M4.1 Draft guideline on fees and penalties for non-compliance related to the EHDS regulation

First section: 4.1.1: Draft guideline on fees related to the EHDS regulation

TEHDAS2 – Second Joint Action Towards the European Health Data Space

17 September 2025

Co-funded by the European Union

M4.1.1 Draft guideline on fees related to the EHDS regulation

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. Document info

Authors

Author(s)	Organisation
Julie Baussand	Health Data Hub, France (Lead)
Inge Franki	Health Data Agency, Belgium
Mervi Siltanen	Finnish Social and Health Data Permit Authority Findata
Ida Møller Solheim	Norwegian Institute of Public Health, Norway
Emer Doyle	Department of Health, Ireland
Minerva Alvarez	Ministry of Health, Spain
Louisa Stuwe	Digital Health Delegation, French Ministry of Health
Katrine Højen Vad	Danish Health Data Authority, Denmark
Julie Richard	Digital Health Delegation, French Ministry of Health
Emilie Passemard	Digital Health Delegation, French Ministry of Health
Zdenek Gütter	Ministry of Health of the Czech Rep.
Emilia Krawczyk	Ministry of Health, Poland

Keywords

Keywords	TEHDAS2, Joint Action, Health Data, European Health Data Space, Health Data Access Bodies, Health Data Holders, Fees, Compensation, Invoice

Document history

Date	Version	Editor	Change	Status
06/03/2025	0.1	Julie Baussand, HDH	Initial document creation	Draft
22/05/2025	0.2	Julie Baussand, HDH	First draft	Modified
13/06/2025	0.3	Julie Baussand, HDH	Second version	Modified
25/06/2025	0.4	Julie Baussand, HDH	Third version	Modified
30/06/2025	0.5	Julie Baussand, HDH	Finalized and submitted to WP4 lead	Final version
30/06/2025	0.6	Louisa Stuwe, DNS	Finalized and submitted to TEHDAS2 consortium and EC	Final version
30/07/2025	0.7	Julie Baussand, HDH	Addressing comments throughout the document from DG SANTE	Modified
21/08/2025	0.8	Julie Baussand, HDH	Addressing additional comments from TEHDAS2 Community	Modified
28/08/2025	0.8.1	Julie Baussand, HDH	Addressing additional comments from TEHDAS2 Community	Modified
05/09/2025	0.8.2	Julie Baussand, HDH	Addressing additional comments from TEHDAS2 Community	Modified

Accepted in Project Steering Group on 2025 September 11th

M4.1.1 Draft guideline on fees related to the EHDS regulation

Table of contents

Abbreviations

Term	Abbreviation
European Health Data Space	EHDS
European Union	EU
Joint Action	JA
Health Data Access Body	HDAB
Trusted Health Data Holder	TDH
Health Data Holder	DH
Health Data user	DU
Secure Processing Environment	SPE
Central Processing Unit	CPU
Graphics Processing Unit	GPU
Finnish Innovation Fund	Sitra
Towards the European Health Data Space	TEHDAS
Work Package	WP

1. Executive summary

This document provides recommendations on the fee structure for accessing health data for secondary use, in accordance with Article 62 of the European Health Data Space (EHDS) regulation. The EHDS establishes a common framework for the secondary use of electronic health data across the EU. The aim of this work is to define and clarify the principles governing fees and invoicing within this framework.

Three key questions are addressed, which are intended to be approached in a coherent and principles-based manner across Member States:

- Which costs can be included in the fees (and which are excluded),
- To whom the fees should be paid, and
- When the fees should be paid by the data user.

Regarding the first question, the document defines eligible cost categories to be compensated for as well as excluded costs. Eligible costs are described based on activities realised at the different steps of the EHDS procedures (data access application, data request). It distinguishes between marginal costs and fixed costs depending on data holders' strategy to anticipate or not data access requests.

Regarding fee recipients and payment procedure, several scenarios are described and motivated recommendations are proposed. The document emphasizes the importance of allowing flexibility to Member States in implementation modalities to ensure that the proposed models for fee recipients and payment procedure can be applied efficiently without disrupting existing mechanisms or contractual arrangements related to fees and invoicing.

The recommendations aim to promote transparency, fairness, and operational efficiency in line with the principles of non-discrimination, proportionality, and competition neutrality set out in the EHDS Regulation

This guidance is non-binding and may be updated as further implementing acts and delegated acts are adopted by the European Commission or guidelines by the EHDS Board.

