M7.4 Draft technical, functional and security specifications of Secure Processing Environments

TEHDAS2 – Second Joint Action Towards the European Health Data Space

12 September 2025

Co-funded by the European Union

0 Document info

0.1 Authors

Lead Author(s)	Lead organisation
Heikki Lehväslaiho	CSC – IT Center for Science, Finland
Helena Lodenius	CSC – IT Center for Science, Finland
Beatriz Barros	Sciensano, Belgium
Alexandre Berna	Health Data Hub, France
Lucas Bréchot	Health Data Hub, France
Zdenek Gütter	Ministry of Health of the Czech Republic
Hans Aage Huru	Norwegian Institute of Public Health, Norway
Yohan Jarosz	Luxembourg National Data Service, Luxembourg
Todor Kondić	Luxembourg National Data Service, Luxembourg
Jaakko Lähteenmäki	VTT Technical Research Centre of Finland Ltd, Finland
Max Martens	BfArM - Federal Institute for Drugs and Medical Devices, Germany
Minerva Alvarez	Spanish Ministry of Health
Juha Pajula	VTT Technical Research Centre of Finland Ltd, Finland
Thomas Sondag	Luxembourg National Data Service, Luxembourg
Christophe Trefois	Luxembourg National Data Service, Luxembourg
Emmi Turunen	HUS Group, the joint authority for Helsinki and Uusimaa, Finland

0.2 Keywords

Keywords	TEHDAS2, Joint Action, Health Data, Health Data Space, Secure Processing Environments, SPE Federation, Federated computing

0.3 Document history

Date	Version	Editor	Change	Status
11/10/2024	0.1	Helena Lodenius, Heikki Lehväslaiho	Table of Contents	Draft
27/06/2025	0.2	Heikki Lehväslaiho, Helena Lodenius, Beatriz Barros, Alexandre Berna, Todor Kondić, Jaakko Lähteenmäki, Juha Pajula	Draft to be reviewed by the Consortium	Draft
12/09/2025	1.0	Heikki Lehväslaiho, Helena Lodenius, Beatriz Barros, Jaakko Lähteenmäki	Document to be submitted for public consultation	Final

Accepted in Project Steering Group on 12 September 2025.

Table of contents

1 Executive summary

This report presents the technical, functional, and security specifications of **Secure Processing Environments (SPEs), a central component of the European Health Data Space (EHDS) as required under Article 73 of Regulation (EU) 2025/327**. SPEs are designed to enable the safe secondary use of electronic health data while ensuring compliance with data protection, confidentiality, and information security obligations.

Based on a thorough analysis, this report defines **a structured set of minimum requirements for SPEs**. It covers core capabilities needed to safeguard any sensitive data, as well as enabling requirements arising from the needs of scientific research principles. **A generic SPE specification** is developed that is flexible enough to fulfil current and future functional requirements. Functional and operational needs of **two main SPE use cases identified to be needed by EHDS** are shown to be derivable from this model.

The report goes beyond the obligatory demands of EHDS to define minimal requirements of interoperability between compatible services that form an **SPE-based federation** that is required when data needs to be transferred between organisations. The SPE federation model is further expanded with a set of tentative **practical implementation requirements for federated computing** that is still an active research area.

Justification of the report focus on high-level functional requirements over technical ones is given in the appendices that cover existing solutions, pitfalls of too narrow approaches, and crucial interplay of SPEs with other services in the ecosystem. Particular attention is given to how various approaches and technical implementations affect the **trade-off between data security and freedom, privacy, and responsibility of actors accessing sensitive data.**

These specifications are intended to support Member States and stakeholders in the design and operation of SPEs and to inform the work of the European Commission in the preparation of the implementing act under Article 73(5) of the EHDS Regulation.

2 Introduction

Advancing health data use in the European Health Union

As part of the European Health Union, the European Union (EU) is advancing the use of health data for secondary purposes, including research, innovation and policymaking. Smooth and secure access to data will drive the development of new treatments and medicines and optimise resource utilisation—all with the overarching goal of improving the health of citizens across Europe.

TEHDAS2, the second joint action Towards the European Health Data Space, represents a significant step forward in this vision. The project will develop guidelines and technical specifications to facilitate smooth national and cross-border reuse of health data, and support health data holders, health data users and the new health data access bodies in fulfilling their responsibilities and obligations outlined in the European Health Data Space (EHDS) Regulation.

TEHDAS2 focuses on several critical aspects of health data use.

- Data discovery: findability and availability of health data, ensuring it is accessible for secondary purposes.
- Data access: developing harmonised access procedures and establishing standardised approaches for granting data access across Member States.
- Secure processing environment: defining technical specifications for environments where sensitive health data can be processed safely.
- Citizen-centric obligations: providing guidance on fulfilling obligations to citizens, such as communicating significant research findings that impact their health, informing them about research outcomes and ensuring transparency in how their data is used.
- Collaboration models: developing guidance on collaboration and guidelines on fees and penalties as well as third country and international access to data.

TEHDAS2 will contribute to harmonised implementation of the EHDS Regulation through the concrete guidelines and technical specifications. Some of these documents and resources will also provide input to implementing acts of the regulation. Hence, the joint action will increase the preparedness for the EHDS implementation and lead to better coordination of Member States' joint efforts towards the secondary use of health data, while also reducing fragmentation in policies and practices related to secondary use.

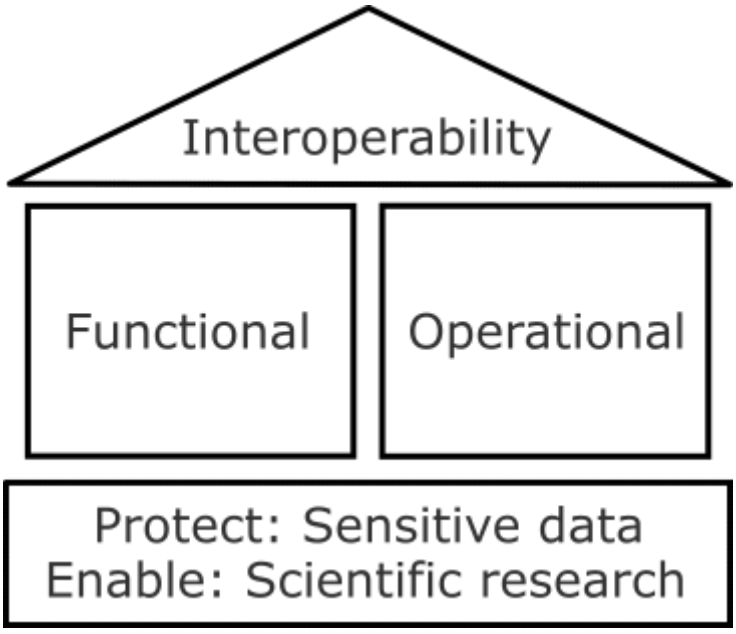
The aim of TEHDAS2 T7.4 milestone is to develop technical specifications for the lawful and effective operation of Secure Processing Environments (SPEs) in accordance with Article 73 of the European Health Data Space (EHDS) Regulation. These requirements are used to describe functional, security and operational capabilities needed to build SPEs and supporting services. SPEs must support multiple use cases for secondary data use. Design and implementation choices will significantly influence their ability to accommodate these use cases in a compliant and effective manner. The impact of a precise definition of an SPE will be critical for the context of secondary use of health data. It will guide the use of technologies and capabilities of sensitive data environments across many domains within the European

Union and potentially beyond. It is therefore important that the definition of SPE focuses on essentials and enabling the full data management life cycle.

Health data is inherently dynamic and evolves over time as individuals generate and share information throughout their lives and during medical care. Initially used for direct patient treatment (primary use), this data can later be repurposed for broader goals like public health, research, and policymaking (secondary use). As data moves through various stages—collection, processing, analysis, and storage—it is transformed and enriched, often combined with other datasets. This dynamic flow of information across systems and organisations requires strong collaboration and clear governance. The EHDS aims to give citizens more control over their data and ensure transparency about its use. Achieving this requires interdisciplinary understanding across legal, organisational, semantic, and technical layers, as outlined in the European Interoperability Framework (EIF).

We structured the SPE requirements into four main categories (Figure 2.1) At the foundation are the core objectives to ensure protection of any sensitive data and make them available for scientific research. This foundation supports two central pillars: functional and operational requirements. These will be framed and influenced by over-arching interoperability requirements.

Figure 2.1 Main categories of SPE requirements.



3 Scope

3.1 Legal requirements for EHDS SPE

The European Union Data Governance Act (DGA, EU 2022/868) provides the definition of an SPE, which is also referenced in the EHDS regulation's Definitions Article.

> ***'secure processing environment' means the physical or virtual environment and organisational means to ensure compliance with Union law, such as Regulation (EU) 2016/679, in particular with regard to data subjects' rights, intellectual property rights, and commercial and statistical confidentiality, integrity and accessibility, as well as with applicable national law, and to allow the entity providing the secure processing environment to determine and supervise all data processing actions, including the display, storage, download and export of data and the calculation of derivative data through computational algorithms****;* (DGA, Article 2)

SPEs can be used in multiple use cases, domains of knowledge, and governance structures; the latest being the EHDS, which requires using SPE for the secondary use of health data. In the EHDS context, the article 73 describes the legal requirements of a secure, controlled infrastructure designed to enable the processing of electronic health data for secondary use while ensuring compliance with data protection and security requirements. More specifically, an EHDS SPE must meet the following requirements:

- Data security: Prevent unauthorised access, ensure data confidentiality, and maintain integrity.
- Restricted access: Allow data users to process data only within the scope defined by their data access permit.
- Controlled outputs: Ensure that only aggregated or anonymised results will be exported, subject to approval by the Health Data Access Body (HDAB).

SPEs are critical to enabling lawful secondary use of health data while safeguarding individual privacy and ensuring compliance with GDPR and EHDS regulations.

3.2 Preliminary life cycle components of SPE

The use of an SPE within the HealthData@EU infrastructure occurs after a health data user (in distinction from the generic SPE user that is called *data user*) has been granted access to a specific dataset and a data permit has been issued by the HDAB. The general life cycle of an SPE, as outlined in TEHDAS1[1], consists of the following steps:

Environment Creation: Once the data permit is issued, the designated SPE operator sets up an isolated environment instance, tailored to the health data access application form. In the current IT landscape, this is typically done by deploying a virtual machine, either in a dedicated cluster or a cloud computing environment.

[1] D7.2. Options for the services and services architecture and infrastructure for secondary use of data in the EHDS https://tehdas.eu/tehdas1/results/tehdas-proposals-for-the-implementation-of-ehds-technical-infrastructure/

Data Reception: Depending on the interface arrangements between the SPE and the data holder(s), data can either be pulled by the SPE operator from the data holder(s) or pushed by the data holder(s) to the assigned SPE storage.

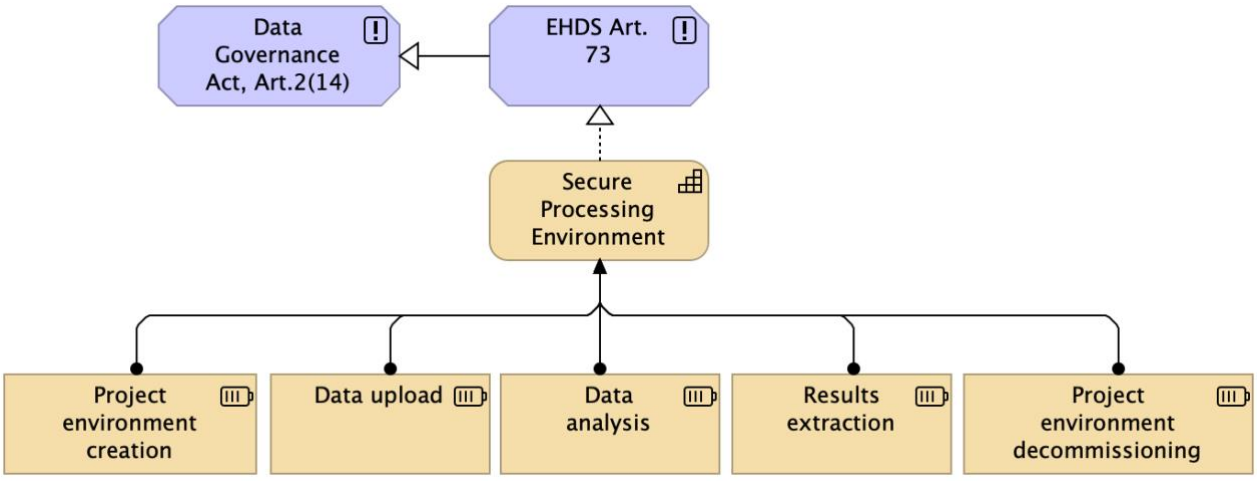
Data Upload: HDABs may use an *intermediate* SPE to prepare the data. This includes anonymisation or pseudonymisation, as well as data linkage (combining the datasets requested by the applicant at individual level). Pseudonymisation ensures that data cannot be linked to an individual without access to a separate key, which only the HDAB holds. Once the data is prepared, it is securely transferred from the intermediate SPE to a user-facing SPE.

Data Analysis: The health data user then processes the uploaded data within the environment using the tools provided in that SPE instance to derive the insights they are seeking.

Results Extraction: Once the analysis is complete, the health data user must request permission to download the fully anonymised aggregated results (whether partial or final) from the SPE. This step involves controlling and monitoring the data that is allowed to leave the secure processing environment by the HDAB.

Environment Decommissioning/Archival: After the project concludes or the data permit expires, the environment may either be decommissioned (with all contents destroyed) or archived for potential future use, under new conditions, such as for reproducibility of results or new project permits.

Figure 3.1. The SPE functionalities as specified in the TEHDAS1 project.



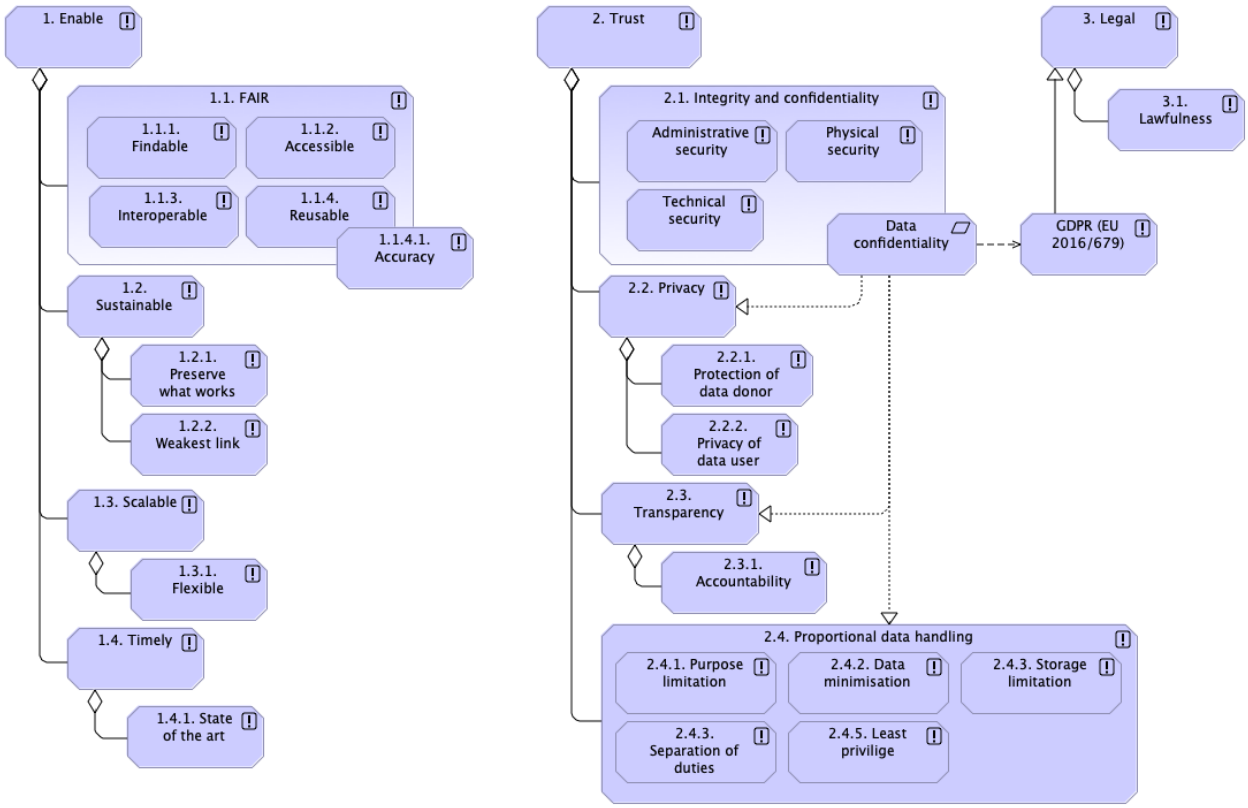
4 Core SPE requirements

4.1 Principles

Principle elements are used in Archimate language to all external needs that affect requirements and capabilities. Here, legal principles are EU-wide laws that guide or define legal requirements. They can be referred to by name or down to a specific paragraph of an article that spell out the specific requirement. The most important laws for SPEs and sensitive data are covered in Annex G: Overview of relevant EU regulations.

With respect to functional requirements, practical and ethical design principles are best categorised as either enabling or trust-dependent ones. Trust covers traditional cybersecurity principles but also wider aspects like privacy and transparency (Figure 4.1).

Figure 4.1. Principles affecting the SPE service provision that separates enabling and trust-dependent principles. GDPR is the main European legislation determining the application of data confidentiality requirements that balances a wide selection of these principles.



Enabling principles relate to the benefits and functionalities that SPEs are expected to deliver. These may vary depending on the perspective — for example, from a societal viewpoint (e.g. scientific progress), a service provider's viewpoint (e.g. performance,

scalability, maintainability), or a data user's viewpoint (e.g. usability, analytical capabilities, cost).

The most important set of enabling principles are FAIR. The combination of findability, accessibility, interoperability and reusability determine the how well data and services support scientific research.

Other enabling aspects gauge if the design is reasonable for current needs (state of the art), it can be maintained with the resources available (sustainable), will be able to respond to ever-changing needs (flexible), and provide services in a timely manner (timely).

Data confidentiality is an application of the aspects of trust-building security principles to data management. The GDPR is the main legislation that lays out rules for data confidentiality and integrity in proportion of perceived risk.

4.2 User stories

In the context of the EHDS framework, two typical user profiles can be distinguished in relation to the use of SPEs, each associated with different roles, restrictions, and levels of responsibility.

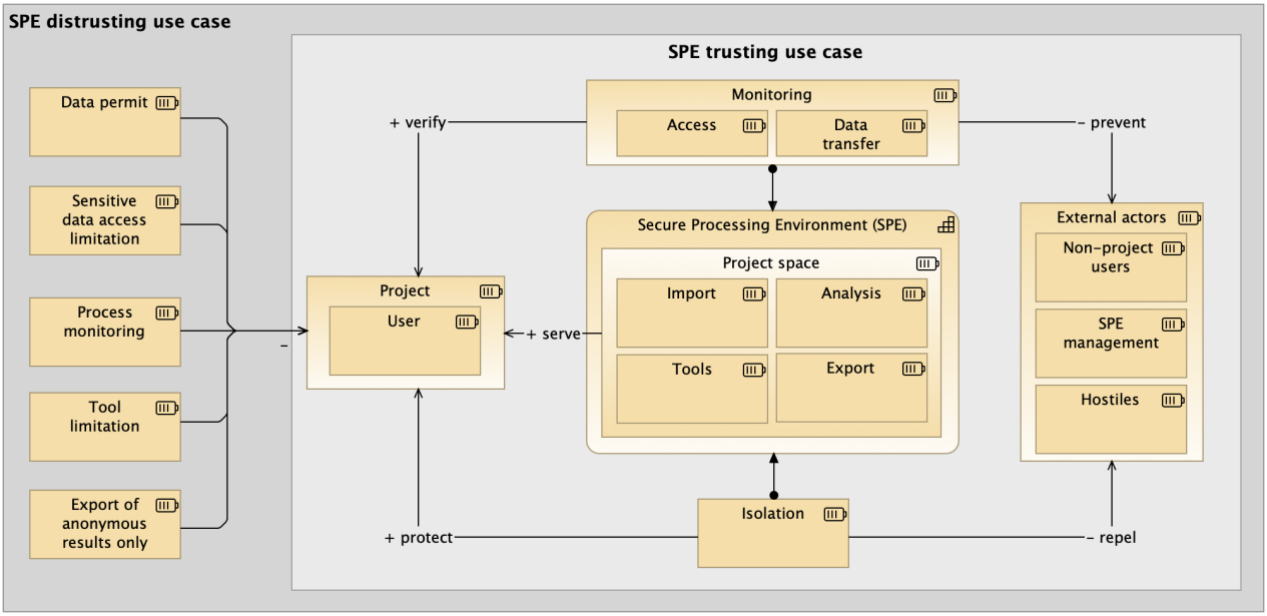
We must differentiate the two uses of the term SPE. For the SPE operator, it is a service that must be maintained to its users. SPE is also used to refer to collective functionalities of SPE user accessible instances. For users, SPE is a project-based environment that is fully separated from other project environments within the service. To distinguish these, the full term for the latter would be **SPE project-based user space**. Any shorter versions of this should include either 'user' or 'project' to be clear what environment is meant.

The main use case is the one where the user is the health data user processing health data for secondary use according to a data permit. The whole process is overseen by an HDAB that provides the prepared sensitive data to the user, and the user is not allowed to export anything but anonymous results out of the SPE.

The second use case defines the HDAB itself as the user. An HDAB employee must do the final processing of the permitted data that in the next stage will be handed over to the health data user's SPE environment. The HDAB user is fully responsible for both managing the import of datasets, possibly from multiple data holders, and export of the final permitted dataset with personal details. This use case utilises SPE solely to protect the data and its processing from non-users. It is subject to general access control monitoring of services, but nothing more. In wider context, this matches the requirements of generic scientific research wanting to protect their sensitive data processing.

These two use cases reflect different trust and control models, which have implications for design of SPEs and how technical and operations measures (TOMs) are implemented. To emphasise the importance of trust, we name these **distrusting and trusting use cases** (Figure 4.2). The challenge is to understand the implications of both use cases to the functional and technical design of SPE.

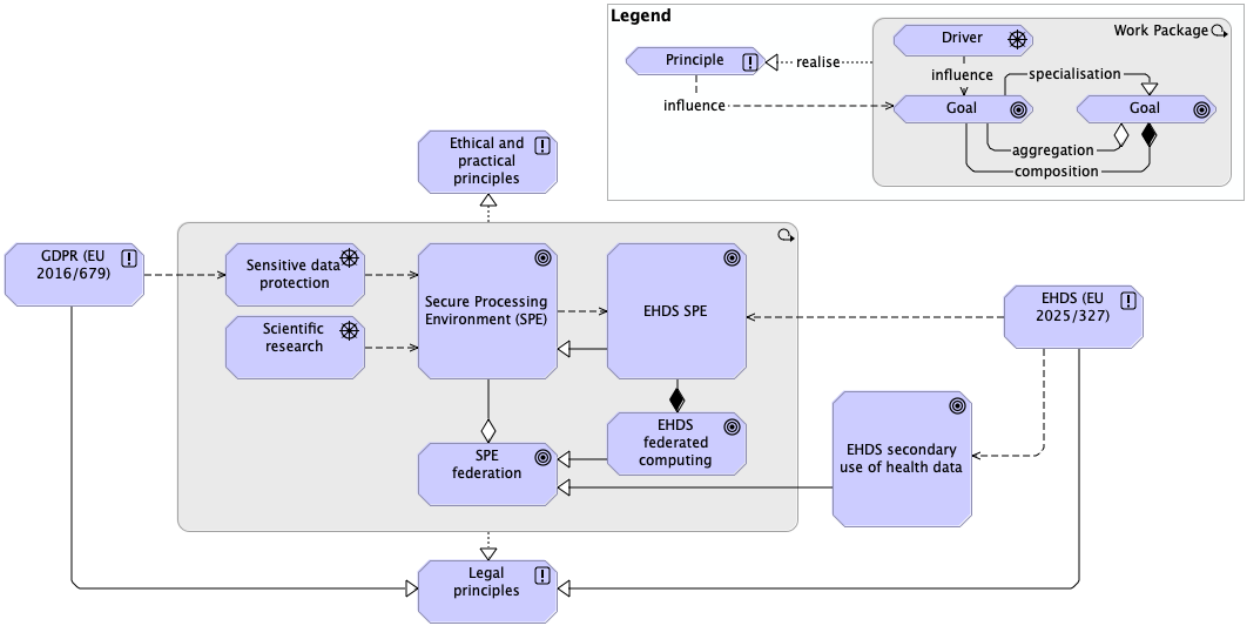
Figure 4.2. SPE trusting and distrusting use cases.



4.3 Functional requirements

This section defines the functional requirements of SPEs based on legal drivers and system-level goals. Our goals are to understand and define the minimum requirements of general-purpose SPEs, what additional requirements EHDS sets for SPEs, minimum requirements for a federation of SPEs, and how those federation concepts can be extended to enable federated computing (Figure 4.3).

Figure 4.3. Drivers and goals

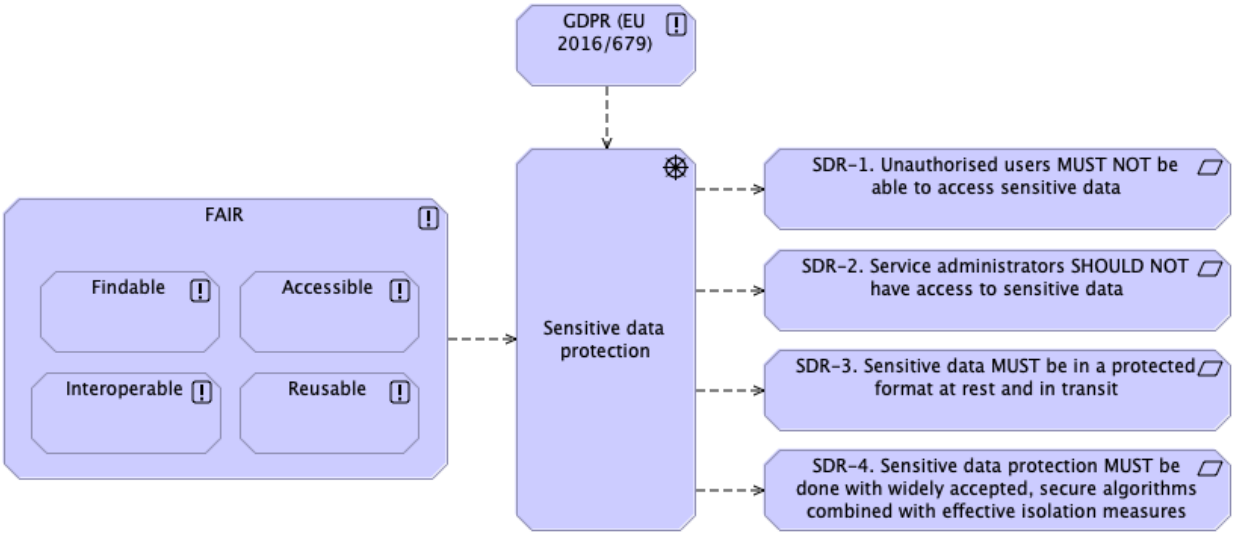


4.3.1 Sensitive data

It is important to see that the requirements and functionalities of SPEs depend equally on the restrictive pressure to **protect sensitive data** and to **enable scientific research**.

The need to protect the privacy of the data donor and data processing as expressed in GDPR constrains how the FAIR principles of that demand open science principles of findable, accessible, interoperable and reusable are applied to sensitive data (Figure 4.4., see also Figure 4.1 for GDPR in the context of general principles).

Figure 4.4. General sensitive data protection requirements. Requirement identifiers follow the convention 'SDR–n', where 'SD' indicates the sensitive data requirements namespace and 'R' denotes requirement.



All processing of sensitive data must be based on the **sensitive data privacy requirements**. The most important concept is the **authorised user**. It implements the principles of segregation of duties and least privilege. All other actors are non-authorised users. Security and privacy measures to protect sensitive data should always mention if they are directed to support or control authorised users, or deter non-authorised users or both.

That gives us our first requirement. The requirements are sequentially numbered and preceded by the namespace acronym followed by the letter 'R'. All requirements in this report are collected to a list in Annex B: SPE and related requirements.

SDR-1. Unauthorised users MUST NOT be able to access sensitive data

Only authorised users are allowed to access sensitive data inside the SPE according to their role that is internal to the project. User identification and measures to authenticate users are the foundation of sensitive data security.

The next requirement emphasises this difference between processing by data users and service administrators.

SDR-2. Service administrators SHOULD NOT have access to sensitive data

Administrators responsible for maintaining the SPE service should not get in the project space of SPE with the sensitive data, except in cases where processing personal data is necessary for technical maintenance purposes (e.g. users explicitly asking to resolve technical issues).

To narrow the chances that sensitive data is available to non-authorised users, it needs to be protected by default. Encryption is the default way to protect digital information, and it lessens needs to isolate and guard the data.

SDR-3. Sensitive data MUST be in a protected format at rest and in transit

Encryption algorithms are constantly improved to answer to new vulnerabilities that technical development brings up. Encryption should rely on peer-reviewed, widely accepted algorithms that are aligned with recognised international standards. Only open encryption algorithms are generally considered safe. Encryption protection is complemented with isolation measures that are either physical, technical or operational.

Additionally, pseudonymisation is a widely used method to protect the privacy of individuals from authorised users when they do not need all information and from accidental exposure to non-authorised users. Pseudonymisation and non-sensitive data types such as anonymised and synthetic data are covered in TEHDAS2 M7.2 Draft guideline on data minimisation, pseudonymisation, anonymisation and synthetic data.

SDR-4. Sensitive data protection MUST be done with widely accepted, secure algorithms combined with effective isolation measures

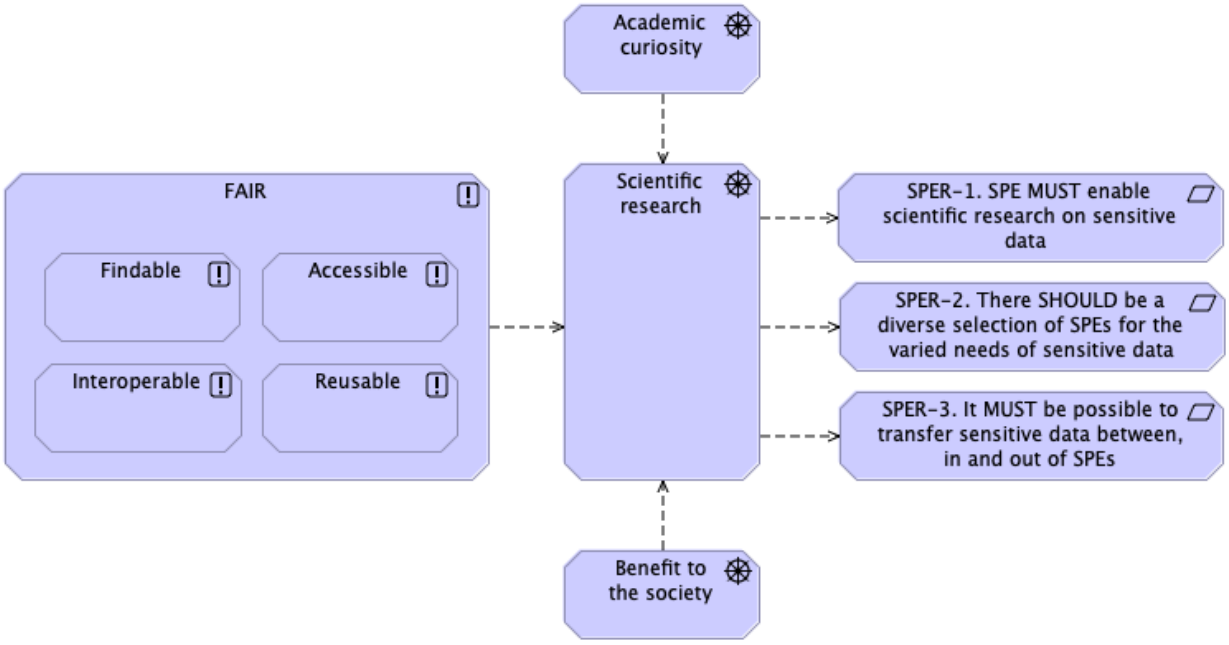
4.3.2 Scientific research

The connection to scientific research helps us to understand the important key difference between sensitive and classified data. Similar methods are used to secure both types of data, but these data types have opposing aims. Secrets are meant to be dangerous facts that society wants to protect and make available for a very limited group of people only when necessary. Protection of sensitive data aims to enable lawful and open-ended research on sensitive data as widely as possible to benefit society.

While sensitive data processing can have many different purposes, their effect on the requirements of SPEs are difficult to tell apart from wider scope of scientific research. We use here scientific research as a catch-all term that emphasises the open-endedness and neutrality of sensitive data processing requirements that includes the idea that it generates value to the society.

Scientific research is a curious combination of society's need to promote the academic curiosity of individuals to reap benefits from their explorations of unknown in a way that unexpected results are common (Figure 4.5).

Figure 4.5. Scientific research requirements. Requirement identifiers follow the convention 'SPER–n', where 'SPE' indicates the general purpose SPE requirements namespace and 'R' denotes requirement.



Openness and interoperability as expressed in the FAIR principles are essential elements of scientific research. In the context of sensitive data processing, they also embody the unpredictability of scientific research results that must be considered in the SPE design. Since it is impossible to determine beforehand what twists and turns empirical scientific research will take, the functional definition of an SPE must be open-ended. In other words, we cannot limit the SPE capabilities *a priori*.

The requirements of general purpose SPEs have by the namespace acronym 'SPE' followed by the letter 'R'. All requirements in this report are collected to a list in Annex B: SPE and related requirements.

SPER-1. SPE MUST enable scientific research on sensitive data

It is impossible for any single SPE to fulfil all present and future requirements. Data capacity and analysis capabilities will differ widely.

SPER-2. There SHOULD be a diverse selection of SPEs for the varied needs of sensitive data research

The scope of EHDS on the secondary use of health data does help to focus on the kind of analysis that is most likely done inside an SPE, but it does not separate these needs from other domains of scientific research.

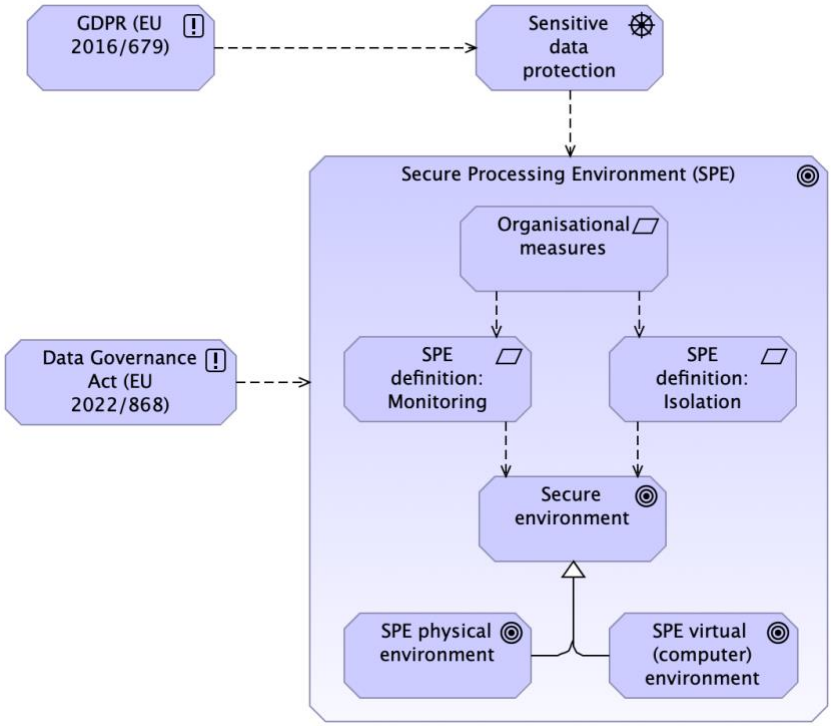
SPER-3. It MUST be possible to transfer sensitive data between, in and out of SPEs

Activities that create sensitive data are out of scope but to make datasets available for analysis they need to be transferred into an SPE. Also, the diversity and changing needs of research will make it necessary that sensitive data must be transferred between SPEs during the analysis.

