader adoption and participation.

The establishment of shared standards also underpins robust data governance and security measures. Clear guidelines for data access, usage, and protection ensure the network operates as a trusted environment for the exchange of sensitive health data, further solidifying its integrity.

In terms of regulatory compliance, standardized AI models align with data protection and privacy regulations, bolstering the credibility of the EHDS. This compliance not only instills confidence within the network but also attracts regulatory support, creating a conducive environment for sustained growth and development.

A patient-centric approach is a hallmark of the network effect, where standardized AI models contribute to personalized and effective patient care. As the network expands, more patients benefit from tailored AI-driven insights and treatments, reinforcing the positive impact on individual healthcare outcomes.

Beyond the local sphere, a well-standardized and interconnected EHDS gains global visibility and influence. Attracting international collaboration, partnerships, and recognition, the network effect amplifies its impact on the global stage, positioning the EHDS as a influential player in healthcare innovation.

In essence, the network effect triggered by the standardization and exchange of AI models and services in the EHDS creates a harmonious and synergistic environment. This interconnected network not only benefits the participating stakeholders but also contributes to advancing healthcare on a broader, global scale.

3.3.5 Proposal for techno-legal specification for generic EHDS secure processing environments

Within the Gatekeeper project, a Data Protection Impact Assessment (DPIA) has been conducted, focusing on the technical security aspects related to data protection within the platform. The findings from this assessment have been integrated with standard legal clauses, creating a cohesive framework that addresses both technical and legal dimensions of data protection. This integrated framework, that we have called the "techno-legal specification of the Gatekeeper platform", has been reused as a common package across different pilot implementations. Local lawyers, providing legal expertise for each pilot site, have reviewed and accepted this common package, confirming its alignment with local legal requirements. As a result, the Gatekeeper platform has been deemed "trustable" for all pilot sites, signifying its reliability, security, and compliance with both technical and legal requirements.

By migrating this approach to EHDS, the concept of techno-legal specifications for the European Health Data Space (EHDS) involves establishing a comprehensive framework that combines both technological and legal elements to ensure the secure and lawful processing of healthcare data. These specifications aim to provide clear guidelines and standards for the design, implementation, and governance of secure processing environments within the EHDS ecosystem.

From a technological perspective, the specifications address key aspects such as encryption, access controls, auditing, vulnerability assessments, incident response planning, data retention policies, and continual monitoring. These measures are crucial to safeguard the confidentiality, integrity, and availability of health data. Encryption is emphasized for both data at rest and in transit, access controls are implemented to restrict sensitive data access, auditing ensures accountability, regular vulnerability assessments identify and rectify security vulnerabilities, and incident response plans are in place to handle security breaches effectively.

On the legal front, the techno-legal specifications align with relevant laws and regulations, with a particular emphasis on data protection laws like the General Data Protection Regulation (GDPR). Compliance with legal and regulatory requirements is deemed essential, and the specifications ensure that the secure processing environment adheres to all applicable laws, including those related to data protection, electronic health records, and other relevant regulations. This alignment is crucial for building trust and ensuring responsible data handling practices.

Additionally, the techno-legal specifications emphasize the need for continual monitoring and improvement of secure processing environments to adapt to evolving threats and technologies. Regular audits should be conducted to assess compliance (section x.x.x secure container building pipeline address this need), and the specifications support ongoing learning and collaboration within the EHDS community.

The integration of both technological and legal elements ensures a holistic approach to secure data processing within the EHDS. It recognizes the importance of not only implementing robust technical measures but also adhering to legal and ethical standards to foster transparency, accountability, and responsible use of health data. Ultimately, these techno-legal specifications contribute to the EHDS's overarching goal of facilitating secure, interoperable, and privacy-preserving healthcare data exchange across the European Union.

4 Intelligent connectors

Intelligent Connector are developed only in the context of T5.4 and are mainly related to the integration of the Medisanté ecosystem.

4.1 Example use case

4.1.1 Update of Intelligent Connected Care Service in Puglia

The intelligent connected care services (ICCS) are an innovative solution to push device data over cellular network into Gatekeeper Data Federation at zero configuration effort for patients (see more detailed description in D5.4.1).

The service has been considered by several pilots for large scale deployment in the early phase of the program. It turns out that many pilots had preferred a Bring-Your-Own-Device (BYOD) approach for the connectivity in deployment – relying on patient's ability to configure and share its vital parameters, instead of automating at zero configuration efforts thanks to medical IOT services. This innovative service is now being validated in RUC 5 and RUC 7 with 200 devices with embedded SIM-Card shipped to Puglia region.

As planned and communicated in the deliverable 5.4.1, efforts have been made to broaden the portfolio of existing medical devices to offer more choices to care teams. In period 1, several devices were already newly integrated and such efforts have been continued in period 2 to successfully add:

- CAT-M weight-scale up to 250 kg (CE mark)
- CAT-M blood pressure monitoring (CE mark)
- 2 new BLE devices (weight scale & glucometer) connecting with the 4G Gateway connected in period 1 (and certified CE mark in period 2 together with OEM manufacturer for deployment in Europe (WP8)

See the illustrations below (figure 12) of all these integrated innovations now ready & certified to support RPM care team across Europe.



Figure 12 –Integrated Hardware connected to Gatekeeper DF

Such solutions have well addressed the pain points communicated by the pilots, such as high cost of hardware and a need of more choices in vital parameters selected. This relevant feedback communicated by several pilots such as Greece or Cyprus will be very much relevant to scale the service across Europe on additional sites post-Gatekeeper project.

In the figure below (figure 13), we can observe the connectors available to the services all via ONE API directed to Gatekeeper Data Federation (see also Deliverable 54), first for devices with an embedded SIM-card and secondly for the Gateway (GW1) with an embedded SIM-Card connecting with BLE devices from one OEM.

Send Test Measurement

Are you sure that you want to send test measurements to
Data Federation Gatekeeper?

Please ensure the IMEIs below exist in webhook before sending any

BG800 (IMEI: 9000000000000001)
BP800 (IMEI: 9000000000000002)
BC800 (IMEI: 9000000000000003)
PMI00 (IMEI: 9000000000000004)
BT005 (IMEI: 9000000000000005)
BTI05 (IMEI: 9000000000000006)
D40G (IMEI: 9000000000000007)
GTEL (IMEI: 9000000000000008)
GW9017 (IMEI: 9000000000000009)
BS-2001-G1 (IMEI: 9000000000000010)
LS802-GP (IMEI: 9000000000000011)
PS300C-RPM (IMEI: 9000000000000012)
BM300C-RPM (IMEI: 9000000000000013)

Send Test Measurement

Are you sure that you want to send test measurements to
Data Federation Gatekeeper?

Please ensure the IMEIs below exist in webhook before sending any
synthetic data

Choose device model

GW9017 (IMEI: 9000000000000009)

Body Temperature
SpO2
Blood Pressure
Body Weight
Ketone
Cholesterol
URIC Acid
Glucose

Figure 13 –ICCS Measurements connected to Gatekeeper DF

Such successful integrations have required a lot of work in different areas. Especially the alignment with the new supplier, which helped to integrate the Cat-M technology into the new cellular Blood Pressure and Weight Scale was a main task which took a lot of effort from M+ side. The device firmware, the production process, the certification of the devices, the definition of the different legal roles like importer/distributor as well as to find a strong partner who can offer the device shipping to an acceptable price have had to all be defined, updated, and verified.

