M7.1 Draft guideline on how to use data in a secure processing environment

TEHDAS2 – Second Joint Action Towards the European Health Data Space

20 January 2025

This project has been co-funded by the 4th EU Health Programme (2021–2027) under Grant Agreement no 101176773.

0 Document info

Disclaimer

Views and opinions expressed in this deliverable represent those of the author(s) only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

0.1 Authors

Lead Author(s)	Lead organisation
Irene Schlünder	TMF e.V., Germany
Dr. Antal Bodi	National Directorate General for Hospitals, Hungary
Victorien Hanché	Health Data Hub, France
Raitis Peculis	Latvian Biomedical Research and Study Centre, Latvia
Léa Rizzuto	Health Data Hub, France
Vita Rovite	Latvian Biomedical Research and Study Centre, Latvia
Inga Selecka	The Centre of Disease Prevention and Control of Latvia (SPKC), Latvia
Lise Skovgaard Svingel	Central Denmark Region, Denmark
Anna Szilagyi	National Directorate General for Hospitals, Hungary
Emmi Turunen	HUS Group, the joint authority for Helsinki and Uusimaa, Finland
Reviewers	
Marianne Benderra	Health Data Hub, France
Pia Brinkmann	BfArM - Federal Institute for Drugs and Medical Devices, Germany
Heikki Lehväslaiho	CSC – IT Center for Science Ltd., Finland
Helena Lodenius	CSC – IT Center for Science Ltd., Finland
Katharina Schneider	BfArM - Federal Institute for Drugs and Medical Devices, Germany

0.2 Keywords

Keywords	TEHDAS2, Joint Action, Health Data, European Health Data Space

0.3 Document history

Date	Version	Editor	Change	Status
01/07/2024	0.1	Irene Schlünder	First draft	Draft
19/12/2024	0.2	Irene Schlünder	Draft to be reviewed by the Consortium	Draft
20/01/2025	1	Irene Schlünder	Document to be submitted for public consultation	Final

Accepted in Project Steering Group by written procedure on 22 January 2024.

Contents

1 Abbreviations

Term	Abbreviation
D	Deliverable
Data Governance Act	DGA
Data Protection Officer	DPO
Directorate-General	DG
European Health Data Space	EHDS
European Union	EU
General Data Protection Regulation	GDPR
Health Data Access Body	HDAB
Joint Action	JA
Multifactor authentication	MFA
Random-access memory	RAM
Secure Processing Environment	SPE
The Finnish Innovation Fund	Sitra
Towards the European Health Data Space	TEHDAS
Work Package	WP

2 Executive summary

The aim of this guideline is to provide support to those who plan to access personal electronic health data for secondary use purposes through the infrastructure set up by the European Health Data Space (EHDS) Regulation, the “HealthData@EU” infrastructure. The guideline is designed to support data users, specifically reflecting their activities from the moment they gain access to the approved datasets within a Secure Processing Environment (SPE) until the completion of their analysis and the export of results.

The guideline takes the perspective of the data user and is intended to be consulted already from the project planning phase, i.e., before a data permit has been granted. This is advisable since certain fees apply already when the data user, as data applicant, submits a data access application to an HDAB; thus, the data applicant will benefit from assessing the feasibility of conducting their analyses in an SPE provided via the HealthData@EU infrastructure at an early stage. In addition, the data applicant may state in the data application which SPE appears to be suitable for analysing the requested data.

The proposal of a certain SPE must be justified in the application and will be part of the data permit.

It should be noted that it is very challenging to draw up a guideline for data users as long as not only the EHDS Regulation is still in the process of being adopted, but also the recommendations on the implementation of the Regulation by the HDABs have not yet been finalised. After all, it is the HDABs that will implement the regulation in the first instance and thus make the first decisions on how the Regulation is to be interpreted. Therefore, this guideline will be adapted immediately after completion of other guidelines and specifications.

3 Introduction

3.1 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 cross-border use of health data, and support data holders, 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 work performed in Work package 7 (WP7) addresses "Safe and secure processing" of electronic health data within the EHDS infrastructure. The goal is to enable secure processing of EU citizen's electronic health data while fostering a secure, interoperable, and efficient health data ecosystem. The output of this work package consists of guidelines and technical specifications that shall inform further decisions and technical frameworks to set up the EHDS.