4.3.3 Minimum SPE requirements

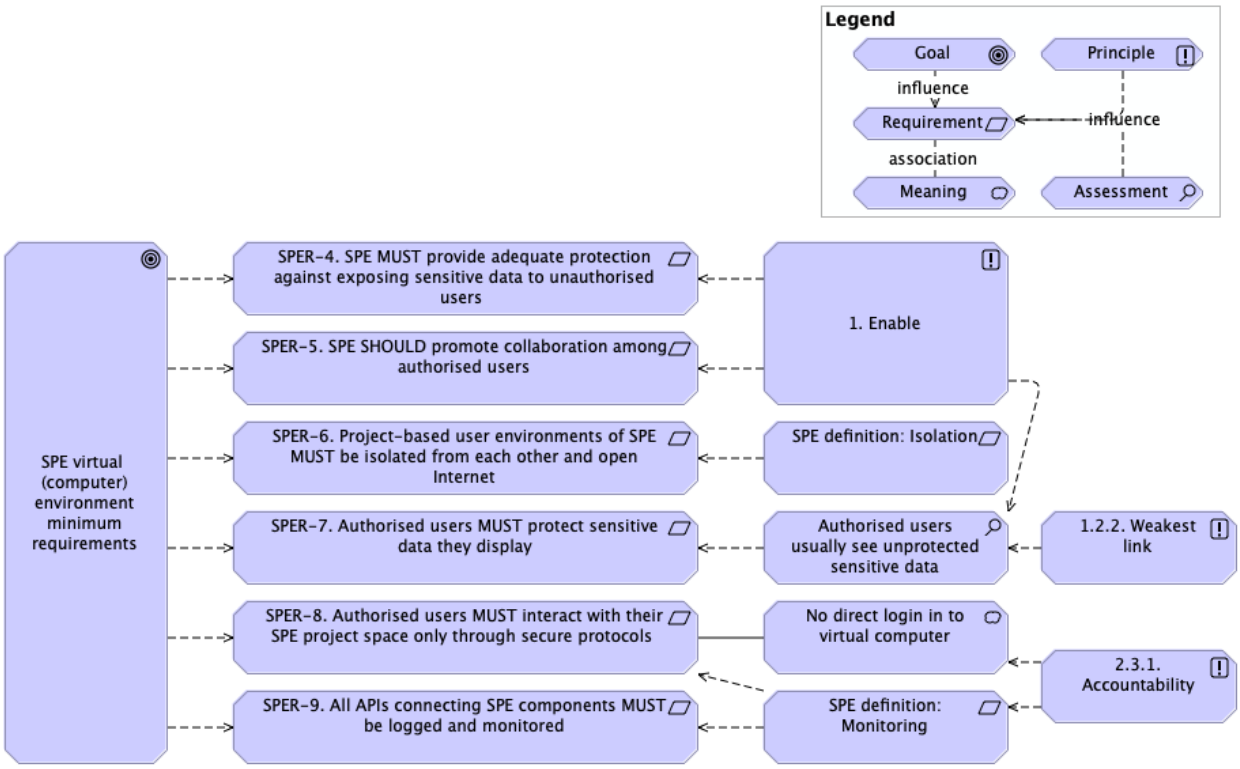
When applying the wide and risk-based requirements from GDPR to SPE, they can be reduced to two main security conditions: monitoring and isolation (Figure 4.6).

Figure 4.6. Legal conditions of SPE



In this context, we can ignore the physical environment and focus on SPE as a virtual computer environment (Figure 4.7).

Figure 4.7. SPE minimum requirements. Requirement identifiers follow the convention 'SPER–n', where 'SPE' indicates the general purpose SPE requirements namespace and 'R' denotes requirement.



SPER-4. SPE MUST provide adequate protection against exposing sensitive data to unauthorised users

The aim of the isolation is two-fold: It protects clear-text sensitive data from non-authorised users and enables authorised users to concentrate on their analysis tasks.

SPER-5. SPE design SHOULD promote collaboration among authorised users

It is not enough that SPE makes it possible to analyse sensitive data, its design and capabilities should actively seek to promote this collaboration effort among authorised users.

SPER-6. Project-based user environments of SPE MUST be isolated from each other and open Internet

SPEs must be designed in such a way that user environments within it are clearly and completely separated from each other and from the open Internet. When data processing includes remote queries or federated computing, this means that user's project environment extends to include them.

SPER-7. Authorised users MUST protect sensitive data they display

With the freedom comes the responsibility. Authorised users bear the responsibility of their actions. For most of the expected SPE use, authorised users display plain-text sensitive data on their computer screens, and they have to be aware what they are allowed to do with that information.

Monitoring displayed sensitive information in modern distributed and virtualised environments is hard. If the use of SPE is made too cumbersome for technical or regulatory reasons, users are easily tempted to take note of the displayed information from the screen rather than follow official ways. In practice, this means that any addition of any security measures that limit the ability of authorised users to process sensitive data need to be evaluated against this danger.

SPER-8. Authorised users MUST interact with their SPE project space only through secure protocols

The effectiveness of all preceding measures must be monitored. The monitoring creates the transparency needed to retroactively ascertain past events and ensures the accountability of all actors.

Replacing of standard computer operating system logging protocols by secure alternatives promotes the overall security and uniformity of sensitive data processing. These usually involve multifactor authentication (MFA) and use of proxies.

SPER-9. All APIs connecting SPE components MUST be logged and monitored

Monitoring and logging requirements that include transparency and data minimisation are included in GDPR.

Expanding monitoring and security requirement to all SPE-related communications and data transfers to use secure application interfaces (APIs) opens the possibility to see discrete sensitive data processing events as part of the same user environment, even if they happen in different physical locations.

The last two requirements mean that the SPE provider is the one responsible for ensuring the SPE technical and organisational safety as per the SPE definition (DGA Art. 2(20)). The wording of this requirement differs from the exact definition of SPE for the reasons deliberated in the section on Monitoring of SPE use.

4.4 EHDS SPE requirements

Article 73 of the EHDS regulation lists baseline legal requirements for SPEs, serving as a foundation and minimum standard for the guidelines. Article 73(5) provides the empowerment for further implementing acts that this report aims to advise.

Article 73 requires that all processing of electronic health data for secondary use purposes must take place in a secure processing environment (SPE). These environments must have robust data protection processes, restricted unauthorised access, and prevent data from being copied or transferred unlawfully. SPE providers must comply with relevant EU laws, ensuring transparency and oversight to maintain trust and privacy.

Requirements derived from Article 73 are given below and in Figure 4.8. The detailed analysis of Article 73 and derivation of its requirements is provided in Annex G: EHDS article 73 analysis to deduce SPE requirements.

The functional and operational requirements of SPEs derived from the EHDS regulation have by the namespace acronym 'EHDS' followed by the letter 'R'. All requirements in this report are collected to a list in Annex B: SPE and related requirements.

EHDSR-1. HDAB MUST grant access to EHD using a data permit

EHDSR-2. EHD MUST be accessed using an SPE

EHDSR-3. Natural persons listed in the data permit MAY access the identified EHD in SPE

EHDSR-4. TOMs MUST minimise the risk of unauthorised EHD access in SPEs

EHDSR-5. Authorised health data users MUST be strongly identified

EHDSR-6. All access and operation logs of SPE MUST be available for verification and auditing

EHDSR-7. All SPE logs MUST identify the actor

EHDSR-8. All SPE logs MUST be kept at least for one year

EHDSR-9. TOMs of SPEs MUST be monitored for security

EHDSR-10. EHD MUST be identified in the data permit

EHDSR-11. Health data holder MUST upload the permitted EHD to be available in an SPE for the health data user

EHDSR-12. Health data user MAY download only non-personal EHD from SPE. Anonymised personal data is non-personal

EHDSR-13. HDAB MUST ensure by reviewing that no personal data is taken out of the SPE by the health data user

EHDSR-14. Regular internal and external security audits MUST be done on SPE TOMs

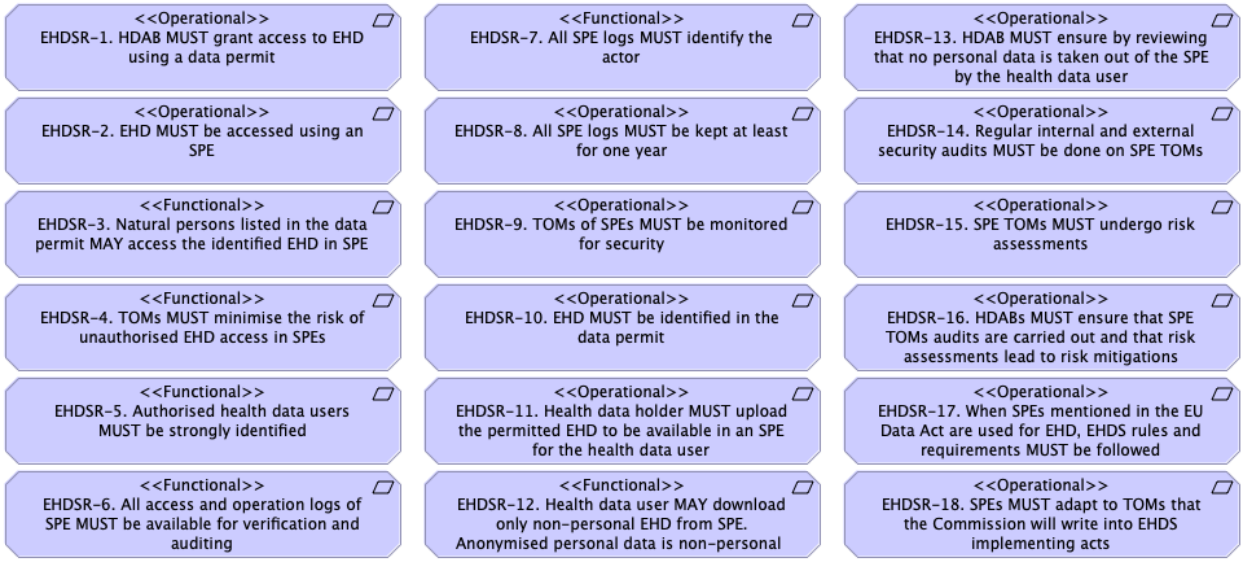
EHDSR-15. SPE TOMs MUST undergo risk assessments

EHDSR-16. HDABs MUST ensure that SPE TOMs audits are carried out and that risk assessments lead to risk mitigations

EHDSR-17. When SPEs mentioned in the EU Data Act are used for EHD, EHDS rules and requirements MUST be followed

EHDSR-18. SPEs MUST adapt to TOMs that the Commission will write into EHDS implementing acts

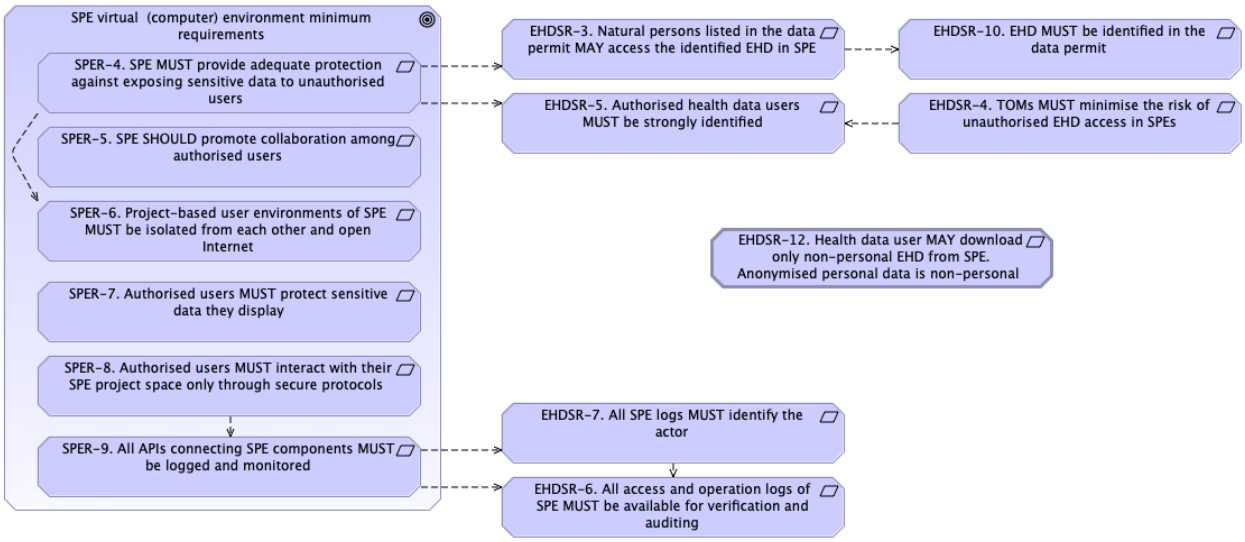
Figure 4.8. SPE requirements derived from EHDS Article 73. Requirement identifiers follow the convention 'EHDSR–n', where 'EHDS' indicates the EHDS requirements namespace and 'R' denotes requirement.



Seven of these 18 requirements are functional in a way directly affecting the functionality of SPEs under EHDS (Figure 4.9). These requirements touch either security of data processing and user identity giving specific demands (SPER-1), or logging requirements (SPER-9) already covered in more general terms in generic SPE requirements. The problematics about logging details in EHDSR-6 already seen coming from the SPE definition is covered in chapter Monitoring of SPE use. The novel requirement of allowing health data users to only export anonymised results from the SPE (EHDSR-12) will be discussed in chapter Data export from SPE.

Figure 4.9. EHDS SPE functional requirements. Requirement identifiers follow the convention: 'SPER–n' for general purpose SPE requirements ('SPE' namespace + 'R' for

requirement) and 'EHDSR–n' for EHDS requirements ('EHDS' namespace + 'R' for requirement).



4.5 Operational requirements

Operational requirements in SPEs refer to the necessary processes, controls and capabilities that ensure systems handling sensitive data operate securely, reliably and in line with the functional obligations defined in Article 73 of the EHDS Regulation. These requirements cover both technical and procedural aspects, such as access management, system configuration, monitoring, incident handling and service continuity. They are primarily derived from the functional requirements laid down in the EHDS Regulation, which establishes the mandatory framework for SPE setup and operation.

However, it is important to consider that the EHDS Regulation enters a broader ecosystem of established standards, directives, and best practices that inform the design and implementation of operational controls. In developing the operational requirements presented in this section, we have analysed several state-of-the-art references to complement the EHDS provisions and provide practical guidance for SPE operators:

- **ISO/IEC 27000 series**[2]:Internationally recognised standards for information security management, providing a structured approach for establishing, implementing, and maintaining an Information Security Management System (ISMS). While ISO 27001/27002 are not explicitly mandated by the EHDS Regulation, they offer widely accepted guidance for implementing robust security processes and controls.

- **NIS2 Directive (Directive (EU) 2022/2555):** A European directive establishing cybersecurity risk management obligations. Importantly, NIS2 obligations may apply to SPE operators depending on their organisational context and national transposition. NIS2 provides guidance on areas such as incident response, access control, supply chain security, and business continuity. Its obligations serve as a reference for state-of-the-art

[2] ISO27001 https://www.iso.org/standard/27001

operational practices, but compliance depends on the scope defined by national legislation.

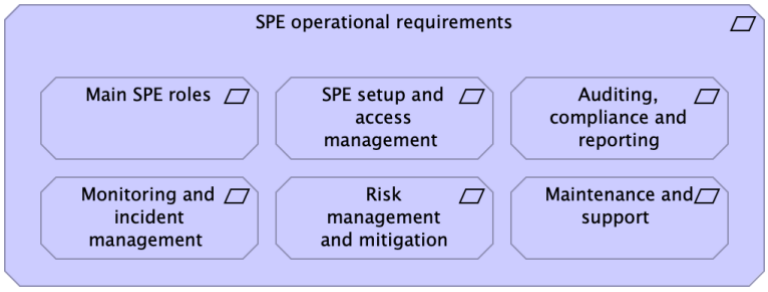
- **FitSM:** A practical framework for IT service management that supports structured service governance, clear role definitions, and disciplined operational processes. FitSM is not a mandatory requirement; rather it can be adapted to an organisation's needs and resources. Implementation can be approached in phases, starting with core processes for operational stability, followed by scalability-focused processes, and culminating in continual improvement to enhance service performance and maturity. FitSM is proposed here as a helpful approach to unify governance and operational practices across SPE providers.

By drawing on these references, HDABs/SPE operators can better understand the practices, protocols, and controls that support secure and resilient operation. None of these references are explicitly mandated by the EHDS Regulation; instead, they provide examples of well-established practices that can help SPEs meet regulatory expectations and implement robust operational management.

We will collect SPE operational requirements under six categories (Figure 4.10):

- Main SPE roles
- SPE setup and access management
- SPE auditing, compliance and reporting
- Monitoring and incident management
- Risk management and mitigation
- Maintenance and support

Figure 4.10. SPE operational requirement categories



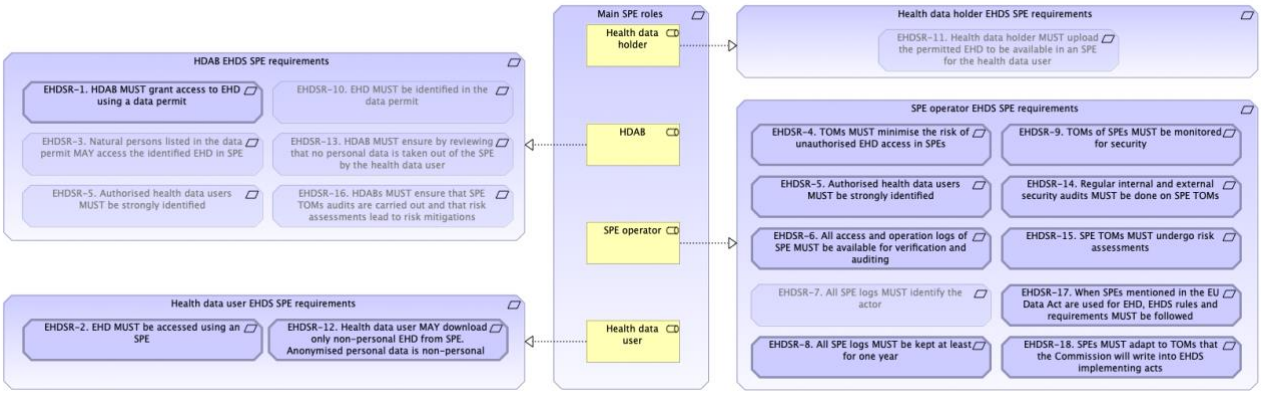
We will mark these requirements in relation to the EHDS regulation, its identified functional requirements, and to the three aforementioned regulations. The conclusions will be discussed in chapter Operational requirements for SPE federation.

4.5.1 Main SPE roles

In the context of the EHDS, operational requirements are shaped by a framework of clearly defined roles that collectively ensure the secure, lawful, and ethical secondary use of electronic health data.

The core roles - Health Data Access Body, SPE Operator, Data Holder, and Health Data User - each carry specific responsibilities that influence or depend on robust operational processes. These roles do not operate in isolation; rather they interact through structured processes, mutual dependencies and shared responsibilities that ensure compliance with legal standards, while also enabling technical interoperability and data protection. Their relation to EHDS SPE requirements is shown in Figure 4.11.

Figure 4.11. The main SPE roles and their relation to identified EHDS SPE requirements. Purely functional requirements have been greyed down. Requirement identifiers follow the convention 'EHDSR–n', where 'EHDS' indicates the EHDS requirements namespace and 'R' denotes requirement.



HDAB: The HDAB is the national authority tasked with overseeing the access to sensitive electronic health data. The HDAB ensures that data provided to users corresponds to the conditions set out in the data permit. The HDAB also oversees SPE compliance, initiates and conducts audits to ensure conformity with data security standards. Beyond their regulatory responsibilities, HDABs may, in practice, function as a central repository of expertise on the secondary use of health data and as the primary point of contact for related information. In this context, it is considered good practice for HDABs to provide guidance on lawful and appropriate data use, deliver training on the handling of sensitive health data, and operate a helpdesk to address user enquiries concerning data access, use, and compliance.

SPE Operator: According to the EHDS Regulation, the responsibility for managing and operating the SPE lies with the HDABs. However, an HDAB may designate a separate entity to act as the SPE operator, in which case the operator functions as a processor on behalf of the HDAB under a data processing agreement. In such arrangements, the SPE operator is expected to oversee all technical, operational, and security measures of the environment, and may additionally provide user support for technical or environment-related issues. When an SPE operator is appointed, the data processing agreement should clearly specify the allocation of duties and responsibilities. The data management activities assigned to HDABs under the EHDS Regulation, such as ensuring log availability and auditability, must remain guaranteed for the HDAB, even if SPE operation is outsourced.

Data Holder: The Data Holder is the organisation or entity that possesses the original health datasets, such as hospitals, research institutions or health registries. Their role is to provide the requested data to the HDAB under appropriate legal and technical safeguards.

Health Data user: The Health Data User is the authorised individual or organisation accessing health data within the SPE for approved secondary use of health data. They are responsible for using the data in accordance with the permit, respecting all ethical and legal obligations and operating within the secure boundaries defined by EHDS.

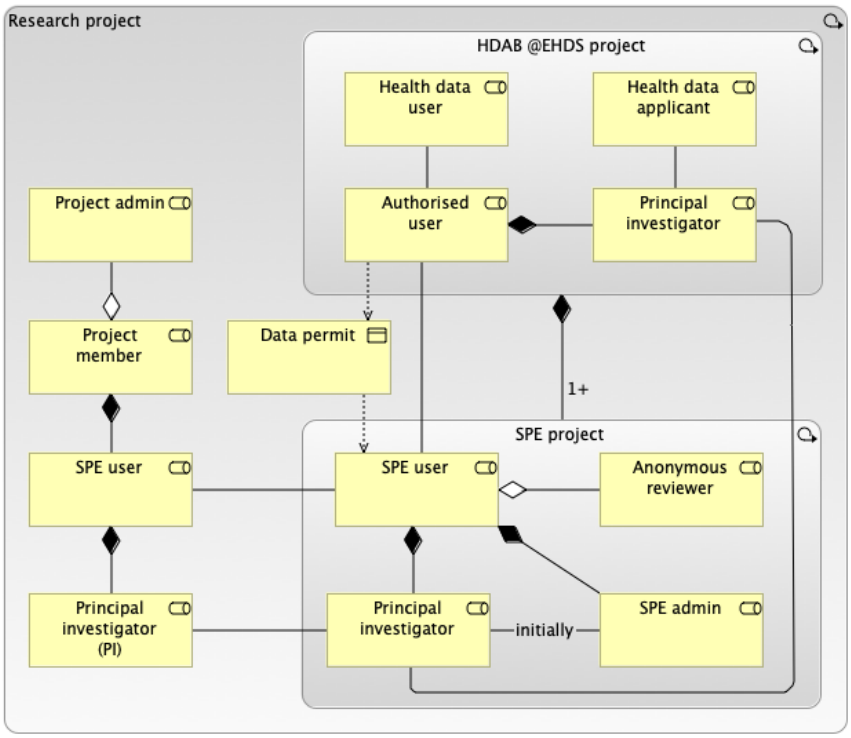
Health data user is here defined from the EHDS perspective as the entity that carries the responsibility for the approved health data. Natural persons mentioned in the health data permit are authorised users. For researchers, the situation is more complicated. To clarify it, we need to specify the correct **project context for their roles (**Figure 4.12).

A **research project** may consist of many members with different roles, and not all of them are accessing sensitive data specified by the data permit. To get the data permit, the principal investigator (or one of the many) acts as the health data applicant to submit a data application. A positive decision creates the data permit that includes identifiers to approved **EHDS project** authorised users, of which the principal investigator is one. The data permit gives access to applied data and to at least one SPE service. Based on the data permit, the principal investigator establishes a project in the SPE to start analysing the data and includes other authorised users to it.

The principal investigator, who started the **SPE project,** might not be technical enough or might be too busy to be continuously involved with the day-to-day activities in the SPE. It would be convenient to specify a role, SPE admin, that the principal investigator can hand over to another capable SPE user who is available. This does not reduce the responsibility of the principal investigator under EHDS Article 68 (10)(d) (See also Recital 62).

Academic publication system has a role of **anonymous reviewer** who might need to access original data and ascertain the analysis has been done correctly before publication is accepted. Since under EHDS, the permitted sensitive data cannot be taken out of the SPE, the reviewer needs to be added to the data permit to be a health data user, but that identity should not be made available to other users in the updated data permit. Further, the SPE operator should be able to obfuscate the identity of the reviewer by using an alias not to leave identifying information inside the SPE project area.

Figure 4.12. Role names in three different projects that are all encompassing research project, that could include multiple EHDS projects and SPE projects.



4.5.2 SPE setup and access management

According to EHDS rules, the setup of each SPE must be tailored to the specific conditions of the data permit (EHDSR-1), which identifies authorised users (EHDSR-3) who may access the specified data (EHDSR-10) and the purpose of data processing inside the named SPE (EHDSR-2).

SPEs must be set up to allow HDABs to make health data available to the health data user within two months of receiving it from the health data holder (Article 68(7)). This means that the SPE must be operational within this timeframe and configured according to the terms of the data permit. Access to the SPE must be restricted to approved users (Article 73(1)) and apply strong security measures to prevent unauthorised access (Article 73(1)(b)). Additionally, all data must be deleted within six months after the data permit expires (Article 68(12)).

To operate effectively and securely, HDABs must have clear procedures for granting and revoking access based on the roles and authorisations set out in the data permit, and limit privileged access to essential personnel only, thereby minimising the risk of internal misuse. They should also ensure that all health data users understand their responsibilities, including best practices for handling sensitive health data and the obligation to comply with the terms defined in the data permit. Furthermore, HDABs must be able to suspend or terminate access to the environment if misuse or security breaches are detected.

By following these requirements, HDABs have to provide a secure and compliant SPE environment for processing health data while protecting privacy and ensuring regulatory compliance.

The identified operational requirements of SPEs have the namespace acronym ‘OP’ followed by the letter ‘R’. All requirements in this report are collected to a list in Annex B: SPE and related requirements.

Table 4.13. lists these requirements and their derivation that is shown in figure 4.13.

Figure 4.13. SPE setup and access management. Requirement identifiers follow the convention: ‘SPER–n’ for general purpose SPE requirements (‘SPE’ namespace + ‘R’ for requirement), ‘EHDSR–n’ for EHDS requirements (‘EHDS’ namespace + ‘R’ for requirement) and ‘OPR–n’ for operational requirements (‘OP’ namespace + ‘R’ for requirement).

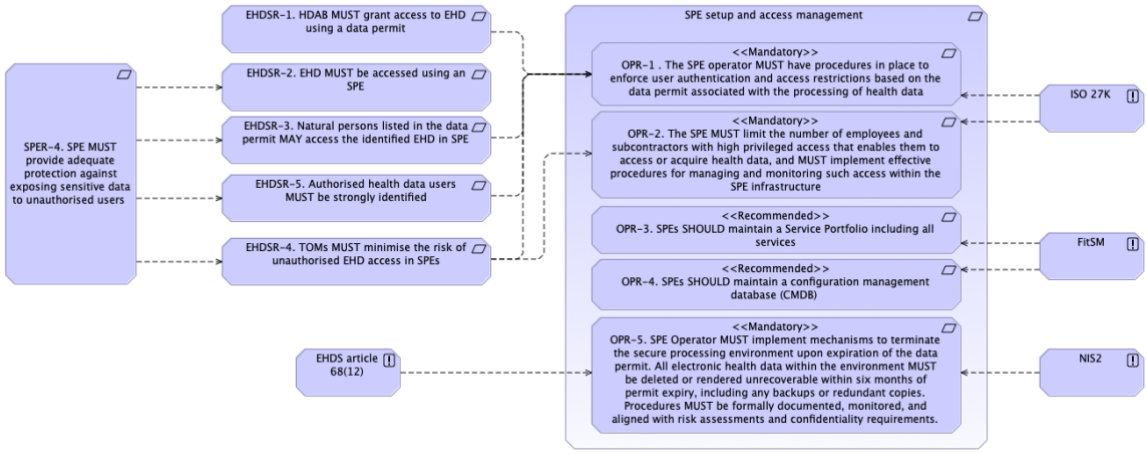


Table 4.13: SPE setup and access management

#	Link	Area	Requirement	Rationale	Importance
OPR-1	EHDSR-1 EHDSR-3 EHDSR-4 EHDSR-5	Access management	**SPE Operator MUST have procedures in place to enforce user authentication and access restrictions based on the data permit associated with the processing of health data.** These procedures shall ensure that only authorised natural persons listed in the data permit can access	Article 73(1)(a) - Restriction of SPE access to authorised natural persons listed in the data permit Article 73 (1)(d) - Use of individual and unique user identities and confidential access modes to assure health data users have access only	Mandatory

#	Link	Area	Requirement	Rationale	Importance
			the SPE, with access rights regularly reviewed and adjusted according to the data permit conditions. User identities shall be unique, persistent, auditable, and non-transferrable. Role-based access control shall be implemented to ensure that users have only the minimum level of access necessary to perform their authorised tasks. If an alias or user ID is used, it must be uniquely attributable to a single individual within the organisation's identity management system. The true identity behind any alias must be known and auditable at all times.	to the electronic health data covered by their data permit. ISO 27001/27002 - Best practices on authentication and access control	
OPR-2	EHDSR-4	Access management	**SPE Operator MUST limit the number of authorised staff and any subcontractors who have high-privileged access enabling them to access or process health data and MUST implement effective procedures for managing and monitoring such access within the SPE infrastructure.** Exceptionally, administrators may obtain privileged access for technical troubleshooting or incident response, provided this is explicitly authorised, time-limited, and subject to enhanced monitoring and documentation.	Article 73(1)(c) - Restrict access to electronic health data in the SPE (input, inspection, modification, and deletion) to a limited number of authorised and identifiable individuals. ISO 27002 recommendation - organisations must strictly limit and manage privileged access by assigning it only when necessary, regularly reviewing permissions, and logging all privileged actions. Elevated access must be controlled: authorised temporarily, detailed in logs, and promptly reviewed for necessity and accountability.	Mandatory

#	Link	Area	Requirement	Rationale	Importance
OPR-3		Set up	**SPE Operator SHOULD maintain a Service Portfolio including all services** (e.g., data ingestion, analysis platforms, audit tools). Include: Descriptions Availability Security levels Eligible user groups (e.g., researchers, data owners) Publish the catalogue internally and regularly update it.	FitSM Service Portfolio recommendation	Recommended
OPR-4		Set up	**SPE Operator SHOULD maintain a configuration management database (CMDB)** that includes: - All virtual machines, services, APIs, and databases - Tag systems based on data classification (e.g., sensitive, anonymised)	FitSM Configuration Management recommendation	Recommended
OPR-5	EHDSR-1	Set up	**SPE Operator MUST implement mechanisms to terminate the secure processing environment upon expiration of the data permit. All electronic health data within the environment MUST be deleted or rendered unrecoverable within six months of permit expiry, including any backups or redundant copies. Procedures MUST be formally documented, monitored, and aligned with risk assessments and confidentiality requirements.**	Article 68(12) - electronic health data in the SPE must be deleted within six months of the data permit's expiry. NIS2 emphasises secure data disposal as part of business continuity and disaster recovery planning, including managing backups and redundant data to prevent unauthorised access or data leaks.	Mandatory

4.5.3 SPE auditing, compliance and reporting

Technical and organisational security measures

To ensure health data is handled securely, SPEs must follow strict technical and organisational security measures. These measures help protect sensitive data, prevent unauthorised access, and comply with legal obligations under the EHDS Regulation and the NIS2 Directive.

The EHDS regulation, particularly Article 73, requires that SPEs minimise security risks, monitor compliance, and undergo regular audits to identify and fix vulnerabilities. At the same time, the NIS2 Directive sets additional cybersecurity rules, including risk assessments, encryption policies, incident response plans, and security checks for suppliers.

One key requirement is that SPE Operators implement an ISMS to establish and maintain security policies aligned with applicable laws, such as the GDPR. The ISMS should also be complemented with backup and disaster recovery plans, employee cybersecurity training and use of encryption to protect data and security assessment of suppliers and external systems.

By following these security measures, SPEs can provide a safe and reliable environment for processing health data while ensuring compliance with EU regulations.

Table 4.14. lists these requirements and their derivation that is shown in Figure 4.14.

Figure 4.14. SPE auditing, compliance and reporting. Requirement identifiers follow the convention: 'EHDSR–n' for EHDS requirements ('EHDS' namespace + 'R' for requirement) and 'OPR–n' for operational requirements ('OP' namespace + 'R' for requirement).

