

Deliverable 4.1

SHIFT-HUB SMART HEALTH SOLUTIONS PLATFORM

• Submission date: 27.12.2023

• Author(s): LEDUC, BRYAN & AMIOT, FRANCK





Funded by the European Union Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them. Project: 101095720



Project SHIFHUB: report/deliverable name

Project acronym	SHIFT-HUB		
Project title	Smart Health Innovation & Future Technologies Hub		
Name	SHIFT-HUB Smart Health solutions platform		
Number	4.1		
Work package	WP4 - SHIFT-HUB platform and demonstration framework		
Due Date	31/12/2023		
Submission Date	27/12/2023		
Lead Partner	СОР		
Contributing partners	S2i		
Author name(s)	Bryan Leduc, Franck Amiot		
Version	1.0		
Status	Final		
Туре:	 R - Document, Report DMP - Data Management Plan DEC - Websites, patent filings, videos, etc 		
Dissemination level:	PU - PublicSEN - Sensitive		







		Document History	
Version	Date	Modified by	Comments
0.1	6/12/2023	Bryan Leduc	First release
0.2	8/12/2023	Franck Amiot	Update
0.3	12/12/2023	Francisco Javier Casado Hebrard	Internal review
0.4	22/12/2023	Bryan Leduc + Franck Amiot	Implement changes
1.0	22/12/2023	Alice Seim	Final review

Abstract

The Deliverable D4.1 "SHIFT-HUB Smart Health solutions platform" presents and explains the structure of the Cleyrop platform, how it works, as well as how it will be used within the SHIFT-HUB project.

Keywords

SHIFT-HUB, Cleyrop, data, dataset, platform





DISCLAIMER

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them.





Table of contents

1	Intro	oduction6
2	SHI	-T-HUB Needs definition for WP47
3	SHI	-T-HUB Requirements definition for WP49
4 plat	The tform	proposed solution: Integration with Cleyrop DataHub, a unified and sovereign covering the entire data life cycle inside Mipih SNDS environment10
5	Stru	cture & functionalities11
5	.1	Functional requirement coverage11
5	.2	Technical requirement coverage15
5	.3	Roles inside the platform16
6	Lon	g term sustainability17
6	.1	Hosting and maintenance17
6	.2	Run Error! Bookmark not defined.
6	.3	Reversibility

List of Figures

Figure 1 The steps of the data within a framework of trust	10
Figure 2 Outlines and principles of the solution	10
Figure 3 The three pillars of the Cleyrop offer	11
Figure 4 Screen for adding a data source	12
Figure 5 Data access management screen	13
Figure 6 Dataflow editing screen	14
Figure 7 Diagram of the functional principles of the platform	15
Figure 8 How GAIA-X data sources are integreted into the platform	16





Abbreviations and Acronyms

Abbreviation, Acronym	Description
ΑΡΙ	An API, or Application Programming Interface, is like a messenger that allows different software applications to communicate and share information. It defines the rules and methods for how these applications can interact, enabling them to work together seamlessly. APIs are crucial for building connections between different software systems, facilitating data exchange and functionality.
DMZ	A DMZ, or Demilitarized Zone, in the context of computer networks, is a segmented or isolated area of a network that acts as a buffer zone between the internal secure network and the external, untrusted network (usually the internet). It serves to enhance security by placing additional barriers between sensitive internal systems and potential external threats.
IDE	Integrated Development Environment. It is a software suite that combines essential tools for software development into a unified platform. An IDE typically includes a code editor, debugger, compiler, and other features to streamline the process of writing, testing, and debugging software applications.
SAAS	Software as a Service. It refers to a cloud computing model where software applications are provided and accessed over the internet on a subscription basis. Users can access the software through a web browser without the need for local installation, and updates and maintenance are managed by the service provider.
SQL	Structured Query Language, is a standardized programming language used for managing and manipulating relational databases. It provides a set of commands for querying, updating, inserting, and deleting data in databases, as well as defining and modifying the structure of the data. SQL is widely used in database management systems, and its syntax is designed to be easily understood and executed by both humans and machines.
VPN	A VPN, or Virtual Private Network, is a technology that enables secure and encrypted communication over the internet by creating a private network connection, often used to protect users' privacy, data security, and anonymity.
WP	Work Package





1 Introduction

The SHIFT-HUB project aims to set up a pilot platform providing a controlled and thus safe environment - a sandbox - for (1) the collection and the access to secure and anonymized data provided by patients and citizens based on criteria defined and implemented under the strict supervision of practitioners and regulation experts, (2) the development and integration of Smart Health applications and repositories and (3) as a demonstration and experimental framework to connect application providers with the professional and private end users (practitioners and respectively citizens).

