



Privacy compliant health data as a service for AI development

Grant Agreement Number: 101095384

D5.5: Integrated Data Services v1

Deliverable Identifier:	D5.5
Deliverable Version:	1.0
Status	Final (F)
Work Package:	WP5 Health Data Hub
Task:	Task 5.4 Service Integration
Author(s) and Organisation:	Rosario Catelli (ENG), Geert Machtelinckx (FJBE), Yves Lepère (FJLU)
Peer Reviewer(s):	Gonçalo Gonçalves (INESC TEC), Kassiani Zafeirouli (AIN)
Deliverable Due Date:	2025/03/31
Deliverable Submission Date:	2025/03/31
Dissemination Level:	PU: Public
Funding Authority:	European Commission
Funding Program:	Horizon Europe Health Work Programme 2021 – 2022
Topic:	HORIZON-HLTH-2022-IND-13-02
Rights:	PHASE-IV-AI Consortium

Document Control History

Version	Date	Edited by	Modification reason
0.1	2025/02/26	Rosario Catelli (ENG)	Table of Contents
0.5	2025/03/14	Rosario Catelli (ENG)	Draft version ready for reviews
0.8	2025/03/24	Rosario Catelli (ENG)	Reviewed version
1.0	2025/03/31	Rosario Catelli (ENG)	Final version

Executive Summary

The PHASE IV AI project aims to develop a privacy-compliant Health Data Hub (HDH) that enables secure and interoperable data exchange for AI development in healthcare. Deliverable D5.5 – Integrated Data Services v1 focuses on defining the technical integration framework necessary to unify various HDH components into a seamless and functional ecosystem.

Task 5.4 of the project has developed a Service Integration Framework that establishes the methodologies, processes, and standards required to ensure interoperability, security, and compliance with regulatory frameworks such as GDPR, EHDS, and the AI Act. The framework relies on federated identity and access management, using Decentralized Identifiers and Verifiable Credentials to enhance secure authentication and authorization. It also defines standardized APIs and data exchange mechanisms, ensuring compatibility across services through RESTful APIs, FHIR-compliant data models, and secure messaging protocols.

A key focus of this deliverable is the integration of multiple services, including the Data Catalogue, Identity Wallet, Payment Wallet, Software Certification, and Onboarding Services. These components work together to enable secure data transactions and AI service deployment. The HDH is built on a decentralized infrastructure, ensuring a scalable and resilient ecosystem that supports peer-to-peer communication, cryptographic security, and automated deployment mechanisms. To maintain compliance and security, the framework incorporates logging, auditing, and observability features, ensuring transparency and robust monitoring.

As the first version of the integration framework, this deliverable lays the foundation for further refinements in subsequent iterations, particularly in D5.6. Future developments will incorporate real-world validation and stakeholder feedback, improving service interactions and optimizing performance. The work carried out in Task 5.4 contributes to the establishment of a trusted, AI-driven health data marketplace, fostering secure data sharing, innovation, and collaboration across the European healthcare ecosystem.

Disclaimer

The content of the publication herein is the sole responsibility of the publishers and it does not necessarily represent the views expressed by the European Commission or its services.

While the information contained in the documents is believed to be accurate, the authors(s) or any other participant in the PHASE-IV-AI consortium make no warranty of any kind with regard to this material including, but not limited to the implied warranties of merchantability and fitness for a particular purpose.

Neither the PHASE-IV-AI Consortium nor any of its members, their officers, employees or agents shall be responsible or liable in negligence or otherwise howsoever in respect of any inaccuracy or omission herein.

Without derogating from the generality of the foregoing neither the PHASE-IV-AI Consortium nor any of its members, their officers, employees or agents shall be liable for any direct or indirect or consequential loss or damage caused by or arising from any information advice or inaccuracy or omission herein.

Copyright message

©PHASE-IV-AI Consortium, 2023-2026. 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. Reproduction is authorised provided the source is acknowledged.

Table of Contents

1. Introduction	8
1.1 Purpose of the document	8
1.2 Reference documents.....	8
1.3 Definitions	8
1.4 Structure of the document.....	9
1.5 List of Acronyms.....	9
2. Background: The Health Data Hub and constituent components	11
2.1 The Health Data Hub.....	11
3. Service Integration Information Collection framework	15
3.1 General Information	15
3.2 API & Communication Details.....	16
3.3 Data Exchange & Interoperability	16
3.4 Identity, Security & Compliance	17
3.5 Infrastructure & Hosting.....	17
3.6 Integration Patterns & Technologies	18
3.7 Testing & Monitoring.....	18
3.8 Additional Comments & Open Issues	19
4. Services Integration Details	20
4.1 Diagram	20
4.2 Data Catalogue Service	22
4.3 Services Deployment Service.....	23
4.4 Service Creation	24
4.5 Onboarding Service	25
4.6 Software Certification Service.....	26
4.7 Identity Wallet Service	27
4.8 Payment Wallet Service.....	28
5. Conclusions	30

Table of Figures

Figure 1: System Context C4 Diagram (Level 1).....	11
Figure 2: Health Data Hub C4 Container Diagram (Level 2)	14
Figure 3: Health Data Hub C4 Components Diagram (Level 3)	21

Table of Tables

Table 1: General Information	15
Table 2: API & Communication Details	16
Table 3: Data Exchange & Interoperability	16
Table 4: Identity, Security & Compliance	17
Table 5: Infrastructure & Hosting	18
Table 6: Integration Patterns & Technologies	18
Table 7: Testing & Monitoring	18
Table 8: Additional Comments & Open Issues	19
Table 9: Data Catalogue Service Integration Details	22
Table 10: Services Deployment Service Integration Details	23
Table 11: Service Creation Integration Details	24
Table 12: Onboarding Service Integration Details	25
Table 13: Services Certification Service Integration Details	26
Table 14: Identity Wallet Service Integration Details	27
Table 15: Payment Wallet Service Integration Details	28

1. Introduction

This introductory section is intended to provide information regarding the purpose of the document and its structure, as well as any references, definitions and acronyms adopted, as outlined in the sections to follow.