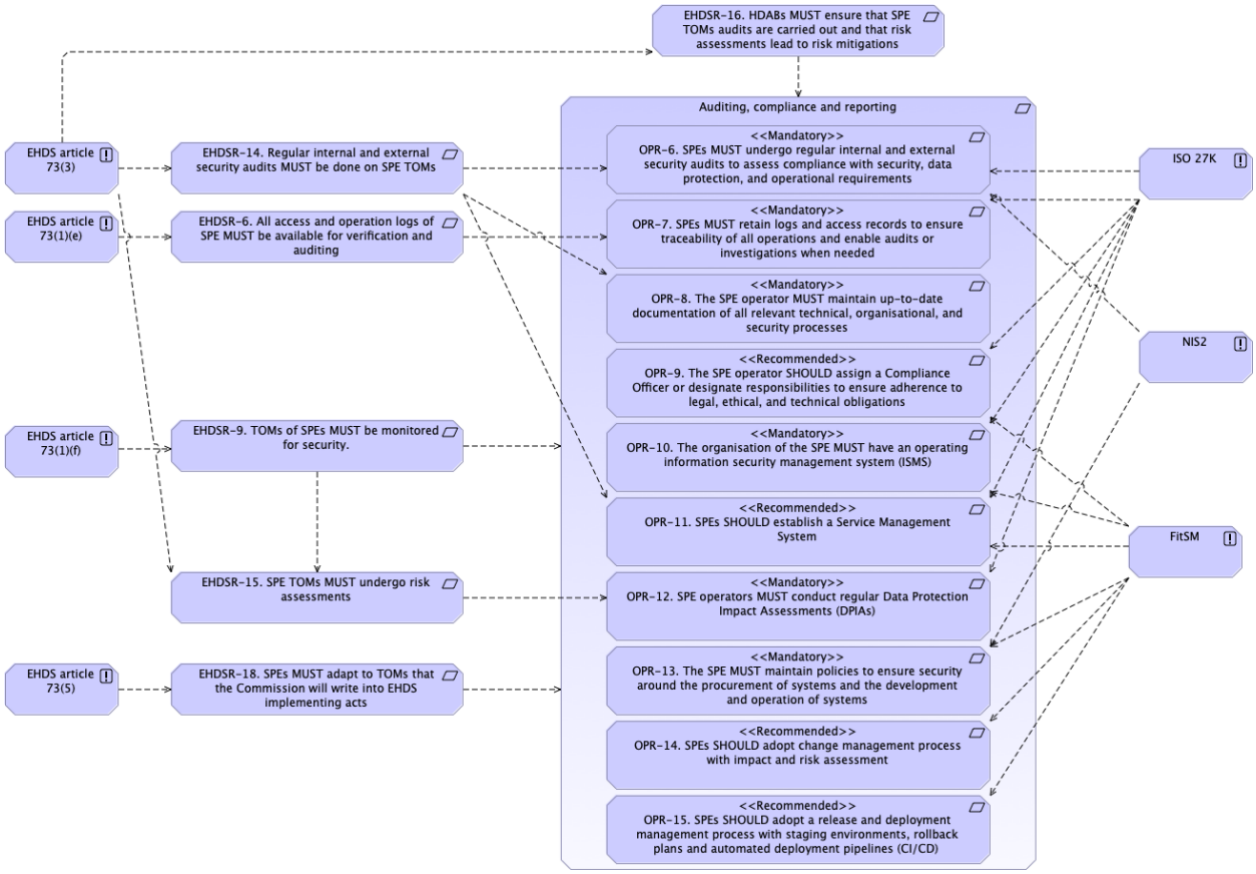


Table 4.14. SPE auditing, compliance and reporting

#	Link	Area	Requirement	Rationale	Importance
OPR-6	EHDSR-14	Auditing	**SPE Operator MUST undergo regular internal and external audits to assess compliance with security, data protection, and operational requirements.**	EHDS Regulation (Article 73(3)) explicitly requires regular audits by internal or third-party entities. NIS2 Directive mandates the evaluation of risk-management measures and corrective actions. ISO/IEC 27001 and 27002 require planned internal audits to verify the ISMS and other controls.	Mandatory

#	Link	Area	Requirement	Rationale	Importance
OPR-7	EHDSR-6	Auditing	**SPE Operator MUST retain logs and access records to ensure traceability of all operations and enable audits or investigations when needed.**	EHDS Regulation (Article 73(1)(e)) obliges SPEs to ensure logs of data access and operations/activities in the SPE. ISO/IEC 27002 defines how logs and access records should be retained and reviewed: • Logs should capture key details such as user IDs, timestamps, and the type of activity performed. They must be protected against tampering, stored in a time-synchronised environment, and regularly reviewed to detect anomalies or unauthorised access. Logs access should be restricted to authorised personnel.	Mandatory
OPR-8	EHDSR-14	Compliance/ Reporting	**SPE Operator MUST maintain up-to-date documentation of all relevant technical, organisational, and security processes.**	ISO/IEC 27001 and 27002 mandate control documentation and maintenance of procedures and records. FitSM supports structured documentation as a basis for consistent compliance and audit-readiness.	Mandatory
OPR-9	EHDSR-14	Compliance	**SPE Operator SHOULD assign a Compliance Officer or designate responsibilities to ensure adherence to legal, ethical, and technical obligations.** This role may be supported by other functions, such as Privacy Officer or, where applicable under the GDPR, a DPO (Data Protection Officer).	ISO/IEC 27002 recommends that roles and responsibilities related to security and compliance be clearly assigned and communicated. FitSM recommends the assignment of roles such as Information Security Manager or Compliance Officer to oversee governance. This ensures accountability and provides a single point of contact for audit, compliance, and regulatory matters.	Recommended
OPR-10	EHDSR-15 EHDSR-16	Compliance	**SPE Operator MUST have an operating information security management system (ISMS)** in line with the state of the art, for example following the requirements of ISO/IEC 27001. Following the recommendations from the FitSM standard, this should include: - End-to-end encryption - Role based access control - Multifactor authentication - Vulnerability assessments and penetration tracking	To have an ISMS in place is already a requirement of other laws (including GDPR). FitSM - Information Security Management requirement ISO 27001/27002 requirement	Mandatory

#	Link	Area	Requirement	Rationale	Importance
			- Document and update security controls		
OPR-11	EHDSR-14	Compliance	**SPE Operator SHOULD establish a Service Management System**: Define the scope of the SMS to include all services and components involved in processing health data. Appoint a Service Management Officer (SMO) responsible for maintaining SMS compliance.	FitSM SMS requirement ISO 27002: Maintain a real-time inventory of assets used to process health data	Recommended
OPR-12	EHDSR-14	Auditing	**SPE Operator MUST conduct regular Data Protection Impact Assessments (DPIAs)**	ISO 27002	Mandatory
OPR-13	EHDSR-14	Compliance	**SPE Operator MUST maintain policies to ensure security around procurement systems and development and operation of systems.** Under FitSM: Require Data Processing Agreements (DPAs) with security and audit clauses. Monitor SLA compliance of suppliers, particularly for critical infrastructure.	This is a NIS2 requirement to ensure security around supply chains and the relationship between the company and the direct supplier. Companies must choose security measures that fit the vulnerabilities of each direct supplier. And then companies must assess the overall security level for all suppliers. FitSM requirements ISO 27002 – security clause within procurement contracts	Mandatory
	EHDSR-18	Compliance	SPE Operator MUST adapt to TOMs that the Commission will write into EHDS implementing acts.	Article 73(5)	Mandatory
OPR-14	EHDSR-15	TOM	**SPE Operator SHOULD adopt change management process with impact and risk assessment**	FitSM requirements	Recommended
OPR-15	EHDSR-15	TOM	**SPE Operator SHOULD adopt a release and deployment management process.** These may include staging environments, rollback plans and automated deployment pipelines (CI/CD)	FitSM requirements	Recommended

4.5.4 Monitoring and incident management

To keep health data safe, SPEs must track and log all access and activities. Logs must be retained for a minimum of one year, as per Article 73(1)(e) of the EHDS, and for longer periods if justified by audit or compliance needs. These logs should be regularly audited to detect risks or vulnerabilities, and any issues identified must be addressed through corrective actions. The SPE Operator is the data controller of the user logs in their system, and they will hand over relevant information to the HDAB to fulfil its legal obligations for data safety.

SPEs should include procedures for the proper handling of security incidents, covering timely detection, alerting, and response. Existing frameworks, such as those described in NIS2 Directive, provide useful examples of state-of-the-art practices for incident management, including notification to the relevant Computer Security Incident Response Team (CSIRT), initial reports within 24 hours, detailed follow-ups within 72 hours, and final reports within one month. Adopting equivalent procedures will ensure that incidents are appropriately handled, reported, and mitigated, in line with current best practices.

As part of incident management, SPEs must have strong backup and recovery systems to protect electronic health data from loss, damage, or unauthorised access. Backups should be stored safely, separate from the original data, to avoid the same risks. The NIS2 requirements also highlight the importance of up-to-date backups and a clear recovery plan to keep IT systems running smoothly during and after security incidents. SPEs should regularly test backup restoration to ensure data can be recovered when needed.

These measures help protect health data, ensure compliance, and improve trust in secure data processing.

Table 4.15. lists these requirements and their derivation that is shown in Figure 4.15.

Figure 4.15. Monitoring and incident management. Requirement identifiers follow the convention: 'EHDSR–n' for EHDS requirements ('EHDS' namespace + 'R' for requirement) and 'OPR–n' for operational requirements ('OP' namespace + 'R' for requirement).

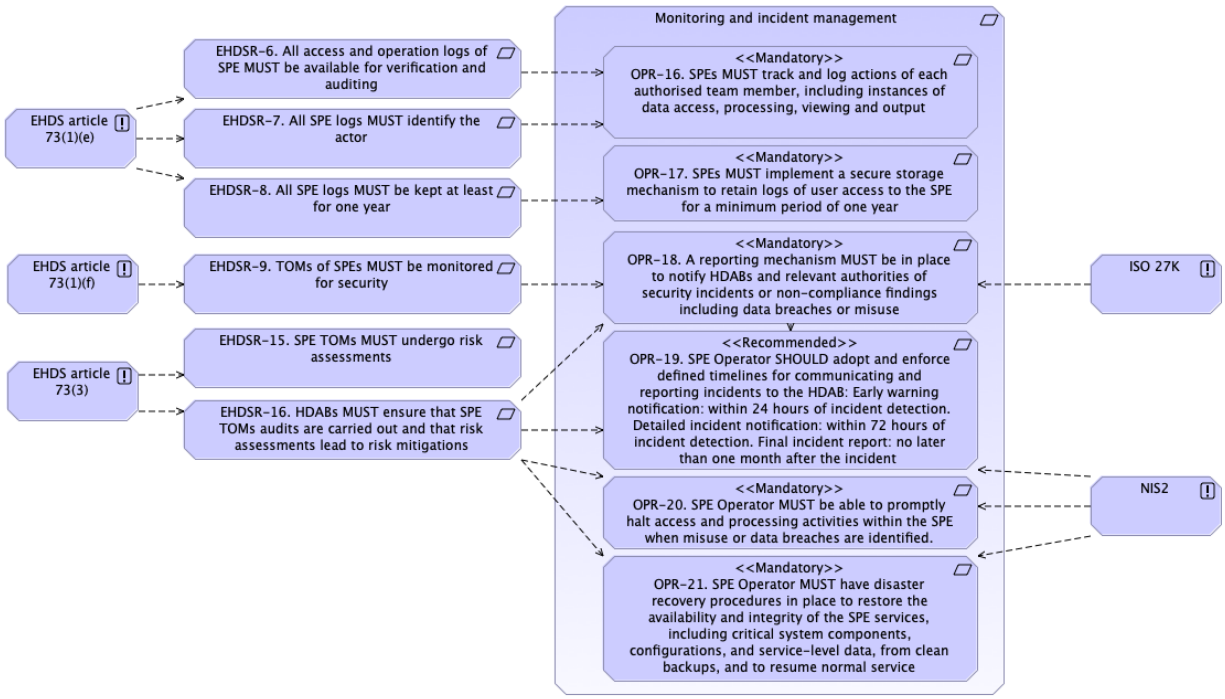


Table 4.15. Detailed operational requirements

#	Link	Area	Requirement	Rationale	Importance
OPR-16	EHDSR-6 EHDSR-7	Monitoring	**SPE Operator MUST track and log actions of each authorised project member, including instances of data access, processing, viewing and output** Under FitSM requirement, this should include: Use an ITSM tool (e.g., Jira Service Management, Freshservice) to log and track all incidents and requests. Classify and prioritise based on data sensitivity and service impact. Provide a clear user-facing incident reporting process. Maintain a 24/7 on-call rota for critical service issues.	Article 73(1)(e) - the keeping of identifiable logs of access to and activities in the secure processing environment for the period necessary to verify and audit all processing operations in that environment. FitSM requirement	Mandatory

#	Link	Area	Requirement	Rationale	Importance
OPR-17	EHDSR-8	Monitoring	**SPE Operator MUST implement a secure storage process to retain logs of user access to the SPE for a minimum period of one year**. While one year is the minimum retention period, longer retention (5 years) is recommended to support effective auditing procedures	Article 73(1)(e) - logs of access shall be kept for at least one year	Mandatory
OPR-18	EHDSR-9 EHDSR-16	Incident management	**A reporting process MUST be in place to notify HDABs and relevant authorities of security incidents or non-compliance findings including data breaches or misuse**	EHDS Regulation (Art. 73(3)) requires that SPEs allow for compliance checks and audits by competent authorities to notify the HDAB of any incidents affecting the integrity or confidentiality of data. ISO/IEC 27002 requirement: Incident reporting procedures should define what must be reported, by whom, how quickly, and through which communication channels. This includes the escalation path, responsibilities, and necessary documentation to ensure timely and effective response.	Mandatory
OPR-19	EHDSR-16	Incident management	**SPE Operator SHOULD adopt and enforce defined timelines for communicating and reporting incidents to the HDAB:** • Early warning notification: within 24 hours of incident detection • Detailed incident notification: within 72 hours of incident detection • Final incident report: no later than one month after the incident	The NIS2 Directive mandates incident notification to the relevant authorities within 24 hours of awareness, with a more detailed incident notification to be submitted within 72 hours (or within 24 hours for certain entities, such as trusted service providers). Additionally, a final incident report must be provided no later than one month after the incident.	Recommended
OPR-20	EHDSR-16	Incident Management	**SPE Operator MUST be able to promptly halt access and processing activities within the SPE when misuse or data breaches are identified.**	Article 73(3) - HDABs shall take corrective action for any shortcomings, risks or vulnerabilities identified. ISO 27001/27002 and NIS2 requirements identify the need for organisations to have procedures to stop or restrict access and processing in response to security incidents or data breaches.	Mandatory
OPR-21	EHDSR-16	Incident management	**SPE Operator MUST have disaster recovery procedures in place to restore the availability and integrity of the SPE services, including critical**	ISO 27001/27002 - Organisations must have a disaster recovery process that includes restoring systems and service	Mandatory

#	Link	Area	Requirement	Rationale	Importance
			system components, configurations, and platform-level functionality, from clean backups, and to resume normal service operations following an incident. Project-specific data recovery may be limited to what is feasible within backup and retention policies.	functionality from backups and return to normal operations. NIS2 requirement - Organisations must have plans for business continuity and disaster recovery, including the ability to restore critical systems and services after an incident.	

4.5.5 Risk management and mitigation

Effective risk management is critical for ensuring the security and reliability of SPEs. While SPEs are designed to safeguard sensitive data, they are vulnerable to a variety of security threats, including hardware vulnerabilities, software exploits and operational risks.

This section details the process for identifying, mitigating, and monitoring risks in SPEs. Based on the identified threats to SPEs (see Annex F: Classification of risks and threats against SPEs), the associated risks can be assessed and appropriate mitigation strategies planned. By implementing a comprehensive risk management framework, SPE providers can protect sensitive data and maintain system integrity.

Operational requirements that contribute to risk management and mitigation were already included in other sections, such as:

OPR-6: SPEs MUST undergo regular internal and external security audits to assess compliance with security, data protection, and operational requirements.
OPR-9: The SPE operator SHOULD assign a Compliance Officer or designate responsibilities to ensure adherence to legal, ethical, and technical obligations. This role may be supported by other functions, such as Privacy Officer or, where applicable under the GDPR, a DPO (Data Protection Officer).
OPR-14: SPEs SHOULD adopt change management process with impact and risk assessment
OPR-18: A reporting process MUST be in place to notify HDABs and relevant authorities of security incidents or non-compliance findings including data breaches or misuse.

Table 4.16. lists additional requirements and their derivation that is shown in Figure 4.16.

Figure 4.16. Risk management and mitigation. Requirement identifiers follow the convention: ‘SPER–n’ for general purpose SPE requirements (‘SPE’ namespace + ‘R’ for requirement),

'EHDSR–n' for EHDS requirements ('EHDS' namespace + 'R' for requirement) and 'OPR–n' for operational requirements ('OP' namespace + 'R' for requirement).

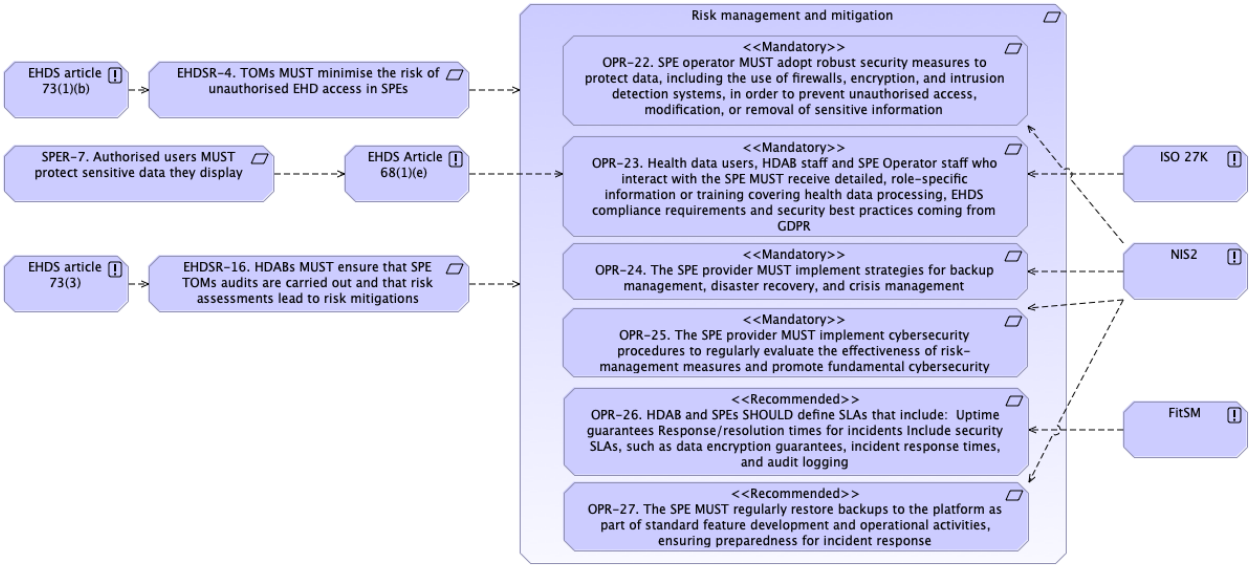


Table 4.16. Risk management and mitigation

#	Link	Area	Requirement	Rationale	Importance
OPR-22	EHDSR-16	Risk mitigation	**SPE Operator MUST adopt robust security measures to protect data, including the use of firewalls, encryption, and intrusion detection systems, to prevent unauthorised access, modification, or removal of sensitive information.**	Article 73(1)(b) - Technical and organisational measures must be in place to minimise the risk of the unauthorised reading, copying, modification or removal of EHD hosted in the SPE. NIS2 Directive explicitly references the implementation of appropriate security measures, such as incident detection and regular security assessments, to protect data and systems. Measures like encryption and multi-factor authentication are considered state-of-the-art practices that can support compliance with these requirements.	Mandatory
OPR-23	EHDSR-16	Risk mitigation	**Health data users, HDAB staff and SPE Operator staff who interact with the SPE MUST receive detailed, role-specific information or training covering health data processing, EHDS compliance requirements and security best practices coming from GDPR.**	Article 68(1)(e) - The health data applicant must demonstrate sufficient technical and organisational measures to prevent data misuse and protect the rights of data holders and individuals. Article 73(1)(b) - Technical and organisational measures must be in place to minimise the risk of the unauthorised reading, copying, modification or removal of EHD hosted in the SPE.	Mandatory

#	Link	Area	Requirement	Rationale	Importance
				ISO 27001/27002 requires role-appropriate security training for all staff and relevant contractors, covering policies, responsibilities, and basic security practices.	
OPR-24	EHDSR-16	Risk mitigation	**The SPE Operator MUST implement strategies for backup management, disaster recovery, and crisis management.**	NIS2 directive stipulates that organisations should have a plan for managing business operations during and after a security incident. This means that backups must be up to date. There must also be a plan for ensuring access to IT systems and their operating functions during and after a security incident.	Mandatory
OPR-25	EHDSR-16	Risk mitigation	**The SPE Operator MUST implement cybersecurity procedures to regularly evaluate the effectiveness of risk-management measures and promote fundamental cybersecurity practices and provide necessary training.**	NIS2 directive requires that employees receive cybersecurity training and practice for basic computer hygiene.	Mandatory
OPR-26	EHDSR-16	Risk mitigation	**SPE Operator SHOULD define SLAs that include: uptime guarantees, response/resolution times for incidents include security SLAs, such as data encryption guarantees, incident response times, and audit logging.**	FitSM Service Level Management recommendation – a service catalogue must be maintained, and SLAs should clearly define service targets such as uptime, response times, and resolution times.	Recommended
OPR-27	EHDSR-16	Risk mitigation	**The SPE Operator SHOULD regularly restore backups to the platform as part of standard feature development and operational activities, ensuring preparedness for incident response.**	NIS2 directive - organisations should have a plan for managing business operations during and after a security incident. This means that backups must be up to date. There must also be a plan for ensuring access to IT systems and their operating functions during and after a security incident.	Recommended

4.5.6 Maintenance and support

SPE providers must ensure that their systems remain up to date, secure, and operational through ongoing maintenance and timely support. This includes having formal processes in place for patch management, system updates, and technical support. Regular maintenance helps prevent vulnerabilities, while reliable support ensures that any technical issues or incidents are handled promptly to maintain data security and availability. The EHDS Regulation highlights that SPEs must apply state-of-the-art technical and organisational measures to maintain security over time. In particular, Article 73(1)(e) specifies that these environments must be regularly updated to prevent unauthorised access or incidents.

Table 4.17. lists these requirements and their derivation that is shown in Figure 4.17.

Figure 4.17. Maintenance and support. Requirement identifiers follow the convention: 'EHDSR–n' for EHDS requirements ('EHDS' namespace + 'R' for requirement) and 'OPR–n' for operational requirements ('OP' namespace + 'R' for requirement).

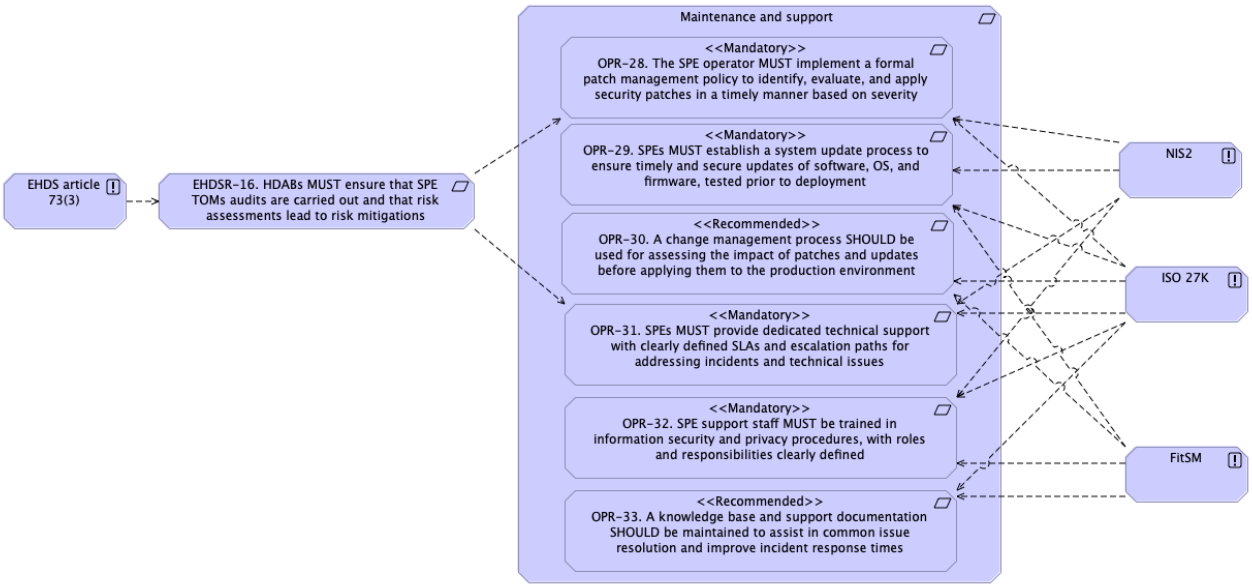


Table 4.17: Maintenance and support

#	Link	Area	Requirement	Rationale	Importance
OPR-28	EHDSR-16	Maintenance	**The SPE Operator MUST implement a formal patch management policy to identify, evaluate, and apply security patches in a timely manner based on severity.**	EHDS Regulation (Article 73(3) requires corrective actions for any shortcomings, risks or vulnerabilities identified un the process of SPE audits. NIS2 Directive mandates organisations to implement risk management measures that include updating systems to protect against known vulnerabilities. ISO/IEC 27001 and 27002 explicitly require a patch management process.	Mandatory
OPR-29	EHDSR-16	Maintenance	**SPE Operator MUST establish a system update process to ensure timely and secure updates of software, OS, and firmware, tested prior to deployment.**	NIS2 Directive stipulates regular system maintenance and software updates to reduce risk exposure. ISO/IEC 27001 and 27002 require that information systems are regularly updated with new versions and patches.	Mandatory

#	Link	Area	Requirement	Rationale	Importance
				FitSM requirement - structured and traceable system update processes must be in place	
OPR-30	EHDSR-16	Maintenance	**A change management process SHOULD be used for assessing the impact of patches and updates before applying them to the production environment.**	ISO/IEC 27001 and 27002 requirements recommend that changes to information systems are controlled and authorised to avoid disruptions or vulnerabilities. FitSM requirement - all changes to services or infrastructure must be reviewed, approved, and documented, ensuring continuity and stability during patching and updates. This reduces the risk of unplanned outages or system compromise.	Recommended
OPR-31	EHDSR-16	Support	**SPE Operator MUST provide dedicated technical support with clearly defined SLAs and escalation paths for addressing incidents and technical issues.**	EHDS Regulation (Article 73(3)) requires SPEs to be resilient and ensure service continuity, implying that support services must be available and responsive. NIS2 Directive: Need for incident handling and response capabilities. ISO/IEC 27001/27002 requires that technical support be available to ensure effective incident detection, response, and recovery.	Mandatory
OPR-32	EHDSR-16	Support	**SPE Operator support staff SHOULD be trained in information security and privacy procedures, with roles and responsibilities clearly defined.**	ISO/IEC 27002 highlights the importance of assigning roles and responsibilities related to security and ensuring adequate training. NIS2 Directive requires that personnel involved in system operation receive regular cybersecurity awareness training. FitSM requirement **-** ongoing education and role-based training should be provided to ensure staff can respond effectively to incidents and manage secure systems.	Recommended
OPR-33	EHDSR-16	Support	**A knowledge base and support documentation SHOULD be maintained to assist in common issue resolution and improve incident response times.**	ISO/IEC 27002 - Documentation of known issues and responses shall be kept to improve incident handling and system stability**.** FitSM requirement - Building and maintaining a knowledge base as part of support operations to enable faster resolution and avoid repetition of known problems. This contributes to efficiency and organisational learning in support environments.	Recommended

5 SPE federation

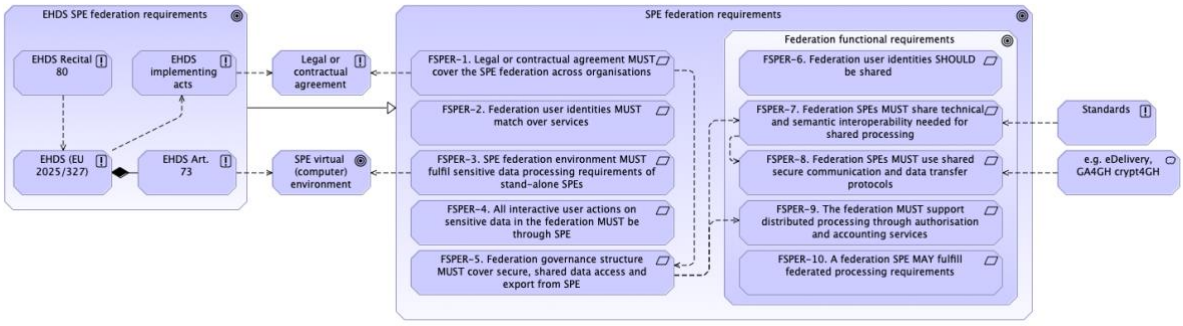
Secondary use in EHDS is a federation of interconnected services sharing information about available health data, as detailed in the EHDS regulation. It provides legal requirements for member states to produce metadata of their health datasets and requires those to be findable centrally though the HealthData@EU infrastructure.

EHDS has a legal obligation for moving sensitive data only within jurisdiction, demanding at least one SPE service for processing the permitted data in each member country and one for the European Commission. The latter will primarily serve European Union's institutions, bodies, offices and agencies (Article 75(2)). Additionally, the member state EHDS official, HDAB, may allow their data to be processed alone or in together with other permitted datasets in any of the approved European SPEs.

Moreover, EHDS calls other European infrastructures like ERICs and EDICs Authorised Participant (AP) (Article 75(4)) that may participate in the HealthData@EU. EHDS data permits will then allow users to access the data based on the legal framework of that AP and get access to the data in the EHDS 73 compliant SPE provided by that AP.

The initial design of EHDS was based on the assumption of a stand-alone SPE maintained by HDAB. Data transfer would have been solely within each jurisdiction and would not need to be included into EHDS. However, the need to combine datasets over national borders and especially the rise of federated computing (see Implementing federated computing) makes it necessary to bridge the gap from stand-alone SPE that works under one organisation by defining general requirements of an SPE federation.

Figure 5.1. SPE federation requirements. These requirements are already embedded in the EHDS structure. Requirement identifiers follow the convention 'FSPER–n', where 'FSPE' indicates the SPE federation requirements namespace and 'R' denotes requirement.



EHDS is already an SPE federation in almost all but in name and technical support. The following general SPE federation requirements are regardless of federation type (Figure 5.1). Their main limitations arise from sensitivity of data. They enable sensitive data processing beyond a single service to cover processing over multiple separate but linked services within one organisation (like database or HPC), as well as enabling processing over organisational borders. These requirements should work for *ad hoc* proprietary data as well as EHDS-based processing.

Federation rules will need to direct the collaboration and establish federation services for management, mutual trust, findability and information sharing. An SPE federation will need findability services to locate datasets and SPEs. EHDS already includes detailed instructions for data findability services, but a similar approach will be needed to SPEs, too, depending on the national setting. Heath data users will need to search and compare a catalogue to find EHDS compliant SPEs nationally across EU.

The requirements of general purpose SPE federation have the namespace acronym 'FSPE' for Federated SPE followed by the letter 'R'. All requirements in this report are collected to a list in Annex B: SPE and related requirements.

FSPER-1. Legal or contractual agreement MUST cover the SPE federation across organisations

Rules of federation must be based on writing for transparency and enable shared sensitive data processing across organisations.

FSPER-2. Federation use identities MUST match over services

An SPE federation needs to identify and authorise its users in all federation services. This can be done by matching local identities over federated identities.

FSPER-3. An SPE federation environment MUST fulfil sensitive data processing requirements defined for stand-alone SPEs

An SPE federation is based on SPEs providing the main sensitive data protection. None of the stand-alone SPE rules should be violated. For EHDS this means that the legal requirements from Article 73 apply equally to stand-alone SPEs and those functioning in the federated context.

The real challenge here is to ensure full technical compatibility between EHDS SPEs in different jurisdictions under EHDS. Without tighter coordination from EHDS implementation acts, there is a danger that establishing cross-border data exchange will be challenging.

FSPER-4. All interactive user actions on sensitive data in the federation MUST be through SPE

While the aim of the SPE federation is to widen the scope and quantity of sensitive data via automated, remote processing, the users will have to log in to an SPE to initiate the processing and export the results.

FSPER-5. Federation governance structure MUST cover secure, shared data access and export from SPE

The federation will provide the governance structure that defines operational procedures for secure, shared processes for data transfer.

FSPER-6. Federation user identities SHOULD be shared

While the federation can function with matching user identities, its authorisation services are significantly strengthened by a federation-wide identity system. This is a nod of approval

towards the shared European eiDAS system. A shared user identity system may enhance service authorisation and auditability to better comply with GDPR principles.

FSPER-7. Federation of SPEs MUST have technical and semantic interoperability needed for shared processing.

Shared processing needs agreements on technical and semantic standards that will enable it. In a European context, the most common health data model for remote data queries is OMOP CDM.

FSPER-8. Federation of SPEs MUST use shared secure communication and data transfer protocols

Security of communication between federation members, especially when transferring sensitive data, is the fundamental requirement. EU eDelivery and GA4GH crypt4GH are examples of these (see chapter Data access management and SPE interoperability).

FSPER-9. The federation MUST support distributed processing through authorisation and accounting services

A federation needs usually centrally managed registries that secure distributed processing. Federation registries need to have infrastructure metadata about its authorised services and service providers, content metadata for dataset discoverability, and governance metadata about data permits that link project and users to datasets and resource usage[3].

For EHDS to fully take up the role of a federation, these services should be provided centrally.

FSPER-10. A federation SPE MAY fulfil federated processing requirements

Federated processing is its own branch of distributed processing that some members of the SPE federation may enable (See chapter Implementing federated computing).

5.1 Operational requirements for SPE federation

The paragraphs of Article 73 of the EHDS regulation contain very specific legal demands for SPEs that function primarily in isolation. We have interpreted these to form 18 practical requirements out of them (See chapter EHDS SPE requirements and Annex G: EHDS article 73 analysis to deduce SPE requirements). Many of them are purely operational, meaning their successful and secure application depend on the interplay of the personnel maintaining the services and the processes they implement, rather than being purely technical. Extending these requirements to fully support federated processing within EHDS, these requirements need to be interpreted and deployed in a uniform manner throughout the federation to maintain the functioning and security of services.

In addition to EHDS, the catch-all security standards applied to high-end security services and organisations are the ISO/IEC 27000 -family of standards (ISO 27K) where ISO/IEC 27002 deals with information security controls. These are not very helpful for determining

[3] DARE UK Federated Architecture Blueprint https://doi.org/10.5281/zenodo.14192786

specific technical requirements for SPEs that go beyond it. The bulk of ISO27K requirements are operational, affecting the organisation maintaining the service.

The NIS2 Directive builds on ISO 27K security demands and requires reporting within member countries and collection of all serious incidents to an EU-wide registry.

We have organised and tabulated various operational demands to the SPE from EHDS, ISO 27K and NIS2 (see Operational Requirements), but have excluded operational requirements that direct HDAB and its reporting requirements to the EU central platform.

Our evaluation also includes the FitSM standard, which we suggest as a lightweight framework to support operation and reporting because we feel it offers significant advantages to the EHDS over ISO 27K that is closed, heavyweight and expensive to certify against.

FitSM is a lightweight, open-source standard for IT Service Management, developed through an EC funded project (FedSM, 2012-2015) and still actively maintained today. It supports any service delivery scenario but is unusual in supporting federated service provision and being much easier to implement in academic and public sector than other heavier traditional frameworks. FitSM has a wider scope than ISO 27K but covers information security management and can be easily integrated with an ISO27K management system. It offers strong advantages in that it can then connect security requirements more thoroughly to service design and delivery, service level management, capacity and availability management, customer relationship management and other IT Service Management processes (e.g. Service Portfolio Management, Incident & Problem Management, Change Management, Service Level Management).

The core FitSM documentation is freely available under a Creative Commons licence, so it can be easily embedded and referenced in organisational and national documentation. This lets it be used as a free and lightweight reference model for managing services and for connecting security and operational requirements across EHDS.

FitSM is proposed here as a practical framework to support coordination and documentation but does not replace the legal obligations established under the EHDS Regulation or related EU legislation.

To ensure mutual trust across the EHDS SPEs, the documentation produced by each MS should demonstrate that auditing processes are implemented consistently and meet the common safeguards required under Article 73. This documentation should enable Member States to recognise each other's SPEs as compliant with the Regulation. Without such transparency and alignment, the SPE federation cannot guarantee a uniform level of protection for sensitive data across borders, thereby undermining its core objective

FitSM should offer a lightweight and cost-effective way for EHDS SPE federation to coordinate its operational requirements to fulfil that goal.

We recommend that EHDS should build its guidance for EHDS SPE federation compliance reporting based on FitSM.

5.2 Cybersecurity of SPE infrastructure

This chapter of the SPE security considerations deal with challenges arising from virtualisation that are wider than SPE.

SPE implementation is typically run as part of a virtualisation service where the aim is to isolate the user's project environment from other user environments and fully control the access to additional services. This isolation is built using nested security structures, where user applications run within virtualised environments that simulate their own operating systems, hosted on a local network. Eventually, the virtualisation layers interface with the physical hardware, where all processes are ultimately executed.

The main aim of cybersecurity guidelines is to determine what is the adequate protection for a given setup balancing both enabling and limiting aspects. Given the aim to tap into creativity of researchers, we have already highlighted the need for openness and information sharing among the project members that requires them to understand their responsibilities and implications of their actions because too onerous technical actions are counterproductive.

In our minimum requirements for SPEs these are expressed by:

> **SPER-4. SPE MUST provide adequate protection against exposing sensitive data to unauthorised users**
>
> **SPER-6. Project-based user environments of SPE MUST be isolated from each other and open Internet**

We may rephrase these sentiments to:

> **SPE users and their programmes SHOULD NOT be able to harm anything else but their own immediate environment**

Unfortunate users of an SPE should be allowed to find themselves in a situation where they exercise their freedom of actions to the extent that they lose the most recent results and their environment must be recreated to re-allow them access to the permitted sensitive data.