The platform will have the mission to offer a sandbox consisting in the adequate infrastructure and the associated tools to support the development of Smart Health products and services by facilitating the access to data and the interactions and solution co-development between the providers and users. Based on tailored scouting activities, relevant Smart Health applications providers and repositories will be identified and invited to join the community and integrate their solutions in the Hub's portfolio that will progressively evolve into a Smart Health solutions marketplace.

On the side of the professionals - Smart Health solution providers and respectively practitioners - the platform will offer a trusted workspace to foster the co-development, testing and uptake of the solutions, while creating awareness, literacy and trust for the patients and citizens and collectively accelerating the learning curve for all the community members.





2 SHIFT-HUB Needs definition for WP4

Within the project, Work Package (WP) 4 is responsible for the development and implementation of the SHIFT-HUB Smart Health solutions platform. To this end, WP4 includes four dedicated tasks:

T4.1: Setting up a Digital Health solutions lighthouse infrastructure and framework

This task aims to design the platform architecture and implement all the modules and features allowing the integration and management of data and Smart Health applications. First, the platform will integrate a Health Data space pilot – a virtual database following a federated model, built in accordance with the vision and guidelines of the future European Health Data Space and by leveraging the Gaia-X architectural reference, standards, and best practices to foster interoperability and allow an easy replication and integration of new nodes in the SHIFT-HUB network. The SHIFT-HUB Smart Health Data Space will integrate mechanisms to certify the data provided by the patients and citizens involved in the activities, in collaboration with the healthcare practitioners, guarantee its traceability and validation along with a high-level security of the data access, offering thus a trustworthy and interoperable experimental environment. On top of the data base, a search engine along with modules and features for standardized data analysis will facilitate its exploitation for the development of applications, while guaranteeing a strict compliance with the data protection legislation and ethical standards according to the existing and future regulatory framework. Based on the current activities of the task leader and the consortium members in collaboration with their direct partners (public and private health institutions in France and Germany, as well as key stakeholders of the European Could and Data community and particularly the ones involved in Gaia-X) the platform will serve as a lighthouse allowing to enroll new nodes across other European countries and regions, progressively expanding the outreach, adoption and uptake.

T4.2: Scouting existing digital health solutions repositories:

This task will identify all relevant Smart Health apps and repositories by browse databases such as the one established by https://digitalhealtheurope.eu where 65 mature digital solutions resulting from EU-funded projects and initiatives that address challenges related to key policy priorities are included. A complete mapping will be elaborated with apps classified by thematic area and will be made available to the community through a catalogue.

T4.3: Linking repositories and integrating digital health:

This task aims to design and develop the tools needed to integrate Smart Health solutions smoothly and seamlessly in the framework, in particular by linking the experimental framework. A particular emphasis will be put at this level on interoperability and cybersecurity, as the applications will involve an access to the Data Hub. All the layers of the stack will be developed to allow the Apps to access needed data, along with ergonomic search tools to the users to identify relevant sources of data.

T4.4: Set up a marketplace for the uptake of Smart Health solutions:

This task focuses on the design and development of the upper layers of the sandbox, with the objective to support the "test before invest" logic and consequently create on the longer term a marketplace to foster the uptake of Smart Health solutions. Based on the integration





mechanisms and the certification process defined in the scope of T4.3, layers including adequate APIs will be implemented for on-boarding Smart Health applications and service providers identified in T4.2 will be implemented based on a highly modular and lightweight architecture. An easy to use and user-friendly Web access will be offered to the community members to integrate and make available their solutions, including the possibility to access the Health Data Space for the implementation of their features.

We can summarize the needs as follows:

- Data providers need to store their data in a central repository.
- The data repositories should be in line with the security expected to host health data.
- Data providers need to control over their data patrimony to keep confidentiality and to be in line with European regulatory norms.
- Smart Health data apps need to access these prepared data respecting permission schemas defined by data providers.
- Smart Health data apps will be accessible to end users using a repository accessible via a web site on which they can download the app, install it, and test it. The underlying mechanism to interconnect the Smart Health Apps with the Smart Health Data Hub should be transparent.