The new Cat-M technology required a lot of testing and validation. Connection tests were performed in different countries to verify that this new technology is working properly. Below (figure 14) you see a verification example performed in Switzerland and Germany.

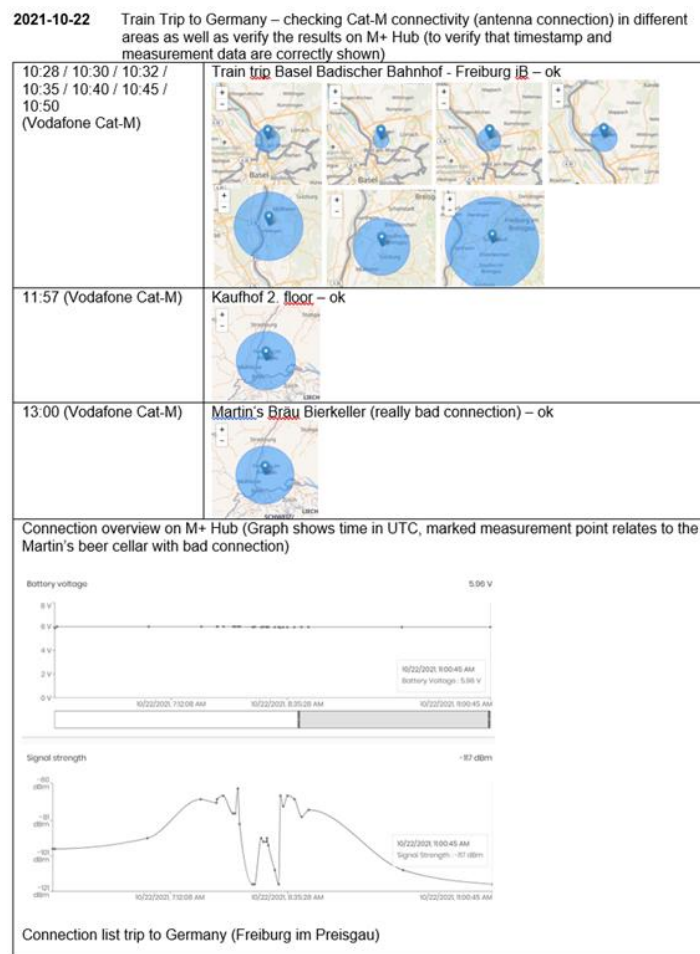


Figure 14 – verification example performed in Switzerland and Germany

The step from Cat-1 to Cat-M was made to reduce the sales price of the cellular devices – the new devices can be now offered for approx. half of the price of the Cat1 devices. Adding new connectors has required additional efforts in the ICCS Hub, in terms, infrastructures security or data harmonization (e.g. new input adaptors in HL7 FHIR Format), so that an increased number of functionalities specific to each hardware integrated could get managed in the cloud (e.g. device time zone, sync, etc.) available for deployment in Gatekeeper. (see more detailed examples in the annexes). During the Gatekeeper project Medisanté has also provided its measurements and has taken feedback to make changes it in collaboration with Gatekeeper. In addition to the measurement format, Medisanté has also provided an architecture summary and security policy (See deliverable 5.4.1).

In terms of architecture, Medisanté had enable to adapt existing architecture to match Gatekeeper's overall architecture requirement. As reminder, the ICCS's method of sending any measurements is via the webhook of the Data Federation. The push API used by the ICCS for measurements is a one-way communication system, more specifically sender-initiated POST request to one or more public subscription URL endpoints using event-driven webhooks. Robust error handling logic exists with an automated retry attempt mechanism where applicable. Undelivered measurements are not included in webhooks if retry attempts are exhausted or unapplicable. Continuous polling (checking API periodically) is not required for requesting new measurements unless retry-attempts for undelivered measurements are exhausted or unapplicable.

Medisanté set up and maintained a new infrastructure for the ICCS in Gatekeeper, which includes a VPN proxy and VPN client inside a virtual private server (VPS). Measurements are still triggered in the same way with webhooks. However, there was also the need for a VPN server (this is outside of the Medisanté system boundary) in addition to the typical webhook endpoint. Referencing stats available by Medisanté, at the time of writing, there are just under 15,000 requests sent to the Gatekeeper proxy, from real devices (200 are listed), i.e. excluding synthetic measurements.

Additionally, work relating to general security of our infrastructure also applies to Gatekeeper. Gatekeeper is considered a tenant of our system, so all general work applies to tenants globally, but this should not be seen with any less consideration or weight because of this.

Recently, work carried out on trying to protect and maintain the security of our infrastructure and service has included a long list of items illustrated in the annexes of this document.

In annexe, efforts regarding development of scripts for data transformation of CVS into data federation have been as well highlighted. While this has been required within Gatekeeper, this is outside of the Medisanté task related to ICCS.

In parallel, efforts have been investigated to explore new integrations of new relevant devices, which could be proposed to care teams in Gatekeeper. Such efforts are related to supplier identification, high level data exchange architecture review, hardware accuracy review and communication coordination. This effort has been lead more specifically for a 4G pulse oximeter, a spirometer, an ECG via patch solution or a cellular Gateway (GW2) connecting via Bluetooth to a broader set of devices from several suppliers.

Medisanté also led efforts to explore new solutions with Gatekeeper partners leveraging this unique eco-system for innovation. Discussions highlighted for instance the challenge for care teams and devices manufacturers to deploy BLE medical devices with certain patients. Patients not owning specific smartphone version used by the medical devices' provider for integration, patients not at-ease with smartphone, or patients not owning a smartphone have been usually exclude of the study cohorts. To overcome such challenges Medisanté explored the idea of building a lock-down tablet used as mobile gateway for BLE devices & managed in the cloud with MDM technology. The proposal (March 2022) would have included Medisanté, Biobeat & Samsung to develop the solution.

Such valuable idea and additional explorations have been set on hold in End of March 2022 to focus on validating of existing technology with patients in Puglia.

In a broader overview, it is important to consider that Europe is lagging far behind the USA, in the deployment of cellular connected devices for remote patient monitoring. It is currently estimated that Europe-USA have a 1:20 factor in term of expertise of such connected medical devices deployed by care team. The industrial delay is due to the lack of incentives & policies to invest into the development of an European digital health infrastructure for virtual care. It is expected to change very rapidly in the next 5 years.

ICCS could be leveraged post-Gatekeeper in additional European Innovative Trial or in local regional care model for secondary preventions.

At Medica in November 2022, Beurer has already announced its willingness to co-invest into such technology with Medisanté for a broad distribution across Europe.

Several European digital health platforms with new digital model of care are scaling up based this innovative approach developed by Medisanté. Oviva (100-million-funded venture fighting against obesity, hypertension, diabetes across Switzerland, UK, Germany, France) or 50+ medical systems have integrated the service. The world largest EHR – EPIC – has as well integrated the service for a trial in Switzerland and additional one in an US hospital.