The results of WP7 are distributed across five tasks. Task 7.1 provides guidance to users about their duties and responsibilities when analysing data in a secure processing environment. Next, technical specifications for data minimisation and de-identification give guidance on how to address the challenges of health data minimisation,

pseudonymisation, anonymisation and the generation of synthetic data (Task 7.2 includes Sub-tasks: 7.2.1, 7.2.2, 7.2.3 & 7.2.4). Specifications for the implementation of a common IT infrastructure (Task 7.3) shall help member states to connect to the EHDS ecosystem. To ensure interoperability, common security requirements applicable to all secure processing environments are defined in addition to functional and technical services that should be part of all secure processing environments (Task 7.4). Lastly, information about data linkage techniques and possibilities of quality control of linked data are collected (Task 7.5).

Here is an overview of the documents that are part of WP7:

- Guidelines for data users on how to use data in a secure processing environment (Task 7.1).
- Technical specifications for Health Data Access Bodies on data minimisation and de-identification (Task 7.2).
- Technical specifications for Health Data Access Bodies on the implementation of the common IT infrastructure (Task 7.3).
- Technical specifications for Health Data Access Bodies on the implementation of secure processing environments (Task 7.4).
- Guidelines for Health Data Access Bodies on linkage of health datasets (Task 7.5).

3.2 Introduction to the guideline on how to use data in a secure processing environment

This guideline belongs to a set of guidelines and technical specifications supporting the implementation the European Health Data Space (EHDS) as provided in the Regulation.

Other TEHDAS2 guidelines and technical specifications that are relevant for data users:

- Deliverable 4.1 Guideline for Health Data Access Bodies on fees and penalties for non-compliance related to the EHDS regulation.
- Deliverable 4.3 Guideline for Health Data Access Bodies on international and third country access and transfer of electronic health data.
- Deliverable 5.4 Guideline for Health Data Access Bodies on enrichment of health datasets.
- Deliverable 6.2 Guideline for Health Data Users on good application practice for data access and data requests.
- Deliverable 7.2 Technical specification for Health Data Access Bodies on data minimisation and de-identification.
- Deliverable 7.4 Technical specification for Health Data Access Bodies on the implementation of secure processing environments.
- Deliverable 8.2 Guideline for Health Data Access Bodies on data altruism in health
- Deliverable 8.4 Guideline for data users on handling research Outcomes.

3.3 Target audience

This guideline is written for data users before they submit an application for data access through the EHDS infrastructure or after they have received a data access permit from a competent HDAB. (For the application process, see D6.2 Guideline for Health Data Users on good application practice for data access and data requests).

3.4 Scope

The aim of this guideline is to provide support to those who plan to access personal electronic health data for secondary use purposes through the infrastructure set up by the European Health Data Space (EHDS) Regulation[1], the "HealthData@EU" infrastructure. It is designed to support data users engaging with the EHDS2 framework, specifically focusing on their activities from the moment they gain access to the selected datasets within a Secure Processing Environment (SPE) until the completion of their analysis and the export of results. Thus, the guideline covers the following steps:

1. What to take into consideration regarding the use of an SPE when planning to access data through the EHDS infrastructure, even before submitting a data application, including:
    - What is an SPE and why and when is an SPE needed?
    - Can the data user choose an SPE?
    - What will be the cost of using an SPE.

2. How to access the approved data in the SPE, once the data user has received the data permit following their application regarding a certain set of personal data, including:
    - How to communicate with the SPE manager/provider?
    - How to get access to the data in the SPE?

3. What needs to be considered when analysing data in an SPE, including:
    - What rules to follow?
    - Who is accountable as data controller?
    - What happens in case of rule violation?

4. What happens after completion of the data analysis and export of results, including:
    - Export of authorised anonymised results in a statistical format.
    - Archiving of the data used to ensure reproducibility of results.
    - Deletion of the data in the SPE.
    - Transparency obligations.

3.5 Legal framework

The main legal act for accessing data through the EHDS infrastructure is the EHDS Regulation (...). But processing personal data for secondary use under the EHDS also falls into the scope of the General Data Protection Regulation (GDPR, Regulation (EU) 2016/679). Its rules remain applicable. Therefore, data users should always involve local data protection officer when processing personal health data even when the users are acting within an SPE.