1.1 Purpose of the document

This document serves as the first version (v1) of the Integrated Data Services deliverable (D5.5) within Task 5.4 – Service Integration of the PHASE IV AI project. Its primary objective is to define the technical framework, methodologies, and processes required to integrate the diverse components of the Health Data Hub (HDH) into a unified, interoperable, and secure ecosystem. Task 5.4 focuses on ensuring that all the services can be seamlessly accessed, managed, and transacted through the HDH. This deliverable outlines the service integration strategy, interoperability mechanisms, data exchange standards, and security compliance measures necessary to facilitate efficient and privacy-preserving transactions within the federated health data infrastructure. Additionally, the document provides an initial validation and testing strategy for integrated services, setting the foundation for future iterations (D5.6) that will refine and enhance the integration framework based on real-world implementation and feedback, outlining working approaches adopted time-by-time to make things work.

1.2 Reference documents

[1] PHASE IV AI deliverable *D5.3: Health Data Hub Design and Data Market v1*

1.3 Definitions

List of Definitions	
Data Hub	A data hub takes care of the flow of data between source/target systems and users. The goal is to indicate exactly what actions need to be performed with the underlying data, where systems can distribute data through the data hub. The data hub concept brings some structure in the integration between peers with whom data needs to be shared.
Data Marketplace	A multi-sided place (i.e., platform) where data providers and data consumers can find each other to stimulate data exchange or access.
Data Platform	An environment that facilitates the exchange of value between two or more parties, with the multiple parties interacting through the platform.
Data Sovereignty	The capability of a person or organization to make all data-related decisions on their own.
Data Space	A decentralized infrastructure for transparent and trustworthy data sharing and exchange in data ecosystems within a certain application domain, based on commonly agreed principles and capabilities, consisting of data platform(s), data marketplace(s), and data sovereignty.
Decentralised Identifiers (DIDs)	They are a type of globally unique identifier that enables an entity to be identified in a manner that is verifiable, persistent (for as long as the DID controller desires) and does not require the use of a centralized registry. DIDs enable a new model of decentralized digital identity that is often referred to as a self-sovereign identity. They are an important component of decentralized web-applications.
DePIN	Decentralised Physical Infrastructure Network, the decentralised technology aiming at creating decentralisation in the setup of a physical hardware network.
Federation	A group of participants (e.g. multiple data producers and consumers, model producers and consumers) interconnected with agreed governance, access, and security rules.

List of Definitions	
Flist	<p>An flist is a script to describe a software workload, so that it can be deployed fast and with high reliability.</p> <p>In a flist, we separate the metadata from the data, where the metadata is a description of the files in that image, providing information about the app/software. Thanks to flist, there is no need to install a complete software program to run properly. Only the necessary files are installed. Zero-OS, the federated ‘DePIN’ operating system that we chose to be used as a foundational layer for the data space can read the metadata of a container and only download and execute the necessary binaries and applications to run the workload when it is necessary.</p> <p>A flist is referred to when registering the deployment of the software on blockchain infrastructure, proving the authenticity of the workload.</p>
Verifiable Credential (VC)	<p>They are digital credentials which follow the relevant World Wide Web Consortium open standards. They can represent information found in physical credentials, such as a passport or license, as well as new things that have no physical equivalent, such as ownership of a bank account. They have numerous advantages over physical credentials, most notably the fact that they are digitally signed, which makes them tamper-resistant and instantaneously verifiable. Verifiable credentials can be issued by anyone, about anything, and can be presented to and verified by everyone. The entity that generates the credential is called the Issuer. The credential is then given to the Holder who stores it for later use. The Holder can then prove something about themselves by presenting their credentials to a Verifier.</p>
Verifiable Presentation (VP)	<p>A Verifiable Presentation is the object, derived from a Verifiable Credential (VC) which is presented to a verifier. It allows an individual (the holder) to selectively share parts of their credentials to prove certain aspects of their identity or qualifications, while maintaining privacy by keeping other information hidden.</p> <p>The verifier can independently validate the authenticity of the credentials using the cryptographic signatures included in the VP.</p>

1.4 Structure of the document

This document serves as a support for service integration operations and is divided into two main sections: section two, which summarizes the key points of the Health Data Hub and its services in need of integration, and section three, which instead emphasizes the information gathered through the technical framework, methodologies and processes adopted to manage the integration phase. Finally, in the conclusions, a possible future evolution scenario will be outlined downstream of what has been done.

1.5 List of Acronyms

List of Acronyms	
ABAC	Attribute-Based Access Control
AES	Advanced Encryption Standard
AI	Artificial Intelligence
API	Application Programming Interface
AWS	Amazon Web Services
DaaS	Data as a Service
DePIN	Decentralised Physical Infrastructure Networks
DID	Decentralized Identifier
EHDS	European Health Data Space

List of Acronyms	
FHIR	Fast Healthcare Interoperability Resources
FList	File List
GCP	Google Cloud Platform
GDPR	General Data Protection Regulation
HDH	Health Data Hub
Iac	Infrastructure as code
IAM	Identity And Access Management
ID	Identity
IoT	Internet of Things
MaaS	Model as a Service
OMOP CDM	Observational Medical Outcomes Partnership Common Data Model
RBAC	Role-Based Access Control
SIIC	Service Integration Information Collection
SSO	Single Sign-On
TFT	ThreeFold Token
TLS	Transport Layer Security
UI	User Interface
VC	Verifiable Credential
VP	Verifiable Presentation

2. Background: The Health Data Hub and constituent components

This differently articulated section provides an essential link to Task 5.2 from which comes the main input to the work needed in Task 5.4, its background. In fact, the project architecture is summarised with reference to the Health Data Hub and the services that are planned to be integrated to ensure full functionality and interoperability within the platform developed in the PHASE IV AI project.

2.1 The Health Data Hub

As a recap, we describe here the different services that will be offered in the Health Data Hub, seen in PHASE IV AI deliverable D5.3: Health Data Hub Design and Data Market v1 [1].

The HDH infrastructure acts as an entry system to the overall PHASE IV AI framework, servicing different stakeholders that each have their role in the ecosystem and need to securely interact.

The level 1 system diagram provides a high-level overview of the HDH, the Identity Management system (also covered in WP5), and its interactions with other systems.