We expected many more innovative projects going in such direction. Many projects across Europe are already in the process to leverage the innovation developed with the ICCS within Gatekeeper with scientific results expected in the coming years: a leading clinical trial with 600 patients at the German Center of Heart Insufficiency convincing ambulatory cardiologists to prescribe patients into remote monitoring program, a regional RPM program led by GPs in remote German countryside, large hospital@home programs in France with centres of excellence in university hospitals. These projects relying all on the ICCS Hub to learn how to scale new model cares – digitally or multichannel – organized around the needs of patients.

5 IoT web connectors

IoT web connectors are developed in many tasks of the project. A part of T5.4, T4.4 data federation and T7.5 technical pilot deployment have been developing many web connectors for integration of data within the pilot environments.

5.1 Example use case

5.1.1 Update of Web Connector implementation

The implementation by BioBeat involves a transformer that facilitates the extraction and transformation of questionnaire data from the Samsung Platform. The transformer operates by obtaining questionnaire data in JSON format from the Samsung Platform and subsequently pushing it into the GateKeeper FHIR Data Federation in FHIR format.

The step-by-step process unfolds as follows:

- **Token Acquisition:**

To initiate requests to the Samsung Platform, the transformer acquires a token. This token is obtained by reading a file named "creds_token.txt," which contains the necessary credentials, specifically the email and password.

- **GETS Request to Samsung Platform:**

Once the token is secured, the transformer initiates a GETS request to retrieve the questionnaire data from the Samsung Platform.

- **Transformation into FHIR Format:**

The received data is then transformed into FHIR format and saved as a JSON file. If the patient associated with the questionnaire response does not exist, the transformer creates the patient and executes a POST request to establish the patient's presence. In the case where the questionnaire response from the patient already exists, the transformer avoids creating a duplicate response.

- **Pushing to GateKeeper's FHIR Data Federation:**

The FHIR JSON text file, now in the appropriate format, is pushed into the GateKeeper FHIR data federation.

- **Conversion to Tabular Format:**

Upon successful submission to the server, the implementation by BioBeat involves an additional transformer responsible for converting the stored FHIR data into a tabular format. This tabular format is then utilized for statistical analysis.

The overall process ensures the seamless flow of questionnaire data from the Samsung Platform to the GateKeeper FHIR Data Federation, with necessary transformations and checks in place to maintain data integrity and prevent redundancy. This structured approach aligns with the architecture and operational flow outlined in Figure 15, facilitating efficient data handling and analysis.

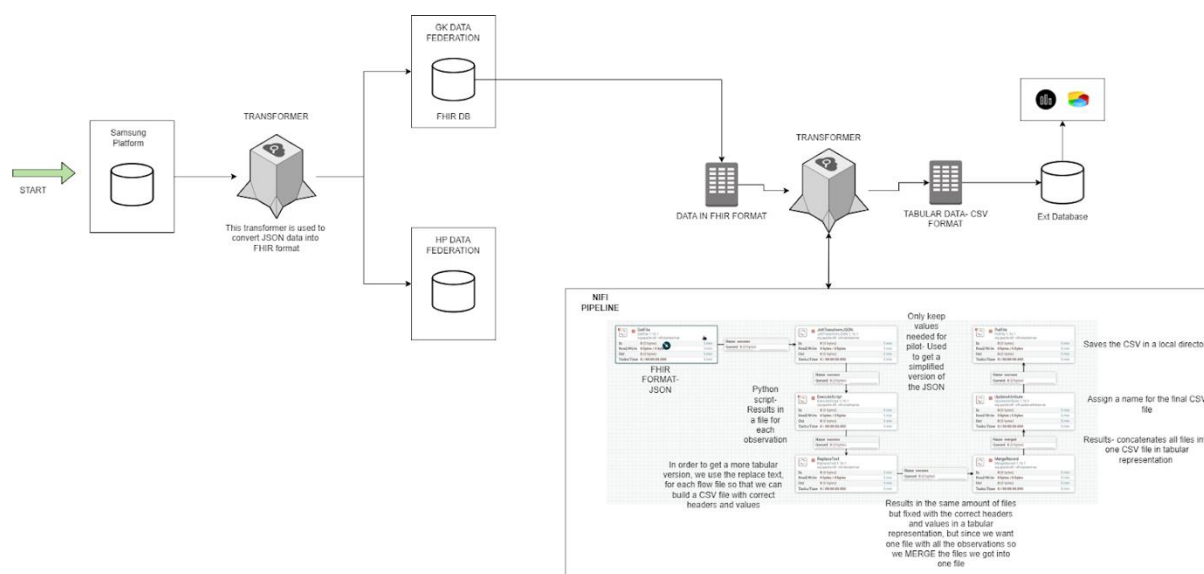


Figure 15 – Architecture of extraction and transformation of questionnaire

Basque Country connectors implementation:

Biobeat developed a module which pulls data from Biobeat's AWS data repository and uses a connector to transform Biobeat's proprietary measurements data to a FHIR patient bundle using LOINC code system. The connector creates a secure connection to the gk-fhir-server using the TOTP based VPN. After a successful connection, the FHIR bundle is posted to the FHIR server.

5.1.2 Update of Web Connector implementation

RUC4: Basque Country StatOn Parkinson data

Sense4Care gathers vital parameters from the StatOn Parkinson's device, for precise data analysis on the Gatekeeper server. This streamlined process enhances insights, supporting healthcare professionals in optimizing Parkinson's care and research.

StatOn data requires synchronization through a mobile device, undergoing an initial transformation into JSON format. The resulting structure follows a defined pattern, ensuring seamless compatibility and facilitating efficient data processing.

Example. Internal Data in JSON format

["2022-10-17":

["hoursDysk":0.0,"hoursINT":1.0,"hoursMonitorized":12.5,"hoursOFF":1.5,"hoursON":5.5,"nEvent":4,"nFOG":54,"nFalls":0,"patientID":"87700"} ...]

The JSON-formatted data undergoes a transformation into a format that aligns with FHIR specifications. By adhering to FHIR format, the proposed standards on the Gatekeeper server are met.

As the application utilized in the project is developed for the Android platform, the chosen connector is HapiFHIR, a versatile open-source Java library. HapiFHIR not only ensures compatibility but also facilitates efficient communication and data exchange, contributing to the overall robustness and effectiveness of the project's implementation. An

observation is generated using pre-defined profiles, converting the earlier JSON data into FHIR format.

Example. Add observation Hours of the patient in dyskinesia in HapiFHIR

```
Observation observationdysk = new Observation();
observationdysk.setSubject(new Reference(patient.id));
observationdysk.setValue(
    new Quantity()
        .setValue("2.0")
        .setSystem("http://unitsofmeasure.org")
        .setCode("h"));
```

Prior to data transmission via Hapi FHIR, for security reasons, it is imperative to acquire a bearer token using a provided key and password. This acquisition is facilitated through a POST request, where the username and password are provided as credentials. The method employed for this involves specifying the grant_type parameter with the value "password". The obtained token serves as the access credential to UPM servers, ensuring a secure and authenticated connection before any data is sent.