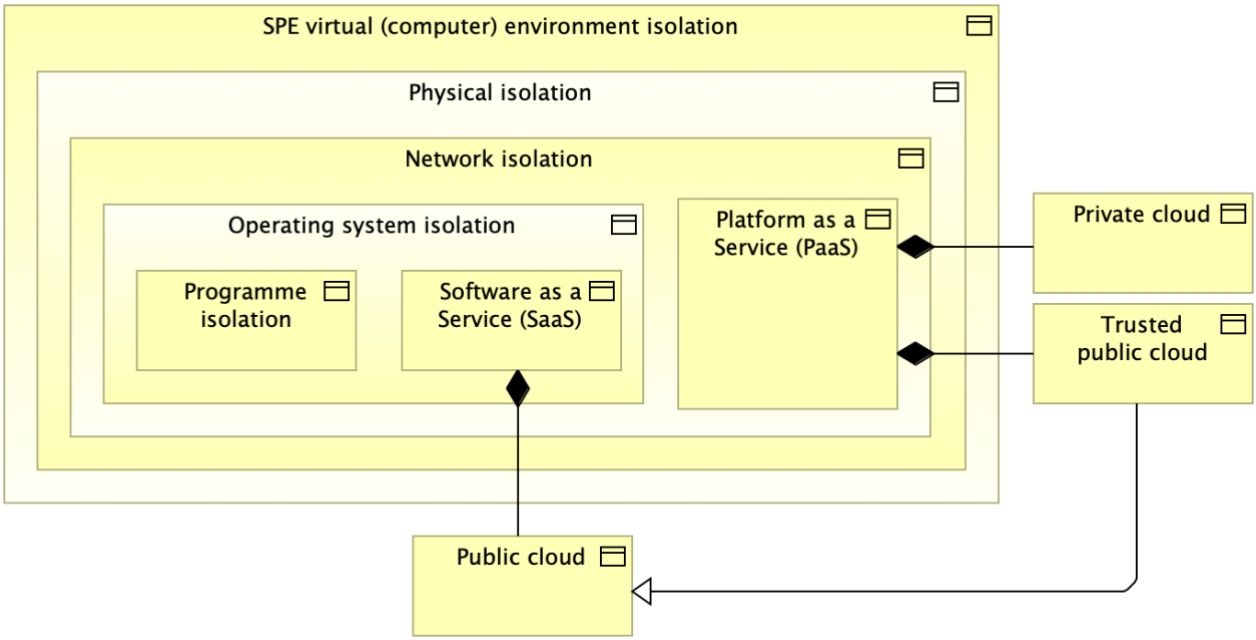
Following the same logic, it should not matter what applications users have access to or what they have installed themselves. The operating system setup should prevent the harm from spreading outside the immediate user environment. If the users or their applications gain elevated access privileges, the network isolation is there for the next layer of protection. They should never reach administrative control of the system.

However, it is not elephants all the way down. Networks and processes run on physical hardware. Secure processing running on the hardware maintained by the SPE service provider is seen as private cloud. In contrast, services running on public cloud necessarily involve trust to the cloud provider.

The figure 5.2 illustrates the nested structure of computer cybersecurity layers and the full chain of trust inside a private cloud should enable the provider to control the operating system layer and offer its users SPE as a Platform as a Service (PaaS) where users have wide freedoms to import and install their own software. In public clouds, the lesser amount of SPE service provider control, especially regarding their network, might force them to approach SPE as Software as a Service (SaaS) and demand them limit their user rights to install

software without supervision to reach the same level of network security as private cloud provides.

Figure 5.2. Nested isolation of SPE processing



A key outcome of the UK community-driven SATRE project is a standard architecture specification for TREs (See Annex H: SATRE). Closely aligned with TEHDAS2, it defines core TRE capabilities and criteria for assessing compliance. With regard to network security and nested isolation, the SATRE specifications include the following mandatory requirement:

SATRE 2.2.9. Your TRE must control and manage all of its network infrastructure

The recent proliferation of public clouds is based on reliance on laws and regulations providing foundations for usual business subcontracts that have been seen enough to elevate public clouds to the same level as private clouds.

Public cloud providers may be subject to extraterritorial legislation (e.g. US CLOUD Act[4]). The suitability of such providers for SPE hosting must therefore be reassessed under EU data protection, cybersecurity, and sovereignty requirements. The recent American governmental policy shift to emphasise US presidential prerogative power has already led to calls for purely European cyber infrastructure to safeguarding sensitive European data[5].

The current political climate has made it clear that cybersecurity requirements for SPEs and their federation will need to be reconsidered to redefine the adequate level of trust for sensitive data computer infrastructure is run on.

[4] CLOUD Act https://en.wikipedia.org/wiki/CLOUD_Act

[5] e.g. EuroStack https://euro-stack.eu/

5.3 Data access management and SPE interoperability

This section outlines the key requirements for interoperable and secure data access in an SPE federation, with emphasis on automation, standardisation, and support for scalable cross-border secondary use of health data under the EHDS Regulation.

The interoperability and scalability of an SPE federation largely depends on the efficiency of data access management. The security of large-scale dataset management needs automated procedures that manual data management cannot match. It has been estimated that to reach similar throughput of data permits in the future European EHDS framework serving the whole EU as the Findata authority did manually in 2024 for Finland, would need European member states altogether hire 10000 persons to handle all the applications (Jaakko Leinonen, CSC, FI, personal communication).

The lack of scalability in manual management of sensitive data life cycles (see Annex D: Sensitive data lifecycles) has been the driver for designing the automated "state-of-art" approaches (see Annex E: Scenarios). It is also the reason why this report defining the requirements and capabilities of SPE has been unable to do it without considering all the sensitive data lifecycle components. They are totally interdependent on each other. Enabling better functionality in one will not be possible without the other components supporting that.

The core functionality of the SPE is to process sensitive data. We already have a good grasp on how it can be done securely and ample approaches to make it happen more efficiently (see Annex E: Scenarios), but they all depend on the capabilities of other components of the federation.

Cornerstones of this scalable infrastructure are:

- Shared identity and authorisation
- Machine-actionable access permits
- Data streaming
- Structured data warehouses

All security and accountability of sensitive data processing is based on identifying users with certainty (see Annex E: Identity and authorisation). The data applicants and health data users should be identified in the beginning of the application process and that information needs to be maintained unbroken throughout. This means the official, human-readable document that is the data permit needs to be converted to a machine-readable access permit. This access permit combines the identifiers of the data permit to users' identities and to the newly created dataset that users will have access to.

An access permit cannot exist before the dataset promised in the data permit has been created and stored with a unique identifier. This new dataset is usually a combination of several source datasets with unnecessary items removed (according to GDPR minimisation principle), and pseudonymised.

Manual data transfer increases operational complexity and potential risk of human error. Automated streaming based on machine-readable access permits strengthens traceability and compliance with Article 73. The data manager will continue to have full control of the dataset and access to it. They can also make user-requested amendments to data if needed.

The primary technical requirement to start an SPE federation is to agree on at least one data transfer protocol (**FSPER-8**). While EHDS has already committed itself to eDelivery as the secure communication protocol for sending metadata records, data access application forms and messages between national contact point and the EU central platform, the data transfer details have not been agreed on. In practice, there is only one set of implemented community standards that cover the needs of sensitive data transfer and data access management for research: the Global Alliance for Genomics and Heaths (GA4GH) standards[6].

GA4GH crypt4GH[7] is an open source, secure, streaming capable encryption algorithm. It is created for huge genomic sequences but applicable to any file type. The encrypted file is separated to header and payload. The payload is encrypted with a key that is stored in the header and never exposed. New user keys can be added and at will. Header can be stored and updated independently of the payload that cannot be decrypted with brute force methods. Sending header and payload separately creates a quantum safe data transfer. Storing crypt4gh headers in a secure database enables automatic user key management for secure data projects. The streaming capability used by GA4GH htsget API[8] can be linked to token-based data access control implementing the GA4GH Passport standard[9].

These standards are not mandated under the EHDS Regulation but could be considered as candidates for future implementing acts or common specifications under Article 73(5).

The use of these standards already forms the basis of plans for the framework for the EU One Million Genomes Initiative[10]. Its framework[11] gives detailed information about the initiative that aims to establish itself as a genomics Authorised Participant for EHDS.

Finally, one of the challenges in manual data management is that many European health data is either not stored in an electronic format at all or not in a standardised format. This makes determining the data applicants' needs and the creation of the needed new dataset difficult. The secondary use of health data should be one more driver towards rapid digitalisation and standardisation of health data at source.

[6] GA4GH implementations https://www.ga4gh.org/our-products/implementations/
[7] Crypt4GH https://www.ga4gh.org/product/genetic-data-encryption-crypt4gh/
[8] GA4GH htsget https://www.ga4gh.org/product/htsget/)
[9] GA4GH passport https://www.ga4gh.org/product/ga4gh-passports/
[10] 1+MG https://digital-strategy.ec.europa.eu/en/policies/1-million-genomes
[11] 1+MG framework https://framework.onemilliongenomes.eu/

6 Implementing federated computing

To ensure clarity and consistency throughout this chapter, some key terms are defined below. These terms are not explicitly defined in the EHDS regulation but serve to frame the technical requirements proposed in this chapter and provide the foundational understanding necessary for interpreting the concepts and discussions that follow.

1. **Federated computing**[12]**.** A decentralised data processing approach where computations occur locally on distributed SPEs rather than being centralised into a single SPE.

   Such approach is encouraged by EHDS regulation (Recital 80) which promotes the principle of "bringing algorithms to the data" to enhance privacy-preserving computation. Federated computing methods enable data to remain closer to their original location while only aggregated results or model updates are shared, enhancing privacy and security.

2. **Federated analysis**. A specific approach for federated computing where statistical results are computed locally on several distributed SPEs.

   This method enables benchmarking, multi-site research, and collaborative analytics while preserving data privacy and security. Only aggregated insights or summary statistics are exported from each SPE ensuring that personal data is not transferred out from the secure environments.

3. **Federated learning.** A specific approach for federated computing where models are trained and validated on distributed of SPEs.

   Raw data resources are not exchanged between SPEs. Instead, only model updates are communicated thereby enhancing data privacy and security. Due to the difficulty of assessing the anonymity of intermediate outputs, federated learning it is essential that computation and information exchange takes place within a network of trusted SPEs.

6.1 General

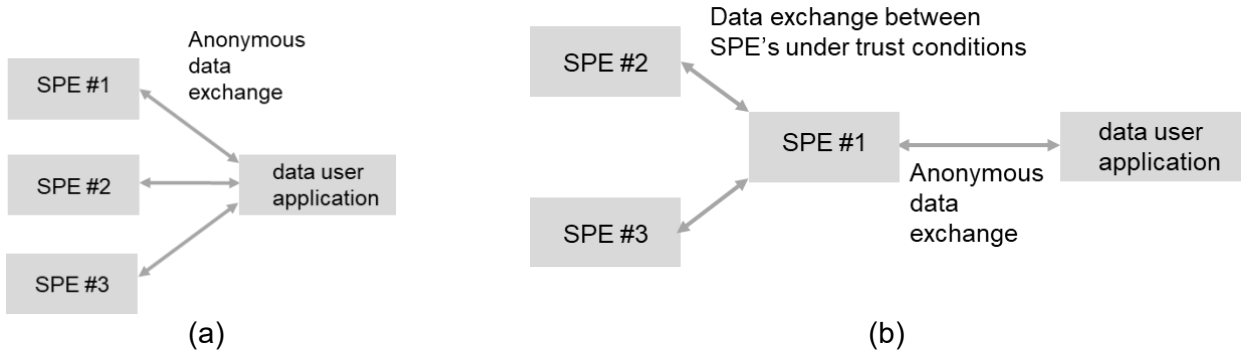
The federated computing approach aligns with Recital 80 of the EHDS Regulation, which encourages bringing questions to the data instead of moving the data. However, the EHDS Regulation does not provide further direction on how to support the implementation of federated computing. Further guidance is expected in the context of implementing acts concerning secure processing environments.

Two settings can be identified. In the first (Figure 6.1.a), the health data user interacts independently with multiple SPEs, retrieving anonymised outputs from each and combining them to compute statistical results. In this case, no communication occurs between SPEs. Figure 6.1.b illustrates a second setting where a master SPE (SPE#1) communicates with

[12] Federated computing is not mandated under the EHDS Regulation, but may be supported under future implementing acts (Article 73(5))

other SPEs to perform a federated computing task. The master SPE is responsible for orchestrating the federated learning process, collecting computation results from the other SPEs to iteratively train or validate a machine learning model. The health data user retrieves the final output (e.g., the trained model) from the master SPE. An SPE federation (see SPE federation) is a prerequisite for the setting of Figure 6.1.b to establish the framework for SPE-SPE communication. Such federation is not required in the setting of Figure 6.1.a, which does not involve communication between SPEs.

Figure 6.1. (a) Data user application collects and further processes anonymous computation results from multiple SPEs. (b) Data user application is connected with one master SPE, which in turn communicates with other SPEs under trust conditions to accomplish a federated computing task.



The scenario illustrated in Figure 6.1.a is relevant for federated analysis use cases, where only anonymous aggregated results need to be transferred to the health data user by each SPE. Figure 6.1.b is relevant for federated learning use cases, where it is difficult to ensure the anonymity of the interim outputs of the SPEs and which thus need to be kept within the network of federated SPEs.

6.2 Scope

These proposed functional requirements focus on the required functionality of SPEs to support federated computing.

The following are **out of scope** for this deliverable:

- The impact of federated computing on the data permit application phase, including how the involved SPEs are selected by the health data user and approved by the HDAB.
- Implications for EHDS governance, such as the authorisation of data transfers between SPEs and the deployment or execution of required software components (see also SPE federation)

- Manual implementation of federated analysis, which does not rely on specific SPE functionalities—for example, cases where the user manually combines results from separate analyses, as permitted by their approved data permit(s).

The technical requirements that support these functional requirements are listed in the Interoperability requirements chapter.

Some requirements for federated computing (e.g., FCR-1) are also useful in a single-SPE setting.

6.3 Assumptions

Proposed functional requirements for supporting federated computing are based on the following assumptions:

- Federated computing is applied in the framework of the EHDS regulation complemented by more detailed specifications (e.g. implementing acts)
- All computations on sensitive personal data are carried out in an SPE environment aligned with EHDS requirements (e.g. Article 73 of the EHDS regulation)
- All SPE services involved in federated computing are EHDS compliant, trusted, audited and under control of HDABs or trusted data holders (directly or through data processing agreements with SPE operators).
- SPEs involved in federated computing, i.e. **federation SPE,** may be located in one or more countries. Federated computing support is a recommended but optional feature of an SPE ("MUST" in a requirement means that the requirement is applicable if federated computing is supported by the SPE).
- An SPE may also support only a subset of requirements, e.g. limiting to federated analysis.

6.4 Overall functional requirements

The requirements of computing performed by a federation of SPEs have the namespace acronym 'FC' followed by the letter 'R'. All requirements in this report are collected to a list in Annex B: SPE and related requirements.

FCR-1. SPE MUST support common data models, such as OMOP CDM and applicable GA4GH standards, to enable technical and semantic interoperability (see also **FSPER-5**).

FCR-2. The HDAB MUST have in place (internally or in collaboration with SPE operators) required processes to deploy additional data models required by health data users.

FCR-3. SPE MUST support the deployment and execution of software components needed to carry out federated computing as described and authorised in the data permit, subject to applicable security controls.

FCR-4. The HDAB MUST have in place required processes to authorise the use of software components needed in federated computing

6.4.1 Functional requirements for federated analysis

FCR-5. SPE MUST enable a data user application to retrieve anonymous results from the SPE to be merged with results retrieved from other SPEs

FCR-6. SPE MUST include functionality to support HDAB in their operations for ensuring the anonymity of results, with applicable methods such as: (1) pre-assessment of software components producing the results, (2) automated anonymity assessment, (3) additional privacy protection mechanisms (such as differential privacy) and (4) manual inspections

FCR-7. The HDAB MUST have established processes for approving computerised access permissions for data user applications in the context of data permit authorisation. (see also **FSPER-5)**.

6.4.2 Functional requirements for federated learning

FCR-8. SPE MUST provide an open and standardised interface needed to exchange information with other trusted SPEs as needed to accomplish federated learning computations (see also **FSPER-5** and technical requirements in the Interoperability requirements chapter)

FCR-9. SPE SHOULD support the use of privacy protection methods (such as differential privacy) for federated learning.

6.5 Interoperability requirements

6.5.1 General

Existing EHDS technical requirements (Specific Contract no.22 documents: D02.03 Requirements Catalogue[13], D03.03 Architecture Artefacts[14], D03.03 System Specifications[15]) are not directly addressing technical interoperability, but we have aligned with these documents where applicable, adopting consistent terminology and using the architecture described in them as a reference for interoperability requirements.

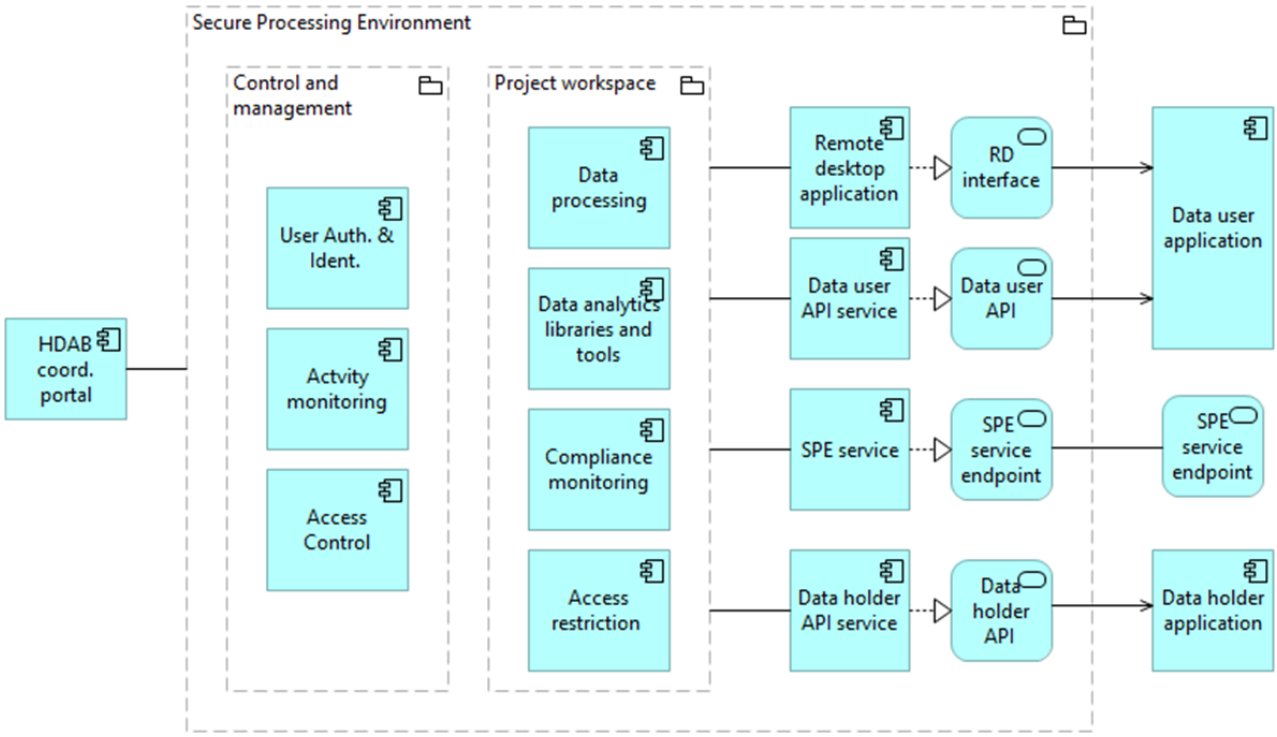
[13] D02.03 Requirements Catalogue for Scale-Up version. Specific Contract no. 22 under Framework Contract SLG.AVT.DI07926 - BEACON – Lot2, Date: 17.12.2024, Version 3.1.
[14] D03.03 Architecture Artefacts – Scale-Up version. Specific Contract no. 22 under Framework Contract DI7925-DI7932 – BEACON – Lot2: Analysis and design of the European Health Data Space infrastructure for secondary use of health data (HealthData@EU), Date: 4.12.2024, Version 1.0.
[15] D03.03 System specifications – Scale-Up version. Specific Contract no. 22 under Framework Contract DI7925-DI7932 – BEACON – Lot2: Analysis and design of the European Health Data Space infrastructure for secondary use of health data (HealthData@EU), Date: 10.12.2024, Version 1.0.

The high-level technical architecture of the SPE is depicted in Figure 6.2. Interfaces towards the following entities are identified:

- Data user applications
- Data holder applications
- External SPEs
- HDAB coordinator portal

This section provides high-level technical specifications for the data user and data holder application interfaces as well as the interface for connecting external SPEs. The specifications are applicable both for federated computing and traditional data processing. The governance and safeguarding processes needed for setting up and managing the interfaces are beyond the scope of this deliverable.

Figure 6.2. High-level technical architecture of SPE (modified from D03.03 Architecture Artefacts).



Data user application is a user-facing interface supported by the SPE infrastructure. It allows users to securely connect to their designated project workspace, which is isolated from the other workspaces. Two technical approaches should be supported by the SPE: (1) remote desktop connection and 2) API based connection.

The remote desktop (RD) interface allows the data user to interactively process data in the project workspace. Remote desktop is the legacy approach widely used in the currently existing SPE installations. The data user application (remote desktop client) allows the user to access the project workspace as if working directly on its host machine, supporting interactive tasks such as file management, software execution, and visualisation.

The API interface enables more versatile user applications, such as web portals and standalone applications provided or authorised by the HDAB. The API based data user application may enable to trigger execution of functions (e.g. data processing scripts), monitor progress of function execution, retrieve anonymous data analysis results, download anonymised data sets, communicate with authorised organisation (HDAB, UHDAS, authorised participant) and upload custom tools, supporting contents (e.g. vocabularies) and additional data sets.

The HDAB is responsible for setting up the required processes and controls to ensure that the exposed API services ensure privacy protection and compliance with EHDS regulations. For instance, it should be possible for the HDAB to manually or automatically verify that the downloaded contents meet anonymity requirements.

Data holder application is the data holder's interface supported by the SPE infrastructure. The application enables the data holder to upload data to the SPE based on an accepted data permit. The data holder application uses an API based connection mechanism.

SPE service endpoint enables connections to external SPEs. Such connections enable federated networks of trusted SPEs to be formed. Options for setting up such networks are discussed in the SPE federation chapter.

HDAB coordinator portal connection enables the relevant organisation (HDAB, UHDAS, authorised participant or SPE operator) user to carry out overall management of the SPE and to execute specific tasks needed to support data usage and ensure regulatory compliance. Technical specifications for connecting the HDAB coordinator portal with SPEs may vary between implementations and are not in the scope of this deliverable.

To fulfil the regulatory requirements (EHDS, Article 73) the HDAB shall implement the RD interface and the data holder API shown in Figure 6.2. Implementation of the data user API and SPE service endpoint is recommended, but not obligatory.

Technical requirements for the interfaces are listed in the following subsections.

6.5.2 Data user and data holder API requirements

Technical interoperability requirements of federated SPE under EHDS regulation have the namespace acronym 'TI' for **Technical Interoperability** followed by the letter 'R'. All requirements in this report are collected to a list in Annex B: SPE and related requirements.

General

TIR-1. Client application. The API services MUST be accessed via data user and data holder applications, which can be dedicated client applications or browser-based applications.

TIR-2. Architectural style. The API MUST follow a web services architecture (e.g., RESTful, GraphQL), enabling stateless request/response interactions and supporting secure file transfer protocols. The implementation MUST adhere to common architectural and security practices, including resource grouping and, where appropriate, the use of microservices to isolate sensitive services.

TIR-3. Communication protocol. The API MUST operate over HTTPS, ensuring compatibility with standard web clients and server implementations. Additionally, it MUST support secure file transfer protocols (e.g. HTTPS, SFTP or FTPS for file operations) where applicable.

TIR-4. Data exchange format. The API MUST use JSON as the primary data format for requests and responses with support for other formats where needed, ensuring secure data transfer via HTTPS or other secure protocols.

TIR-5. File transfers. The API MUST be capable of uploading and downloading files of all standard file types. It MUST be possible to set restrictions on allowed file types and transfer direction (upload/download) through configuration settings separately for each API and workspace. The API must support efficient handling of large files, including resumable transfers, asynchronous processing where necessary, and configurable timeout settings to accommodate long-lasting operations.

TIR-6. Comprehensive API documentation. API SHOULD have clear, machine-readable documentation (e.g., OpenAPI/Swagger).

API Security and Access control

TIR-7. Authentication and authorisation. All requests to data user and data holder API functions MUST be authenticated and authorised using standard methods (e.g., OAuth 2.0 with JWT, API keys). The system MUST enforce role-based access control, limiting access based on user roles and privileges.

TIR-8. Data encryption. API communication MUST use encryption (e.g., TLS) to protect data in transit. Additionally, the API MUST support application-level encryption (e.g., AES-256) before transmission, with enforcement by the HDAB, to provide an extra layer of security and ensure data remains protected even if stored by the API. If content encryption is used, keys MUST be securely managed and stored according to industry best practices (e.g., using a dedicated key management system).

TIR-9. Input validation and sanitisation. The API MUST validate all incoming data to prevent injection attacks (e.g., SQLi, XSS) and malformed requests, responding only to predefined and approved requests with valid parameters.

TIR-10. Error handling and security. Error responses MUST avoid exposing sensitive details (e.g., stack traces, internal error codes) while providing meaningful messages for debugging.

TIR-11. Network access control. API access MAY be restricted based on network-level controls, including firewall rules, IP whitelisting, or Virtual Private Network (VPN) restrictions.

API service management and monitoring

TIR-12. Logging and auditing. Requests, responses, and errors MUST be logged for monitoring and compliance.

TIR-13. Monitoring and alerts. API usage, failures, performance metrics and service availability MUST be monitored, with alerts for anomalies.

TIR-14. API version management. Mechanisms MUST be in place for approving and deploying new API versions and phasing out old API versions.

API performance and Scalability

TIR-15. Response time. The API SHOULD meet defined latency and response time SLAs.

TIR-16. Load handling. The system SHOULD handle expected and peak loads efficiently, with rate limiting and throttling mechanisms in place if necessary.

TIR-17. Data Validity. The API SHOULD ensure the accuracy and consistency of requests and responses by applying relevant format validators and other appropriate validation mechanisms, thereby preventing the delivery of erroneous, corrupted or inconsistent data.

6.5.3 Data user remote desktop interface

General

TIR-18. Functionality. The remote desktop interface MUST enable the data user to interactively access and process data with a remote desktop application (“SPE UI” in D03.03 Architecture Artefacts).

TIR-19. Data user application. The data user MAY use a standard browser or a dedicated client to access the remote desktop environment.

TIR-20. Platform compatibility. The remote desktop solution MUST support remote desktop clients running on standard operating systems including (e.g., Windows, macOS, Linux).

TIR-21. Protocol support. The remote desktop solution MUST use a secure remote access protocol (e.g., RDP, VNC, or SSH with X11 forwarding) to establish a connection between the client and server. The communication protocol MUST support encryption to protect data in transit.

Security and Access control

TIR-22 Authentication and authorisation. Access MUST require multifactor authentication (MFA). The system MUST enforce role-based access control, limiting remote access based on user roles and privileges.

TIR-23. Network Access Control. Access MAY be restricted based on network-level controls, including firewall rules, IP whitelisting, or VPN restrictions.

TIR-24. Session management. Idle sessions MUST be automatically terminated after a predefined time to prevent unauthorised access. Users MUST be logged out or locked after a period of inactivity.

TIR-25. Data transfer restrictions. Clipboard sharing and file transfers MUST be restricted by default to prevent unauthorised export of data from the SPE. Changes to or removal of these restrictions MUST be configurable at the discretion of the HDAB.

Service management and monitoring

TIR-26. Logging and auditing. All remote sessions MUST be logged, including authentication attempts, connection times, and actions performed during the session.

TIR-27. Monitoring and alerts. Service usage, failures, performance metrics and service availability MUST be monitored, with alerts for anomalies.

6.5.4 SPE service endpoint

General

TIR-28. Functionality. The service endpoint MUST enable communication between trusted SPEs to support federated learning.

TIR-29. Communication Protocol. The interface MUST support protocols needed to support client-server and bidirectional streaming communication (gRPC with HTTP/2 or an equivalent). Connections MUST be secured using TLS.

Security and access control

TIR-30. Authentication and authorisation. All connections MUST be authenticated using appropriate methods such as mutual TLS (mTLS) or token-based authentication while ensuring all data is transmitted over an encrypted channel.

TIR-31. Client whitelisting. Connecting clients MUST be preapproved and have valid certificates.

TIR-32. Network access control. The interface MAY be configured to restrict access using firewall rules, IP whitelisting, VPN access, other network-level policies and holistic approaches such as virtual closed networks.

Service management and monitoring

TIR-33. Logging and auditing. All requests, responses, and streaming events MUST be logged securely.

TIR-34. Monitoring and alerts. Real-time monitoring of streaming sessions, errors, latency, and security threats MUST be implemented. Alerts MUST be generated based on anomalous activity (e.g., failed authentication attempts, unusual data patterns).

Performance and scalability

TIR-35. Response time and streaming performance. The interface MUST meet defined latency and response time SLAs.

TIR-36. Load handling. The system MUST implement flow control to manage varying loads and prevent overloads, using rate limiting, throttling, and adaptive resource allocation as needed.

6.5.5 Relation to existing specifications

Existing specifications (D03.03 Architecture Artefacts) use the term “SPE UI”, which seems to refer to a remote desktop type interface used in legacy SPE implementations. In this deliverable, we have proposed two types of data user applications: (1) the legacy remote desktop approach and (2) a new approach where the data user application interacts with the SPE via a web API interface. Approach (1) provides a familiar, typically browser-based, method for data users. Approach (2) enables new ways to enable more efficient ways to exchange information between the health data user and the SPE environment, including the possibility to execute federated analysis in multiple SPEs.

Similarly, we propose the API approach to be used also for the data holder application, where such approach would enable a more efficient process with less manual interaction needed.

We have included the interface for connecting external SPEs, which is not covered in D03.03 Architecture Artefacts. This interface enables advanced federated learning scenarios requiring direct information exchange between SPEs. These scenarios need careful governance and implementation to comply with EHDS regulation which are not in the scope of this deliverable.

Annex A: Glossary

Table A.1. Key terminology

Term	Description
Access permit	Machine-actionable data structure that contains the information from data permit in a standardised format that can be securely transferred and acted on by computer services
Anonymisation	The process by which personal data is altered in such a way that a data subject can no longer be identified directly or indirectly. (Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information, Recital 52; EHDS Regulation, Recital 92)
Authorised user	An authorised natural person listed in the data permit, giving them the rights to process sensitive data inside an SPE
API	Application Programming Interface
Data permit	An administrative decision issued to a health data user by a Health Data Access Body to process certain electronic health data specified in the data permit for specific secondary use purposes based on conditions laid down in Chapter IV of the EHDS regulation (Article 2(2v)).
Data holder application	A software application that provides the data holder with secure digital access to the Secure Processing Environment (SPE). Its core functions include facilitating the upload and download of data in accordance with the data holder's responsibilities under the EHDS regulation.
Data processor	The data processor should handle data exclusively in the manner prescribed by the controller. A data processor acts under the detailed instructions of the data controller only, by processing personal data on his behalf. (GDPR, Article 4(1)(8))
Data user application	A software application that provides the data user with secure, computerised access to their workspace within the SPE. Its primary functions include facilitating the

Term	Description
	upload and download of data while ensuring robust authentication and authorisation mechanisms to prevent unauthorised access.
EDIC	European Digital Infrastructure Consortium
EHD	Electronic health data means personal or non-personal electronic health data (EHDS Article (2)(c))
EID	European Interoperability Framework
ERIC	European Research Infrastructure Consortium
Federated analysis	A decentralised approach to data analysis where statistical results are computed locally on distributed data resources rather than aggregating raw data centrally. This method enables benchmarking, multi-site research, and collaborative analytics while preserving data privacy and security. Only aggregated insights or summary statistics are shared between nodes, ensuring compliance with data protection regulations.
Federated learning	A decentralised machine learning approach where models are trained and validated on distributed data resources without transferring raw data. Instead, only model updates or gradients are exchanged between nodes, enhancing data privacy and security. This method enables collaborative model development across multiple organisations or devices while maintaining local data sovereignty and regulatory compliance.
Federated processing	A decentralised data processing approach where computations occur locally on distributed nodes rather than being centralised. This method enables data to remain on local devices or servers while only aggregated results or model updates are shared, enhancing privacy and security. It is commonly used in machine

Term	Description
	learning ("federated learning"), analytics ("federated analysis"), and secure data collaborations across multiple organisations.
Federation SPE	Secure processing environment (SPE) that is engaged in federated computing.
FTPS	Secure File Transfer Protocol
gRPC	Google Remote Procedure Call
Health Data Access Body (HDAB)	Member state-designated authority that facilitates the secondary use of electronic health data. HDABs assess the information provided by the health data applicant and decide on health data requests and access applications, authorise and issue data permits, obtain data from data holders and make data available in Secure Processing Environments. HDABs systematically track the data request and data access applications received and the data permits issued. As per Article 58 of the EHDS, HDABs are required to publicly list information on the data permits issued. (EHDS Article 55 and Recital 52)
Health data user	A natural or legal person, including Union institutions, bodies, offices or agencies, which has been granted lawful access to electronic health data for secondary use pursuant to a data permit, a health data request approval or an access approval by an authorised participant in HealthData@EU. (EHDS Article 2(2u))
High Performance Computing (HPC)	HPC is the use of advanced and not commonly available computational infrastructure — such as supercomputers or compute clusters — to solve highly complex and resource intensive computational problems.
HTTP/2	Hypertext Transfer Protocol Version 2
HTTPS	Hypertext Transfer Protocol Secure

Term	Description
JSON	JavaScript Object Notation
JWT	JSON Web Token
OAuth	Open Authorisation
OMOP CDM	A standardised, common data model (CDM) specification originally developed by the Observational Medical Outcomes Partnership (OMOP) and now maintained by the Observational Health Data Sciences and Informatics (OHDSI) community. It defines a consistent structure and a set of standardised vocabularies for observational health data, enabling researchers to perform large-scale, reproducible analyses across diverse databases.
Pseudonymisation	The processing of personal data in such a way that the data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organizational measures to ensure non-attribution to an identified or identifiable person. (GDPR Article 4(5))
RDP	Remote Desktop Protocol
RESTful API	A RESTful API is an application programming interface that follows the principles of Representational State Transfer (REST), using standard HTTP methods (GET, POST, PUT, DELETE) to perform operations on resources identified by URLs. It emphasises stateless interactions, meaning each request contains all necessary information without relying on server-side sessions. RESTful APIs are widely used for building scalable and interoperable web services.
Secure Processing Environment (SPE)	An environment in which access to electronic health data can be provided in following a data permit. An SPE is subject to technical and organisational measures and security and interoperability

Term	Description
	requirements. Specifically allowing access to only those persons listed in the permit, as well as user authentication, authorisation, restricted data handling, logging and the compliance monitoring of respective security measures. (EHDS Regulation, Article 73). Used to mean the SPE service run be the operator and to refer to collective functionalities of SPE user accessible instances.
sFTP	Secure File Transfer Protocol
SLA	Service Level Agreement
SPE project-based user space	Project-based environment within an SPE service that users access. Any shorter versions of this should include either 'user' or 'project' to be clear what environment is meant.
SSH	Secure Shell
SQLi	SQL Injection
TLS	Transport Layer Security
TOMs	Technical and Operational Measures
Trusted Research Environment (TRE)	TREs emerged from the UK health research sector, shaped by community-led principles and structured around flexible, function-based zones. They aim to create trusted, auditable access to sensitive data, often under national governance frameworks.
UDAS	User and Data Access Services
URL	Uniform Resource Locator
Virtual environment	Virtual environment is a networked application that gives users of simulated experience of something that doesn't directly correspond to underlying hardware. Commonly used for scalability reasons to maintain simpler environments using vastly more complex hardware.
VNC	Virtual Network Computing

Term	Description
VPN	Virtual Private Network
XSS	Cross-Site Scripting

Annex B: SPE and related requirements

List of all the main requirements in this report. The nominative list is maintained in a separate document “SPE_requirements.xlsx”.

Requirements have identifiers in the form ‘`{category}R-#`’.