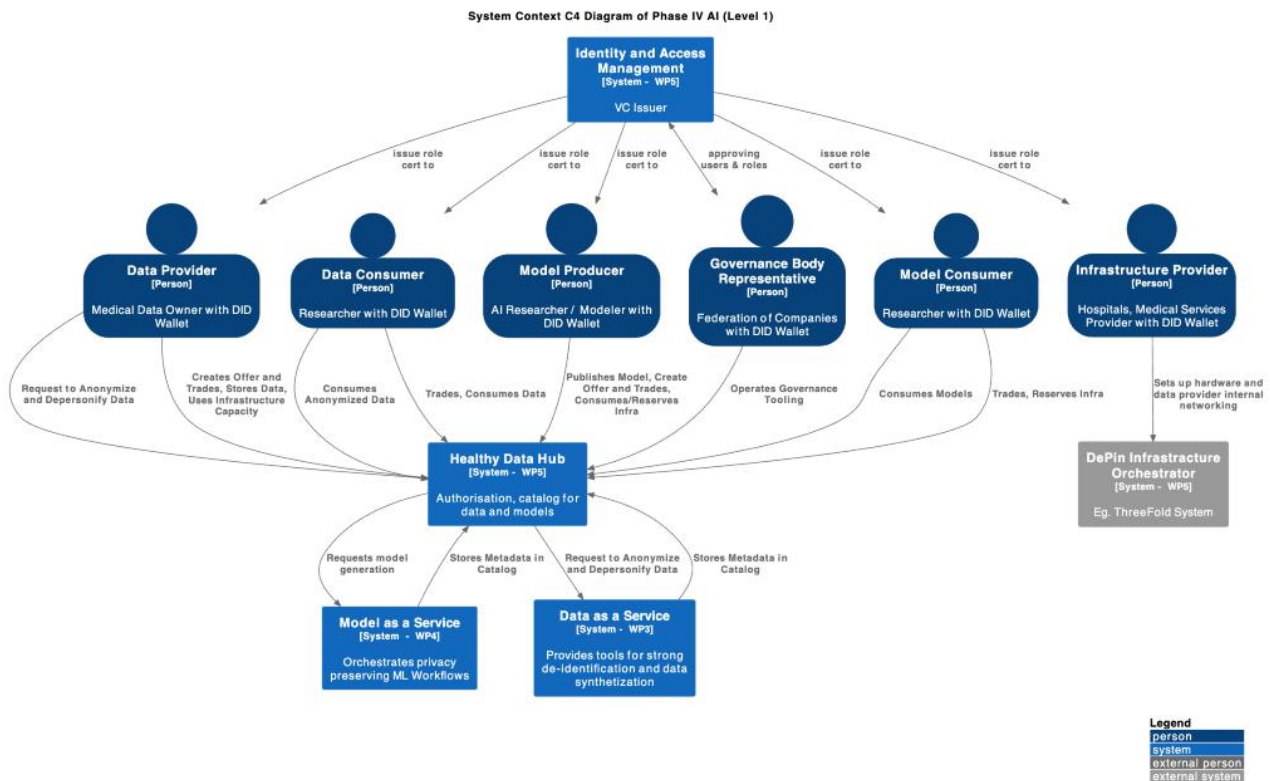


Figure 1: System Context C4 Diagram (Level 1)

Service integration needs to happen between the Identity and Access Management (IAM) subsystem and the HDH subsystem, with the VCs stored in a user's Identity Wallet used to manage authentication and authorisation across the HDH components.

The IAM service needs to be integrated, servicing the following identified stakeholders:

- Data providers / data holders (as stated in the EHDS regulatory framework)
- Data consumers

- Model producers
- Model consumers
- Governance bodies
- Service providers such as intermediators in the process, ex. Health Data Access Bodies (HDABs) as stated in the EHDS regulatory framework.
- Infrastructure providers (to be checked whether this is a specific service provider role)

All these actors interoperate inside a Health Data Space, and therefore, need to be clearly identified. To do so, actors need to be onboarded through an Identity Management approach that is common for data spaces, through Verifiable Credentials (VCs).

The services that need to be integrated from the IAM subsystem are:

- User onboarding/offboarding
- Identity wallet management
- Payment wallet management
- Authentication and authorisation services to get access to a HDH service

The HDH subsystem contains the following services to be integrated into the framework:

- **Catalogue App, containing 2 catalogues:**
 - **Data Catalogue:** a centralized component exposing relevant data sets.

Users: data consumers (which can be the same as model producers)

- **Service Catalogue:** a centralized component containing services which are available for deployment. These services can be of different kinds: de-identification services, data harmonisation services (as implemented under WP3-DaaS), AI training model services (as implemented under WP4-MaaS), etc. Data offering services, providing of channel infrastructure to keep data available for view purposes (through ex. data streaming) and for which monetization through subscription models is set up, are also services registered in a service catalogue.

Users:

- DaaS and MaaS Service providers to upload the content and/or metadata of their services into the catalogue
- Model consumers wishing to run these services on the infrastructure of their choice, running on one instance or in a federated manner (combining different instances in a secure way)
- Data consumers that make use of offered services to consume the data Software

During the framework integration design, we will explore ways to combine the data catalogue and service catalogue into one user portal.

- **Data and services federated marketplace:** a federated component allowing providers of data and services to put a price on their offer and allow for monetization. It is a component that connects an offering to a payment wallet.

Users:

- Data and service providers ('sellers')
- Data and service consumers ('buyers')

- **DePIN service**, enabling **Software creation** and **Software deployment services**, facilitating the easy deployment of a software (providing a service) on any premise.
 - A software creation service is the ‘Flist hub’, able to mount flists. The concept has been described in D5.3, paragraph 4.6.4. It will be further analysed whether the flist hub can be accessed through the same portal as the software certification service, as both are linked.
 - The Software Deployment Service is merely an Infrastructure catalogue and orchestrator, a centralized component
 - holding all available infrastructure to run software
 - containing the IaC scripts to deploy the software

Users:

- Services providers (both providers of DaaS and MaaS services), Model and Data Consumers to manage the deployment of the proprietary or catalogue services on self-chosen infrastructure
- Service provider to generate the flist workload, to be used as a reference in the service catalogue
- HDAB / governance bodies: to read the flist and refer to it for certification

The integration of all HDH services relies on a combination of components: the HTTP server, the Portal UI and the Portal API. These elements serve as main containers for service interactions. Chapter 3 of this deliverable provides a detailed description of their behaviour and functionality.

Another service is linked to both the identity management and the HDH

- **Software certification service**: a federated service to issue VCs on software.

The certificates are born as a claim by software developers, linked to the DID linked to their identity in the IAM, certified by governance bodies in the form of Verifiable Credentials, stored in the software developer’s DID wallet, and presented to third parties as a Verifiable Presentation (VP).