Once this security token is obtained, the request can be made using the client provided by HapiFHIR itself: a REST call through the same library that facilitates the sending of a bundle of patient and observation data.

Example. Sending the FHIR Bundle with all the observations to Gatekeeper servers

```
IGenericClient client = FHIR_CONTEXT.newRestfulGenericClient(url_to_gatekeeper);
client.transaction().withBundle(bundle_to_send).execute();
```

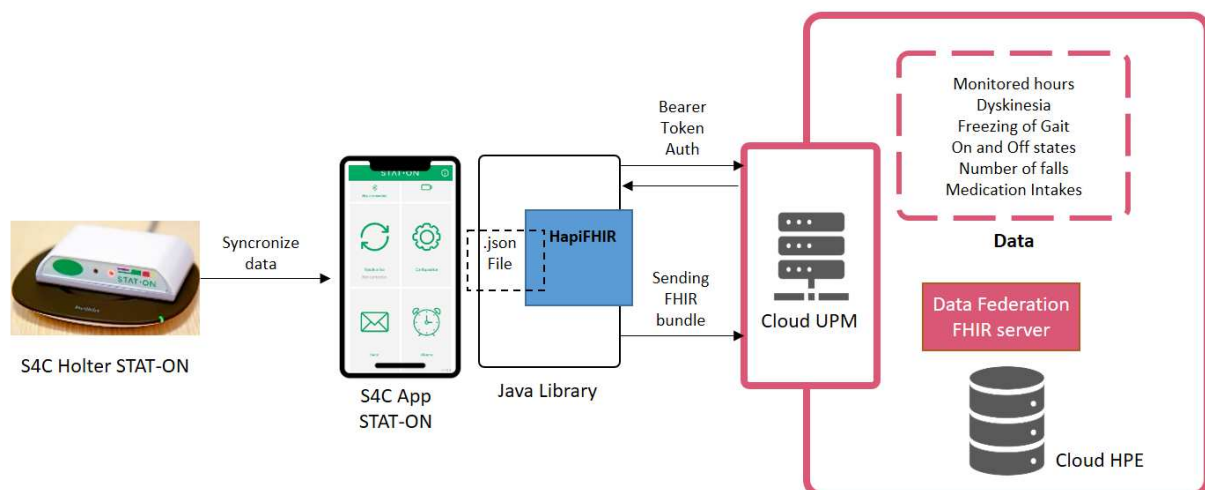


Figure 16 – Web Connector flow RUC4 Parkinson Basque Country

The data flow (figure 16) from the StatOn Parkinson's device to the Gatekeeper data federation involves a streamlined process orchestrated by Sense4Care. Initially, vital parameters are gathered from the StatOn device, and the data undergoes synchronization and transformation into JSON format through a mobile device. To ensure compatibility

and facilitate efficient processing, the JSON data is then transformed into a format adhering to Fast Healthcare Interoperability Resources (FHIR) specifications.

The Hapi FHIR connector, an open-source Java library, is employed to convert the JSON-formatted data into FHIR format. This transformation involves the generation of observations using predefined profiles, exemplified by creating observations such as hours of dyskinesia. To secure the data transmission process, a bearer token is acquired through a POST request using a provided key and password, with the `grant_type` parameter set to "password."

Subsequently, the data, now in FHIR format, is transmitted to Gatekeeper data federation. This is achieved using the Hapi FHIR client, which executes a transaction with a FHIR bundle containing all the observations. The process ensures a secure, standardized, and efficient exchange of data, supporting healthcare professionals in optimizing Parkinson's care and research.

5.1.3 Connectors for MAHA and Data Federation

This section describes the ETL engines that have been developed in Aragon and Basque Country to migrate the custom data collected by Maha app into the GATEKEEPER Data federation following the GATEKEEPER FHIR implementation guide.

The MAHA application defines a data model for managing health-related data, particularly in the context of patient information, clinical activities, observations, symptoms, prescribed medications, comorbidities, and forms/questionnaires. The key components of the MAHA data model are described in Annex A. The model is designed to capture comprehensive health-related data for effective healthcare management and analysis.

The data model is mapped to FHIR standard by using the MAHA connectors.

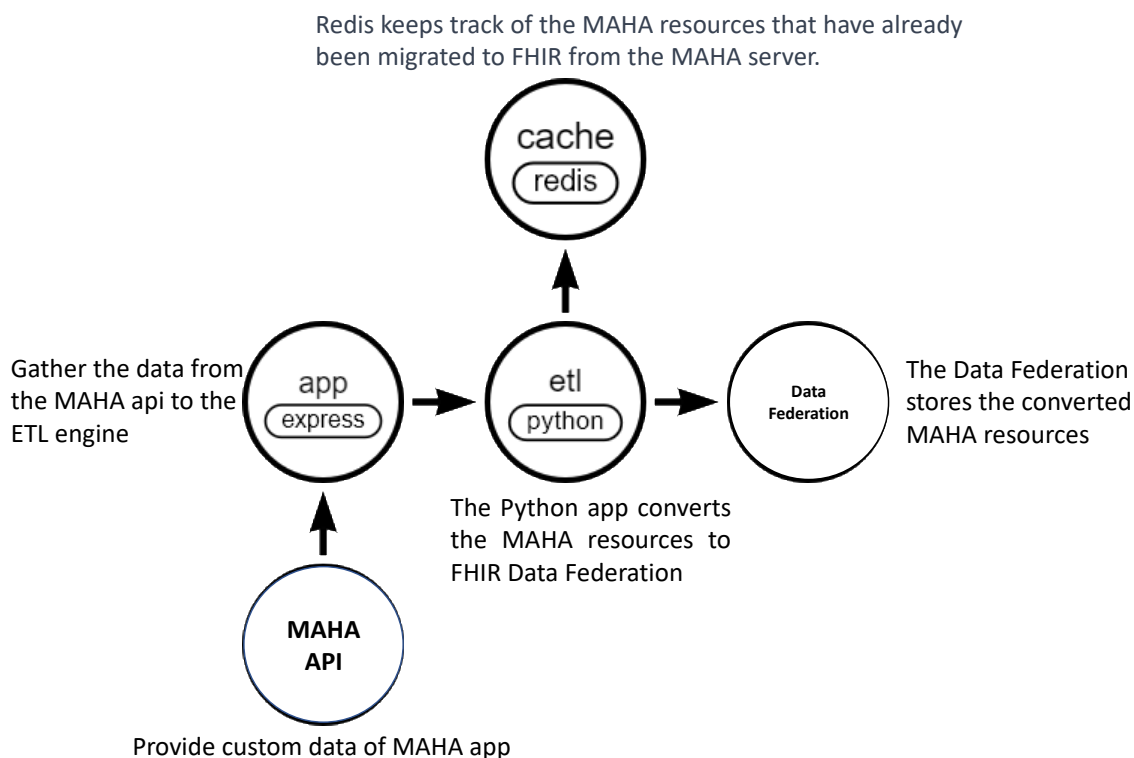


Figure 17 – Maha connectors for data integration into data federation

Following the **Error! Reference source not found.** the connector includes 3 main components:

- An Express app that gathers the data from MAHA apis
- A Redis DB that keeps track of the execution of the ETL engine and the MAHA resources that have been already migrated.
- A Python app that is the core of the ETL engine that converts the MAHA data into FHIR.