Sensitive data (SD) requirements

SDR-1. Unauthorised users MUST NOT be able to access sensitive data

SDR-2. Service administrators SHOULD NOT have access to sensitive data

SDR-3. Sensitive data MUST be in a protected format at rest and in transit

SDR-4. Sensitive data protection MUST be done with widely accepted, secure algorithms combined with effective isolation measures

Secure Processing Environment (SPE) requirements

SPER-1. SPE MUST enable scientific research on sensitive data

SPER-2. There SHOULD be a diverse selection of SPEs for the varied needs of sensitive data research

SPER-3. It MUST be possible to transfer sensitive data between, in and out of SPEs

SPER-4. SPE MUST provide adequate protection against exposing sensitive data to unauthorised users

SPER-5. SPE design SHOULD promote collaboration among authorised users

SPER-6. Project-based user environments of SPE MUST be isolated from each other and open Internet

SPER-7. Authorised users MUST protect sensitive data they display

SPER-8. Authorised users MUST interact with their SPE project space only through secure protocols

SPER-9. All APIs connecting SPE components MUST be logged and monitored

European Health Data Space (EHDS) SPE requirements

EHDSR-1. HDAB MUST grant access to EHD using a data permit

EHDSR-2. EHD MUST be accessed using an SPE

EHDSR-3. Natural persons listed in the data permit MAY access the identified EHD in SPE

EHDSR-4. TOMs MUST minimise the risk of unauthorised EHD access in SPEs

EHDSR-5. Authorised health data users MUST be strongly identified

EHDSR-6. All access and operation logs of SPE MUST be available for verification and auditing

EHDSR-7. All SPE logs MUST identify the actor

EHDSR-8. All SPE logs MUST be kept at least for one year

EHDSR-9. TOMs of SPEs MUST be monitored for security

EHDSR-10. EHD MUST be identified in the data permit

EHDSR-11. Health data holder MUST upload the permitted EHD to be available in an SPE for the health data user

EHDSR-12. Health data user MAY download only non-personal EHD from SPE. Anonymised personal data is non-personal

EHDSR-13. HDAB MUST ensure by reviewing that no personal data is taken out of the SPE by the health data user

EHDSR-14. Regular internal and external security audits MUST be done on SPE TOMs

EHDSR-15. SPE TOMs MUST undergo risk assessments

EHDSR-16. HDABs MUST ensure that SPE TOMs audits are carried out and that risk assessments lead to risk mitigations

EHDSR-17. When SPEs mentioned in the EU Data Act are used for EHD, EHDS rules and requirements MUST be followed

EHDSR-18. SPEs MUST adapt to TOMs that the Commission will write into EHDS implementing acts

Operational (OP) EHDS SPE requirements

OPR-1. The SPE operator MUST have procedures in place to enforce user authentication and access restrictions based on the data permit associated with the processing of health data

OPR-2. SPE Operator MUST limit the number of authorised staff and any subcontractors who have high-privileged access enabling them to access or process health data and MUST implement effective procedures for managing and monitoring such access within the SPE infrastructure

OPR-3. SPE Operator SHOULD maintain a Service Portfolio including all services

OPR-4. SPE Operator SHOULD maintain a configuration management database (CMDB)

OPR-5. SPE Operator MUST implement mechanisms to terminate the secure processing environment upon expiration of the data permit. All electronic health data within the environment MUST be deleted or rendered unrecoverable within six months of permit expiry, including any backups or redundant copies. Procedures MUST be formally documented, monitored, and aligned with risk assessments and confidentiality requirements.

OPR-6. SPE Operator MUST undergo regular internal and external security audits to assess compliance with security, data protection, and operational requirements

OPR-7. SPE Operator MUST retain logs and access records to ensure traceability of all operations and enable audits or investigations when needed

OPR-8. The SPE operator MUST maintain up-to-date documentation of all relevant technical, organisational, and security processes

OPR-9. The SPE operator SHOULD assign a Compliance Officer or designate responsibilities to ensure adherence to legal, ethical, and technical obligations

OPR-10. SPE Operator MUST have an operating information security management system (ISMS)

OPR-11. SPE Operator SHOULD establish a Service Management System

OPR-12. SPE operators MUST conduct regular Data Protection Impact Assessments (DPIAs)

OPR-13. SPE Operator MUST maintain policies to ensure security around procurement systems and development and operation of systems.

OPR-14. SPEs SHOULD adopt change management process with impact and risk assessment

OPR-15. SPE Operator SHOULD adopt a release and deployment management process

OPR-16. SPE Operator MUST track and log actions of each authorised project member, including instances of data access, processing, viewing and output

OPR-17. SPE Operator MUST implement a secure storage process to retain logs of user access to the SPE for a minimum period of one year.

OPR-18. A reporting process MUST be in place to notify HDABs and relevant authorities of security incidents or non-compliance findings including data breaches or misuse

OPR-19. SPE Operator SHOULD adopt and enforce defined timelines for communicating and reporting incidents to the HDAB: Early warning notification: within 24 hours of incident detection. Detailed incident notification: within 72 hours of incident detection. Final incident report: no later than one month after the incident

OPR-20. SPE Operator MUST be able to promptly halt access and processing activities within the SPE when misuse or data breaches are identified.

OPR-21. SPE Operator MUST have disaster recovery procedures in place to restore the availability and integrity of the SPE services, including critical system components, configurations, and platform-level functionality, from clean backups, and to resume normal service operations following an incident.

OPR-22. SPE operator MUST adopt robust security measures to protect data, including the use of firewalls, encryption, and intrusion detection systems, in order to prevent unauthorised access, modification, or removal of sensitive information

OPR-23. Health data users, HDAB staff and SPE Operator staff who interact with the SPE MUST receive detailed, role-specific information or training covering health data processing, EHDS compliance requirements and security best practices coming from GDPR.

OPR-24. The SPE Operator MUST implement strategies for backup management, disaster recovery, and crisis management.

OPR-25. The SPE Operator MUST implement cybersecurity procedures to regularly evaluate the effectiveness of risk-management measures and promote fundamental cybersecurity practices and provide necessary training.

OPR-26. SPE Operator SHOULD define SLAs that include: uptime guarantees, response/resolution times for incidents include security SLAs, such as data encryption guarantees, incident response times, and audit logging

OPR-27. The SPE Operator SHOULD regularly restore backups to the platform as part of standard feature development and operational activities, ensuring preparedness for incident response.

OPR-28. The SPE operator MUST implement a formal patch management policy to identify, evaluate, and apply security patches in a timely manner based on severity

OPR-29. SPE Operator MUST establish a system update process to ensure timely and secure updates of software, OS, and firmware, tested prior to deployment.

OPR-30. A change management process SHOULD be used for assessing the impact of patches and updates before applying them to the production environment

OPR-31. SPE Operator MUST provide dedicated technical support with clearly defined SLAs and escalation paths for addressing incidents and technical issues

OPR-32. SPE Operator support staff SHOULD be trained in information security and privacy procedures, with roles and responsibilities clearly defined.

OPR-33. A knowledge base and support documentation SHOULD be maintained to assist in common issue resolution and improve incident response times

Federated (FSPE) requirements

FSPER-1. Legal or contractual agreement MUST cover the SPE federation across organisations

FSPER-2. Federation use identities MUST match over services

FSPER-3. SPE federation environment MUST fulfil sensitive data processing requirements defined for stand-alone SPEs

FSPER-4. All interactive user actions on sensitive data in the federation MUST be through SPE

FSPER-5. Federation governance structure MUST cover secure, shared data access and export from SPE

FSPER-6. Federation user identities SHOULD be shared

FSPER-7. Federation of SPEs MUST share technical and semantic interoperability needed for shared processing

FSPER-8. Federation of SPEs MUST use shared secure communication and data transfer protocols

FSPER-9. The federation MUST support distributed processing through authorisation and accounting services

FSPER-10. A federation SPE MAY fulfil federated processing requirements

Federated computing (FC) requirements

FCR-1. SPE MUST support common data models, such as OMOP CDM and applicable GA4GH standards, to enable technical and semantic interoperability

FCR-2. The HDAB MUST have in place (internally or in collaboration with SPE operators) required processes to deploy additional data models required by data users.

FCR-3. SPE MUST support the deployment and execution of software components needed to carry out federated computing as described and authorised in the data permit, subject to applicable security controls

FCR-4. The HDAB MUST have in place required processes to authorise the use of software components needed in federated computing

FCR-5. SPE MUST enable a data user application to retrieve anonymous results from the SPE to be merged with results retrieved from other SPEs

FCR-6. SPE MUST include functionality for ensuring the anonymity of results, with applicable methods such as: (1) pre-assessment of software components producing the

results, (2) automated anonymity assessment, (3) additional privacy protection mechanisms (such as differential privacy) and (4) manual inspections

FCR-7. The HDAB MUST have established processes for approving computerised access permissions for data user applications in the context of data permit authorisation

FCR-8. SPE MUST provide an open and standardised interface needed to exchange information with other trusted SPEs as needed to accomplish federated learning computations

FCR-9. SPE SHOULD support the use of privacy protection methods (such as differential privacy) for federated learning.

Technical Interoperability (TIR) requirements

TIR-1. The API services MUST be accessed via data user and data holder applications, which can be dedicated client applications or browser-based applications.

TIR-2. The API MUST follow a web services architecture (e.g., RESTful, GraphQL), enabling stateless request/response interactions and supporting secure file transfer protocols. The implementation MUST adhere to common architectural and security practices, including resource grouping and, where appropriate, the use of microservices to isolate sensitive services.

TIR-3. The API MUST operate over HTTPS, ensuring compatibility with standard web clients and server implementations. Additionally, it MUST support secure file transfer protocols (e.g. HTTPS, SFTP or FTPS for file operations) where applicable.

TIR-4. The API MUST use JSON as the primary data format for requests and responses with support for other formats where needed, ensuring secure data transfer via HTTPS or other secure protocols.

TIR-5. The API MUST be capable of uploading and downloading files of all standard file types. It MUST be possible to set restrictions on allowed file types and transfer direction (upload/download) through configuration settings separately for each API and workspace. The API must support efficient handling of large files, including resumable transfers, asynchronous processing where necessary, and configurable timeout settings to accommodate long-lasting operations.

TIR-6. API SHOULD have clear, machine-readable documentation (e.g., OpenAPI/Swagger).

TIR-7. All requests to data user and data holder API functions MUST be authenticated and authorised using standard methods (e.g., OAuth 2.0 with JWT, API keys). The system MUST enforce role-based access control, limiting access based on user roles and privileges.

TIR-8. API communication MUST use encryption (e.g., TLS) to protect data in transit. Additionally, it MUST be possible to enforce application-level encryption (e.g., AES-256) before transmission to provide an extra layer of security. If content encryption is used, keys MUST be securely managed and stored according to industry best practices (e.g., using a dedicated key management system).

TIR-9. The API MUST validate all incoming data to prevent injection attacks (e.g., SQLi, XSS) and malformed requests, responding only to predefined and approved requests with valid parameters.

TIR-10. Error responses MUST avoid exposing sensitive details (e.g., stack traces, internal error codes) while providing meaningful messages for debugging.

TIR-11. API access MAY be restricted based on network-level controls, including firewall rules, IP whitelisting, or Virtual Private Network (VPN) restrictions.

TIR-12. Requests, responses, and errors MUST be logged for monitoring and compliance.

TIR-13. API usage, failures, performance metrics and service availability MUST be monitored, with alerts for anomalies.

TIR-14. Mechanisms MUST be in place for approving and deploying new API versions and phasing out old API versions.

TIR-15. The API SHOULD meet defined latency and response time SLAs.

TIR-16. The system SHOULD handle expected and peak loads efficiently, with rate limiting and throttling mechanisms in place if necessary.

TIR-17. The API SHOULD ensure the accuracy and consistency of requests and responses by applying relevant format validators and other appropriate validation mechanisms, thereby preventing the delivery of erroneous, corrupted or inconsistent data.

TIR-18. The remote desktop interface MUST enable the data user to interactively access and process data with a remote desktop application

TIR-19. The data user MAY use a standard browser or a dedicated client to access the remote desktop environment.

TIR-20. The remote desktop solution MUST support remote desktop clients running on standard operating systems including (e.g., Windows, macOS, Linux).

TIR-21. The remote desktop solution MUST use a secure remote access protocol (e.g., RDP, VNC, or SSH with X11 forwarding) to establish a connection between the client and server. The communication protocol MUST support encryption to protect data in transit.

TIR-22. Access MUST require multifactor authentication (MFA). The system MUST enforce role-based access control, limiting remote access based on user roles and privileges.

TIR-23. Access MAY be restricted based on network-level controls, including firewall rules, IP whitelisting, or VPN restrictions.

TIR-24. Idle sessions MUST be automatically terminated after a predefined time to prevent unauthorised access. Users MUST be logged out or locked after a period of inactivity.

TIR-25. Clipboard sharing and file transfers MUST be restricted by default to prevent unauthorised export of data from the SPE. Changes to or removal of these restrictions MUST be configurable at the discretion of the HDAB.

TIR-26. All remote sessions MUST be logged, including authentication attempts, connection times, and actions performed during the session.

TIR-27. Service usage, failures, performance metrics and service availability MUST be monitored, with alerts for anomalies.

TIR-28. The service endpoint MUST enable communication between trusted SPEs to support federated learning.

TIR-29. The interface MUST support protocols needed to support client-server and bidirectional streaming communication (gRPC with HTTP/2 or an equivalent). Connections MUST be secured using TLS.

TIR-30. All connections MUST be authenticated using appropriate methods such as mutual TLS (mTLS) or token-based authentication while ensuring all data is transmitted over an encrypted channel.

TIR-31. Only pre-approved clients with valid certificates MUST be allowed to connect.

TIR-32. The interface MAY be configured to restrict access using firewall rules, IP whitelisting, VPN access, other network-level policies and holistic approaches such as virtual closed networks.

TIR-33. All requests, responses, and streaming events MUST be logged securely.

TIR-34. Real-time monitoring of streaming sessions, errors, latency, and security threats MUST be implemented. Alerts shall be generated based on anomalous activity (e.g., failed authentication attempts, unusual data patterns).

TIR-35. The interface MUST meet defined latency and response time SLAs.

TIR-36. The system MUST implement flow control to manage varying loads and prevent overloads, using rate limiting, throttling, and adaptive resource allocation as needed.

Annex C: Historical context and legacy models

Sensitive data is a barely a decade old concept introduced in GDPR. We are still reeling from the effort to interpret and adapt to it. Sensitive does not mean secret and information is no longer contained in physical documents. The first technical solutions designed for analysing small, structured datasets like statistical information are no longer sufficient. Opening up sensitive data to distributed use brings to the fore the human aspects of these activities that are difficult to contain.

Legacy approach: Secret

The topic of securing sensitive data immediately evokes popular images of clandestine operations or military secrets handled behind closed doors. If that sounds familiar, you already have got the wrong end of the stick. Confidential information is where a huge amount of effort is put to keep information as tightly limited as possible to as few people as possible. In contrast, health data for secondary use is meant to be utilised as efficiently as possible for the benefit of the society at large and specifically for future health care choices. The ideas contained and the implications of findings and decisions are meant to be discussed, debated, and published as widely as possible.

Unfortunately, the technology and therefore vocabulary that we use obscures the distinction between the two. It must be remembered that research on health data has always been done, and it is being done at many different approaches, situations, and purposes. A well-intended regulation that is guided by oversimplification can lead to serious loss of institutional knowledge.

Legacy approach: Physical isolation

Confidential legacy is also strongly linked to paper-based approach in information handling. The powerful imagery of red "TOP SECRET" stamped to the corner of the paper permeates popular imagination. Most European national legislations only started changing from data-based definition of official information in the 1990s and the effects of that are still seen in practice. The idea of a physical entity holding specific information sits fast.

In the brave new world of electronic data storage and communication, the realities are quite different. Information can be searched and accessed much faster, flexibly and scalably. Secure storage requirements are conceptually different and need abilities to back up and recover. Sharing, collaboration, and tracking of changes can be automated. Data security becomes paramount to preserve data integrity, prevent unauthorised access, allow auditing and tracking.

These are the very challenges that carefully designed SPEs and supporting services face when they strive to combine security for improved efficiency, productivity, and collaboration.

Legacy approach: Statistical analysis of registries

The first implementations of electronic environments tended to focus on statistical analysis of large amounts of social and demographic information from national registries. These analyses have been typically done using standard statistical packages like SAS along with

R. The source data are mostly ordered columns of textual and numeric values that require modifications to conform to the needs of the analysis appropriate for the question.

In other kinds of research, data types can be more varied, datasets significantly larger, the duration of the study much longer, and types of analysis more complex with a need for the development of completely new algorithms and software implementations. For these, the single closed environment created for the purpose of one analysis is a poor fit. A long-term clinical study or genomic study of a rare disease does challenge the boundaries of this approach. They need a more dynamic and distributed environment. That is the current problem of defining how to allow for a controlled and secure ecosystem that combines secure and precise communication to sensitive data services. We will have to draw from the research and practices of large-scale federated data management and combine these to the special needs of secondary use of health data in the context of EHDS.

Technical solutions can never be completely secure on their own

Development of sensitive data processing computer environments is driven by increasing security requirements. Although proper risk assessment is supposed to balance the impact of a threat with its probability, there is a clear pressure to keep increasing limiting security features to services. Adding them have the tendency to increase the cost-of-service maintenance and reduce the user experience. The usefulness of new security measures must be considered against the least secure existing aspects of the service.

In SPE, the graphical interface of the computer that allows the authorised health user to see unprotected health data is that element of least security. It is also the most important for users. Any technical security measures that users do not see the benefit of or make the routine use of the service harder for the user discourages them to use the service at all or prompts them to invent new ways to circumvent these obstacles.

This implies correctly that non-interactive means of accessing sensitive data can be more secure than interactive ones. This builds a case for query-based sensitive data services where queries and replies can be unequivocally recorded.

Together, these mean that regulation of authorised users will have to build largely on trust and accountability, i.e. logs. Trust is based on the health data users' willingness and ability in their professional capacity to follow the data access permit. Clear separation of internal and external threat issues and their countermeasures should solve many governance and scalability issues these services face.

Interoperability between SPEs will have a major unifying impact on services

The only way secure sensitive data interoperability can be ensured is through shared user and project identity. Only with these, can usage and the flow of information be tracked across a network of services.

The establishment of secure and trusted networks of communication between services will make evident the benefits of sharing resources. Many of the resources that SPEs need are available from the Internet but such resources need additional assurances to clear them for sensitive data and a secure communication layer. These in place, these resources will start

serving a larger community of services. A prime example is software libraries for commonly used programming languages.

Annex D: Sensitive data life cycles

The security of data processing inside SPE is tightly dependent on data transfer and supporting services. **Sensitive data access management and interoperability requirements** link SPEs to wider sensitive data cycles (Annex D: Figure 1) that are common to all sensitive data types and uses, including sensitive intellectual property and consent-based data processing.

This means that the whole ecosystem where SPEs function must be designed to support modularity and flexibility in order to accommodate different governance models and data flows (Principle 1.3.1. Flexible).

Both sensitive data management and interoperability depend on reliable and universally applicable **identity management**, which must extend to all participants and services in an interconnected environment. All these services also rely on stable and secure **network connectivity** to ensure availability and communication.

The generic, high-level use of SPEs (Annex D: Figure 1) has three use cases for sensitive data processing:

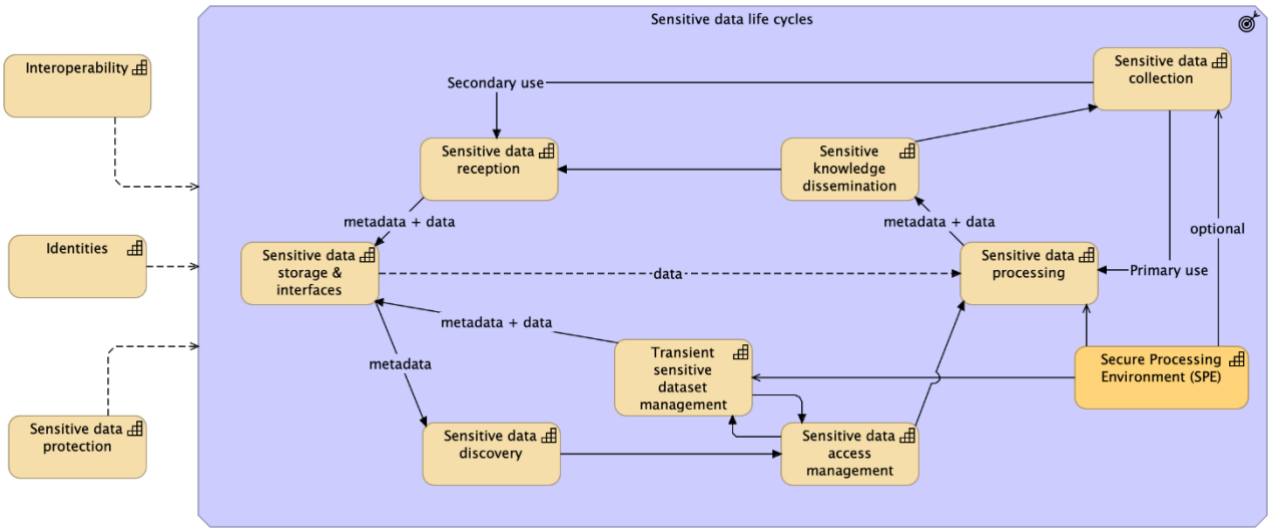
1. Primary use
2. Secondary use
3. Preparation of transient datasets for secondary use

The primary use might use SPEs for protecting the processing, even in the absence of regulatory obligation. SPE use for safeguarding sensitive data for general use is a major driver for SPE design. SPE should be a convenient way to fulfil the general GDPR and national legislation requirements of due diligence when wanting to protect any kind of sensitive data.

In the secondary use of sensitive data, the datasets must have descriptive and governance metadata that make them findable and enable the decision-making process of granting access to an identifiable dataset using a data permit.

When the permitted data is not handed over to the user exactly as it is in the data repository, like when data is first pseudonymised, combined or minimised, the dataset preparation adds another processing step (Transient sensitive data management) before it is ready for the user (TEHDAS2 M7.2 Draft guideline on data minimisation, pseudonymisation, anonymisation and synthetic data).

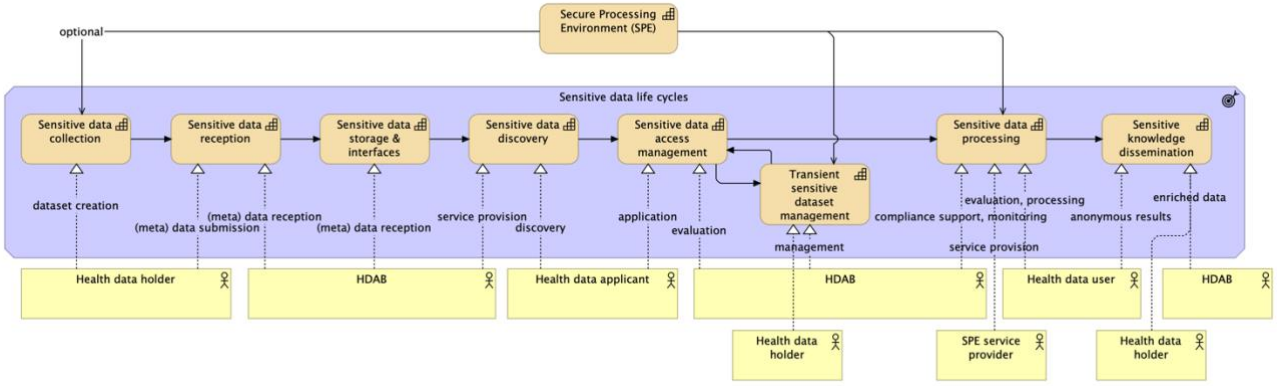
Annex D: Figure 1. Key capabilities of sensitive data life cycles.



To be maintainable and economically feasible, the specific requirements in EHDS for the secondary use of health data need to fit into these wider SPE ecosystem requirements.

Annex D: Figure 2 presents these sensitive data life cycles in a more commonly used linear use case format for easier reading. Capabilities are here directly linked to actors with role names and the functions they perform in EHDS context.

Annex D: Figure 2. EHDS sensitive data life cycles in linear format with actors.

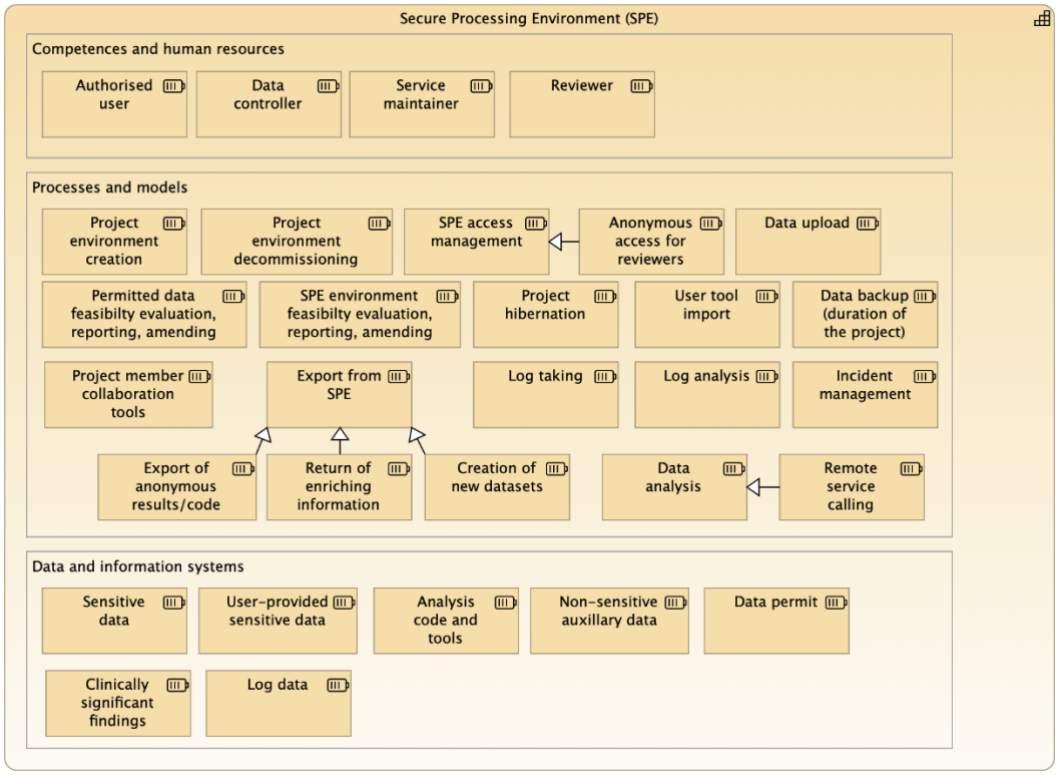


Annex E: Design considerations and expert commentary

Functional requirements of the SPE cannot be fully separated from the sensitive data life cycles ecosystem (Annex D: Figure 1). Some of its functional requirements must be shared with the whole ecosystem. This becomes clearer when we consider the needs of data transfer and interoperability that arise from the needs of distributed and federated computing.

In architecture, a stand-alone SPE concept is abstracted to a capacity that has resources (Annex E: Figure 1). We will follow the order of active resources to discuss the various aspects of SPE functionalities.

Annex E: Figure 1. Resource map of SPE capability.



Identity and authorisation

The absolute first requirement of any secure system is that all its users are uniquely and reliably identified (see **SDR-1**) as stated in the EHDS Article 16. That requirement includes that the chain of trust created by this identification is carried uninterrupted throughout the data application and use.

EHDS needs national and cross-border authorisation of users on a scale not yet realised within the EU. Therefore, its success will depend heavily on the widespread adoption of European electronic identification system with its cross-border eIDAS network connecting national electronic IDs and their efficient use in the European Digital Identity (EUDI) Wallet.

The use of eID for sensitive data processing should require high level of assurance that is equivalent to initial registration of turning up in person and authenticating with an official document[16].

Implementing acts and secure protocols underlying the EU Wallets are still under development, raising a possibility that EHDS services will initially need to allow the use of alternative authentication methods that should not be lower than the aforementioned assurance level. Also, the aim of the EUDI Wallet system is to provide identity services first for official uses and the needs of education and research are only fourth in priority list.

In addition to authorisation, EHDS needs to enable authentication of access to sensitive data with equally high level of assurance. This would allow converting the official, human-readable data permit to machine-actionable access permit that can be used to channel permitted health data to the health data user's project space inside SPE.

In Europe, automated research data access authentication services are offered only by Life Science Login (LS LOGIN) by EOSC and ELIXIR that implement GA4GH passports and visas[17] that are secure protocols able to give remote access to electronic resources.

GA4GH visas allow for additional attributes to be added to the person and the permit (as required in EHDS article 68(1)(d) about needed professional qualifications). Users' ability to access the data could be linked to their affiliation (organisation, position) as well as their abilities to handle sensitive data (see the next chapter Priority of user training). The LS Login currently has "de facto researcher" attribute that ensures that the person has the research status in their current organisation that is not available to EUDI Wallet. The word is out that data spaces will be responsible for implementing the attributes and services they will need to as Wallet verifiable credentials.

It should be a high priority for EHDS to implement these approaches. Luckily, EU Digital Wallet and GA4GH passport protocols are both based on OAuth and OpenID Connect specifications, making the work easier.

We may derive more specific requirements that go beyond minimum requirements and bring in precision, flexibility and convenience to SPE environments. The lesser status of them in this report are indicated with decimal numbering:

SDR-1.1. Authorised users MUST be identified reliably

SDR-1.2. Authorisation of users MUST require high level of assurance

SDR-1.3. Multifactor authentication MUST be supported

SDR-1.4. Single Sign-On (SSO) SHOULD be supported, using secure and standards-based protocols

SDR-1.5. User identities SHOULD be linkable to allow the use of existing logins

[16] eIDAS https://ec.europa.eu/digital-building-blocks/sites/display/DIGITAL/eIDAS+Levels+of+Assurance
[17] GA4GH passports https://www.ga4gh.org/product/ga4gh-passports/

Priority of user training

There are three main principles governing the implementation of SPE:

1. Enabling (Principle 1.)
2. Privacy (Principle 2.2.)
3. Accountability (Principle 2.3.1)

Various technical and operational ways ensure the privacy of data and the accountability of its processing. The enabling SPE requirement **SPER-7** separates the sensitive data from classified information and creates the weakest link in the SPE usage: Users in all practical cases are required to see, learn from its details, and reorganise the sensitive data before it can be effectively analysed.

This has two major consequences. Firstly, users need to understand the responsibility of accessing sensitive data and conversely that understanding needs to be demonstrated to representatives of data controllers before they are given access to the data. Secondly, all security measures directed to data users need to take this weakest link into account. Overdoing measures against authorised users are waste of effort and money, as well as increase the security risk due to user irritation and frustration.

In other words, if data processing is based on users learning new insights, the system must have trust to the users. This needs to be earned by demonstrating ability to shoulder the responsibility based on clearly defined curriculum. The modular SPE certification framework could be formalised to have several modules covering different levels of necessary knowledge, e.g. sensitive data protection, user responsibilities and penalties under EHDS, specifics of the SPE user. The licence could answer part of the requirements from EHDS Article 68(1) points (d) and (e) for heath data applicant to demonstrate their qualifications to perform the intended research (see also previous chapter Identity and authorisation).

Without naming the exact requirement, we could note down a provisory requirement:

> **SPER-7.1 Users SHOULD demonstrate understanding of sensitive data processing requirements**

SPE as collaboration area

The other aspect of the enabling principle is that SPE needs to promote communication between project members to do the data analysis (**SPER-5**). The requirements of sensitive data demand that all that communication should happen within the SPE. A prudent principal investigator would immediately set up a shared file that project members will be using as the project logbook. That would allow project members to know exactly what is happening in the project without the risk of revealing sensitive project data to outsiders.

We need to take that approach even further: SPE should be designed to enable information transfer and communication among project members. While each researcher needs their own private space within the SPE, the default should be that actions and files need to be fully shared. Ideally, all record keeping tools should allow for concurrent editing. This thinking should be extended to ongoing analysis runs so that project members working in different physical locations and time zones stay up to date with the project. The SPE should offer these tools. Similarly, the immediate communication in the form of chats, and phone and

video calls should happen within the SPE when the technical capabilities make that possible, like in generic virtual desktops.

From this point of view, the SPE is a new kind of social experiment in computing. The isolation of the SPE recreates the open and protected environment of isolated mainframes of the late 1960s, when file and user project space protection was in infancy, but enthusiastic users explored the potentials new technology in. This era also saw an explosion of new ideas getting implemented and information was freely exchanged. The roots of the open-source movement are in those times.

We are still so early in developing the SPE concept that these ideas have not yet turned into guidelines or implementations[18].

Finally, the project logbook is a good example of the privacy problems of sensitive data processing. It allows project members to focus on data analysis and use freely identifying pseudonyms in their data. The accountability of their work depends on the accuracy of their log keeping. This raises open questions about how accountability and anonymisation obligations intersect. Further guidance will be needed on retention of project-level collaborative metadata.

Data analysis

Data processing involves the software and frameworks provided within the SPE that enable data users to manipulate, transform and extract actionable insights from data. These tools enable the SPE to fulfil its core function, ensuring compliance to secure sensitive data processing rules and protecting against threats.

By default, in an isolated multipurpose SPE that is based on the secure virtual desktop concept, the data analysis itself should not be that different from normal desktop computing -- apart from lack of direct and open network access.

SPE environments that are built for specific analysis using pre-installed software for specific purposes, based for example on Jupyter Notebook, can be more straightforward to manage and run.

Tilting the balance towards more advanced technologies is the reality that building SPEs and their supporting services is an expensive and time-consuming effort. The cost is so high that it would be unrealistic to build a low technology solution for EHDS and develop and maintain simultaneously a higher-end SPE for other purposes. Low-end solutions are quick to set up but have higher running costs with more error-prone human steps. Deploying low-end SPE solutions may lead to higher long-term costs and reduced scalability, making upgrade paths a critical consideration for EHDS infrastructure planning.

This is where we move the focus of this treatise from abstract to more concrete. Different use cases and implementation choices will strongly affect the flexibility, scalability and capabilities of SPE-based sensitive data processing.

[18] Lehväslaiho, H. 2025-06-09 Secure Processing Environment in search of a metaphor. https://research.csc.fi/2025/06/09/secure-processing-environment/

SPE environment management

The definition of an SPE regardless of the use case is based on isolation and accountability. The isolation extends to the separation of roles to manage the SPE environment separately from data access and other services maintained within the same organisation. Additionally, it requires that data from different projects are not mixed at any point and a new SPE instance is launched for each of them.

The trusted use case: HDAB

The trusted use case (see chapter User stories) of SPE gives the user the freedom to select the environment and modify its details to meet their needs. For the HDAB, this means that the tools that the HDAB data managers need can be largely predicted and pre-installed to the environment where the HDAB data managers bring in the source datasets provided by the health data holders, process them, bring them out of the SPE, and place them for importing or streaming into the distrusted use case SPE for processing according to the data permit.

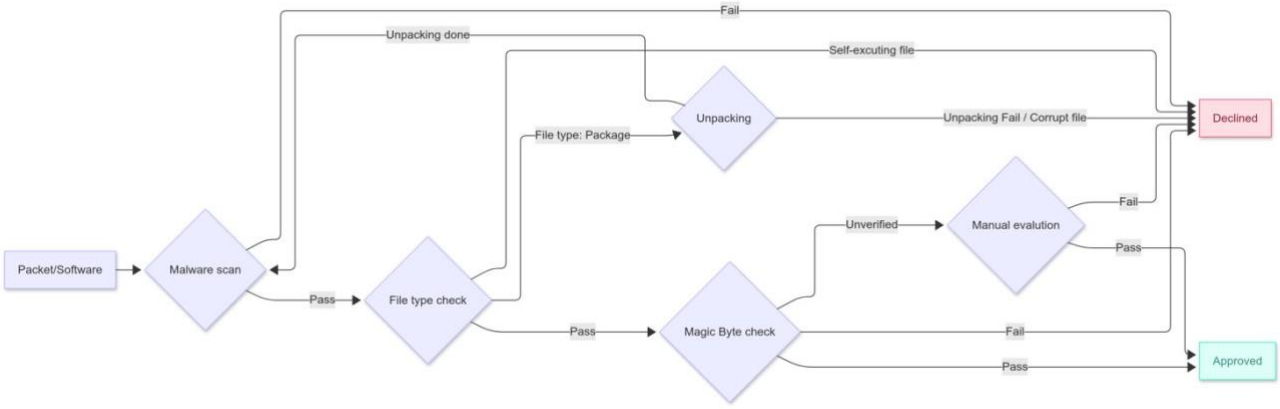
The distrusted use case: Health data users

In the distrusted use case of EHDS, various limitations and restrictions are imposed on the health data user. Under EHDS (Article 61), the HDAB designates the SPE for the health data user based on availability and user needs expressed in the data access application form. The health data user is required to pre-declare their own datasets containing personal data needed in the environment. Depending on the SPE setup, some needed resources (applications, libraries, datasets) will be provided into the SPE environment by the SPE provider (Article 67(2)(i)). In any case, users will need to import various kinds of files into their SPE project space. In an ideal case, there should be no limitations for user imports, but this can depend on the SPE setup (see chapter Cybersecurity of SPE infrastructure).