3 SHIFT-HUB Requirements definition for WP4

The main requirements for the SHIFT-HUB Smart Health solution platform are:

3.1 Functional requirements

On one side the SHIFT-HUB Data Space will allow data providers to create with their data a data repository on which they will have fine grained permission mechanism to keep control. On the other side, those data repositories will be exposed to end users via Smart Heath Apps that will consume prepared data following a permission schema decided by the data providers. This data manipulation journey will require the following features:

- Storage: data providers should have a secured and controlled data storage
- **Data Cleaning & Preparation**: data providers should be able to prepare or clean their data via low code / no code transformation for a better use on Smart Health App side.
- **Data governance:** data providers should have an advanced mechanism to manage data validation, accessibility, and protection.
- **Data search engine**: data providers and app developers should have a catalog that allows indexing & classification of the data to have a strong follow up of data patrimony and to make efficient searches.
- **Data lineage**: data providers who are comfortable with data distribution should have a guaranty on the traceability of data manipulation & usage.
- **Data access:** Smart Health App developers should have an easy access to the list of datasets and to the data.

3.2 Technical requirements

- **Data exchange interoperability**: The data should be pushed directly from the data providers, but the system should be compatible with dataspace like IDSA or Gaia-X to extend the SHIFT-HUB data sources to a wide range of data spaces.
- Smart App integration: Smart Health Apps (Android, IOS or Web) should have access to the SHIFT-HUB data space via a standard way (token + API).
- **Sovereignty:** Data providers should not be under the US Cloud Act.
- **Trusted environment**: Patient & medical data are highly sensible, hosting should respect SNDS & GDPR regulations.
- **Industrial:** the consortium of European companies working together should have access to a reliable, highly available and easy-to-maintain system.
- The SHIFT-HUB Data Space should be accessible with pre-configured network options





4 The proposed solution: Integration with Cleyrop DataHub, a unified and sovereign platform covering the entire data life cycle inside Mipih SNDS environment

WP4 proposes to create a platform with the Cleyrop Data Platform deployed inside the Mipih SNDS environment. The Cleyrop Data Platform offers a full data processing journey from injection to exposition with integrated governance and data ops (Figure 1), within a framework of trust.

The data platform is composed of several layers sharing the data organization setup by the data platform manager. All these layers are deployed on top of a Kubernetes network running on several VMs hosted inside the Mipih Data Centre offering SNDS certificate. All modules of the solution will be deployed automatically using infra as code approach.



Figure 1 The steps of the data within a framework of trust



Figure 2 Outlines and principles of the solution





5 Structure & functionalities

Globally, the proposed solution articulated around the Cleyrop data platform hosted on a Mipih SNDS environment will cover the requirements expressed in Chapter 3.

On one side, it will offer data providers all necessary tools to manage their data with a high degree of control and security in a simple, intuitive, and easy to adopt industrial solution. On the other side, security aspects will be covered, too via the industrial environment running inside Mipih data centre.



Figure 3 The three pillars of the Cleyrop offer

5.1 Functional requirement coverage

Cleyrop is offering turnkey functionalities.

Collect & Storage:

Data providers will create a Data Source and will create data sets to store their data.

The benefits of a single, unified, and secure access point to data are:

- Collect data industrially with a library of connectors and multiple entry points
- Easily add sources and datasets in just a few clicks to get started with the first use cases
- Simplify the management of the Data architecture and control costs thanks to integrated storage infrastructure







I

»	Ajo Suive	ut d'une source de doni z simplement les 4 étapes pour compl	née léter le processus d'ajout d'une source de données.		
くこれる第100	(c) (c) (3) (4)	Choix du hype de connecteur Sedenmer, l'instrume adu hype de donnée Configuration de la consection Configuration de la source consection Configuration de la source consection de la source consection de la source consection de la source consection de la source de données consection consection	Les champs suivis d'un * sont obligatoires. Nom de la source de donnée * Le rom attribué sera utilité sur la platforme Cleyrop Responsable Courre Sayag Sensibilité des données * Publique Sensibilité des données * Eubel * X	Description	
9 2		votre source de données à la plateforme Cleyrop.	Annuler		Précédent Suivant

Figure 4 Screen for adding a data source

Governance & Collaboration

Data providers will organize their data patrimony via security policy setup and catalog management. Data providers can work in a collaborative manner to create transformations if necessary.

Freeing up uses and organizing data assets within a framework of trust enables:

- Collaboration in watertight project spaces with tagging and notification capabilities.
- Easily understanding and exploring data with a secure catalog system.
- Guarantee of organization's security policies through detailed management of rights and authorizations.