App Express

The App Express, is a web application developed in Node.js and Express.js, that serves as a versatile tool for implementing gateway functionalities and REST APIs for the MAHA services. Its key features include a flexible routing system, middleware support, and the ability to act as a reverse proxy. In the context of the connectors, The App Express enables the creation of routes for different endpoints, allowing for seamless handling of tasks such as authentication, request/response transformation, and load balancing. Middleware functions play a pivotal role in implementing features like authentication, logging, and error handling, ensuring consistent and customizable behavior. The app facilitates proxying requests to various backend services, enabling dynamic routing based on conditions like URL paths or headers. Additionally, it supports load balancing and CORS handling, making it well-suited for managing diverse services and clients. The App Express is a robust and adaptable gateway, tailoring it to the specific requirements of the connectors architecture.

Redis DB

Redis, a versatile in-memory data store, it plays a crucial role in optimizing system performance and managing task execution in the MAHA connectors. It serves as a caching layer for frequently accessed data from REST-APIs of the MAHA services, reducing latency by storing API responses in memory. Additionally, Redis manages the status of ETL python engine connector execution, offering real-time updates on tasks such as background jobs or asynchronous processes. Acting as a message broker, Redis facilitates reliable task queue implementation, and its data structures enable real-time analytics for quick analysis of data retrieved from APIs. Furthermore, Redis serves as efficient session storage in web applications, enhancing overall system responsiveness and scalability. Through these capabilities, Redis proves instrumental in improving the efficiency and reliability of systems interacting with RESTful APIs and handling diverse processing scenarios.

ETL python engine

The Python engine can be divided into 2 main parts:

1. Connecting to a Redis database and retrieving data from the local MAHA services
2. Processing and sending data to Data Federation

Connecting to a Redis database and retrieving data from local MAHA services:

This part starts by connecting to the Redis server using the redis library. The ping function is called to check if the connection is successful. If the connection is successful, a message is printed. Otherwise, an error message is printed.

Next, data is retrieved from the server specified in url_get_patient. The requests library is used for this purpose. If an error occurs during the request, the appropriate error message is printed.

Then, a loop is executed over the retrieved data. For each element, a FHIR resource is created and sent to the server specified in `url_post_patient`. The `json` library is used to create a JSON payload. The payload contains a `resourceType`, an identifier, a `generalPractitioner`, and a period. After sending the payload, the ID of the created resource is stored in Redis.

Finally, various values are extracted from the retrieved data and stored in variables.

Processing and sending data to Data Federation:

This part involves processing and sending data to another server. First, the FHIR resource for the observation data is created using the `json` library. The resource contains various values such as subject, status, and code. Then, the resource is sent to the server specified in `url_post_observation`. If an error occurs during the request, the appropriate error message is printed.

After sending the observation data, a loop is executed over the questionnaires specified in `config`. For each questionnaire, the FHIR resource is created using the `json` library. The resource contains various values such as subject, status, and code. Then, the resource is sent to the server specified in `url_post_ques`. If an error occurs during the request, the appropriate error message is printed.

6 GATEKEEPER IoT gateway

The main purposes of the GATEKEEPER IoT gateway are a) to collect events and data from IoT sensors (i.e. MYSPHERA's "LOCS home monitoring sensors"), make simple computation with received data and transmit them to the cloud IoT platform via 4G; b) to collect data from BLE standard healthcare profile medical devices and wearable devices and transmit them to the cloud IoT platform via 4G; c) to integrate security and AI capabilities at the edge.

6.1 Development of IoT Gateway

The original development plan of the IoT gateway has suffered a long delay due to the semiconductor crisis occurred in 2021-2022. Various of the needed components had estimated delivery times of more than a year at the time of acquisition. The updated development plan is shown in the table below:






Table 1: Gantt chart of the development plan of the GATEKEEPER IoT gateway

		May'21 Nov'21	Jun'21 Dec'21	Jul'21 Jan'22	Aug'21 Feb'22	Sep'21 Mar'22	Oct'21 Apr'22	Nov'21 May'22	Dec'21 Jun'22
HloTGW v1	Prototype HW								
	Firmware develop.								
	Integration, testing and reworking								
	Validation for industrial prototyping								
		Sep'22	Oct'22	Nov'22	Dec'22	Apr'23...Sep'23		Nov'23	Dec'23
HloTGW v2	Specifications								
	HW design								
	Firmware design								
	Prototype HW								
	Firmware develop.								
	Integration, testing and reworking								
	Validation for industrial prototyping								

The first delay affected the development and validation of version 1 of the gateway, which was produced and tested in the laboratory with five samples.

The second delay impacted the fabrication of version 2, which was planned to deliver the 30 prototypes to be deployed in the pilot site. At the time of provisioning the components for the fabrication, the following table shows the estimated delivery times we received from the providers.