The SPE operators might want to agree on minimal security procedures that will allow them to share the burden of preparing them. Annex E: Figure 2 gives an overview of a proposal under preparation.

To summarise, security measures should strike a balance between protection and usability. They must not impose unnecessary restrictions or unnecessarily impede researchers' work. A streamlined process is needed to ensure tools can meet users' needs in a fast and efficient way.

Annex E: Figure 2. Resource-checking procedure for SPE operators, currently a draft best practice under development by the SPE Community of Practice subgroup (personal communication, Miikka Kallberg, CSC, FI).



The first task for users getting access to a new SPE is to evaluate its appropriateness to its intended use. Data controllers, represented by the HDAB, are responsible that the provided data is according to the data permit. The SPE provider should be approached in technical questions about the environment. There should be written instructions for most common problems (as detailed in the chapter Priority of user training).

The goals, approaches, and timetables of a project tend to change. Some of these changes will need changes to the data permit. Managing the timely onboarding and offboarding of researchers presents a significant security and operational challenge for projects, which can be addressed effectively with automated identity and affiliation management. It would be convenient if the management of the list of authorised participants could be amended directly by the principal investigator.

The specifics that are mentioned in the data permit will need to be passed on to the HDAB for recording or approval. The principal investigator of the project carries the main onus that official requirements are fully followed (EHDS recital 62). The requirement of destroying user enriched data from SPE at the end of the project (after 6-month grace period) will cause requests for long extensions to keep the project alive at least until its main findings are published (see further discussion in the chapter Data export from SPE). A related complication is that HDABs and SPEs will need to prepare their procedures for including pseudonymous access to the SPE to cater for anonymous review (anonymous for project members) of data and results that might be needed prior to publication. Let's define this case of using user aliases as a specific, lower number requirement for effective collaboration (SPER-5):

SPER-5.1 SPEs SHOULD have way to add project members without revealing their identity to other project members

Ending and pausing the user SPE

At the end of the project, or in case of a detected security incident, the HDAB and the SPE Operator will need to be able to cut access to data or SPE.

The health data users also need to be able to control the data and processing for security and cost reasons. The project area of SPE always has two main components: storage and analysis. Depending on the SPE implementation, these can be tightly or loosely coupled. A very loose coupling is exemplified by the system where the secure storage is fully independent of the processing component of which there can be multiple instances accessing data (see Operational SPEs in Europe: CSC SD Services)

Users might need an *SPE backup facility* to safeguard some intermediate results from their temporary storage area into the secure storage.

An extended pause in the activities might put a severe strain on the budget of a project. SPE operators should offer an *SPE hibernation* service to maintain the project in inactive state with significantly lower cost. This same sentiment is expressed in EHDS article 68(12) suggesting that HDAB could store permitted datasets in a system with lower cost to the user.

Monitoring of SPE use

Accountability (Principle 2.3.1) is the way to ascertain that the permitted health data user's actions correspond to the rules for sensitive data processing and specifically to their granted data permit.

Monitoring is the way to implement accountability. User actions are recorded by the SPE operator and monitored for irregularities. However, the legal responsibility of investigating them is on the HDAB. The sharing of executive actions on incidents between the SPE operator and the HDAB needs more clarification, especially when they belong to different organisations.

The DGA's definition of an SPE and again the EHDS legislation stress that all user actions must be recorded and be available for auditing. They are usually seen to be the **access log, the transfer log, and the activity log**. The first two record who accesses what service and when, and what is being transferred in or out of the service. They are seen as standard procedures for all services and make perfect sense to monitor those for sensitive data processing SPE.

The inclusion of obligation to record user activity within the service is highly contentious. It means that each and every action of the user within a service needs to be recorded, regardless if it has to do with sensitive data processing. Very few existing services are known to implement this. We feel this severely violates the user privacy. Their credentials for accessing sensitive data have already been vetted and approved. They have been given access to sensitive data they applied for, making them the data controllers of the approved data limited by conditions in the data permit in an environment that should protect them and sensitive data from outside perusal in addition to promote collaboration among project

members. Collecting unspecified personal data on user actions inside the SPE is the kind of activity that GDPR was written to prevent.

Not only is collecting all user actions a violation of privacy, but it also generates huge amounts of data that is very difficult to analyse. Further, graphical user programmes may do whatever with the data in their memory without leaving any sensible trace in the logs, undermining the usefulness of action logs. The increase in log volume is also significant cost issue.

The task of logging and monitoring needs to balance scalability with accountability requirements. Continuous monitoring tasks can see the most obvious and often unintentional breaches of requirements. The other end of the spectrum arises from the novelty of research results and the changing sensitivity of information over time. The latter is near impossible and time consuming to define. The cost is that arises either from human labour or automated tools is prohibiting.

Rather than focusing on action logs of dubious legality and utility, its efforts should be spent on ensuring that exported information is properly identified and anonymised. The chapter 'Data export from SPE' below proposes that all exported data from the SPE should be kept for an extensive period in case suspicions of their appropriateness or accuracy arises later.

The division of monitoring responsibilities between SPE operator and HDAB should clearly be stated in the implementation act. These cannot be left to be determined separately between each HDAB and SPE nor should they depend and vary on every data permit. SPE provider should do the base monitoring of logs. When an incident is detected, SPE operator should evaluate the severity and pass their significant findings to the HDAB that will take responsibility of the proceedings and ask for more details when needed.

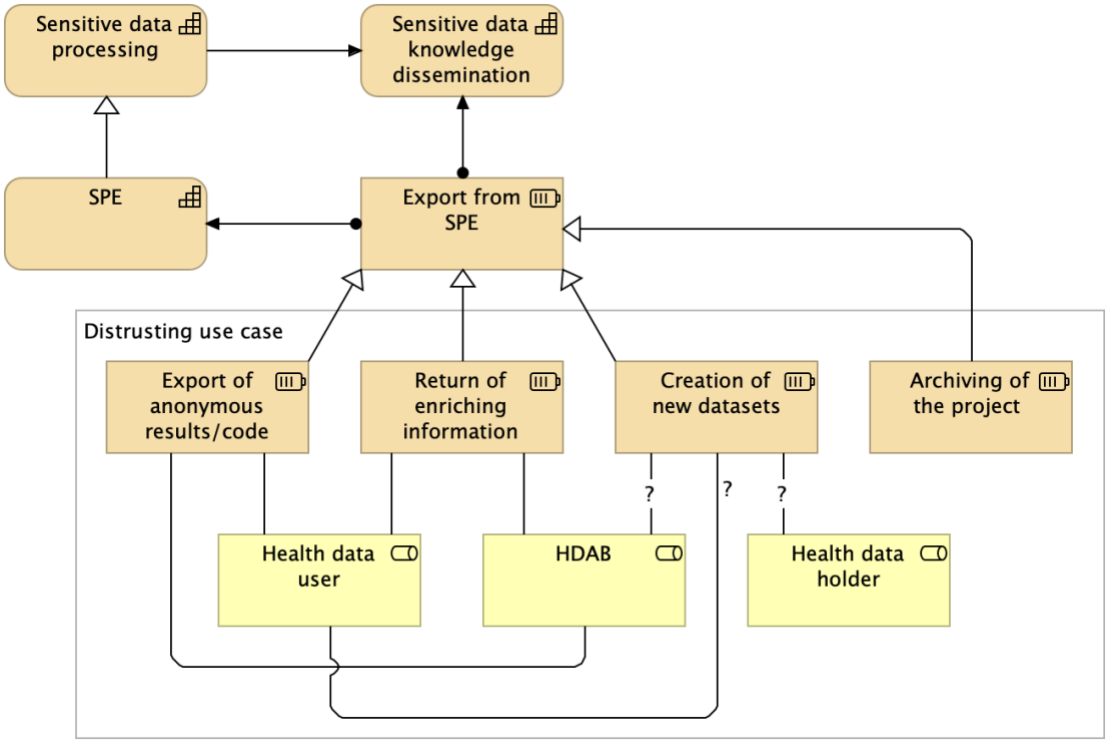
Data export from SPE

Data export from the SPE is not a problem for the trusting use case. HDAB staff are trusted to keep sensitive data secure outside SPE. They can be trusted to store and distribute the sensitive data responsibly in encrypted format and pass it on to a secure processing environment for further analysis by health data users.

The untrusted SPE use case has split the general export functionality into three discrete cases (Annex E: Figure 3):

1. Export of anonymous results
2. Returning of clinically significant results to the original data holders
3. Creation of new datasets of enriched research data

Annex E: Figure 3. Export from SPE



Export of anonymous results

EHDS requires HDABs to verify that exported results are anonymised, but it does not specify the method for performing such verification (Article 73(1)(f)). Member States or HDABs may develop tools or guidelines to support this obligation. The implementing acts under Article 73 may also provide common EU-level requirements or procedures to ensure consistency. Pseudonymisation and non-sensitive data types such as anonymised and synthetic data are covered in TEHDAS2 M7.2 Draft guideline on data minimisation, pseudonymisation, anonymisation and synthetic data.

Export of anonymous results may be divided to two major steps: (1) generation of anonymous results from analysed data and export of it by the health data user and (2) monitoring of the exported files for anonymity by the HDAB (EHDSR-13).

The feasibility of these tasks ranges from straightforward (exported information in code or large numbers of register data in aggregated format) to outright impossible (complex mix of data types).

It should be noted that health data users can be expected to export not only final analysis results for publication, but also intermediary results and the frequency and volume of these ones can be high.

Most obvious mistakes in anonymisation can and should be semi-automatically detected before the export is allowed to happen. Tools and guidelines are currently being developed

(see e.g. Anonymity verification tool (Finland)). Related efforts include the SACRO-project[19], which explores semi-automated checking of research outputs to identify potentially sensitive content.

Ultimately, the requirement of checking all output from open-ended scientific research, even by reviewing, can be extremely difficult. HDABs would need personnel who are experts in all health research domains who would spend long time trying to understand the risks of exports.

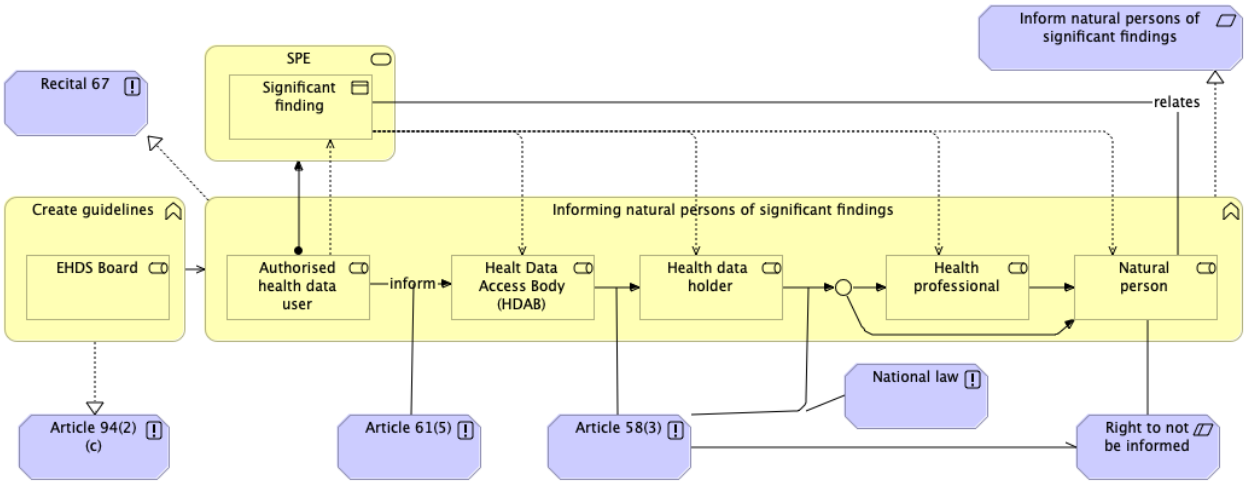
A more scalable approach would be to focus on helping researchers to anonymise their results effectively by training and automated feedback tools and store all output results as if they were logs of activities. Analysis of results anonymity would be a part of the required continuous activity of HDABs that can be subjected to more thorough reviewing in case of a security concern.

Returning of clinically significant results

Clinically significant findings (EHDS Article 65(5)) are an exception to the rule that health data users can export only anonymous data from the SPE (Annex E: Figure 4). While the process is to export out of the SPE, sensitive data is not seen to leave the EHDS secondary use domain as it is to be transferred securely to the HDAB who has the responsibility to inform the original data holders for the ultimate purpose of letting the citizen to know. The details of this process will be decided by member states.

Still, from the SPE service implementation perspective, this is an export of personal information from the closed environment that differs from other exports only in having a different recipient.

Annex E: Figure 4. The EHDS outline for the procedure for clinically significant findings from SPE.



[19] SACRO https://dareuk.org.uk/how-we-work/previous-activities/dare-uk-phase-1-driver-projects/sacro-semi-automated-checking-of-research-outputs/

Creation of new datasets of enriched research data

EHDS does not give any guidance on reusing datasets cleaned and enriched by health data users, and this is generally seen to be outside the basic EHDS use case. Enrichment is mentioned in the EHDS recitals.

Two main arguments may be raised in favour of reusing of datasets enriched in research. The effort of cleaning, harmonising and combining of research data for analysis is often the most time-consuming part of the work. From the practical point of view, throwing this away is this massive waste of effort will have to be again for any updated analysis. At a more philosophical level, this goes against the principle of sharing and the results of enriched data for future studies (Principle 1.1.4. Reusable in FAIR).

The **TEHDAS2 Deliverable D5.4: Short Guide for Data Enrichment for HDABs, Data Holders, and Data Users** will have guidance for returning enriched data to data holders. In addition, member states are recommended to consider enabling national legislations for storing and reusing of their health datasets.

Enriched datasets can be seen to cover two distinct situations:

Enhanced versions of original datasets – This refers to datasets whose quality or structure has been improved (e.g. through cleaning, deduplication, harmonisation), without combining personal data from multiple sources. These may, under certain conditions, be returned to the original data holder within the EHDS framework, especially if no new personal data are added or created.

Newly created datasets resulting from linkage or analysis – These are research outputs combining data at the individual level from multiple sources. Distributing would raise significant concerns about data controllership, re-use conditions, and compliance with the EHDS secondary use rules. These datasets cannot be exported or reused unless they are fully anonymised or covered by a clearly defined legal exception.

If these ELSI aspects can be solved, the practical solution for restricted access datasets are offered by the European Federated EGA services[20]. Federated EGA requires that researchers depositing datasets are members of research organisations that guarantee their identity and trustworthiness and will act as the broker for re-use requests of datasets.

Scenarios

We will somewhat arbitrarily divide available implementation options to three that we call local, state-of-the-art, and distributed (Table E.1.) that will enable progressively more capabilities and choices for data processing. Most existing SPE environment solutions already combine their features across these simplistic categories.

[20] FEGA https://ega-archive.org/about/projects-and-funders/federated-ega/

Table1E.1. Implementation options

SPE Capabilites	Local	State-of-the-art	Distributed
User identity	Local	Federated	Federated
Data protection	Full	Full	Full
Key management	Manual	Automatic	Automatic
Data access	Local copy	Managed, streamed	Managed, streamed
Interface	Virtual machine or fit for purpose restricted access	Virtual machine	Federation services for distributed service provision across organisations
Service type	SaaS	PaaS	

The local scenario represents a situation where the hosting organisation starts from scratch building an SPE from a virtual desktop by securing its operating system and deploys an automation tool inside a secured local network. Users need to be initially authenticated manually and assigned a local identity that will be enabled to a designated virtual desktop. The desktop image will need to be installed with most or all user tools and data need to be copied to the desktop once service providers and possibly users have installed their software products, and it has been secured. The requirement to keep sensitive data encrypted in storage and during transfers causes complications in encryption key management.

The local scenario can reach a high level of data security but at higher maintenance cost arising from the cost of manual labour. More people involved always increases the administrative cost of minimising human error. It is likely that a manual solution cannot effectively enough secure the SPE user environment to allow users to manage their own runtime software installations, making them more dependent on the service provider. This classifies the service as Software as a Service (SaaS).

The state-of-the art SPE system uses more effectively supporting services. User identification to high level of assurance is outsourced to their host organisation or country of origin (See Identity and authorisation). Access control to the SPE and project data is controlled directly by the data controller. The human-readable data permit is converted to a machine-readable access permit that automatically links permitted users and data to the project. Data is stored independently of its use and streamed by demand to the SPE project space. The lack of error-prone human intervention to SPE management and high degree of data protection by design and by default (GDPR Article 25) allow for Platform as a Service (PaaS), where users

are able to manage their application environment themselves (See Cybersecurity of SPE infrastructure).

To future-proof sensitive data processing, we need to carefully think what current sensitive data services can do, what is just becoming reachable based on research activities on public data computation, and what needs to be added to them to make sensitive data processing possible in them.

The first step is clear: We need to widen our thinking of secure processing environment (SPE) beyond a single predetermined virtual desktop. **The concept needs to include all types of processing of special categories of sensitive data** that follows the defined principles, regardless of place or time:

SPE implements sensitive data processing

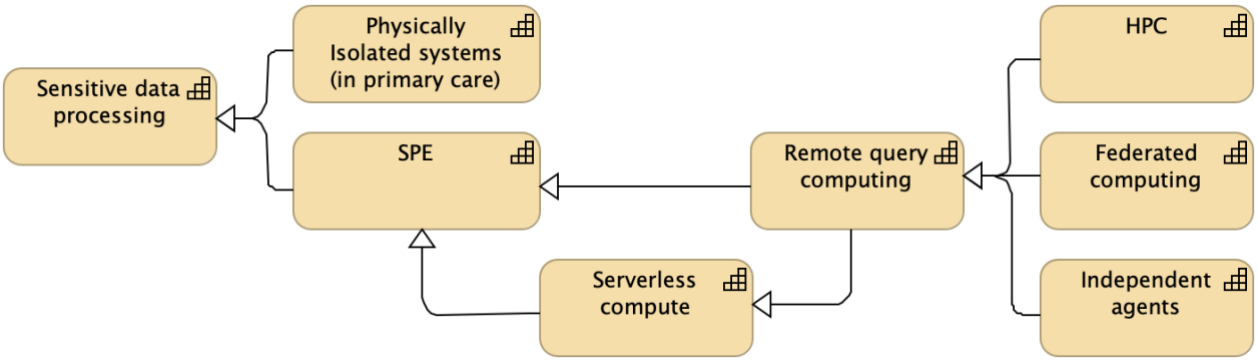
According to this, the desktop is the place where the user interacts with sensitive data, either directly or indirectly.

For the EHDS health data user experience, this means that the user is allowed to initiate processing tasks that have been audited to be part of the EHDS federation or explicitly mentioned in the data permit.

What this means in practice will need a more detailed look into various processing options and modern cryptographic abilities.

Annex E: Figure 5 provides one way to classify different processing approaches. The current SPE concept can be detached from physical location by implementing the desktop into ephemeral containers that can be executed in multiple compatible platforms (Serverless compute).

Annex E: Figure 5. Sensitive data processing approaches.



Allowing a user to send an interactive query to a remote server (Remote query computing), can dissociate the user from the externally managed sensitive data. This increases the security but may distance the user from immediate experience of data. The main questions to ask from a remote query are 1) the permissibility of the query and 2) are returned answers to query sensitive.

High Performance Computing (HPC) utilises massively parallel and interconnected computer systems that by design have been built to be shared simultaneously with multiple users. The challenge is to allow for isolated queuing and executing of sensitive data jobs without leaking any personal data. The implementation must ensure that the job submitter has permission to the processed data (see an example of HPC sensitive data processing).

Queries can be automatically sent to more than one target processor and results returned to the sender or the query itself can be the model to be trained, opening increasingly complex possibilities (see Implementing federated computing) with their own requirements. Federated computing will be the final section of this report (see Implementing federated computing).

Finally, while the processing in an SPE may be anything from simple command line tool execution to complex black box of a large language model (LLM) implementing artificial intelligence (AI), the demand for ever more complex reasoning is driving the development for independent and interacting agents that look for and evaluate information (Independent agents).

Annex F: Classification of risks and threats against SPEs

The table below (Table F.1) is sourced from the French Digital Health Agency and is part of a report designed to help healthcare facilities implement the Politique Générale de Sécurité des Systèmes d'Information de Santé[21], the legislative framework governing IT security in the health sector.

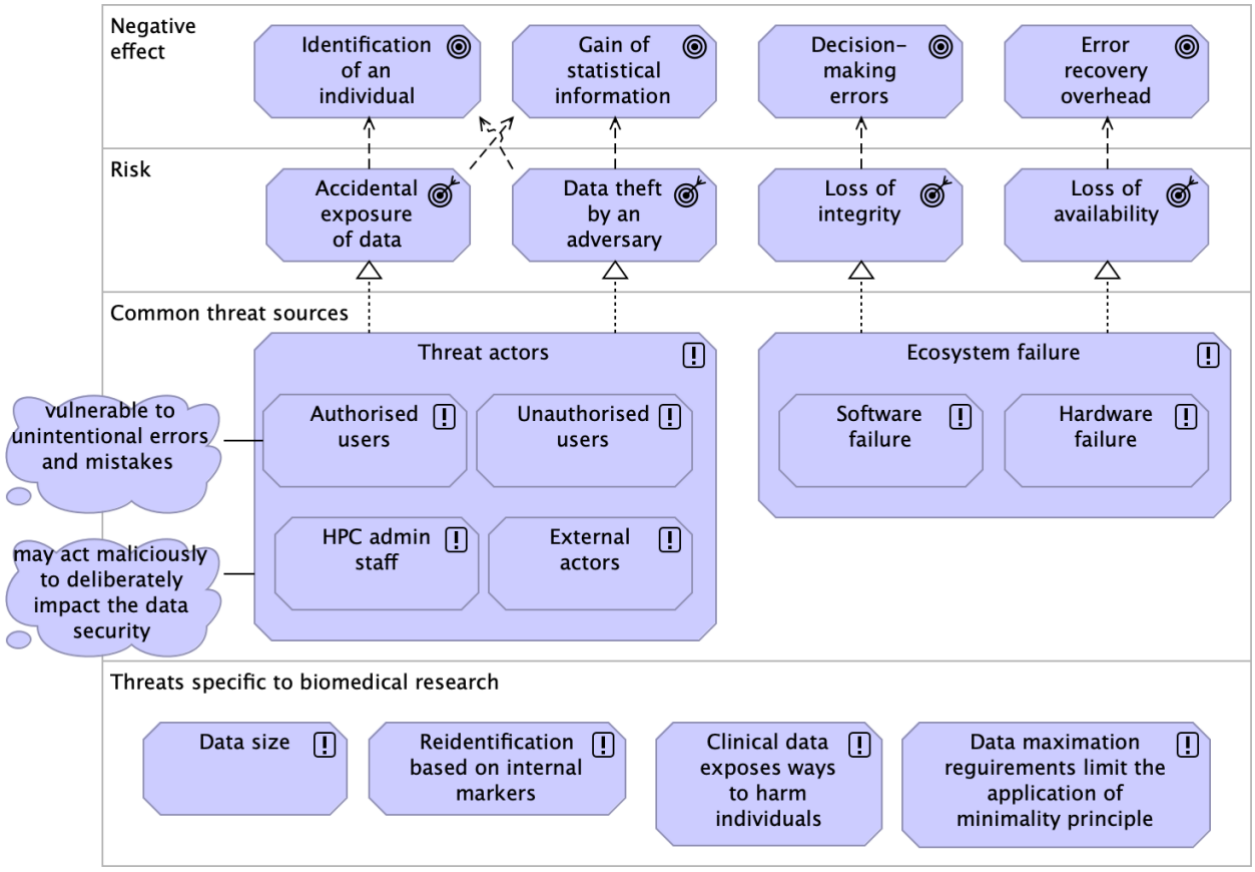
Table F.1. Classification of risks and threats against SPEs

Event	Threat scenarios	Reference	Level of risk
Unavailability of functions (medical application, connected device, etc.) or information (personal health data or personal data of a user) that may lead to service disruption, a negative impact on the image among users, and/or a user risk	Malicious code (virus, Trojan horse); Degradation or interruption of the network (local network, WAN and Internet access, WiFi network); Power supply failure; Failure of server room air conditioning; Handling error by IT staff; Lack of maintenance;	R-01	High
	Application saturation; Uncontrolled modification of software.	R-02	Medium
	Damage to computer equipment (fire, water damage, etc.); Unavailability of personnel (pandemic, health crisis, difficulties accessing buildings, etc.).	R-03	High
Alteration of functions (medical application, connected device, etc.) or information (for example, personal health data or personal data of a user) that may lead to service disruption, a negative impact on the image among users, and/or a patient risk.	Malicious code (virus, Trojan horse); Input or command error by an IS user; Lack of maintenance; Equipment failure.	R-04	High
	Uncontrolled modification of software (software update or configuration/parameter changes); Misuse of software's intended purpose (abuse of system or application rights, direct access to application data, etc.).	R-05	Medium
	Computer intrusion.	R-06	Medium
Alteration of evidence elements generated and stored by the IS (e.g., application logs, etc.) that may increase the legal risk for the organisation in case of litigation.	Misuse of a software's intended purpose (abuse of system or application rights, direct access to application data, etc.).	R-07	Medium
	Computer intrusion.	R-08	Medium
Access to personal health data or personal data of a patient by an unauthorised third party, constituting a violation of privacy and/or professional secrecy.	Loss or uncontrolled removal of equipment (laptop, removable storage media, etc.).	R-09	High
	Misuse of a software's intended purpose (abuse of system or application rights, direct access to software data).	R-10	Medium
	Computer intrusion.	R-11	Medium

[21] PGSSI-S https://esante.gouv.fr/sites/default/files/media_entity/documents/pgssi-s_guide_pssi_non-expert-ssi-v-1.0.pdf

PerMedCoE[22], the HPC/Exascale Centre of Excellence for Personalised Medicine in Europe project, published a report[23] 2022 on about guidelines for HPC systems on cybersecurity. The architecture graph of their main findings is in Annex F: Figure 1.

Annex F: Figure 1. Personalised medicine applications HPC security breach risks and impacts.



[22] PerMedCoE https://permedcoe.eu/
[23] D5.3 Derivation of general guidelines on data protection and privacy preservation of a use-case independent method/software development that will be exascale ready (September 2022) https://permedcoe.eu/wp-content/uploads/2024/10/PMC-D5.3.pdf

Annex G: Overview of relevant EU regulations

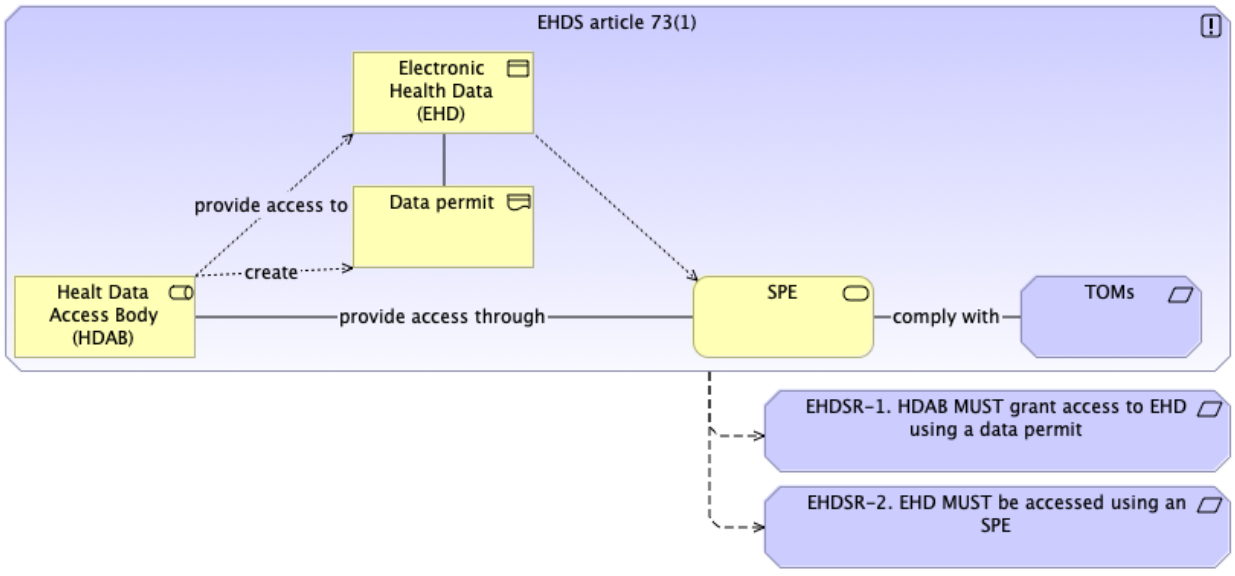
The EHDS will provide a trustworthy setting for secure access to and processing of a wide range of health data. It builds further on the General Data Protection Regulation (GDPR), Data Governance Act (DGA) and Network and Information Systems Directive (NIS2).

These regulations collectively shape the landscape for SPEs, influencing their design, implementation, and operation in several critical ways. The following sections explore in greater detail how EHDS, GDPR and NIS2 shape the architecture and operational principles of SPEs, ensuring that they can securely and efficiently support the goals of the EHDS and GDPR. Notably, EHDS Article 73 serves as the primary reference for the general requirements of SPEs within the EHDS framework.

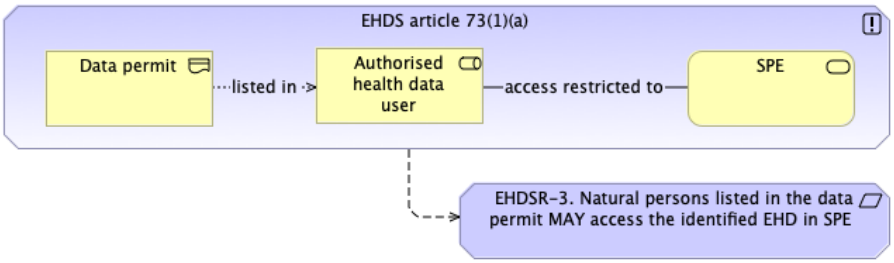
EHDS article 73 analysis to deduce SPE requirements

The logic of creating EHDS-specific requirements is presented as architecture graphs with the source paragraph as the figure legend.