2. Introduction

2.1 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.

TEHDAS2 Work Package (WP) 4 "Collaboration Models", coordinated between the Danish Health Data Authority and the Digital Health Delegation of the French Ministry of Health, aims at defining guidelines for the organizational implementation of the EHDS for Health Data Access Bodies in the EU. More specifically Task 4.1 is divided into two sub-tasks: 4.1.1. on fees led by the French Health Data Hub, and 4.1.2 on penalties, led by the Italian Ministry of Health.

2.2 Purpose of the guideline

The purpose of the guideline is to provide guidance to Health Data Access Bodies (HDABs), Trusted Health Data Holders (TDHs) and Health Data Holders (DHs) on policies related to health data access fees. The guidelines concern financial practices related to data access application, data request and simplified procedures with TDHs. Please note that any reference to DH shall be understood to include TDH, which are DHs assigned with specific missions only in the context of simplified procedures.

Data access application (Article 67) refers to an application seeking to access individual-level electronic health data for secondary use in an anonymised or a pseudonymised format for the purposes referred to in Article 53 of the EHDS Regulation.

Data request (Article 69) refers to the procedure where the applicant may submit a health data request for the purposes referred to in Article 53 of the EHDS Regulation with the aim of obtaining an answer only in anonymised statistical format.

The simplified procedure with TDH (Article 72) applies to both data access application and data request that concern only electronic health data held by the TDH. In such cases, the TDH may assume certain responsibilities normally carried out by the HDAB.

2.3 Problem being addressed

Sharing health data is a major priority for European health policies. One of the objectives of the EHDS is to create a common secure framework at European level to facilitate and supervise the reuse of health data for research, innovation and public policy purposes. The problem addressed by the guideline is the need to support coherent and transparent practices across the EU regarding fee practises for health data access. The objective is to promote consistent financial practices across Member States in line with the principles of transparency, non-discrimination and competition neutrality set out in the EHDS regulation.

2.4 Targeted audience

This guideline is written for HDABs, TDHs and DHs to help to align practices at the EU level regarding fees that can be claimed from data users (DUs) by contributors in the health data access procedure (HDAB, DH, TDH).

3. Proposal for fee guidelines under the EHDS

In the context of the subtask 4.1.1 on fees, a comparative analysis of existing structures and pricing policies has been conducted and workshops organized to provide an overview of current approaches and experiences across Europe. The guidelines provided here state recommendations on which activities and costs may be considered eligible for fee calculation, and how financial flows between stakeholders may be structured under the EHDS frameworks. The work reported in this document aims to characterize the fees and invoicing

that go with this framework while recognising that existing national mechanisms or contractual agreements may remain in place, provided they are consistent with the principles set out in the EHDS Regulation.

3.1 Scope of the work

This guideline is part of a series developed under TEHDAS2 to operationalise the EHDS regulation, specifically addressing Chapter IV on secondary use of health data. Article 62 addresses the fee aspect of data access. HDABs, including the Union health data access service, as well as TDHs, may charge fees for making electronic health data available for secondary use (Article 62.1). Fees may include compensation for the costs incurred by the DH for compiling and preparing the electronic health data to be made available for secondary use (Article 62.2).

The EHDS regulation identifies several steps in the DU journey that can be categorised as presented in Annex 1:

- **Data discovery**: the DU needs to investigate whether the data needed is available, and whether it is available in the necessary format for the secondary use purpose.
- **Data access**: the DU fills in and submits a data access application form or a data request to a HDAB.
- **Data preparation**: the DH delivers the necessary data to the HDAB, which starts to prepare the data for secondary use in accordance with the conditions set in the data permit or data request approval. In the case of a simplified procedure, the data are prepared and delivered by the TDH.
- **Use of data**: the DU performs analysis based on the received data for the purpose defined in the application phase.
- **Finalisation**: the DU performs the analysis and publishes the results.

The scope of this guideline begins with the data discovery phase, once the applicant has identified the datasets of interest in the dataset catalogue and finishes at the end of the use of data. Recommendations proposed here focus on compensation of costs related to the data access, data preparation and use of data phases. This document provides guideline for the following questions:

What costs are considered in the fees and what costs are not: several activities are identified as necessary to answer requests for secondary use of health data. Those activities generate costs based on human and IT resources. The question addressed here concerns which costs can be included in the fees, costs necessary to fulfil access requests can be either marginal (costs initiated by the request) or fixed (costs generated prior to requests to anticipate and optimize data access).