 - Users:
 - Service providers: VC certification request service, VP publishing service (to be attached to the service catalogue)
 - Governance bodies / HDABs: VC certification issuing service

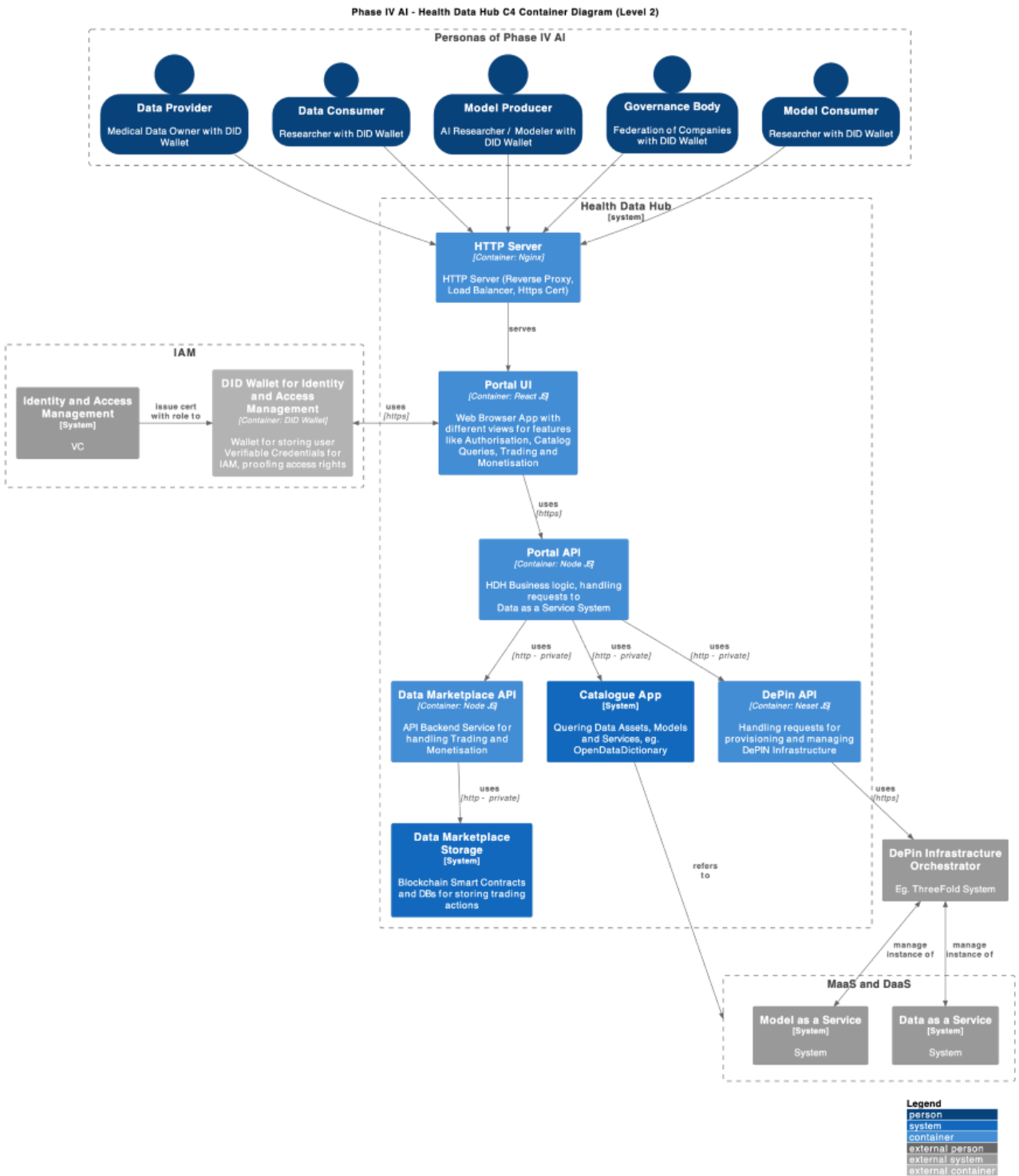


Figure 2: Health Data Hub C4 Container Diagram (Level 2)

3. Service Integration Information Collection framework

This chapter focuses on what is the core of Task 5.4, namely the integration of the HDH components, acting as counterpart to the previous chapter and introducing the integration framework for HDH tools and services. To this end, a framework for service integration has been developed, of which the first key step lies in gathering requirements and information about the services to be integrated. Specifically, hereafter it is introduced the information-gathering framework which condenses technical framework, methodologies, and processes.

The technical framework for service integration within the Health Data Hub (HDH) is built upon a modular, API-driven, and federated architecture, ensuring interoperability between diverse data services, repositories, and marketplace functionalities. This framework defines standardized APIs, data exchange mechanisms, identity management solutions, and security protocols to enable seamless interaction across distributed health data services while taking up possible compliance needs to regulations such as GDPR, EHDS, and AI Act.

The methodologies employed in this integration process follow a service-oriented and event-driven approach, leveraging microservices architecture for scalability and flexibility. Data interoperability is facilitated through FHIR-compliant data models, RESTful APIs, and secure message queues (e.g., Kafka, RabbitMQ) to support real-time transactions, and with particular attention to identity and access management (IAM) mechanism to be enforced.

The processes governing the integration lifecycle include service onboarding, validation, orchestration, and monitoring. Services undergo standardized onboarding procedures, including compliance checks and metadata registration. Automated validation and testing pipelines (CI/CD) ensure that integrated services meet performance, security, and scalability requirements before deployment. The framework also incorporates observability and monitoring mechanisms (e.g., Prometheus, ELK stack) to track service interactions, detect anomalies, and ensure the reliability of data transactions within the federated ecosystem.

All the information related to the technical framework, methodologies and processes are condensed within this section that introduces the SIIC template, or Service Integration Information Collection template, which is specifically designed for collecting information about the Health Data Hub services that need integration. It consists of 8 tables that seek to provide a high-level view of the key information needed by those who practically need to carry out service integration, taking into consideration various aspects ranging from how software tools communicate to infrastructure and security.

3.1 General Information

The first table shown below collects basic details about the service to be integrated. This includes the service name, the responsible partner(s) managing or developing it, and a brief description of its purpose. The primary function field should specify what role the service plays in the Health Data Hub (e.g., data storage, anonymization, AI processing). Additionally, partners should list dependencies, meaning other services that their service interacts with. This section helps establish a high-level overview of the service before diving into technical details.