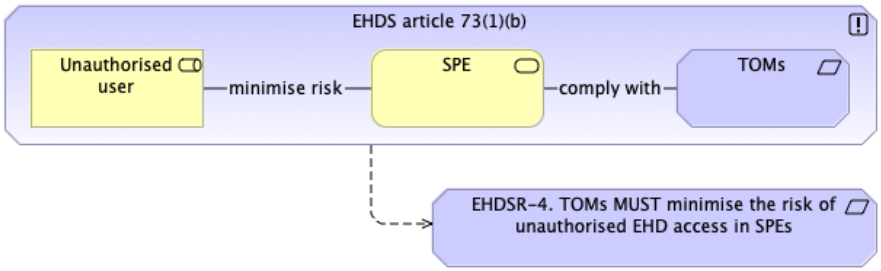
Annex G: Figure 1. EHDS article 73(1). Health data access bodies shall provide access to electronic health data pursuant to a data permit only through a secure processing environment which is subject to technical and organisational measures and security and interoperability requirements. In particular, the secure processing environment shall comply with the following security measures:



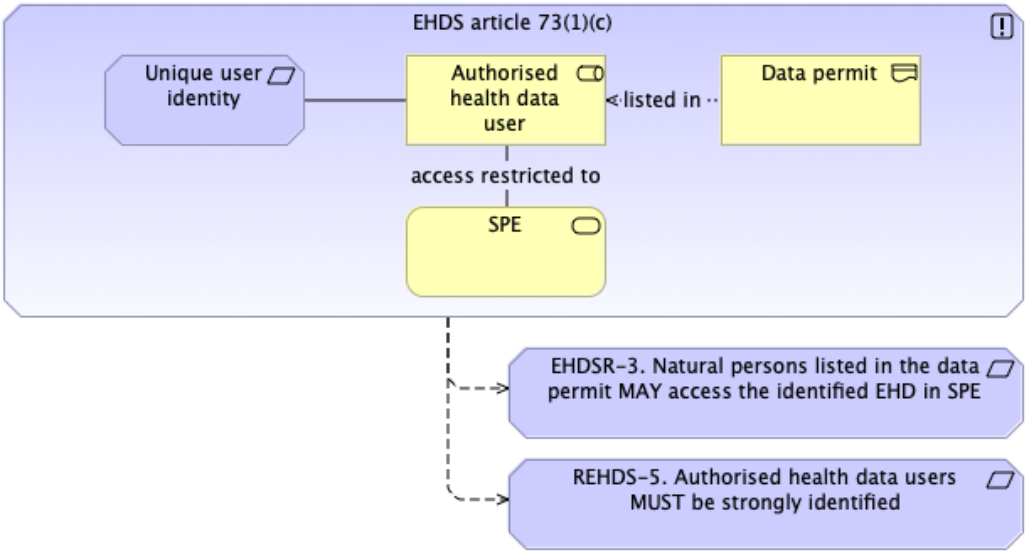
Annex G: Figure 2. EHDS article 73(1)(a). *The restriction of access to the secure processing environment to authorised natural persons listed in the data permit issued pursuant to Article 68;*



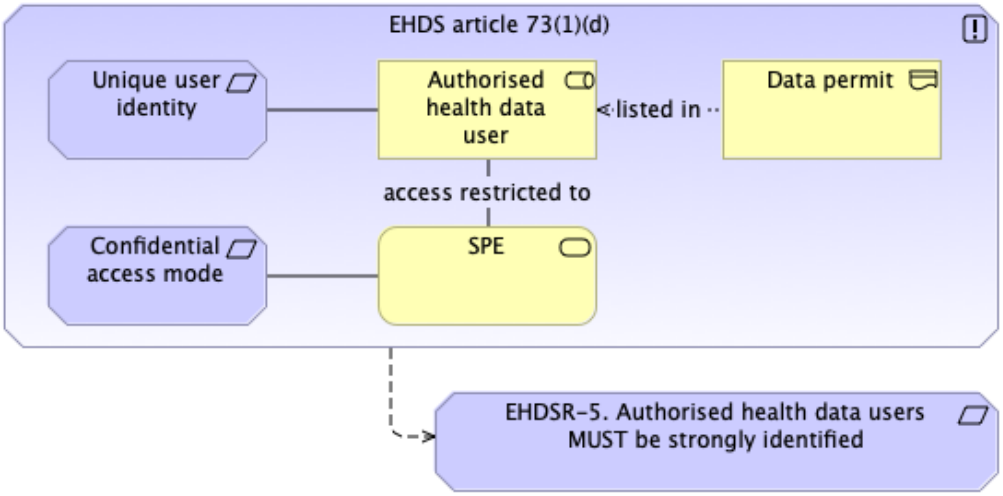
Annex G: Figure 3. EHDS article 73(1)(b). *The minimisation of the risk of the unauthorised reading, copying, modification or removal of electronic health data hosted in the secure processing environment through state-of-the-art technical and organisational measures;*



Annex G: Figure 4. EHDS article 73(1)(c). *The limitation of the input of electronic health data and the inspection, modification or deletion of electronic health data hosted in the secure processing environment to a limited number of authorised identifiable individuals;*

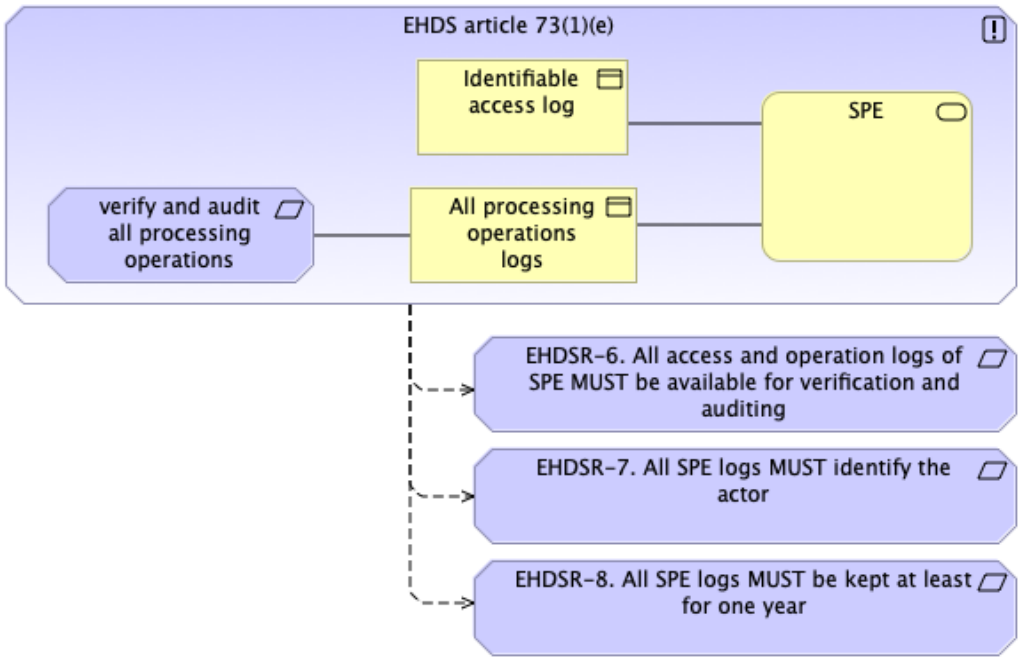


Annex G: Figure 5. EHDS article 73(1)(d). *Ensuring that health data users have access only to the electronic health data covered by their data permit, by means of individual and unique user identities and confidential access modes only;*

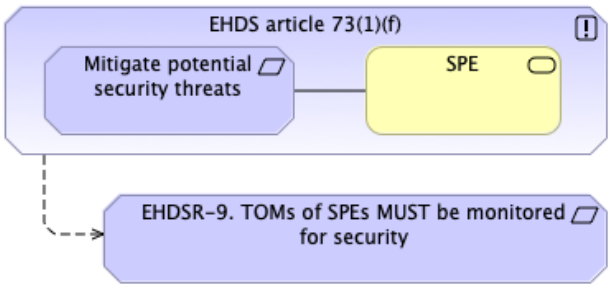


Annex G: Figure 6. EHDS article 73(1)(e). *The keeping of identifiable logs of access to and activities in the secure processing environment for the period necessary to verify and*

audit all processing operations in that environment; logs of access shall be kept for at least one year;



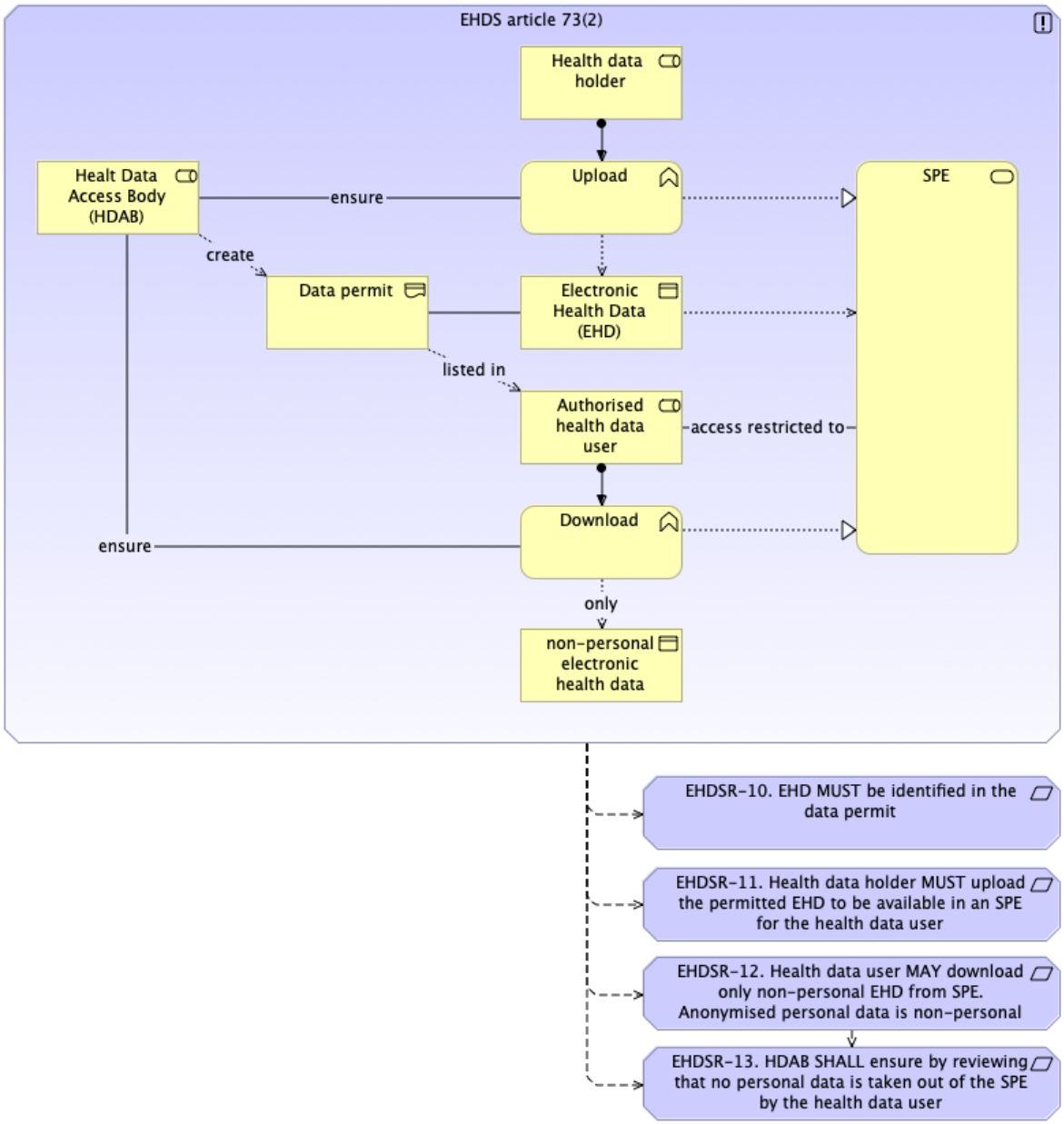
Annex G: Figure 7. EHDS article 73(1)(f). *Ensuring compliance and monitoring the security measures referred to in this paragraph to mitigate potential security threats.*



Annex G: Figure 8. EHDS article 73(2). *Health data access bodies shall ensure that electronic health data from health data holders in the format specified in the data permit can be uploaded by those health data holders and can be accessed by the health data user in a secure processing environment.*

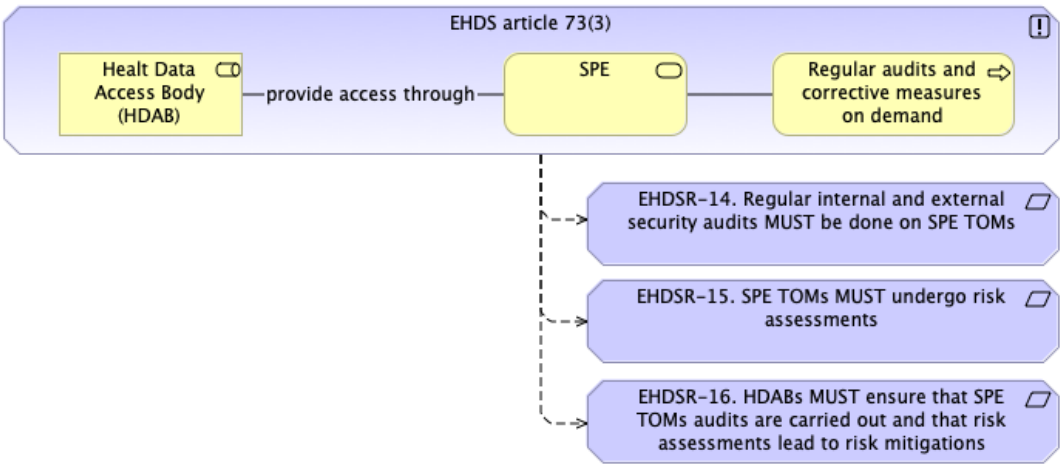
Health data access bodies shall review the electronic health data included in a download request to ensure that health data users are only able to download non-personal electronic

health data, including electronic health data in an anonymised statistical format, from the secure processing environment.

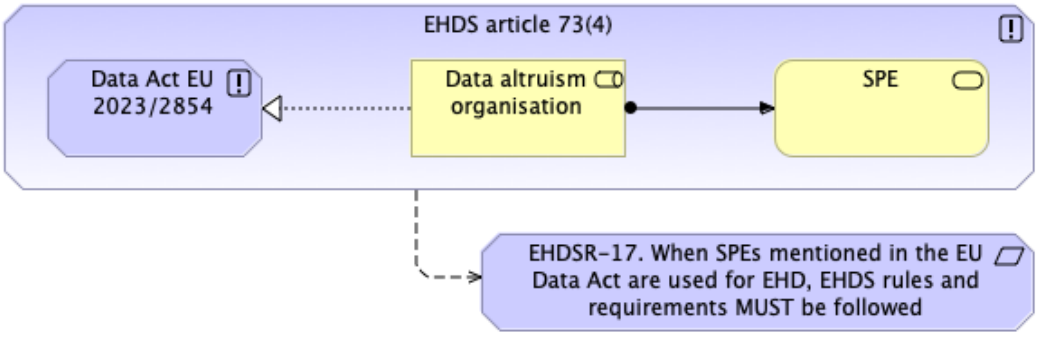


Annex G: Figure 9. EHDS article 73(3). *Health data access bodies shall ensure that audits of the secure processing environments are carried out on a regular basis, including*

by third parties, and shall take corrective action for any shortcomings, risks or vulnerabilities identified by those audits in the secure processing environments.

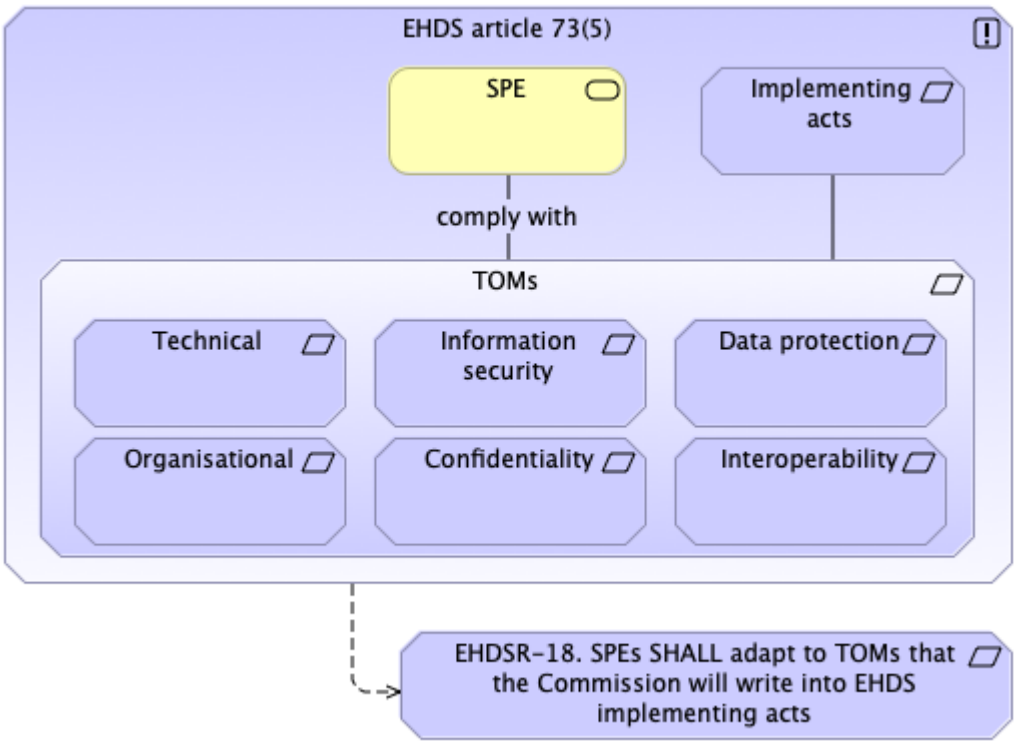


Annex G: Figure 10. EHDS article 73(4). *Where recognised data altruism organisations under Chapter IV of Regulation (EU) 2022/868 process personal electronic health data using a secure processing environment, those environments shall also comply with the security measures set out in paragraph 1, points (a) to (f), of this Article.*



Annex G: Figure 11. EHDS article 73(5). *By ... [two years from the date of entry into force of this Regulation], the Commission shall, by means of implementing acts, lay down the technical, organisational, information security, confidentiality, data protection and interoperability requirements for the secure processing environments, including with regard to the technical characteristics and tools available to the health data user within the secure*

processing environments. Those implementing acts shall be adopted in accordance with the examination procedure referred to in Article 98(2).

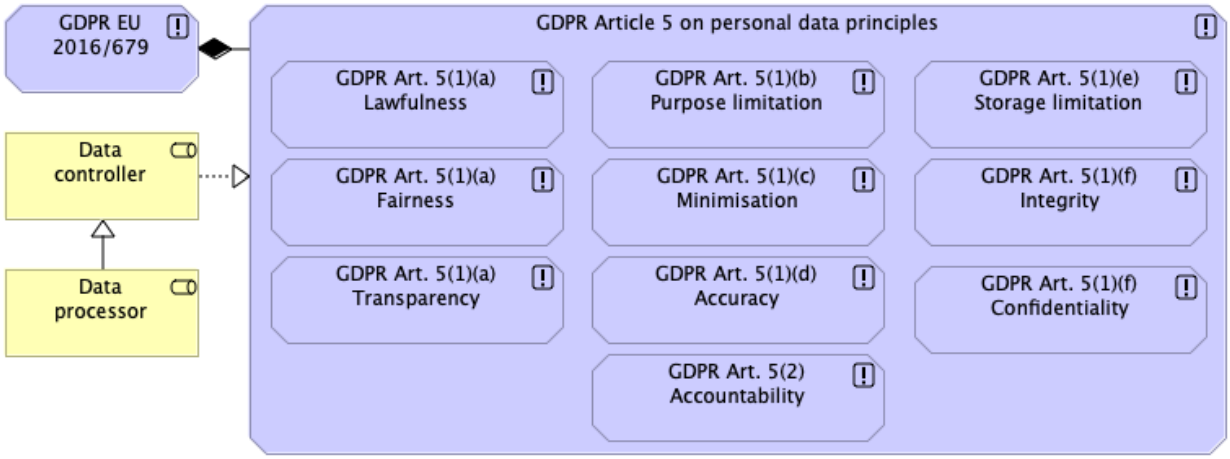


Key GDPR data and processing requirements

When developing technical specifications for SPEs, several GDPR articles must be considered to ensure compliance with data protection requirements. These articles include:

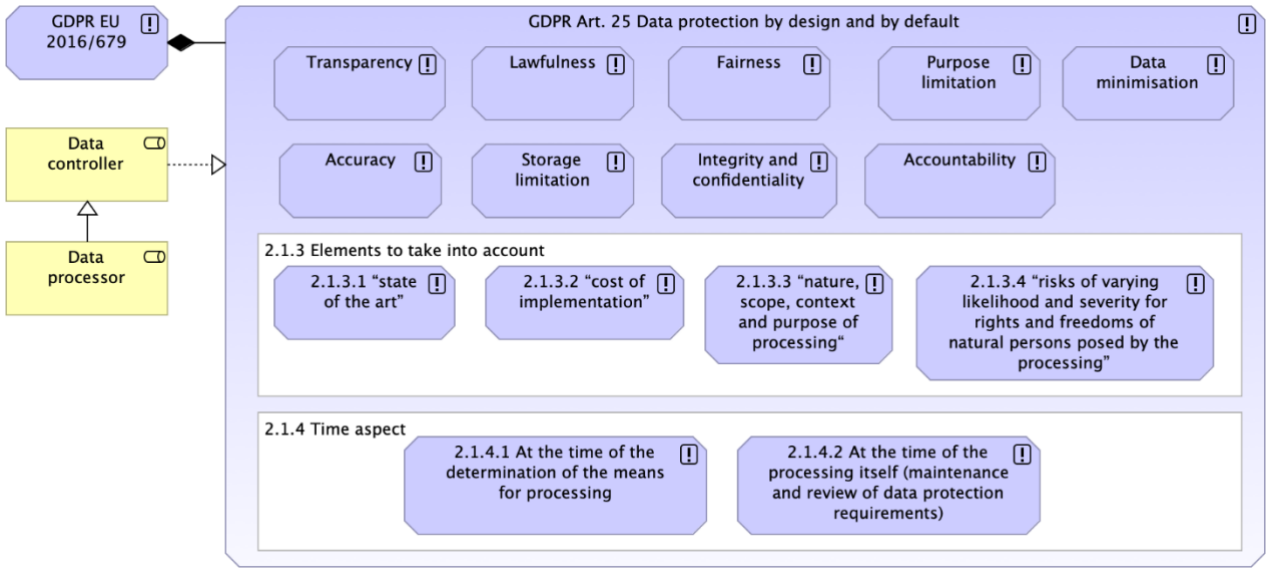
Article 5 – Principles relating to processing of personal data: This article outlines the core principles of data processing, including data minimisation, purpose limitation, accuracy and storage limitation. SPEs must be designed to process only the necessary data for defined purposes and ensure data is accurate and not retained longer than necessary. (Annex G: Figure 12)

Annex G: Figure 12. GDPR Article 5 – Principles relating to processing of personal data



Article 25 – Data protection by design and by default: This article requires incorporating data protection measures into the design of systems, ensuring that privacy is considered from the outset and that the SPE is configured to minimise data exposure by default (Annex G: Figure 13).

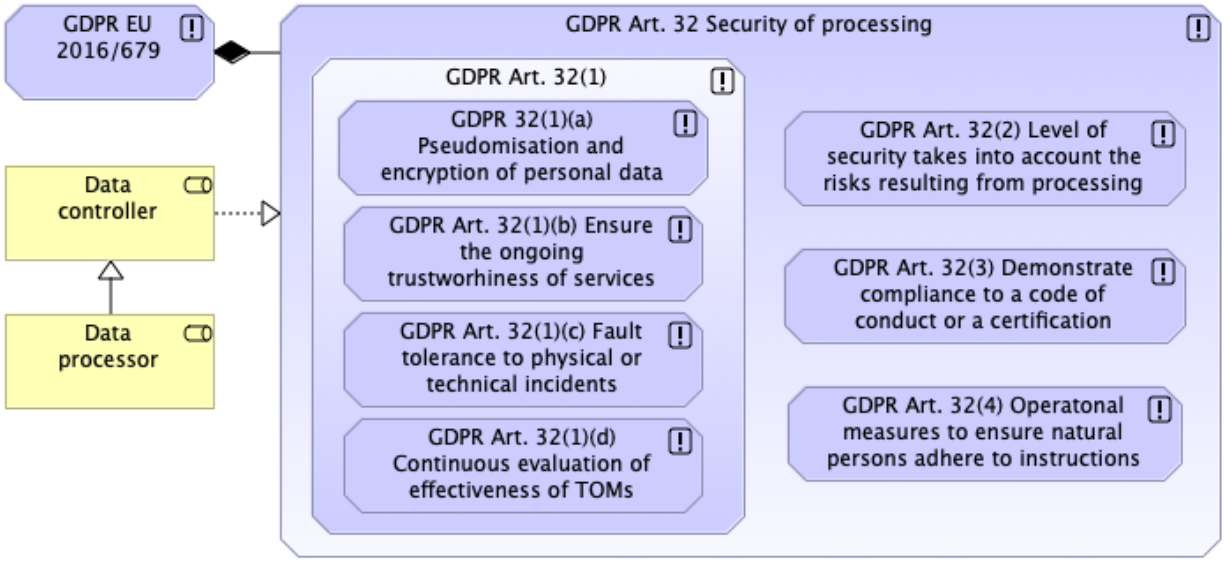
Annex G: Figure 13. GDPR Article 25 – Data protection by design and by default



Article 32 – Security of processing: This article mandates that appropriate technical and organisational measures, such as encryption, access control, and regular security

assessments, are implemented to ensure the security of personal data processed within SPEs (Annex G: Figure 14).

Annex G: Figure 14. GDPR Article 32 – Security of processing



The Article 32 requires specific measures to protect sensitive health data and ensure that it is processed securely.

Firstly, data protection by design and by default, as specified in Article 25, is essential. SPEs should isolate health data from other environments to prevent unauthorised access. By separating sensitive data, SPEs limit exposure to only those with verified access permissions. Additionally, role-based access control systems must be implemented to ensure that only authorised individuals, based on their specific roles and responsibilities, can access the data. This ensures that access is tightly controlled and compliant with GDPR's requirements for data security.

In addition to isolation and access control, encryption is a cornerstone of the SPE security protocols, directly supporting Article 32(1)(a), which mandates appropriate security measures. All health data within the environment should be encrypted both in transit and at rest, to prevent unauthorised access and helping to protect data integrity and confidentiality.

The integration of privacy-preserving technologies, such as anonymisation and pseudonymisation, aligns with Article 32(1)(a) and Article 25(1). These techniques reduce the risk of re-identification and protect data subjects' privacy while enabling secure data processing. The guidelines for implementing these privacy-preserving technologies will be outlined as part of T7.2 in TEHDAS2.

SPEs should also undergo regular security assessments and penetration testing to identify any vulnerabilities and maintain compliance with Article 32(1)(d), which requires ongoing evaluation of the effectiveness of security measures. Robust incident detection and response protocols must be in place to swiftly manage any potential data breaches.

Finally, accountability and documentation are crucial to maintaining GDPR compliance. In accordance with Article 5(2), which emphasises the accountability principle, and Article 30, which requires detailed records of processing activities, SPEs should maintain comprehensive records of all security measures and access logs. Regular internal and external audits should be conducted to verify that the SPE is consistently meeting GDPR's privacy and security standards.

NIS2 Directive

The NIS2 Directive (Directive (EU) 2022/2555) is the EU's key cybersecurity legislation, aimed at enhancing the cybersecurity posture across the Union by improving the resilience and incident response capacities of essential and important entities in critical sectors.

The NIS2 Directive establishes stronger cybersecurity requirements for critical infrastructure across the EU, with the specific obligations defined through national transpositions. SPEs handling sensitive health data will generally fall within scope, requiring compliance with stringent security measures such as endpoint protection, intrusion detection systems, and incident response protocols. They must also demonstrate resilience against cyberattacks and operational risks, ensuring continuity of service and the secure handling of data.

Role of CSIRTs under NIS2

Under the NIS2 Directive, each EU Member State is required to establish its own Computer Security Incident Response Team (CSIRT). These national CSIRTs act as the single point of contact for receiving notifications related to cybersecurity incidents, threats, and near misses within their respective countries.

Cybersecurity risk-management measures under NIS2

According to NIS2, entities must implement cybersecurity risk-management measures based on an all-hazards approach, designed to protect both network and information systems and their physical environment from incidents. At a minimum, these measures must include:

(a) policies on risk analysis and information system security

(b) incident handling

(c) business continuity, such as backup management and disaster recovery, and crisis management

(d) supply chain security, including security-related aspects concerning the relationships between each entity and its direct suppliers or service providers

(e) security in network and information systems acquisition, development and maintenance, including vulnerability handling and disclosure

(f) policies and procedures to assess the effectiveness of cybersecurity risk-management measures

(g) basic cyber hygiene practices and cybersecurity training

(h) policies and procedures regarding the use of cryptography and, where appropriate, encryption

(i) human resources security, access control policies and asset management

(j) the use of multi-factor authentication or continuous authentication solutions, secured voice, video and text communications and secured emergency communication systems within the entity, where appropriate

Annex H: Existing solutions for secure processing

The Annex H gives a listing of SPEs in Europe and brief explanations of some past and current projects. They provide a solid groundwork for SPEs in the EHDS, emphasising a balance between security, computational performance, and interoperability. These insights will guide the development of SPE specifications in TEHDAS2 and EHDS, ensuring that health data can be processed securely and efficiently while complying with European legal standards.

Contents:

Operational SPEs in Europe

Assessing existing SPEs provides a basis for developing technical specifications. Examining national implementations helps identify best practices for EU-wide adoption while also uncovering gaps, inconsistencies, and areas for improvement. This ensures that the specifications address emerging challenges and strengthen interoperability between SPEs.

15 operational SPEs across Europe that process individual-level health data for secondary use have been identified and are listed below (Table H.1). Environments limited to internal organisational use or specific projects have been excluded. Other relevant SPEs may exist but were not identified within the scope of this analysis.

Table H.1. Examples of operational SPEs in Europe

Organisation	Country	Name	Access type	Website	Notes
The Danish Health Data Authority	Denmark	The Secure Research Platform	Virtual Desktop	https://sundhedsdatastyrelsen.dk/da/english/health_data_and_registers/research_services/secure_res	

Organisation	Country	Name	Access type	Website	Notes
				earch_platfo rm	
Social and Health Data Permit Authority (Findata)	Finland	Kapseli	Virtual desktop	https://findat a.fi/en/kapse li/	
The wellbeing services county of Southwest Finland (Varha)	Finland	Atolli	Virtual desktop	https://www. auria.fi/tietop alvelu/en/ato lli/index.html	
HUS Helsinki University Hospital	Finland	HUS Acamedic	Virtual desktop	https://www. hus.fi/en/res earch-and-education/hu s-acamedic-secure-operating-environment	
CSC – IT Center for Science	Finland	SD Desktop	Virtual desktop	https://resea rch.csc.fi/-/sd-desktop	
Esior Ltd.	Finland	SPESiOR	Virtual Desktop	https://esior.f i/en/spesior/	
Statistics Finland	Finland	Fiona	Virtual desktop	https://stat.fi/ tup/tutkijapal velut/fiona-etakayttojarj estelma_en. html	

Organisation	Country	Name	Access type	Website	Notes
Health Data Hub	France	HDH Technological platform	Virtual desktop	https://www.health-data-hub.fr/page/faq-english	
Central Statistics Office	Ireland	Researcher Online System for Applications (ROSA)	Virtual desktop	https://www.cso.ie/en/media/csoie/aboutus-new/dataforresearchers/Researcher_and_RMF_Contact_FAQs.pdf	Process for applying and accessing health survey data provided by The Health Research Data Centre
Statistics Netherlands (CBS)	Nether-lands	Remote access environment	Virtual desktop	https://www.cbs.nl/en-gb/our-services/customised-services-microdata/microdata-conducting-your-own-research	Used for medical data (Based on the TEHDAS country visit report)
The Statistical Office of the Republic of Slovenia (SURS)	Slovenia	Remote access environment	Virtual desktop	https://www.stat.si/StatWeb/en/StaticPages/Index/For-Researchers	Used by the National Institute of Public Health (NIJZ)

Organisation	Country	Name	Access type	Website	Notes
NHS Digital	UK	Secure Data Environment (SDE)	Virtual desktop	https://digital.nhs.uk/services/trusted-research-environment-service-for-england	Previously called "Trusted Research Environment service for England"
NTNU – Norwegian University of Science and Technology	Norway	HUNT Cloud	Virtual desktop	https://about.hdc.ntnu.no/	
University of Oslo (UiO)	Norway	Services for sensitive data (TSD)	Virtual desktop	https://www.uio.no/english/services/it/research/sensitive-data/index.html	
University of Bergen (UiB)	Norway	SAFE (secure access to research data and e-infrastructure)	Virtual desktop	https://www.uib.no/safe	

.

TEHDAS1

TEHDAS1 deliverable D7.2[24] provides an in-depth analysis of SPEs as defined in the EHDS legislative proposal, along with guidelines covering technical, information security and interoperability requirements.

The report indicates that the preferred architectural model for SPEs is a decentralised system where multiple SPE providers across different EU Member States can offer compliant environments. This model also supports a federated learning approach, where data stays within national boundaries, but models and insights can be shared across borders. Each SPE

[24] D7.2. Options for the services and services architecture and infrastructure for secondary use of data in the EHDS https://tehdas.eu/tehdas1/results/tehdas-proposals-for-the-implementation-of-ehds-technical-infrastructure/

would adhere to common EU-wide standards, ensuring interoperability and trust among Member States. This approach also facilitates cross-border research while maintaining national control over data.

Additionally, the report emphasises that SPEs are not only security-focused but must also be flexible enough to handle a wide range of computational tasks. This includes everything from basic data analysis to more complex tasks like deep learning and artificial intelligence (AI). Given the increasing reliance on advanced computational techniques in health research, SPEs should be equipped to utilise high-performance computing (HPC), GPUs, and other advanced computing resources. Systems must be scalable to accommodate large datasets and resource-intensive processes, without compromising security.

To fulfil the requirements of the services outlined in Article 73 of the EHDS Regulation, the key functional capabilities for SPEs identified in TEHDAS1 are as follows:

- **Data Processing Capabilities**: Advanced analysis tools to handle sensitive data, including statistical software, AI libraries, and version control systems for code management.
- **Interactive Access:** Secure, remote access options such as remote desktop and secure shell connections. Some SPEs may also offer API-based access to support federated analysis.
- **Strong Access Control**: Comprehensive access management (data holders for data deposition, data users for data analysis, and system administrators for SPE operations)
- **Controlled Communications**: Data imports, exports, and other outbound communications.
- **High Security Standards**: Adherence to stringent security measures to protect data integrity and confidentiality.
- **Defined Operational Protocols:** Clear, well-documented procedures governing the operation and management of the system.

SPEs must also integrate harmonised security measures across the EU, allowing for consistent and reliable processing environments. This approach promotes cross-border data collaboration, with interoperability as a key consideration. To this end, SPEs must support standardised interfaces and protocols for seamless data sharing and access management. This ensures that data can be processed and analysed across different countries while maintaining a high level of security and legal compliance. Aligning security standards with recognised frameworks such as ISO 27001 and ENISA guidelines is recommended.

Five Safes

The concept of Five Safes — a framework for planning confidential data governance and management solutions has emerged and evolved in the United Kingdom over the past two decades.

The Five Safes framework has been initially conceived to help establish a virtual microdata research data centre at the UK Office for National Statistics in 2003[25]. Since then, it rose to

[25] Green, E., & Ritchie, F. (2023). The Present and Future of the Five Safes Framework. Journal of Privacy and Confidentiality. doi: https://doi.org/10.29012/jpc.831

international prominence and has been adopted as governance/data management standard for introducing (or retrofitting) various confidential data access solutions in the UK, the statistical offices in Canada, Australia, New Zealand, Norway, as well as Eurostat. Moreover, the framework has been built into or has influenced legislation such as Digital Economy Act in the UK and several state laws in Australia.

The Five Safes defines five dimensions of confidential data management: safe projects, safe people, safe settings, safe data and safe outputs. These are often formulated as "key questions" (Table H.2).

Table H.2. The Five Safes and key questions (Green & Ritchie, 2023)

Element	Typical question	Example of problems being addressed
Safe projects	Is this appropriate use and management of the data?	- What is the purpose of the access request? - Is this an ethical and lawful use of the data? - What is the benefit to society or to the organisations sharing data? - Is there a data management plan in place? - What happens to the data at the end of the project?
Safe people	How much can I trust the users to use the data appropriately?	- Do the users have the necessary technical skills? - Do the users need training inhandling confidential data? - Are users likely to follow procedures?
Safe settings	How much protection does the physical environment afford to the data?	- How are data stored? - Are there physical restrictions on the users? - Does the IT prevent unauthorised use? - Are mistakes by authorised users likely to be detected?
Safe outputs	How much risk is there in the outputs of the access breaching confidentiality?	- If the aim of access is to produce statistics, is there any residual risk by, for example, showing outliers? - If the aim of the access is to produce data for onward transmission, how do we make sure that the released data are appropriate for the next use?
Safe data	Is the level of detail in the data appropriate?	- Is there sufficient detail to allow the project to go ahead? - Is there excessive data not necessary for the project?

As Green & Ritchie underline, the dimensions are not limits, they are scales. What "safe" means, or how restrictive the safety requirements of each dimension are, is dependent on

the context. For example, for open data, only the Safe Data dimension must be controlled. In a secure compute environment, data is safer, perhaps it needs to be deidentified only, while there is a higher degree of control across other dimensions. This flexibility goes hand in hand with "principle-based" approach to designing governance and legal measures.

In modern understanding, the Five Safes are an aid in identifying structures and goals of a particular confidentiality solution, not a rigid set of guidelines or rules. The right way to approach a design of a system according to the Five Safes is an analysis of the approach to the design problem, then specification of broad principles and aims of the solution, followed by using the Five Safes to provide the structure and, finally, identifying the good/best practices in each area of Five Safes.

Trusted Research Environments

Based on the Five Safes framework, the past two decades have seen the rise and adoption of the Trusted Research Environments (TREs), particularly in the United Kingdom, as a safer computer environment to protect the identity of human genomic sequences and then expanded to cover health and social data. The TREs are now seen as the preferred approach (DARE UK, 2024) to ensure the safety of processing of any sensitive data. While there exists a set of community-defined requirements and a derived high-level architecture, interestingly, TRE does not have an agreed definition. It has been called by many alternative names like Secure Data Environment (SDE by the England National Health Service) or Data Safe Heaven (Scotland), and the TRE specification allows it to be organised in many ways.

The DARE UK Federated TRE Blueprint[26] clarifies the concept by specifying three functional zones that any TRE may contain in any combination:

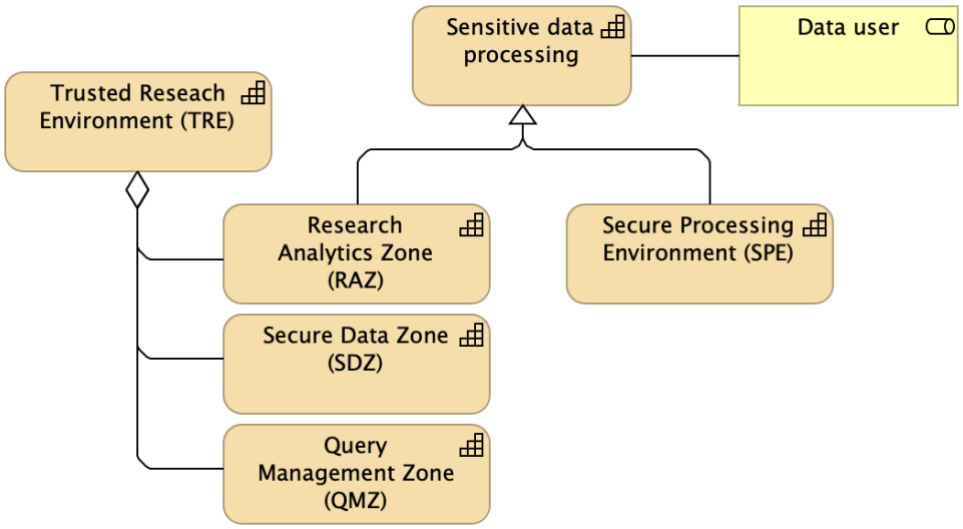
- **Research Analytics Zone (RAZ)** for project-specific data processing by users
- **Secure Data Zone (SDZ)** for active data management roles of data governance
- **Query Management Zone (QMZ)** to provide secure remote data access services

In an extreme case, a TRE can be composed only of RAZ making it impossible to draw any functional conclusions from the name. Any practical discussion about TRE functionalities should now indicate the zone.