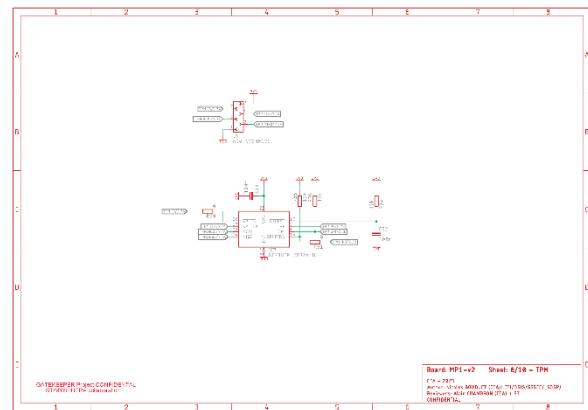
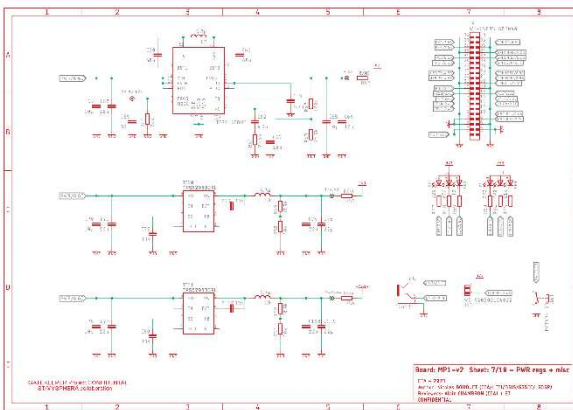
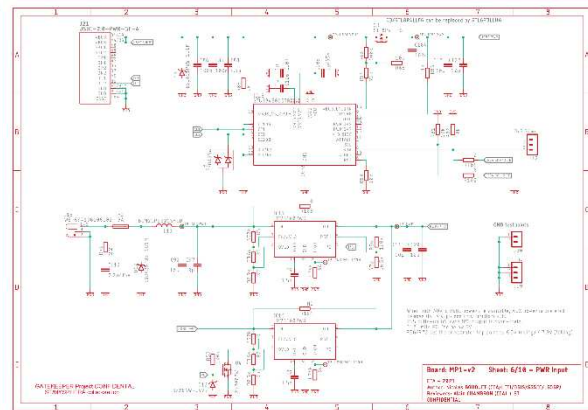
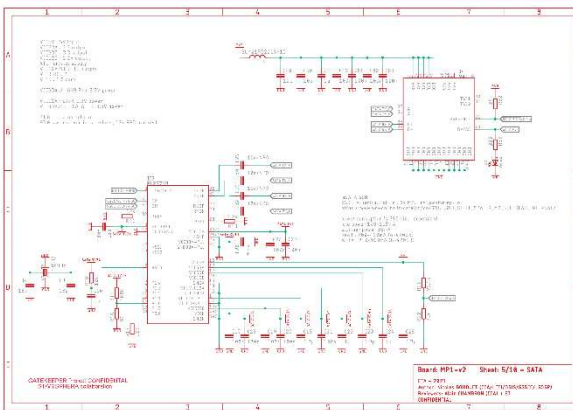
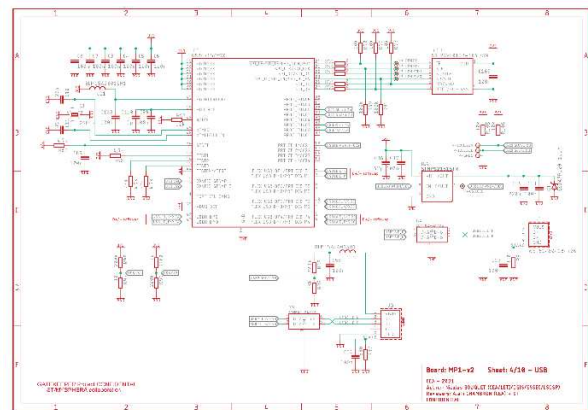
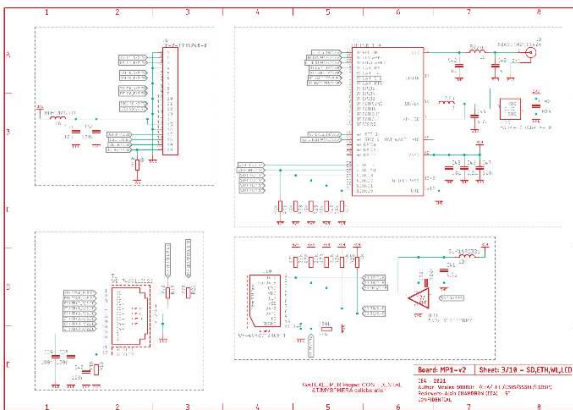
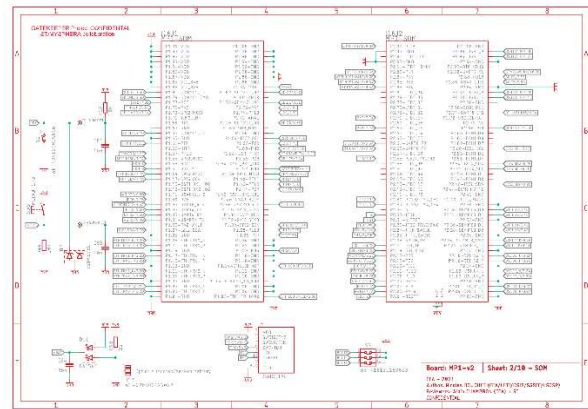
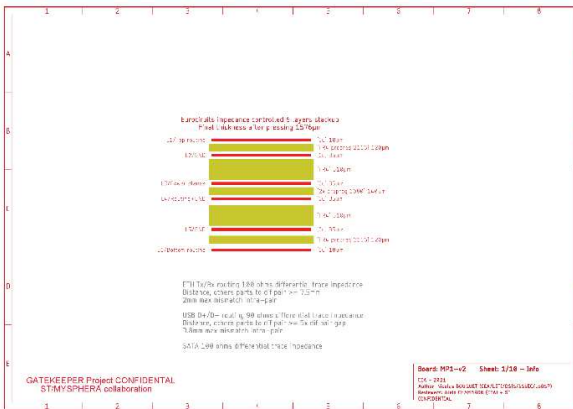
Table 2: Estimated delivery times of GATEKEEPER IoT gateway components

Reference	Qty	Manufacturer	Full part number	Estimated delivery (ordered in nov'21)
REF54	1	MICROCHIP	USB4715/Y9X	14/04/2022
REF55	2	ST	ECMF02-2AMX6	10/04/2022
REF56	1	PROLIFIC	PL2571B	15/06/2022
REF57	1	ST	STMPS2151STR	28/07/2022
REF58	1	Murata	LBEE5KL1DX-883	09/06/2022
REF59	2	Hirose	DF40HC(3.0)-100DS-0.4V(58	28/07/2022
REF60	1	ST	ST33HTPH2032AHD1	12/09/2023 
REF61	3	TI	TPS54202DDCT	08/11/2022
REF62	1	ST	STM32WB5MMGH6TR	05/09/2023 
REF63	1	MICROCHIP	SST26VF016B-104I/SN	29/07/2022
REF64	1	ST	EMIF03-SIM02M8	06/09/2022
REF65	1	ST	STUSB4500QTR	12/09/2023 
REF81	1	Emcraft	SOM-STM32MP1	12/09/2029 
REF83	1	Quectel	LPWA BG96 Mini PCIe	15/09/2023 

At the time of writing this deliverable, we have completed the integration of the prototypes and their validation in the laboratory. Unfortunately, testing them in the pilot site won't be possible.

6.2 GATEKEEPER IoT Gateway Reference Design

A reference design that can be tailored to different needs and requirements has been developed to improve the reusability and exploitation potential. Below, a series of schematics of the hardware design is presented.