Table 1: General Information

Field	Details (to be filled by the partner)
Service Name	
Responsible Partner(s)	
Service Description	
Primary Function (e.g., Data Storage, Anonymization, AI Processing)	

Service Dependencies (Other services this integrates with)	Linked Input Services: <ul style="list-style-type: none"> Service Name and communication method Linked Output Services: <ul style="list-style-type: none"> Service Name and communication method
--	---

3.2 API & Communication Details

The table shown below focuses on how the service exposes and communicates data through APIs. Partners should specify whether the API follows REST, GraphQL, SOAP, or another format. The API documentation URL (or contact person) is crucial for others who need to integrate with this service. The authentication method (e.g., OAuth2, API key) ensures security in service interactions. Finally, API endpoints, rate limits, and error handling mechanisms should be provided to guide integration teams in managing requests effectively.

Table 2: API & Communication Details

Field	Details (to be filled by the partner)
API Type	<ul style="list-style-type: none"> REST GraphQL SOAP Other (specify)
API Documentation URL or Contact	
Authentication Method	<ul style="list-style-type: none"> OAuth2 API Key JWT Other (specify)
API Endpoints (List key ones if available)	
Rate Limits (if applicable)	
Error Handling Mechanisms	

3.3 Data Exchange & Interoperability

The table shown below details how the service handles data. The data inputs field should describe what kind of data the service requires to function (e.g., raw patient records, anonymized datasets, model outputs). The data outputs field should specify what the service generates or provides. Supported data formats (e.g., JSON, XML, FHIR, HL7) should be listed to verify interoperability. If the service supports data transformation or mapping, partners should indicate so, as this affects compatibility with other services. If a standardized data schema (such as OMOP CDM for health data) is used, this should be documented to ensure consistency in data representation. Finally, any data validation rules (e.g., required fields, value constraints) should be listed.

Table 3: Data Exchange & Interoperability

Field	Details (to be filled by the partner)
Data Inputs Required (What kind of data does the service consume?)	
Data Outputs Provided (What does the service generate?)	
Supported Data Formats	<ul style="list-style-type: none"> JSON XML FHIR

	<ul style="list-style-type: none"> <input type="radio"/> HL7 <input type="radio"/> CSV <input type="radio"/> Other (specify)
Supports Data Transformation/Mapping?	<ul style="list-style-type: none"> <input type="radio"/> Yes (specify) <input type="radio"/> No
Is there a standardized schema?	<ul style="list-style-type: none"> <input type="radio"/> Yes (specify) <input type="radio"/> No
Data Validation Rules	

3.4 Identity, Security & Compliance

Security and compliance are critical for handling sensitive health data. The table shown below captures how access is controlled (e.g., Role-Based Access Control – RBAC, Attribute-Based Access Control – ABAC). Data encryption measures should be specified, including whether the service uses TLS (for data in transit) or AES (for data at rest). The compliance standards followed should align with legal and regulatory frameworks such as GDPR, EHDS, and the AI Act. Partners should also outline logging and auditing mechanisms to ensure accountability. Finally, details on user authentication and identity federation (e.g., Single Sign-On, OpenID Connect) should be included to manage access securely.

Table 4: Identity, Security & Compliance

Field	Details (to be filled by the partner)
Access Control & Authorization Model	<ul style="list-style-type: none"> <input type="radio"/> Role-Based (RBAC) <input type="radio"/> Attribute-Based (ABAC) <input type="radio"/> Other (specify)
Data Encryption	<ul style="list-style-type: none"> <input type="radio"/> TLS <input type="radio"/> AES <input type="radio"/> Not Applicable
Compliance Standards Followed	<ul style="list-style-type: none"> <input type="radio"/> GDPR <input type="radio"/> EHDS <input type="radio"/> AI Act (AIA) <input type="radio"/> Data Governance Act (DGA) <input type="radio"/> Data Act (DA) <input type="radio"/> Other (specify)
Logging & Auditing Mechanisms Available?	<ul style="list-style-type: none"> <input type="radio"/> Yes (provide details) <input type="radio"/> No (provide details)
User Authentication & Identity Federation Support	<ul style="list-style-type: none"> <input type="radio"/> SSO <input type="radio"/> OpenID Connect <input type="radio"/> Other (specify)

3.5 Infrastructure & Hosting

The table shown below provides details on where and how the service is hosted. Partners should indicate whether the service runs in a cloud environment (e.g., AWS, Azure, GCP), on-premises, or in a hybrid setup. The networking considerations section should specify whether the service is publicly accessible, requires a VPN, or uses a private network. Finally, scalability features (such as load balancing, auto scaling, and horizontal scaling) should be described, as these impact the service’s ability to handle increased traffic and demand.

Table 5: Infrastructure & Hosting

Field	Details (to be filled by the partner)
Hosting Environment	<ul style="list-style-type: none"> <input type="radio"/> Cloud (AWS, Azure, GCP) <input type="radio"/> On-Premises <input type="radio"/> Hybrid <input type="radio"/> DePIN infrastructure <input type="radio"/> Self-contained (relevant for wallets)
Networking Considerations	<ul style="list-style-type: none"> <input type="radio"/> VPN <input type="radio"/> Private Network <input type="radio"/> Public API
Scalability Features	<ul style="list-style-type: none"> <input type="radio"/> Horizontal Scaling <input type="radio"/> Load Balancing <input type="radio"/> Auto-scaling

3.6 Integration Patterns & Technologies

The table shown below captures the integration methods used for service communication. Partners should specify whether the service integrates via API calls (synchronous request-response), message queues (Kafka, RabbitMQ), or event-driven mechanisms (webhooks, publish-subscribe models). If the service uses a middleware or API gateway, such as Kong or Apigee, this should be listed to understand traffic management. Finally, CI/CD pipeline availability should be documented, including tools used (e.g., Jenkins, GitHub Actions), to ensure smooth deployment and updates.