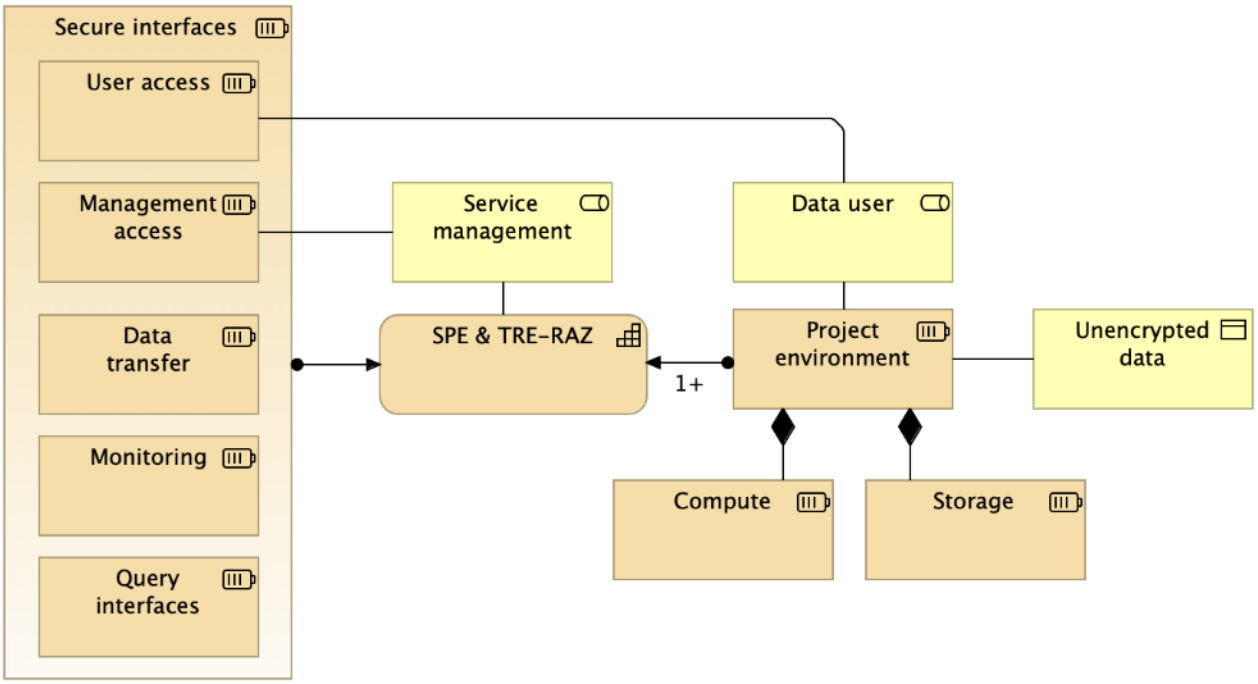
Both Research Analytics Zone of TRE and SPE are based on securing sensitive data privacy within a secure computer environment for the exclusive use by data users. Both isolate users by project and demand clearly defined, limited, and secure interfaces out of them[27] (see Annex H: Figure 1 and Figure 2).

[26] DARE UK Federated TRE Blueprint https://zenodo.org/records/14192786
[27] Lehvaslaiho H. (2025) SPE and TRE terminology for sensitive data processing https://zenodo.org/records/15696511

Annex H: Figure 1. Equivalence of SPE and TRE-RAZ



Annex H: Figure 2. SPE and TRE-RAZ may contain multiple isolated project environments that have only secure and limited access to the outside



DARE UK

The DARE UK (Data and Analytics Research Environments UK) project is a UK-wide initiative aimed at developing a coordinated, trustworthy, and secure national data research infrastructure. It is designed to support research using sensitive data (like health and administrative data) in a way that is safe, ethical, and for public benefit. In its first phase (2021-2024), it focused on developing the technical and governance foundations. The outputs of the projects[28] were categorised according to four of the Five Safes:

- Safe data: Semi-automated Risk Assessment of Data Provenance and Clinical Free-text in TREs (SARA) aimed at using machine learning to better understand privacy risks in the free-text data.
- Safe outputs: Semi-automated Checking of Research Outputs (SACRO) focused on introducing efficiencies into checking research results for disclosure risk before they leave TRE.
- Safe projects: Maintaining the safety of projects spanning multiple TREs (TRE-FX and TELEPORT projects).
- Safe settings: the SATRE project set out to assimilate the essential features of TREs into a common specification and provide a first blueprint for new TRE builders.

The currently ongoing phase two of the project has yielded a blueprint for a federated TRE architecture[29]. For this work, of particular interest are the SATRE specification and the federated architecture blueprint.

SATRE

> *SATRE (Standardised Architecture for Trusted Research Environments) started as a DARE UK Driver Project working to standardise access to secure data in trusted research environments, and included University of Dundee, Alan Turing Institute, UCL, Ulster University, and Research Data Scotland. (*SATRE website[30]*)*

A major outcome of the project has been a standard architecture specification for Trusted Research Environments. The specification is pertinent to TEHDAS2 specification as it defined key capabilities of a general TRE and a comprehensive list of criteria that can be used to evaluate compliance against the framework of any given TRE solution.

The four key capabilities — or pillars in SATRE parlance — are[31]:

- Information Governance
- Computing Technology and Information Security
- Data Management

[28] DARE UK (Data and Analytics Research Environments UK). (2024). The 2023 DARE UK Driver Projects: Summaries and lessons learned. Zenodo. https://doi.org/10.5281/zenodo.11443328
[29] DARE UK. (2024). DARE UK Federated Architecture Blueprint (2.2).
[30] SATRE https://satre.uktre.org/en/page/about/
[31] SATRE Specification. https://satre-specification.readthedocs.io/en/stable/

- Supporting Capabilities

The *Information Governance* pillar consists of requirements ensuring information risk is measured and managed to an acceptable level (EHDS Article 73).

The *Computing Technology and Information Security* pillar lays out a set of technical requirements related to the systems used to secure, manage and provide compute capabilities to *Data Consumers*.

The *Data Management* pillar is concerned with managing data assets while they exist (temporarily or permanently) within a TRE. Note that some of the requirements, particularly those that concern data management from the point of view of a data holder are beyond the scope of an EHDS compliant SPE.

The *Supporting Capabilities* pillar is only indirectly associated with a SPE provider. Most of the requirements are related to HDAB itself.

SATRE Roles and their Relation to the EHDS SPE

SATRE Project also defined roles belonging to different stakeholders involved with the data lifecycle[32]. The roles may fit well the *health data permit-based* project concept as data processing roles are tied to natural persons, as well as the kind of expertise needed to run a SPE provider.

The role of a *Data Consumer* aligns with *health data users* as defined in the EHDS legislation[33]. In particular, *Project Manager* maps to the role of Principal Investigator under EHDS (Article 68). The role of *Data Analyst* corresponds to any authorised natural person who accesses the data with the purpose of processing.

The Data Management roles are broadly applicable to organisations which curate the data in addition to providing it in a SPE. In this sense, they may be beyond the scope of an EHDS SPE directly --- except for *Output Checker*, who may need to be listed on a health data permit with the responsibility of ensuring that no confidential data exists in the outputs of SPE project, the data which is supposed to be released out of the secure environment.

The SATRE roles concerned with infrastructure management describe the governance needs of an *SPE Operator*, in order to be applicable to the broadest community of existing SPE providers.

[32] SATRE specification role https://satre-specification.readthedocs.io/en/stable/roles.html#roles
[33] Directorate-General for Health and Food Safety. (2025, 03 05). Regulation (EU) 2025/327 of the European Parliament and of the Council of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847. Retrieved from European Commission: https://health.ec.europa.eu/publications/regulation-eu-2025327-european-health-data-space-and-amending-directive-201124eu-and-regulation-eu_en

EOSC-ENTRUST

Launched in March 2024, EOSC-ENTRUST[34] is a three-year initiative that brings together partners from 15 European countries to advance the development and interoperability of TREs. These environments play a critical role in enabling secure and compliant access to sensitive research data.

At the heart of the project is the development of an Interoperability Blueprint[35], a foundational framework that outlines how sensitive data can be accessed and analysed across a federated network of composable TREs. By addressing both technical and policy-level interoperability, EOSC-ENTRUST supports the broader goals of the European Open Science Cloud (EOSC) and is working under both the EOSC and ELIXIR umbrellas. The project also maintains strong ties with other ELIXIR and European Commission-funded initiatives and is aligned with the EuroHPC infrastructure to explore high-performance computing for secure data analysis.

A key objective of EOSC-ENTRUST is the establishment of a Provider Forum, a community platform for SPE/TRE providers to exchange knowledge, share best practices, and collaborate. This forum is designed to continue beyond the project's lifetime, ensuring sustained support and onboarding of new providers into the ecosystem.

The project is also developing complementary tools and resources, including:

- A set of Training Packages for TRE providers and users,
- A machine-readable TRE Provider Catalogue that maps available secure environments within EOSC, along with their capabilities.

EOSC-ENTRUST's work is driven by diverse use cases, or "Drivers," which include:

- Federated human genomics,
- Social science data sharing,
- Clinical research data interoperability,
- Public-private collaboration involving health and environmental data.

To maximise relevance and impact, EOSC-ENTRUST closely aligns with related European initiatives. In particular, the project collaborates with TEHDAS2 in support of the EHDS, and with the Genomic Data Infrastructure (GDI) for genomics-related use cases. These connections help ensure that EOSC-ENTRUST complements ongoing efforts and contributes meaningfully to a cohesive European data ecosystem.

Throughout the project, EOSC-ENTRUST will also follow the project outcomes of its sister projects SIESTA (Secure Interactive Environments for SensiTive data Analytics) and TITAN (Trusted envIronments for confidenTiAl computiNg and secure data sharing), which are

[34] EOSC-ENTRUST https://eosc-entrust.eu/
[35] Sætrom, P., Lehväslaiho, H., Stansberg, C., Awan, H., & Hesam, A. (2024). EOSC-ENTRUST D13.4 Year one version of EOSC-ENTRUST Blueprint & Interoperability Framework. https://doi.org/10.5281/zenodo.14362388

funded under the same HORIZON 1.3 call on Trusted environments for sensitive data management in EOSC.

HealthyCloud

The HealthyCloud project aimed to lay the foundation for a European Health Research and Innovation Cloud (HRIC) by promoting the secure, ethical, and efficient sharing and reuse of health data across Europe. It brought together a wide range of stakeholders to define a strategic agenda and develop a practical framework to support cross-border health research and innovation, while ensuring compliance with legal and ethical standards.

Within this context, the objective of Deliverable D5.1 (confidential and unpublished) was to identify common patterns in the design and operation of SPEs, focusing on key requirements that could inform the development of reference guidelines for the planned HRIC. The report draws on in-depth assessments of over 13 representative SPEs, supported by interviews with experts familiar with their implementation and operation. These examples were selected from a broader inventory of SPEs across Europe.

The report concentrates on identifying typical models and recurring patterns, along with the associated technical and governance requirements.

The report highlights that the fragmented policy landscape has resulted in a similarly fragmented architectural landscape, leading to the emergence of various SPE designs, each with distinct objectives, including:

- Secure access to controlled data
- Secure collaboration
- Distributed computational approaches
- Data lakes

Among the 13 infrastructures examined in detail, most focus on enabling secure collaboration, while a few focus on secure access to controlled data, and one each provides “compute-to-data” capabilities and data lake functions.

HealthyCloud deliverable D7.4[36], on the other hand, gathered information on security policies and breach response protocols related to the same infrastructures examined in D5.1. This report presents an overview of current practices related to identification of users and access control, data processing, managing and monitoring the environment, as well as organisational policies and procedures. Given that these environments are built and operated in various ways, they each have distinct privacy requirements. Nevertheless, there is a strong consensus on the establishment of SPEs. The key findings are summarised as follows:

- Most infrastructures employ federated authentication and view multi-factor authentication (MFA) as crucial for effective identity and access management.
- Each project should operate within a dedicated environment that is technically and logically isolated from others, grouped by project rather than user.
- Infrastructures managing permits automatically lock user environments once data permits expire and regularly verify the validity of access.

[36] HealthyCloud deliverable D7.4 https://zenodo.org/records/10225422

- For health data processing, pseudonymised data with minimised variables is used, transferred in encrypted form, and must be audited before leaving the infrastructure.
- All infrastructures implement various technical and organisational measures for security management, guided by institutional policies for federated infrastructures.
- Most respondents have internal and external policies for users and staff, with certified sites offering more comprehensive documentation and staff training as a common best practice.

Global Alliance for Genomics and Health (GA4GH)

The Global Alliance for Genomics and Health (GA4GH)[37] develops open, interoperable standards to facilitate secure, ethical, and scalable genomic and health data sharing. These standards provide technical frameworks for authentication, encryption, access control, federated analysis and data governance, which are also key elements for SPEs.

Among these, **Crypt4GH** is a secure encryption standard specifically designed for genomic data. It ensures data remains protected both at rest and in transit through multi-layer encryption, allowing authorised users to decrypt only the portions they are permitted to access. Another standard is **GA4GH Passports**, which streamline authentication and authorisation by providing a federated identity management system. This allows users to securely access SPEs using credentials from their home institutions while enforcing fine-grained access controls.

SPE providers benefit from a user-friendly, secure, interoperable, and privacy-preserving toolkit designed to support the management and processing of sensitive biomedical data.

GDI and 1+MG

The *Genomic Data Infrastructure* (GDI)[38] project is enabling access to genomic and related phenotypic and clinical data across Europe. It is doing this by establishing a federated, sustainable and secure infrastructure to access the data according to the 1+ Million Genomes Initiative[39]

The GDI project places emphasis on federated processing of genomic data which – in case the data in question falls under EHDS – provides a strong driver to solve SPE interoperability and federation in a standardised manner. Additionally, the scale of compute and storage resources needed for genomic data will impact the high-end SPE requirements, for example, driving the high-performance and cloud compute systems used for GDI data to comply with EHDS SPE requirements.

Access to categories of data in focus of the GDI project can be split in three tiers: *open, registered* and *controlled*. These permissions are designed to be machine readable (Ga4GH passports).

[37] GA4GH https://www.ga4gh.org/
[38] GDI https://gdi.onemilliongenomes.eu/
[39] 1+ Million Genomes Initiative https://digital-strategy.ec.europa.eu/en/policies/1-million-genomes

The 1+MG Framework[40] is a series of components based on the output of the 1+MG projects that provide guidance on ELSI, data quality, data standards, and technical infrastructure standards and APIs.

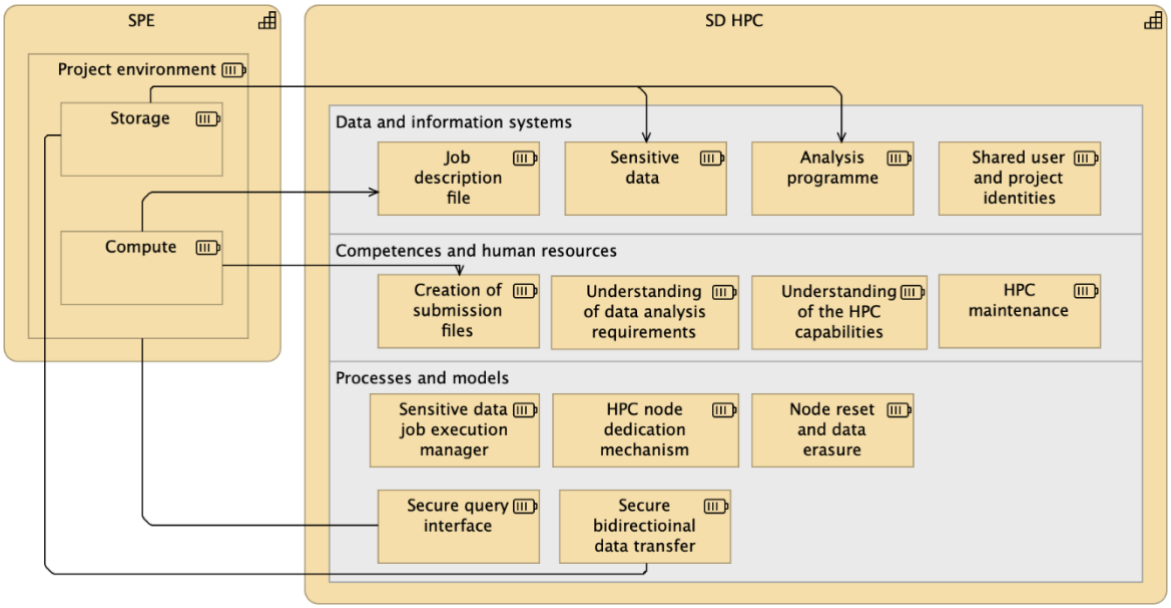
Sensitive Data HPC strategy (CSC, Finland)

SD HPC is the Sensitive Data High Performance Computing solution from CSC – IT Center for Science Ltd., Finland. It is based on the secure transfer of batch job requests from its Secure Processing Environment (SPE) SD Desktop. It contains a batch job management process not visible to other users, secure transfer of user data into a dedicated compute node when the job enters the execution, isolation of the compute node during the execution, and return encrypted results to the SPE user storage space.

The SD HPC service is currently deployed in the CSC Puhti supercomputer. It is in alpha user testing phase.

SD HPC receives job description files from users in the SD Desktop SPE project environment (Annex H: Figure 3). The HPC system must recognise the user and their project. In writing the job description file, the user must consider the capabilities and limitations of the HPC. The dedicated sensitive data execution manager queues and launches the job in a dedicated node. It also copies user resources from SPE storage before isolating the node for the duration of the job execution. It returns the execution results to the user project environment and cleans any trace of the execution event from the node.

Annex H: Figure 3. CSC's sensitive data HPC strategy view



[40] 1+MG framework https://framework.onemilliongenomes.eu/

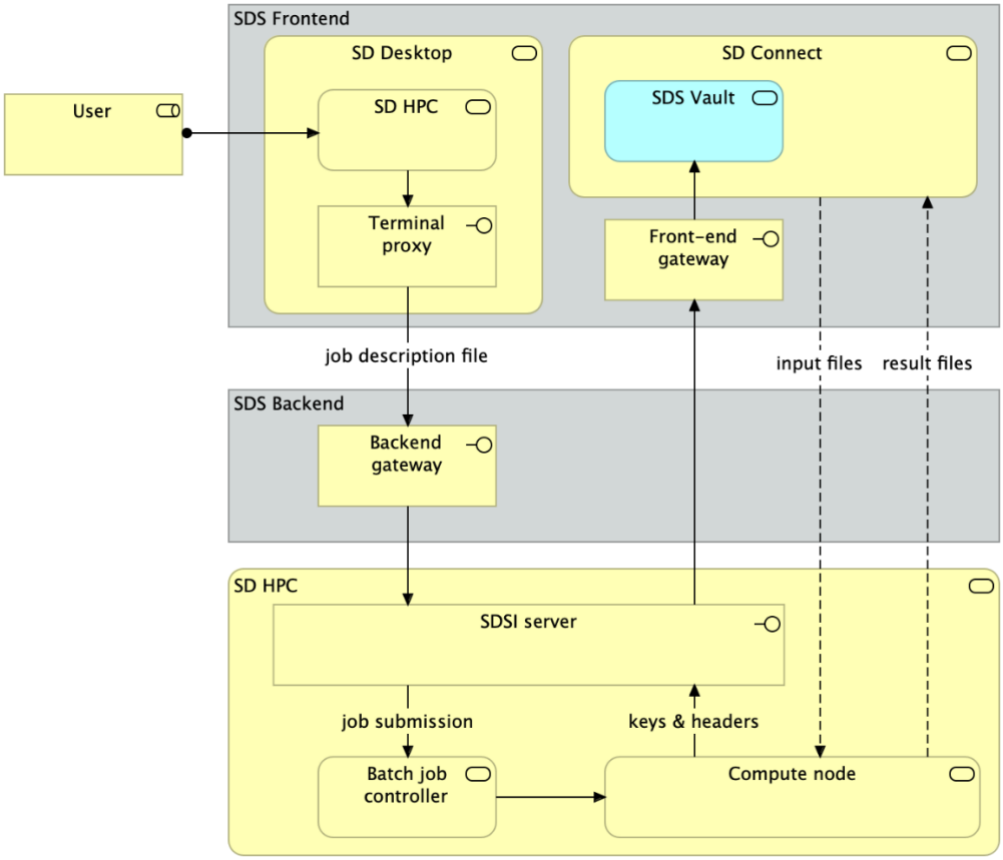
The job request message from the user is passed through the backend gateways to the main HPC gateway called SDSI server (Annex H: Figure 4). The acronym comes from Sensitive Data Slurm Interface indicating the used batch job controller. The SDSI server has complete control of the job management once the job request has been submitted to the HPC system. This isolates the client side cleanly from the HPC execution. Expansion of SD HPC to other HPC systems requires only the adaptation of the SDSI server to their specific capabilities and architecture that needs extensive involvement of their system maintainers.

Inside the SD HPC service, the SDSI server formats the job submission to the requirements of the batch job controller (Annex H: Figure 4). Once the job enters the execution, it fetches the header parts of the user input files from the internal vault component of SD Connect. Before dispatching the headers, the vault re-encrypts the headers with job-specific encryption key according to the SDSI server instructions. These headers are combined with the encrypted payload from SD Connect before passing them into the HPC compute node reserved for the job. The separation of headers and their separate encryption from the payload is an envelope encryption capability of the GA4GH Crypt4GH[41] file container format.

Once all files have been copied to the job node, all access to the node is cut off. Communication to the node is resumed only after the execution is finished, decrypted input files have been removed, and the result files are encrypted with the project key. The SDSI server then copies all result files to a new folder in the SD Connect project space, removes all user files, resets the HPC compute node, and returns it to the pool.

[41] Global Alliance for Genomics and Health (GA4GH) Cryp4GH https://samtools.github.io/hts-specs/crypt4gh.pdf

Annex H: Figure 4. Service diagram showing the user initiating the SD HPC call. Solid arrows represent message directions and dashed lines file transfers.



Using the language of the DARE UK Federated Architecture Blueprint[42], the SD HPC is implemented as an indirect query. It also fulfils other requirements of data privacy and accountability for secure federated analysis: Identity, source and target of all transfers are known, validated, and accounted for, and all data and communication are encrypted in transit. The guaranteed ephemerality of personal data and isolation of processing within the HPC system makes SD HPC a special case of federated analysis with exceptionally strong data privacy guaranties.

NORTRE, infrastructures for sensitive data in Norway

NORTRE[43] (Norwegian Trusted Research Environments) is a collaboration between the three main institutional research infrastructures for sensitive data in Norway, TSD[44] (services for sensitive data) at University of Oslo (UiO), HUNT Cloud[45] at the Norwegian University of

[42]DARE UK Federated Architecture Blueprint https://zenodo.org/records/14192786)
[43] NORTRE https://nortre.no/
[44] Tjenester for Sensitive Data (TSD) https://www.uio.no/tjenester/it/forskning/sensitiv/
[45] HUNT Cloud https://about.hdc.ntnu.no/

Science and Technology (NTNU) and SAFE[46] (secure access to research data and e-infrastructure) at University of Bergen (UiB). The three partners share knowledge and expertise so scientists and data controllers from Norway and around the world can collect, analyse, store, share and collaborate on sensitive data in an optimised and trustworthy manner.

In accordance with EHDS2, Norway has an ongoing project SPUHiN[47] which, among others, aims to prepare NORTRE for EHDS regulation. This includes work such as drafting a requirements list, closely aligned with the requirements of the EHDS regulation, which will be a goal for the three SPEs to achieve. SPUHiN is also creating a GAP analysis/list of points to be followed up with NORTRE to close these gaps.

Secure data transfer solutions (Finland)

In Finland, CSC has developed Supertunneli ("Super Tunnel")[48], an advanced solution designed to address limitations in current data transfer services, particularly for large datasets. Based on the widely adopted S3 interface, Supertunneli significantly improves capacity compared to earlier systems, which were limited to a 4 GB maximum file size per transfer. This enhancement reduces manual workload and minimises the risk of errors typically associated with managing large datasets, particularly for projects utilising AI or machine learning.

Supertunneli enables the secure transfer of large datasets in a single operation and initially supports transfers of at least 1 terabyte (1 TB). Transfers exceeding this threshold are also supported but require prior notification to Findata, Finland's data permit authority. Additionally, the process can be optimised by automating data encryption and decryption.

The connection used in the transfer of the files is encrypted and requires private/public keys exchange to allow the transfer. The transferred data is encrypted with Crypt4GH -encrytion tools provided by Global Alliance for Genomics and Health.

In compliance with Finland's Secondary Use Act, Supertunneli will serve as the standard interface for transferring datasets to approved secure processing environments. This ensures both regulatory alignment and improved security during data exchange.

Further developments include leveraging the S3 interface to support the anonymisation and submission of research results. A complementary system, Tulostunneli ("Results Tunnel"), is under development and will enable secure environments to automate the delivery of published outputs directly to Findata. This replaces the current manual process, traditionally the responsibility of researchers, thereby enhancing security and ensuring consistent anonymisation of disseminated results.

[46] SAFE https://www.uib.no/safe
[47] SPUHiN https://www.helsedirektoratet.no/om-oss/forsoksordninger-og-prosjekter/fair-secure-provision-and-use-of-health-data-in-norway-spuhin
[48] Supertunneli https://findata.fi/en/news/supertunneli-launching-in-may-2025-as-part-of-transfer-service-update/

For smaller or individual data transfers, the existing Tunneli service will remain in operation, providing a flexible, tiered approach to data exchange infrastructure in Finland.

Anonymity verification tool (Finland)

In Finland, Findata is launching a new tool called Portti[49] in Kapseli to simplify and speed up the process of verifying the anonymity of research results. With this update, users no longer need to complete separate summary forms.

Portti—Finnish for *Gateway*—allows users to send results directly for anonymity verification. It enhances data security and streamlines the process for both users and Findata staff.

Portti is strictly for transferring anonymous results. Personal data or non-anonymised information must not be submitted via the tool. To use Portti, users create a new transfer, upload the necessary files, fill in the required details, and select the result type. The results are then submitted to Findata for verification.

Findata processes submissions as quickly as possible, and within a maximum of five working days. Once approved, results are automatically delivered to the user's workspace, where they can be downloaded. Approved results remain available in the workspace for six months.

If anonymity issues are found, Findata will send instructions for corrections via email and may request additional information if needed before final approval.

The tool is being developed as part of the FinHITS project, co-funded by the European Union.

Building a secure health data network (Norway)

In Norway, Norwegian Directorate of Health (NDH), National Institute of Public Health (FHI) and NORTRE[50] are working together in the SPUHiN[51] project to ensure that sensitive data transfers within Norwegian data holders, SPEs and Authorities (Coming HDABs) are secured and guaranteed.

This will be done by building a HealthData@NO network based on the 4-corner model for the eDelivery model[52]. The 4-corner model ensures that data are encrypted, signed and addressed before leaving their safe environments. The 4 corner models addressing system, implemented by among other EU central Servies, guarantees that the data can only be delivered to an approved organisation registered at a central exchange register.

The network aims also to be HealthData@EU compatible by following the eDelivery standard when there are cross border data transfers.

[49] Portti https://findata.fi/en/news/new-tool-for-kapseli-result-submission-will-launch-in-september/
[50] NORTRE
[51] FAIR Secure Provision and use of health data in Norway (SPUHiN) - Helsedirektoratet
[52] How does eDelivery work

Annex I: Methodology

The specifications outlined in this report are the result of an in-depth analysis of the EHDS regulatory text, as well as findings from the PwC–Sopra Steria analysis and design of the European Health Data Space infrastructure for secondary use of health data, specific contract under FRAMEWORK CONTRACT N° DI7925-DI7932 European Commission 2022-2024, existing solutions and project outcomes, and the discussions and feedback received from the Subgroup on SPEs of the Community of Practice. This foundation ensures the specifications are aligned with both regulatory and practical implementation needs.

A key starting point was the thorough examination of the EHDS regulation, especially its Article 73 on secure processing environment (SPE) which underpins Task 7.4 and provides essential guidance for the work. This analysis was further enriched by a detailed review of existing SPE solutions and related project outcomes, ensuring a holistic approach to requirements and capabilities.

Several past and ongoing projects are addressing various aspects of SPEs, providing valuable insights that were analysed to extract relevant requirements and identify reusable approaches. Task partners were requested to suggest relevant projects. The examined projects were primarily selected for their assessment of existing SPEs and their surveys of current infrastructure providers, which eliminated the need for a new survey in this task. Additionally, some projects focus on developing a blueprint for SPEs or are piloting the infrastructure. Including these projects in the evaluation was crucial to ensure alignment among them. These projects are presented in Annex H: Existing solutions for secure processing.

The major contributors of Task 7.4 analysed the collected material and conducted in-depth research on various topics according to their specific areas of expertise and interest. The findings were then presented and discussed with all task partners.

Following the analysis of project results and current requirements for existing SPEs, the next step was to identify gaps and areas for improvement within the current landscape. These results were utilised to establish the minimum functional, operational, security and interoperability requirements for future SPEs.

Enterprise Architecture

The analysis of SPE environment in this report is presented using The ArchiMate® Enterprise Architecture Modeling Language[53] version 3.2 as implemented by open-source Archi[54] application version 5.6.0.

ArchiMate® has not been widely used in biomedical context despite its advantages over other graph generating approaches. ArchiMate® offers a rich selection of conceptual elements that cover from high abstraction level motivation elements through strategy, business, application levels down to technology. The richness of this vocabulary and restricted relations linking

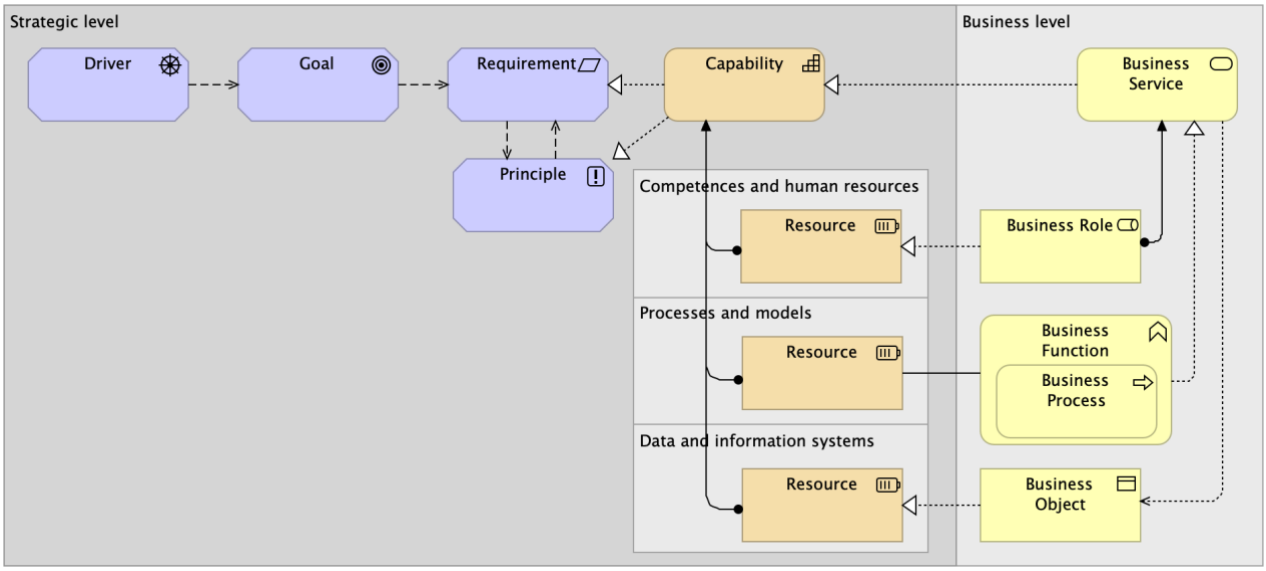
[53] The ArchiMate® Enterprise Architecture Modeling Language https://www.opengroup.org/archimate-forum/archimate-overview
[54] Archi https://www.archimatetool.com/

elements together form an expressive graphical framework enforcing the separation of abstraction levels.

Annex I: Figure 1 presents a simplified metamodel highlighting the business-level elements (in yellow), which represent real-world user-facing functions. These elements are linked to their strategic capacities (orange) and motivational drivers (purple), reflecting how real-world services relate to abstract capabilities and regulatory motivations.

In architecture terminology, this report focuses on determining the strategic level minimum requirements and describing the capabilities needed for SPEs in different implantation scenarios.

Annex I: Figure 1. A simplified ArchiMate® metamodel showing the elements used to show how motivational elements (purple) influence strategic elements (orange), as abstract representations of services and their component roles, processes, and objects (yellow).



Modality of requirements

Requirements are the a priori conditions that capabilities fulfil. The modality of the requirement is given according to the Internet convention IETF BCP 14[55] as summarised in Table I.1.

[55] S. Bradner, B. Leiba; BCP14; The Internet Engineering Taskforce Best Current Practice; https://www.ietf.org/rfc/bcp/bcp14.html

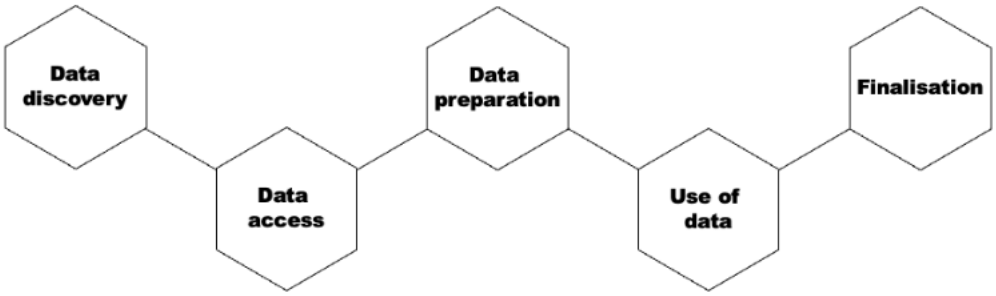
Table I.1. Words used in capital letters to indicate the modality of the requirement.

Adjective	Modal verbs positive	Modal verbs negative
REQUIRED	MUST, SHALL	MUST NOT, SHALL NOT
RECOMMENDED	SHOULD	SHOULD NOT
OPTIONAL	MAY	

Annex J: User journey

When a data user applies for electronic health data for secondary use purposes, such as research and innovation activities, education, and policy-making, within the European Health Data Space (EHDS), the user journey consists of several stages (see Annex J: Figure 1). Access for certain purposes (public or occupational health, policy-making and regulatory activities, and statistics) is reserved for public sector bodies and Union institutions (see Chapter IV, Art. 53(1) and 53(2)).

Annex J: Figure 1. EHDS user journey consists of five main phases: data discovery, data access, data preparation, use of data and finalisation.



Data discovery

Before being able to use the data, the user needs to investigate whether the data needed is available, and whether it is available in the necessary format for the secondary use purpose. This phase is called data discovery. Datasets available in the EU can be found in a metadata catalogue at https://qa.data.health.europa.eu/. Once the data discovery is completed, the user can begin the process of applying for the data.

Data access

In the data access phase, the user fills in and submits a dedicated and standardised data access application form or a data request to a health data access body (HDAB). The user must complete the information required in the form, upload necessary documents, and provide justifications as needed.

Data access application form is used when the user seeks to use personal level data. **Data request** is for cases when the user wants to apply for anonymised statistical data.

Data preparation

During this phase, the data holder(s) deliver(s) the necessary data to the HDAB, which starts to prepare the data for secondary use. Techniques for pseudonymisation, anonymisation,

generalisation, suppression, and randomisation of personal data are employed. The data minimisation principle (as per the GDPR) must be respected to ensure privacy.

Use of data

In this phase, the user performs analyses based on the received data for the purpose defined in the application phase. Analysing personal level data must be performed in a secure processing environment. The duration of this phase is specified in the Regulation (Art 68(12)).

Finalisation

This last phase of the user journey concerns data user's duties regarding analysis outcomes derived from secondary use of data. Data user must publish the results of secondary use of health data within 18 months of the completion of the data processing in a secure processing environment or of receiving the requested health data. The results should be provided in an anonymous format. The data user must inform the health data access body of the results. In addition, the data user must mention in the output that the results have been obtained by using data in the framework of the EHDS.