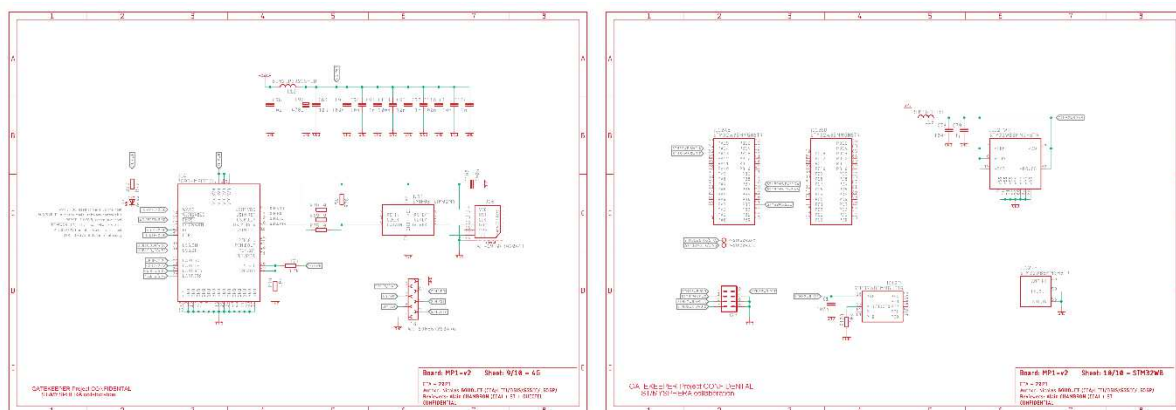


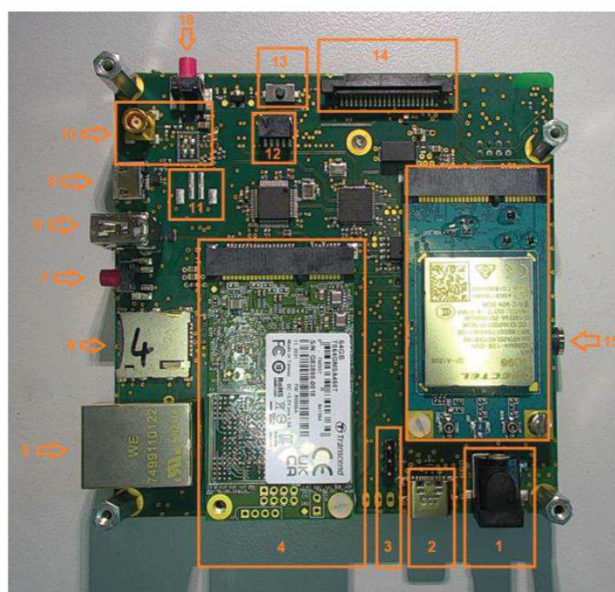
Figure 18: Set of schematics of the GATEKEEPER IoT gateway

The firmware installed is the OpenSTLinux distribution based on the OpenEmbedded build Framework and it runs on the Arm Cortex-A7 processors. This allows to control all the functions of the hardware and deploy any java application in a virtual java machine.

6.3 GATEKEEPER IoT Gateway prototype production

The reference design has been implemented and tested to ensure its feasibility for larger production.

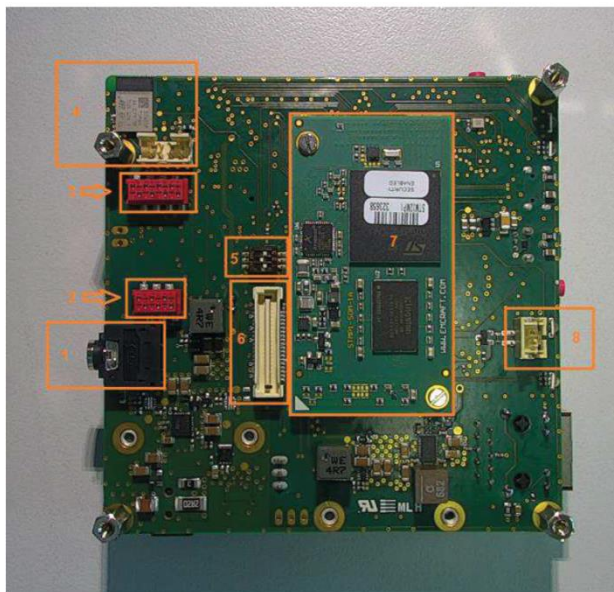
An overview of the board is presented below:



Top view:

- 1- Power supply jack
- 2- USBC power
- 3- J12 (STUSB4500 I2C)
- 4- SSD
- 5- Ethernet
- 6- Micro SD card slot
- 7- User push button
- 8- USB host connector
- 9- USB device connector
- 10- WiFi/BT chip and MCX connector
- 11- GPIO/tamper input
- 12- JTAG connector (MP1)
- 13- Reset button (MP1)
- 14- DSI connector
- 15- 3.5mm stereo jack for FTDI USB cable (linux console)
- 16- Wake up button

Figure 19: GATEKEEPER IoT Gateway top view



Bottom view:

- 1- 3.5mm stereo jack for FTDI USB cable (linux console)
- 2- WR-MM 6 pin connector for external TPM
- 3- WR-MM 8 pin connector for external SIM
- 4- ST WB module and prog. Connector
- 5- BOOT switch config
- 6- Extension connector
- 7- SOM module
- 8- Backup battery connector

Figure 20: GATEKEEPER IoT Gateway bottom view

6.3.1 Functional description

6.3.1.1 MPU device (bottom 7 / S2-M1)

The main component of the board is a dual-core ARM Cortex-A7 (650MHz) + Cortex M4 (209MHz) referenced STM32MP157CAA. This device is located on a 32x59mm mezzanine board manufactured by Emcraft, also called the "SoM" board. This SoM also include 1GB of DDR3L and a 4GB eMMC chip. It is connected to the main board through a pair of 100 pin connectors (IC6).

More information about the SoM can be found on the Emcraft website:

<https://emcraft.com/products/1062>

6.3.1.2 Ethernet (Top 5 / S3-T1)

The board is equipped with a 1Gb/s Ethernet connector. IP can be attributed by a DHCP server or forced by user.

6.3.1.3 USB hub (S4-IC1)

The SoM offering only one USB host port, a hub device (Microchip USB4715) is located on the board to dispatch USB signals to the SATA/USB SSD bridge, wireless modem and USB host connector.

The hub also features a USB to UART interface. The UART is available on the extension connector.

The hub has been configured (programming of a tiny EEPROM attached closed to the chip) during the board testing process.

6.3.1.4 USB host connector (Top 8 / S4-J1)

This connector is connected to the SoM through the USB hub.

6.3.1.5 USB device connector (Top 9 / S4-J5)

This connector is directly connected to the SOM.

6.3.1.6 SSD (Top 4 / S5-J4)

For applications that require a lot of data storage, the board offers an mSATA connector. This connector allows using any kind of SSD. The following models have been tested:

- Transcend TS64GMSA450T (64GB) from RS (#186-4638)
- Swissbit SFSA030GU2AK1TO-I-5S-236-STD (30GB) from Farnell (#3527094)

Using SSD is an option and the card can be used with or without it. Since SSD signaling is SATA-type and the SoM only offer USB, a PL2571B bridge (S5-IC3) is used for signal and protocol translation. The USB signals are then transferred to the SoM through the onboard USB hub.

An orange LED (S5-D2) indicates the status of the disk. Although the LED behavior depends on the disk manufacturer, the most common behavior is as follows:

- OFF: no disk
- ON: disk inserted
- Blinking: disk active

A status signal is wired from the SSD to the MP1.

6.3.1.7 Wireless modem interface (Top 16 / S9-J14)

The board offers an mPCIe connector allowing connection of standardized 3G/4G wireless data modem.

The board has been tested with a Quectel BG96 module.

Some of the signals connecting the modem to the SoM are:

- USB through the onboard hub
- Full duplex with RTS/CTS UART
- WAKE, PERST, RI, DTR signals

A yellow LED (S9-D5) shows the status of the modem.