Table 6: Integration Patterns & Technologies

Field	Details (to be filled by the partner)
Integration Method	<ul style="list-style-type: none"> <input type="radio"/> API Calls <input type="radio"/> Message Queue (Kafka, RabbitMQ) <input type="radio"/> Event-Driven (Webhooks) <input type="radio"/> Other
Middleware or API Gateway Used?	<ul style="list-style-type: none"> <input type="radio"/> Yes (specify) <input type="radio"/> No
CI/CD Pipeline Available?	<ul style="list-style-type: none"> <input type="radio"/> Yes (specify tools like Jenkins, GitHub Actions) <input type="radio"/> No

3.7 Testing & Monitoring

Testing and monitoring are essential for maintaining service reliability. Within the table shown below partners should indicate whether unit and integration tests have been implemented and if performance testing has been conducted to evaluate scalability. Monitoring tools such as Prometheus, Grafana, or the ELK stack (Elasticsearch, Logstash, Kibana) should be listed if in use. Additionally, services should specify if they have error logging systems, ensuring that integration teams can track and debug failures efficiently.

Table 7: Testing & Monitoring

Field	Details (to be filled by the partner)
Unit & Integration Tests Available?	<ul style="list-style-type: none"> <input type="radio"/> Yes <input type="radio"/> No
Performance Testing Conducted?	<ul style="list-style-type: none"> <input type="radio"/> Yes

	<input type="radio"/> No
Monitoring & Alerting Tools Used	<input type="radio"/> Prometheus <input type="radio"/> Grafana <input type="radio"/> ELK <input type="radio"/> Other (provide details)
Error Logging System in Place?	<input type="radio"/> Yes (specify tools) <input type="radio"/> No

3.8 Additional Comments & Open Issues

The final table shown below allows partners to provide any additional insights, challenges, or future plans related to their service. If there are known limitations (e.g., dependency on legacy systems, API constraints), they should be documented here. Planned improvements can be listed to inform the broader integration roadmap. Any other relevant information that could help integration teams should be included to ensure transparency and effective collaboration.

Table 8: Additional Comments & Open Issues

Field	Details (to be filled by the partner)
Known Challenges or Limitations	
Future Improvements Planned	
Other Relevant Information	

4. Services Integration Details

In this chapter, details of the services to be integrated are given and for each of the services there is a single all-inclusive table of the information introduced in the previous section with the SIIC template, keeping only relevant information. In particular, the tables for the following services follow:

- Data and Services Catalogue
- DePIN services for ‘Service Creation’ and ‘Service Deployment’
- Onboarding Service (DID/VCS)
- Software Certification Service (VCS)
- Identity Wallet Service
- Payment Wallet Service

4.1 Diagram

The following diagram provides a high-level representation of the key service interactions within the Health Data Hub. These services work together to facilitate identity management, payments, onboarding, service cataloguing, deployment, and certification.

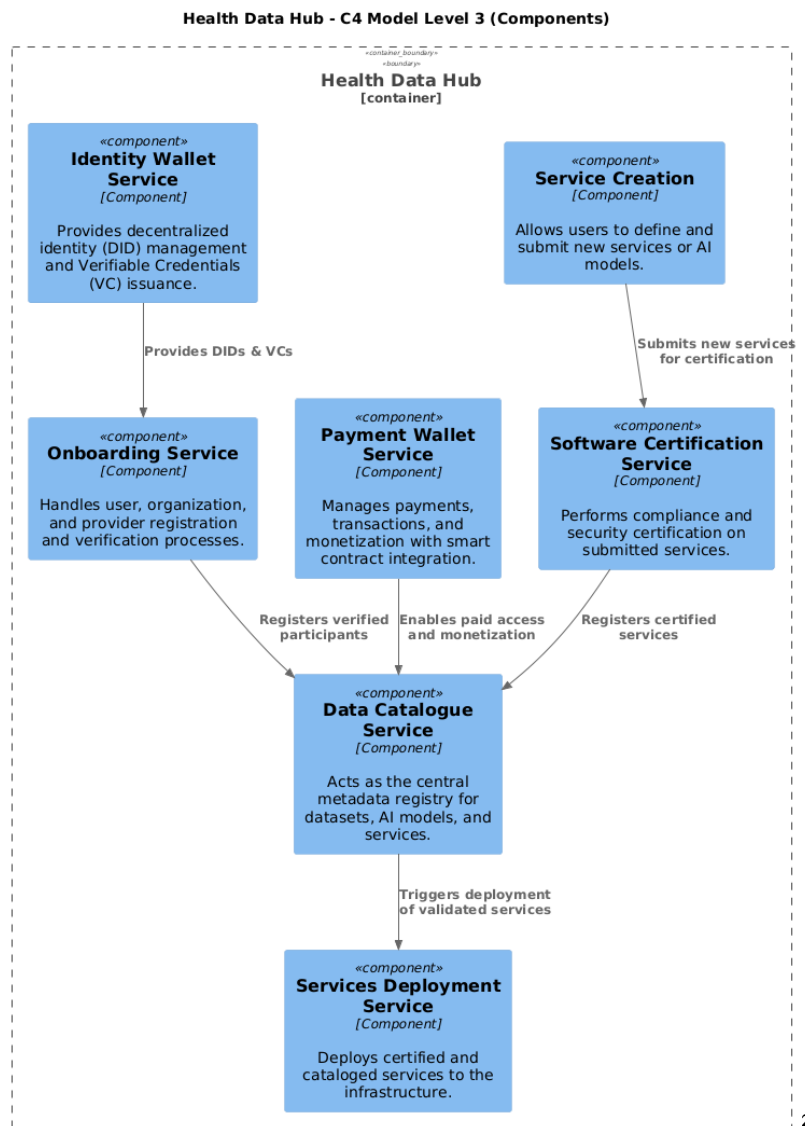


Figure 3: Health Data Hub C4 Components Diagram (Level 3)

At the core of the system is the Data Catalogue Service, which acts as the central hub for managing and organizing service-related information. Various services interact with the Data Catalogue Service, each fulfilling a distinct role:

- The Identity Wallet Service enables user authentication and identity verification, feeding information into the Onboarding Service to facilitate new user registration.
- The Payment Wallet Service ensures financial transactions and integrates directly with the Data Catalogue Service to manage payment-related data.
- The Onboarding Service works as a bridge between identity verification and the Data Catalogue Service, ensuring that users are properly registered.
- The Service Creation process feeds the Data Catalogue Service after validation made by the Software Certification Service.
- Finally, the Services Deployment Service utilizes the information stored in the Data Catalogue Service to deploy services.