Hub France	Gestion des approbations				
Aub Funce	Approbations en attente Historique				
Configuration	O Rechercher				
유 Utilisateurs					
Jeux de données	Nom 0	Type de demande 🗢	Sensibilité 0	Demandeur 0	Action
↓1 transterns		Publication au catalogue	Besterlete	Arlene McCov	G
오 Datasources	Rolling Forecast	Publication au catalogue	Restreinte	Dianne Russell	G
Gestion approbations	ERP_Actions Fournisseurs	Publication au catalogue	Interne	Courtney Henry	6
		Publication au catalogue	Restreinte	Robert Fox	6
		Demande d'accès au dataset	Sensible	Cody Fisher	6
	Pilotage RH	Demanae a acces au aaraser	Sensible		G
Notifications					
Administration					
名 Martin Dupont					
🥣 cleyrop v 1.8					

Figure 5 Data access management screen

Treatment & Exploration

The data providers will likely need to rework their original data to provide the best possible data quality for Smart Health Applications. To do it they need exploration & transformation tools.

By facilitating the work of Data teams and the use of businesses via integrated end-to-end services, the users can

- Explore and analyze data with Python and SQL queries.
- Work in parallel, develop models and process scripts, and version changes.
- Design, organize and manage jobs and pipelines via Data Flow.







← → C	:
» < Nom du dataflow 🙆 🛇 😫	Agir sur la proposition
Proposition 123 En attente de validation	
€ J2 Main	
%	Titre de l'output
Δ	
8	
=	
ا لا	
© Transformation	Titre de l'output
Titre du dataset	
©	
0 0 0	

Figure 6 Dataflow editing screen

Analysis & Exhibition - Serving

Data Sets stored by the data providers will be accessible via the serving layer.

Data and analyses can be made available to everyone via dashboards or transport layers such as REST API or push via messaging. The Smart Health Apps will each have a token, generated by the application creator directly in the Datahub interface. It represents the authorization of access data to one or more datasets.

By providing this token during all exchanges with the API, the Smart App will be authorized to consume the data but will also allow consumption to be monitored.

The API will be requested in the following way: URL-DATAHUB/TOKEN/ACTION

The token can be revoked at any time and in this case, the API will no longer respond.









Figure 7 Diagram of the functional principles of the platform

5.2 Technical requirement coverage

The proposed solution relying on the Cleyrop data platform will provide by design standards:

- Principles of security and privacy-by-design to enable the processing of highly sensitive data in compliance with French and European regulations (GDPR, Data Act, etc.).
- Principles of traceability of data usage based on one side on security policies approach and on data lineage on the other side. Because managing sensitive data on highly distributed environment is quite challenging, the platform can provide additional protection on data lineage using DLT.
- Technical observability
- Standardized connector like GAIA-X / IDSAto to connect to the different dataspace







Figure 8 How GAIA-X data sources are integreted into the platform

The industrial deployment on Mipih infrastructures offer additional guarantees.

- SNDS trusted cloud.
- 100% European technology guaranteeing impermeability to extra-territorial laws.
- 100% of the technical operations are run automatically using infra as code (MCO includes)

5.3 Roles inside the platform

There will be three users of the platform.

Two will use the interface directly and one will retrieve the data via the API

Data Providers can easily:

- add data to the platform
- manage access to data that belongs to them
- update data via a processing and orchestration tool

Application developers are able to:

• List and find the data available in the platform in a simple and efficient manner

Application users will:

• access transparently and efficiently the data necessary for the application they use







6 Long term sustainability

6.1 Hosting and maintenance

Cleyrop and Mipih will host and maintain on-line the data platform during the 3 years of the SHIFT-HUB project.

6.2 Continuity of the platform

Before the end date of the phase of the project, Cleyrop, Mipih and the consortium should reach an agreement about the continuity of this platform. This agreement should cover:

- Sizing of the target platform
- Cost of the platform
- Service Level agreement

6.3 Reversibility

The central part of the SHIFT-HUB Data repository is the Cleyrop DataHub, in this DataHub the data providers will store data using the original format and they can use if needed the development interface to develop transformation useful to clean or to model their data sets.

At the end of the project, Cleyrop is engaged to:

- remove original set of data
- package written script and data set provided using text files and csv files







SHIFT-HUB SMART HEALTH SOLUTIONS PLATFORM



Follow us on LinkedIn See more



Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency (HADEA). Neither the European Union nor the granting authority can be held responsible for them. Project: 101095720