A nanoSIM connector (S9-J16) is placed on the main board, below the modem. An external SIM card can be connected on the connector S9-J18 (Bottom 3) of the board. This connector is a Würth Electronic, 8 pins micro-module (WR-MM) connector (#690367280876). Matching connector is #690357280876. This connector share the signals of the internal SIM card.

6.3.1.8 SD card (Top 6 / S3-J10)

Micro-SD card can be connected in the adequate slot. The board has been validated with a 50MHz SDIO clock. The SD uses SDIO module 1 (SDMMC1) of the STM32MP1. The board also features a card detect signal. A 100kΩ pull-up resistor is implemented on the board.

6.3.1.9 Murata WiFi/BT (Top 10 / S9-IC5, J2)

The board was designed to use a Murata LBEE5KL1DX WiFi/BT chip. When using this chip, an external antenna must be connected on the MMCX gold-plated connector located near the Murata chip. All 2.4GHz / 50Ω matched antennas are compatible.

The high-speed interface of the chip is connected to the MP1 using the SDMMC2 port.

6.3.1.10 ST wireless device (Bottom 4 / S10-IC12, J29)

A STM32WB5MMGH6TR is located on the top-left of the board (bottom side). This chip allows concurrent Bluetooth and Zigbee communications. This chip is connected to the SoM using a 2-line UART.

6.3.1.11 DSI connector (Top 14 / S3-J6)

A DSI display can be connected using connector J6 (TE connectivity 2-1734248-0). This connector includes 2 DSI lanes, a DSI clock, an I2C interface (connected to I2C1, shared with USB hub) and four signals.

6.3.1.12 LEDs, switches, tamper input and console

2 RGB leds (S7-D4, D6) and a push-button (Top 7 / S7-SW1) are located on the side of the board. A connector wired to a GPIO (Top 11 / S7-J11) is available and can be used as a tamper detection input.

The UART4 is wired to a 3.5mm audio jack connector (Top 15, Bottom 1, S7-J3) or is intended to be used with a FTDI cable (#TTL-232R-3V3-AJ). This UART is the Linux console (115200 bauds).

6.3.2 Final product

After the different iterations and validation tests, a 3D printed case has been designed to hold the board, achieving the final result, ready for larger production. Below some pictures of the final GATEKEEPER IoT Gateway.

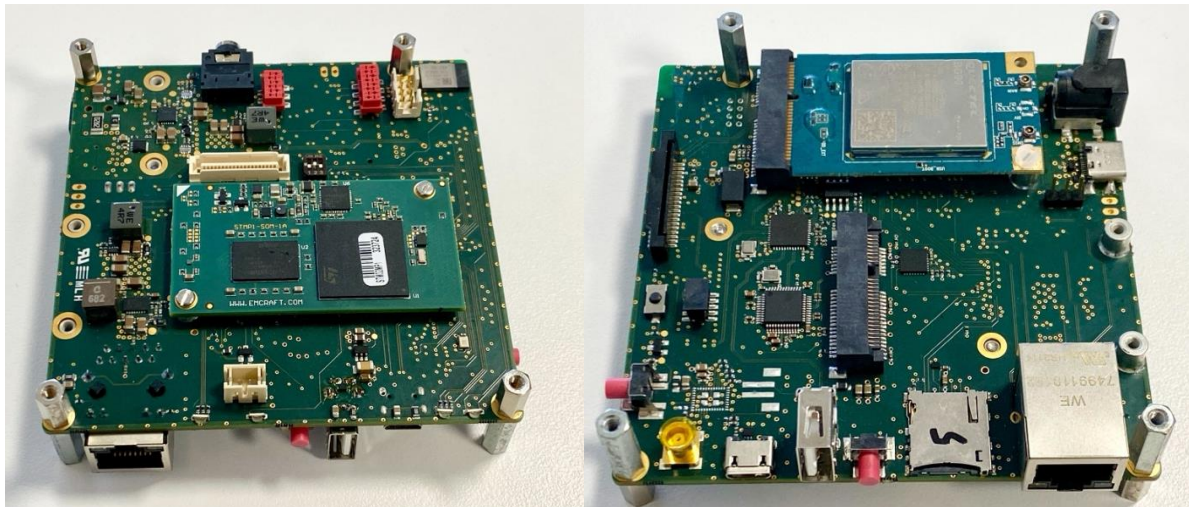


Figure 21: GATEKEEPER IoT Gateway final prototype top and bottom view



Figure 22: GATEKEEPER IoT Gateway final prototype with 3D printed case

7 Conclusions

In conclusion we have significant insights into the successful development and deployment of the GATEKEEPER project, contributing to the broader landscape of the European Health Data Space (EHDS). The adoption of a Bring-Your-Own-Device (BYOD) approach for connectivity is recognized as a pivotal strategy, showcasing its validation in real-world scenarios.

The integration of diverse medical devices, ranging from CAT-M weight-scale to blood pressure monitoring, demonstrates a commitment to providing a comprehensive suite of options for care teams. Our finding emphasizes the importance of addressing challenges, such as high hardware costs and the necessity for a variety of vital parameters, in response to feedback from pilot programs.

The transition from Cat-1 to Cat-M technology is positioned as a strategic move to enhance accessibility by reducing the sales price of cellular devices. The successful integrations and validations in regions like RUC 5 and RUC 7 underscore the scalability and adaptability of the developed technologies.

Medisanté's role is acknowledged, particularly in adapting architecture and security measures, but the broader collaboration with various partners, such as Beurer, signifies the project's collective impact. The integration of the GATEKEEPER service into platforms like Oviva and EPIC, along with announcements from major events like Medica, reflects the project's influence on a global scale.

The document concludes with optimism about the future of remote patient monitoring, anticipating a transformative shift in the next five years. The GATEKEEPER project is positioned as a catalyst for this change, not only in terms of technological innovation but also in fostering a shift towards European digital health infrastructure. The overall success and integration of GATEKEEPER's technologies into existing healthcare ecosystems contribute significantly to the realization of EHDS goals and the advancement of patient-centric, digital healthcare models.

Appendix A MAHA Data Model

Here's a breakdown of the key components of the MAHA data model:

Patient/Practitioner Information:

Identifiers, birthdate, gender, entry/exit dates, primary disease, health area, and death information.

Social Assessment:

Observations related to social factors, such as the number of people living in the household and marital status.

Habits:

Observations related to habits like tobacco use, alcohol consumption, and physical activity.

Clinical Activities (Admissions – Hospitalization):

Information about hospital admissions, including admission and discharge dates, reasons for admission, planned/unplanned admission, and destination after discharge.

Clinical Activities (Consultation):

Information about outpatient consultations, including the date of contact, the person making the intervention, type of contact, and service type.

Prescribed Medication:

Details about prescribed medications, including the status, start date, end date, active principle, dosage instructions, and unit of measurement.

Clinical Variables Value:

Observations capturing clinical variables, their values, and associated units.

Symptoms:

Observations related to symptoms, including their intensity.

Form and Questionnaire (PROMS):

Responses to forms and questionnaires, specifying completion status, completion date, record identification, questionnaire identification, question identification, and corresponding answers.

Comorbidity:

Information about comorbidities, including the date of record, comorbidity description, onset date, end date, and clinical status.