4.2 Data Catalogue Service

Table 9: Data Catalogue Service Integration Details

General Information	
Service Name	Data Catalogue Service
Service Description	<p>Data Catalogue Service is the central metadata repository within the Health Data Hub (HDH) ecosystem, providing a structured and searchable registry of datasets, AI models, and deployable services.</p> <p>It enables efficient data discovery, governance, and controlled access, ensuring that only authorized users can query, retrieve, or utilize available data assets.</p>
Primary Function (e.g., Data Storage, Anonymization, AI Processing)	Dataset Discovery/Catalogue Metadata Management Services Catalogue Integration with Payment & Monetization Services Role-Based Access Control (display, access) Deploy a Service Manage Deployed Services Lifecycle
Service Dependencies (Other services this integrates with)	Linked Input Services: <ul style="list-style-type: none"> • Onboarding Service • Service Creation • Wallet Payment Service Linked Output Services: <ul style="list-style-type: none"> • Services Deployment Service
API & Communication Details	
API Type	o REST
Authentication Method	o OAuth 2.0 / OpenID Connect (OIDC) with JWT token (KeyCloak)
Data Exchange & Interoperability	
Supported Data Formats	o JSON
Identity, Security & Compliance	
Access Control & Authorization Model	o Role-Based (RBAC)
Data Encryption	o TLS
Compliance Standards Followed	o GDPR o EHDS
Logging & Auditing Mechanisms Available?	o Yes
User Authentication & Identity Federation Support	o OpenID Connect
Infrastructure & Hosting	
Hosting Environment	o Under control of governance body

Networking Considerations	<ul style="list-style-type: none"> o Decentralized infrastructure o Peer to Peer communication o Crypted communication
Scalability Features	<ul style="list-style-type: none"> o Horizontal Scaling o Load Balancing o Auto-scaling

4.3 Services Deployment Service

Table 10: Services Deployment Service Integration Details

General Information	
Service Name	Services Deployment Service
Service Description	<p>Services Deployment Service is responsible for automating and managing the deployment of various services within the Health Data Hub ecosystem.</p> <p>It ensures that newly created or updated services instances are properly deployed, configured, and integrated into decentralized infrastructure.</p> <p>This service orchestrates deployments across environments, ensuring scalability, security, and compliance with regulatory standards.</p>
Primary Function (e.g., Data Storage, Anonymization, AI Processing)	Automated Service Deployment Monitoring Deployed Infrastructure
Service Dependencies (Other services this integrates with)	<p>Linked Input Services:</p> <ul style="list-style-type: none"> • Data Catalogue Service • Identity Wallet Service (auth and rbac) • Payment Wallet Service (billing) <p>Linked Output Services:</p> <ul style="list-style-type: none"> • Infrastructure deployment on Threefold
API & Communication Details	
API Type	<ul style="list-style-type: none"> o REST
Authentication Method	<ul style="list-style-type: none"> o OAuth 2.0 / OpenID Connect (OIDC) with JWT token (KeyCloak) o Web3 Identity Providers
Data Exchange & Interoperability	
Supported Data Formats	<ul style="list-style-type: none"> o JSON
Identity, Security & Compliance	
Access Control & Authorization Model	<ul style="list-style-type: none"> o Role-Based (RBAC)
Data Encryption	<ul style="list-style-type: none"> o TLS
Compliance Standards Followed	<ul style="list-style-type: none"> o GDPR o EHDS

Logging & Auditing Mechanisms Available?	<input type="radio"/> Yes
User Authentication & Identity Federation Support	<input type="radio"/> OpenID Connect
Infrastructure & Hosting	
Hosting Environment	<input type="radio"/> Under control of the governance body
Networking Considerations	<input type="radio"/> Decentralized infrastructure <input type="radio"/> Peer to Peer communication <input type="radio"/> Crypted communication
Scalability Features	<input type="radio"/> Horizontal Scaling <input type="radio"/> Load Balancing <input type="radio"/> Auto-scaling

4.4 Service Creation

Table 11: Service Creation Integration Details

General Information	
Service Name	Service Creation
Service Description	<p>Service Creation Service is responsible for the implementation of new services within the Health Data Hub (HDH) ecosystem.</p> <p>It enables users (data provider, researchers, ...) to define, build, and submit new applications, data processing tools, or AI models before they are deployed and made available for use.</p>
Primary Function (e.g., Data Storage, Anonymization, AI Processing)	Service Definition & Templating Development & Configuration Tools Integration with Software Certification Service Metadata Registration in the Data Catalogue Service Approval & Submission for Deployment Version Control Management
Service Dependencies (Other services this integrates with)	Linked Input Services: Linked Output Services: <ul style="list-style-type: none"> • Software Certification Service
API & Communication Details	
API Type	<input type="radio"/> REST
Authentication Method	<input type="radio"/> OAuth 2.0 / OpenID Connect (OIDC) with JWT token (KeyCloak)
Data Exchange & Interoperability	
Supported Data Formats	<input type="radio"/> JSON
Identity, Security & Compliance	
Access Control & Authorization Model	<input type="radio"/> Role-Based (RBAC)
Data Encryption	<input type="radio"/> TLS
Compliance Standards Followed	<input type="radio"/> GDPR

	o EHDS
Logging & Auditing Mechanisms Available?	o Yes
User Authentication & Identity Federation Support	o OpenID Connect
Infrastructure & Hosting	
Hosting Environment	o Under control of the governance body
Networking Considerations	o Decentralized infrastructure o Peer to Peer communication o Crypted communication
Scalability Features	o Horizontal Scaling o Load Balancing o Auto-scaling

4.5 Onboarding Service

Table 12: Onboarding Service Integration Details

General Information	
Service Name	Onboarding Service
Service Description	<p>Onboarding Service is responsible for registering new users or verifying existing users or organizations within the Health Data Hub (HDH) ecosystem. It ensures that participants undergo identity verification, compliance checks, and credential validation before gaining access to HDH services.</p> <p>This service is tightly integrated with the Identity Wallet Service, leveraging Decentralized Identifiers (DIDs) and Verifiable Credentials (VCs) to provide a secure, privacy-preserving, and decentralized identity onboarding process.</p>
Primary Function (e.g., Data Storage, Anonymization, AI Processing)	User and Organization Registration Integration with the Identity Wallet Service Control & Role Assignment Audit Trail & Logging
Service Dependencies (Other services this integrates with)	Linked Input Services: <ul style="list-style-type: none"> • Identity Wallet Service Linked Output Services: <ul style="list-style-type: none"> • Data Catalogue Service • Payment Wallet Service • Services Deployment Service
API & Communication Details	
API Type	o REST
Authentication Method	o OAuth 2.0 / OpenID Connect (OIDC) with JWT token (KeyCloak)