The EHDS regulation, which forms the decisive legal framework for this guideline, is only in the process to be adopted. It is therefore still quite open as to what different

implementations there will be in detail. Thus, it is not yet possible to make a detailed recommendation on all points. Any gaps will have to be closed over time.

This guideline is based on the English language version. In fact, however, all language versions are equally binding.

Data users are required to carefully review the data permit issued by the HDAB to understand the specific terms and conditions for using the data in an SPE. The permit will adhere to a template established by an implementing act,[2] making it non-negotiable and unchangeable. This guideline offers a high-level overview but should always be read alongside the data access permit and relevant agreements to ensure compliance with the EHDS Regulation and GDPR.

It should be noted that the EHDS exists alongside other mechanisms for getting access to data for research and innovation purposes. It does not replace traditional data sharing mechanisms or existing data sharing agreements. Users can continue using them. If they do so, their rules apply. They may differ from those in the EHDS (see also D6.2 Guideline for Health Data Users on good application practice for data access and data requests).

Researchers and innovators from third countries can also access data for secondary use through the EHDS, but some additional requirements might be applied (see D4.3 Guideline for Health Data Access Bodies on international and third country access and transfer of electronic health data)?

4 What is an SPE and why and when do data users need an SPE?

An SPE is a highly protected digital workspace where authorised users can analyse health data for purposes defined in the EHDS Regulation. It is designed to keep the data safe and ensure privacy by allowing access only to approved users and tools and restricting what can be taken out of the environment, allowing only non-personal electronic health data, including electronic health data in an anonymised statistical format, to be exported from the secure processing environmentonly aggregated anonymised results. The core aim of an SPE is that the personal data itself cannot be downloaded by the data user but is analysed within the SPE environment. This guarantees that the use of the data is subject to tight control, which ensures that the data cannot be passed on for unauthorised purposes or linked to other data beyond those specified in the underlying data permit. In addition, the data in the SPE is deleted or temporarily archived (see below) after the specified utilisation period.

In general, the EHDS regulation foresees health data provision within an SPE, whereas only anonymous data, including data in a statistical format, may be provided outside SPEs, as a result of a user's analysis or a successful data access request. Please consult the technical specification for Health Data Access Bodies on the implementation of secure processing environments to find out more about what data are anonymous and in which format data will be delivered in an SPE.

[2] see Art. 74 (3) EHDS Regulation.

Thus, SPEs are critical to enabling lawful secondary use of personal electronic health data while safeguarding individual privacy and ensuring compliance with the EHDS standards. Oversight by the HDAB is essential for an SPE to be used for data processing under the EHDS Regulation. Therefore, the HDABs are responsible to ensure that regular audits of the SPEs are conducted, including those conducted by third parties and to take corrective actions for any shortcomings, risks or vulnerabilities identified.

The term SPE is defined in Article 2 (20) of the Data Governance Act (DGA)[3] to which the EHDS Regulation refers in Art. 2 (1) (c). In addition, the requirements for the secondary use of health data are specified in the provisions of the EHDS Regulation; in particular Art. 73: SPE refers to a technical solution for making health data available for statutory purposes in a way that fulfils high data security requirements. Thus, an SPE must meet the following criteria:

- 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 non-personal electronic health data, including electronic health data in an anonymised statistical format can be exported, subject to approval by the Health Data Access Body (HDAB).

Services are not directly defined by law. However, the European Commission will provide further details for the technical, organisational, information security, confidentiality, data protection and interoperability requirements for the SPEs, including the technical characteristics and tools available to the health data user within the SPE (see Technical specification for Health Data Access Bodies on the implementation of secure processing environments).

5 How to suggest the appropriate SPE?

Data applicants are advised to contact the HDAB to which the application for data access is to be submitted at an early stage to consult regarding the general availability of SPE(s). In the process of filling in the data access application, the applicant must specify the requirements for the SPE for it to meet the expected computational needs (e.g., RAM, processors, storage, software, and tools) (Article 67 (2) of the EHDS Regulation). In the data application, applicants may also state any potential request for provision of a specific SPE, providing any additional reasons, e.g., financial, for this request (see also D6.2 Guideline for Health Data Users on good application practice for data access and data requests, subchapters 3.5, 4.4.8, and 4.4.9). This information will help the HDAB identify the SPE best fit for purpose.

[3] Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act) https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R0868

If the users want to change the SPE they need to contact the HDAB again. Any changes to the SPE such as change of the SPE itself, modifications of user access, data availability or changes to study group members must be approved by the HDAB through an amendment of the data permit on request of the user (Art. 68 (13) of the EHDS Regulation).

6 What are the fees to use the SPE?

Under Article 62 of the EHDS Regulation, HDABs are permitted to charge fees for making data available for secondary use. These fees are intended to cover the costs of services, including SPE provision, provided by the HDAB and data holders. Fees must be transparent, proportionate and non-discriminatory, with potential reductions available for specific user groups, as outlined in D4.1 Guideline for Health Data Access Bodies on fees and penalties for non-compliance related to the EHDS Regulation.

HDABs may charge fees for the following services:

- Evaluating the data access application: Includes reviewing and assessing the data permit application or data request.
- Preparing the dataset: Covers costs such as pseudonymisation, anonymisation, data linkage and consolidating datasets from multiple data holders.
- Using the SPE: Includes access to the SPE, user training, statistical software licenses (if required), processing capacities, data storage/archiving and ongoing support for data use.

The expected fees must be communicated to the data user before the permit is issued, allowing the data user to decide whether to withdraw or proceed with the application. All applicable fees, corresponding to the services provided, must be paid at the time the specific service is delivered, as detailed in the data permit.

If a data user requests changes after the permit has been issued (e.g. extending the permit duration or modifying the dataset), additional fees may apply to cover the costs of these adjustments.

For detailed information on fee structures and conditions, refer to D4.1 Guideline for Health Data Access Bodies on fees and penalties for non-compliance related to the EHDS Regulation.

7 Communication with SPE provider

After access to the data has been granted by the HDAB, the HDAB will provide the users with credentials to identify them as being authorised to access the data within the SPE referred to in the data permit. All communication goes via the HDAB, since the SPE is managed through the HDAB, if not stated otherwise by the HDAB.

8 How to get access to the SPE

The access is subject to the permit to use data issued by the HDAB. Study group members are named individually.

Access to the SPE is a strictly regulated process that ensures and guarantees the protection of health data as a special category of data. Access is managed and controlled at several levels as follows:

First, the individuals who should have access to the data and are therefore named in the data permit must enrol before getting access. The enrolment phase consists of thorough identification measures. Cybersecurity measures follow Zero Trust Policy: Trust no one, identify everything.

Access to an SPE needs to be based on personal login credentials (login ID) and strong password. Strong password should only be changed in the case of suspected leakage or other reason (e.g. forgotten). The SPE management can require a new password to be set by the user if there is evidence of compromise of authentication. The password should be unique, software specific (succession for user, cannot be measured). Multifactor authentication (MFA) will be implemented for the login to an SPE. Only strict MFA is acceptable: this may include something the users know (password), something they have (token or smart card) and something they are (biometric authentication such as fingerprint or facial recognition). A unique device based on possession may be more useful than biometrics. If the ID is misused, the tokens can be replaced. Biometric credentials that have been compromised cannot easily be reissued.

Persons who have access to data in the SPE must log in at any time they access the SPE. Continuous monitoring will then take place. Changes in authorised staff should be reported to the HDAB and in all cases re-approved in accordance with the documentation. Unauthorised persons or persons with unverifiable identities will not have access to the SPE.

9 How to analyse data within the SPE

SPEs are designed with robust security features to assist data users in complying with legal requirements under the EHDS Regulation. These measures include multi-layered cybersecurity protocols to protect sensitive data and prevent breaches. Users should never attempt to bypass or disable SPE features, as they are critical for maintaining compliance and safeguarding data. While different SPEs may offer varying levels of functionality and user-friendliness, all EHDS-compliant SPEs adhere to strict standards for security and accountability.

Data analysis:

- Selecting analytical tools – Using only pre-approved tools provided within the SPE. If custom scripts are necessary, they are submitted for a pre-approval by the HDAB.