Data Exchange & Interoperability	
Supported Data Formats	o JSON
Identity, Security & Compliance	
Access Control & Authorization Model	o Role-Based (RBAC)
Data Encryption	o TLS
Compliance Standards Followed	o GDPR o EHDS
Logging & Auditing Mechanisms Available?	o Yes
User Authentication & Identity Federation Support	o OpenID Connect
Infrastructure & Hosting	
Hosting Environment	o Under control of the governance body
Networking Considerations	o Decentralized infrastructure o Peer to Peer communication o Crypted communication
Scalability Features	o Horizontal Scaling o Load Balancing o Auto-scaling

4.6 Software Certification Service

Table 13: Services Certification Service Integration Details

General Information	
Service Name	Software Certification Service
Service Description	<p>Software Certification Service is responsible for evaluating, validating, and certifying software applications, data processing services, and AI models before they are integrated within the Health Data Hub.</p> <p>This service ensures that software components meet security, compliance, interoperability, and quality standards.</p>
Primary Function (e.g., Data Storage, Anonymization, AI Processing)	Compliance Validation Security Validation Compatibility Validation
Service Dependencies (Other services this integrates with)	Linked Input Services: <ul style="list-style-type: none"> • Service Creation Linked Output Services: <ul style="list-style-type: none"> • Data Catalogue Service
API & Communication Details	
API Type	o REST
Authentication Method	o OAuth 2.0 / OpenID Connect (OIDC) with JWT token (KeyCloak)

Data Exchange & Interoperability	
Supported Data Formats	o JSON
Identity, Security & Compliance	
Access Control & Authorization Model	o Role-Based (RBAC)
Data Encryption	o TLS
Compliance Standards Followed	o GDPR o EHDS
Logging & Auditing Mechanisms Available?	o Yes
User Authentication & Identity Federation Support	o OpenID Connect
Infrastructure & Hosting	
Hosting Environment	o Under control of the governance body
Networking Considerations	o Decentralized infrastructure o Peer to Peer communication o Crypted communication
Scalability Features	o Horizontal Scaling o Load Balancing o Auto-scaling

4.7 Identity Wallet Service

Table 14: Identity Wallet Service Integration Details

General Information	
Service Name	Identity Wallet Service
Service Description	<p>Identity Wallet Service is a decentralized identity management system that enables users, organizations, and services within the Health Data Hub (HDH) ecosystem to securely create, store, and manage their digital identities.</p> <p>Built on Decentralized Identifiers (DIDs) and Verifiable Credentials (VCs), this service provides a self-sovereign identity framework, allowing users to authenticate, prove their credentials, and control access to sensitive data and services without relying on centralized identity providers.</p>
Primary Function (e.g., Data Storage, Anonymization, AI Processing)	Decentralized Identity Management Verifiable Credentials Management Authentication Service Authorization Service Audit Logging & Transparency
Service Dependencies (Other services this integrates with)	Linked Input Services: Linked Output Services: <ul style="list-style-type: none"> Onboarding Service
API & Communication Details	

API Type	o REST
Authentication Method	o Web3 Identity Providers
Data Exchange & Interoperability	
Supported Data Formats	o JSON
Identity, Security & Compliance	
Access Control & Authorization Model	o Role-Based (RBAC)
Data Encryption	o TLS
Compliance Standards Followed	o GDPR o EHDS
Logging & Auditing Mechanisms Available?	o Yes
User Authentication & Identity Federation Support	o OpenID Connect
Infrastructure & Hosting	
Hosting Environment	o Under control of the wallet owner
Networking Considerations	o Decentralized infrastructure o Peer to Peer communication o Crypted communication
Scalability Features	o Horizontal Scaling o Load Balancing o Auto-scaling

4.8 Payment Wallet Service

Table 15: Payment Wallet Service Integration Details

General Information	
Service Name	Payment Wallet Service
Service Description	<p>Payment Wallet Service is responsible for managing financial transactions, payments, and monetization within the Health Data Hub (HDH) ecosystem.</p> <p>It enables secure, decentralized, and verifiable financial operations, ensuring that data providers, service developers, and consumers can seamlessly transact, access, and trade datasets and AI services.</p>
Primary Function (e.g., Data Storage, Anonymization, AI Processing)	<p>Transaction Processing & Data Monetization</p> <p>Integration with Verifiable Credentials (VCs)</p> <p>Audit Logging & Transaction History</p>
Service Dependencies (Other services this integrates with)	<p>Linked Input Services:</p> <p>Linked Output Services:</p> <ul style="list-style-type: none"> Enables monetization in the Data Catalogue Service
API & Communication Details	
API Type	o REST
Authentication Method	o Web3 Identity Providers

Data Exchange & Interoperability	
Supported Data Formats	o JSON
Identity, Security & Compliance	
Access Control & Authorization Model	o Role-Based (RBAC)
Data Encryption	o TLS
Compliance Standards Followed	o GDPR o EHDS
Logging & Auditing Mechanisms Available?	o Yes
User Authentication & Identity Federation Support	o OpenID Connect
Infrastructure & Hosting	
Hosting Environment	o Under control of the wallet owner
Networking Considerations	o Decentralized infrastructure o Peer to Peer communication o Crypted communication
Scalability Features	o Horizontal Scaling o Load Balancing o Auto-scaling

5. Conclusions

This deliverable establishes the technical foundation for service integration within the HDH, defining the necessary architecture, methodologies, and processes to enable seamless interaction between diverse health data services. By implementing standardized APIs, federated identity management, and secure data exchange mechanisms, the integration framework ensures interoperability, scalability, and compliance with regulatory requirements such as GDPR, EHDS, and the AI Act. The work conducted in Task 5.4 paves the way for a trusted and efficient marketplace for health-related data services, allowing stakeholders to securely access, manage, and transact health data while preserving privacy and security.

This document represents the first iteration of the integration framework, serving as a baseline for further refinement and validation in future project phases. The next steps will focus on testing and improving service interactions, addressing integration challenges, and optimizing performance based on real-world implementation and stakeholder feedback. By continuously evolving the integration strategy, the PHASE IV AI project moves closer to its vision of an AI-driven, privacy-preserving, and interoperable health data ecosystem that fosters innovation and collaboration across the European healthcare landscape.