- Executing analytical tasks - Conducting analyses as defined in the data permit. All analyses are being subject to review by the HDAB for compliance and security.
- Generating intermediate outputs - Storing all intermediate results within the SPE. Outputs remain visible to the HDAB for oversight. Outputs with identifiable elements cannot be exported.

Result preparation and export:

- Validating outputs to ensure their accuracy and compliance with the terms of the data permit. Results must not enable the re-identification of individuals.
- Applying EHDS-compliant anonymisation or aggregation methods to prepare export-ready files.
- Downloading results via the SPE's secure export mechanism. Data users are responsible for ensuring secure handling of data after export.

Art. 74 (1) of the EHDS Regulation stipulates, that the data user analysing the data within an SPE is acting as a data controller (see also chapter 9). This role requires full accountability and responsibility to ensure compliance with GDPR and the EHDS framework. The following measures are recommended to support users in fulfilling their obligations:

Accountability: Granting access to the SPE is done by qualified and trustworthy staff who have undergone specific training on handling of health data as a special category of data and data protection. This ensures that the persons handling the data are aware of the legal, ethical and security requirements under GDPR and EHDS.
Control: Defining clear roles for all personnel accessing the SPE, such as data analysts or researchers, and ensuring that access is limited to what is strictly necessary for their tasks. These roles should be assigned when applying for data access.

Users need to ensure proper conduct of personnel while analysing data within the SPE. This "proper conduct" includes but is not limited to the use of personal mobile devices while working in an SPE or around persons analysing data in an SPE (no video calls, it is forbidden to record computer screens in photos and videos) and taking notes about data, especially specific individuals.

Additional challenges might arise when a team of multiple researchers plans to collaborate in an SPE on the same project. The team organisation should be sorted out: clear contracts with responsibility areas should be negotiated, implementation of precise team hierarchy and roles help to compartmentalise sensitive information and knowledge and determine data access level for each team member. SPE may limit the use of popular collaboration tools like Slack or Microsoft Teams due to security concerns. Real-time collaboration features, such as simultaneous editing of documents, workflows and results, may be restricted or have limitations.

The local data protection officer (DPO) should be involved when preparing to access the SPE to ensure that all measures have been taken.

In addition, all restrictions in the data access permit issued by the HDAB must be observed. These restrictions include, for example, limiting access to individuals explicitly

listed in the data permit, prohibiting any attempts to re-identify individuals from any provided data, confining analysis to the purposes specified in the permit, etc.

Users must ensure that these restrictions are observed at all times. The SPE's security features, such as access controls and logging, are designed to enforce compliance and prevent unauthorised activities. Any breach of these restrictions may result in penalties under the EHDS Regulation and GDPR.

10 Who is accountable as data controller?

The EHDS Regulation imports the definition of the term "controller" from the GDPR. A data controller is a person or organisation that determines the use of personal data. The role of the data controller can be shared by many people or organisations. The controller is accountable and responsible for establishing a lawful data processing workflow and observing the rights of data subjects.

The data processor should handle data exclusively in the manner prescribed by the controller.

The EHDS Regulation assigns the role of several actors in the EHDS with regard to certain processes. The main Articles in this respect are Art. 74 and 75 of the EHDS Regulation.

Table 1: Summary of controllership for a simple exemplary scenario (*data holder/HDAB/data user without intervention of a trusted data holder and data intermediation entity*)

Processing activity	**Data controller**
Data preparation (data targeting, quality control, additional pseudonymisation, etc.)	Data holder
Data matching	HDAB
Transfer of data to the SPE	Data holder or HDAB in case of matching and/or compilation of data from several sources
Ingestion and additional pseudonymisation	HDAB
Validation of data on the SPE (assessing if the data requested are correct, complete and fit for purposes in relation to content of the data access permit)	Data user
Pseudonymisation between workspaces, if foreseen by national regulation	HDAB

Import of data in the project's dedicated workspace on the SPE	HDAB
Data analysis	Data user
Export of anonymised data from the technological platform from the SPE to the data user, after verification of the degree of aggregation and anonymisation by HDAB	Data user
Data archiving at the request of the data user	Data user
Security measures on the SPE with regards to data processing including back-up, maintenance and traceability	HDAB
Deletion	HDAB

Where two or more national contact points or authorised participants put electronic health data in the secure processing environment managed by the Commission, they shall be joint controllers, and the Commission shall be processor for the purpose of processing data in that environment (Art. 75 (9) and (10) of the EHDS Regulation).

Where the EHDS Regulation does not provide for specific rules regarding controllership, the general rules from the GDPR apply.

If the applicant/data user consists of more than one individual or entities (e.g. a research consortium), all members of the group could be joint controllers.

11 What happens in case of rule violation?

While using SPEs, the data user's activities are monitored and logged, and data users must strictly adhere to the obligations and conditions outlined in the EHDS Regulation and their data permit. These rules are essential to ensure that secondary use of health data does not harm individuals, groups or society.

Prohibited actions include (Articles 54 and 61 of the EHDS Regulation):

- Using data to take decisions that have detrimental effects on individuals (e.g. related to insurance, employment or banking).
- Engaging in advertising or marketing activities.
- Developing harmful products or services.
- Conducting activities in conflict with ethical provisions laid down in national law.
- Attempting to re-identify individuals or groups from pseudonymised data.
- Providing access to data or the SPE to third parties not listed in the data permit.

Data users must report any breaches or security incidents to the HDAB immediately. Failure to do so may result in additional penalties.

Non-compliance with the rules may result in penalties under the EHDS Regulation. These penalties could include:

- Fines: Based on the severity and nature of the violation.
- Suspensions: Revocation of the data permit or exclusion from the EHDS for secondary use.
- Damage compensation: Liability for damages caused to natural persons or SPE providers.
- Legal actions: Depending on the severity of the breach, further legal or criminal actions may be pursued in accordance with national law.

If a finding of non-compliance indicates a possible breach of the GDPR, the HDAB is required to immediately inform the supervisory authorities and provide them with all relevant information regarding this finding.

Detailed information about enforcement measures is outlined in D4.1 Guideline for Health Data Access Bodies on fees and penalties for non-compliance related to the EHDS Regulation.

The HDAB will give the health data user an opportunity to state their views on suspected breaches within a reasonable period that does not exceed four weeks (Article 63 of the EHDS Regulation and D4.1 Guideline for Health Data Access Bodies on fees and penalties for non-compliance related to the EHDS Regulation).

12 What happens after finishing data analysis?

Export of data is limited to non-personal electronic health data, including electronic health data in an anonymised statistical format. The HDAB will check the data for re-identification risks before they leave the SPE.

There might be situations where the data user has completed the necessary analyses and only needs to access the SPE occasionally, e.g. for checking reproducibility and during review processes. In such situations, while storage is still required, the need for computing capacities could be reduced and/or be less time critical. The SPE might provide for an archiving function to enable the review and reproducibility of research outcomes at reduced capacities. This in turn might reduce the costs. It should be noted that this ‘archiving’ function is intended for part of the duration of the data permit and will not extend the overall storage period.

According to Article 68 (12) of the EHDS Regulation, a data permit shall be issued for the duration necessary to fulfil the requested purposes, which shall not exceed 10 years. This duration may be extended once, at the request of the data user, based on arguments and documents justifying this extension submitted one month before the expiry of the data permit. The extension period cannot exceed 10 years.

The electronic health data within the SPE shall be deleted within six months following the expiry of the data permit. Upon request of the data user, the formula on the creation of the requested dataset may be stored by the HDAB.

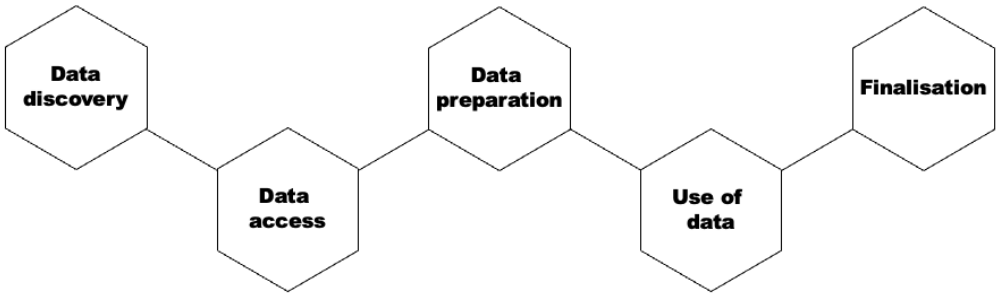
HDABs will have to ensure transparency by publishing information about data access applications, data requests and permits granted (D8.3 Guideline for Health Data Access Bodies on informing natural persons about the use of health data - “Citizen Information Point”). In addition, data users must make public the results of their electronic health data use and inform the HDABs of any significant findings relevant for the health of individuals (D8.4 Guideline for data users on handling research outcomes).

Annex 1: EHDS user journey description

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 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)).

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)ii. 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)iii 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.

Annex 2: Glossary

Table 2: Preliminary glossary according to TEHDAS2. It will be aligned across all TEHDAS2 deliverables in a next step. Please note, that the current version of the glossary is not exhaustive.

Term	Description
Dataset	Means a structured collection of electronic health data. (EHDS, Article 2(2)(w))
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 organisational measures to ensure non-attribution to an identified or identifiable person. (GDPR Article 4(5))
Anonymisation	The process by which personal data is irreversibly altered in such a way that a data subject can no longer be identified directly or indirectly, either by the data controller alone or in collaboration with any other party. Anonymised data falls outside the scope of data protection laws such as GDPR. (GDPR Recital 26)
Secure Processing Environment (SPE)	'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(20); EHDS, Article 2 (1)(c))
Synthetic Data	Data that is artificially generated rather than obtained by direct measurement. Synthetic data can be created using statistical models, machine learning algorithms, or other generative processes to reflect the characteristics of real data while ensuring that no real individual can be identified from the data.
Data Linkage	The process of combining data from different sources that relate to the same entity (e.g., individual, institution) to create a more comprehensive dataset. This can be done using unique identifiers, probabilistic methods, or a combination of techniques.

Re-identification Risk	The potential that anonymised or pseudonymised data could be matched with other data sources to re-identify an individual. Mitigation strategies include robust anonymisation techniques and regular risk assessments. Reference: GDPR Article 6.
Data Minimisation	Principle that mandates that only the minimum necessary amount of personal data should be collected and processed for a specific purpose. This principle is fundamental under GDPR and relevant to the tasks outlined in EHDS. (GDPR Article 5(1)(c))
Data Provenance	The history and origins of a dataset, including the methods and transformations applied to the data throughout its lifecycle. Understanding data provenance is crucial for ensuring data quality and integrity. (Relevant to GDPR's accountability principle Article 5(2))
Actors (Roles)	
Health data access body (HDAB)	[…] providing access to health data through the involvement of health data access bodies, […] In addition, the health data access body should assess the information provided by the health data applicant, based on which it should be able to issue a data permit for the processing of personal electronic health data pursuant to this Regulation that should fulfil the requirements and conditions set out in Chapter IV of this Regulation. […] (EHDS, Recital 52)
Health Data holder	'health data holder' means any natural or legal person, public authority, agency or other body in the healthcare or the care sectors, including reimbursement services where necessary, as well as any natural or legal person developing products or services intended for the health, healthcare or care sectors, developing or manufacturing wellness applications, performing research in relation to the healthcare or care sectors or acting as a mortality registry, as well as any Union institution, body, office or agency, that has either: i. the right or obligation, in accordance with ▌ applicable Union or national law and in its capacity as a controller or joint controller, to process personal electronic health data for the provision of healthcare or care or for the purposes of public health, reimbursement, research,

	innovation, policy making, official statistics or patient safety or for regulatory purposes; or ii. the ability to make available non-personal electronic health data through the control of the technical design of a product and related services, including by registering, providing, restricting access to or exchanging such data; (EHDS, Article 2(2)(t))
Health Data user	'health data user' means 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(2)(u))
Data life cycle	
Data access	Processing by a data user of data that has been provided by a data holder, in accordance with specific technical, legal, or organisational requirements, without necessarily implying the transmission or downloading of such data. (DGA, Article 2(8),(9)&(13))
Data permit	'data permit' means 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 this Regulation; (EHDS, Article 2(2)(v))