To whom the fees are to be paid: health data access procedures may involve multiple stakeholders depending on the type of application/request, and these parties need to be compensated for their activities and services. The question addressed here concerns how these contributors coordinate to invoice DUs and collect the fees.

When the fees are paid: health data access procedures involve several successive phases which may be subject to financial compensation. The question addressed here is when the payment of fees is due by the DU.

3.2 Current situation - Feedback from survey and interviews

3.2.1 Findings: common patterns and divergences across Member States

The results of the survey, conducted following the methodology described in Annex 2, indicate that nearly half of the respondents (46%) reported the absence of formalised pricing practices for data provision, while 63% stated that there is no regulatory framework governing the application of fees for the sharing of health data. The varied responses to the survey questions reflect the differing levels of maturity on this matter among countries and underscore the heterogeneity of practices both between and within countries. For instance, while some respondents report full break-even (e.g., Sweden), others offer free access to data (e.g., Slovakia). In certain cases, price lists are available for data applicants (e.g., Statistics Denmark, Dutch Statistics Office, Statistics Lithuania), although these lists are limited to specific entities and do not represent the practices across the entire country. Where no national pricing framework exists, fees are typically set by DHs with varying degrees of transparency and standardisation.

Similarities were observed in the application of fees. Fees are generally stated to be based on the effort required for data preparation and mainly based on hourly rates. In addition to personnel salaries, these rates may include material, operation and infrastructure costs. Computational resources and disk space costs, as well as administration overheads, are also frequently reported to be included in fees. Some entities offer fixed-price packages, drawing from their experience with use cases.

The profile of DUs is a common factor observed to modulate fees, and reduced fees are frequently reported with a different rate for academics, students, patient organisations and small enterprises. The type of project is another factor for reduced fees depending on how the project is funded or whether it includes collaboration with the DH. Other factors reported to influence the fees are the type of data, the required support for data analysis or the volume of data, which impact the time spent and resources consumed.

3.2.2 Challenges

The most frequently reported challenge anticipated by the respondents of the survey is the capacity to align fees and fee practices within countries (regions, entities) and among Member States. Some Member States already have implemented practices, future HDABs may have a different legal status with specific constraints and different resources. The EHDS framework will need to account for these specificities in promoting coherent implementation of fee practices, without undermining legitimate national arrangements. Moreover, various economic situations exist among Member States and cost recovery is based on salaries and technological provision costs which differ significantly between Member States. This disparity will directly affect fees, potentially creating an imbalance in the distribution of the HDAB

burden across Europe when data are substitutable among countries. It may also impact cross-border access to health data for certain populations and undermine the representativeness of some countries in European studies. Such a phenomenon (limiting cross-border access to health data for users from certain Member States and reducing the participation of researchers from certain countries in European research) would undermine the intended effects of implementing the EHDS and may exacerbate differences between Member States.

In a more general manner, the capacity to maintain preferential pricing schemes for academics and small companies is perceived as an important matter to ensure a level playing field for research and innovation in the EU. Finally, several Member States highlighted the importance of cultural and organisational change, as well as the need for further guidance, capacity-building, and coordination mechanisms.

3.3 Principles guiding the fees

Article 62 sets out the principles that must guide the fees charged on DUs. These include transparency, non-discrimination, proportionality, and competition neutrality.

Transparency involves that fees are related to eligible costs incurred to reply to the DU's request for health data access. Before issuing a data permit or providing a response to a health data request, the HDAB shall inform the health data applicant of the estimated fees (Article 62.5). When fees are claimed, the DU is provided with an invoice describing the amount of the fee (with the currency used) and what it covers.

Non-discriminatory practice involves equal treatment between actors. Non-discrimination requires that DUs of the same category are treated equally, regardless of nationality, and are charged equivalent fees in proportion to the actual costs incurred. Certain categories of DUs located in the EU may benefit from reduced fees, such as public sector bodies or Union institutions, bodies, offices and agencies with a legal mandate in the field of public health, university researchers or microenterprises (Article 62.1). The category of DUs eligible to such reduction is defined at the discretion of each Member state.

The fees must be proportionate to the cost of making the data available and they shall not restrict competition. It implies that the conditions for application must not be used to distort or restrict competition in the market. This includes ensuring that fee structures do not create undue barriers to any categories of eligible DUs regardless of their size.

3.4 What needs to be paid in fees

Fees should reflect the eligible and necessary costs actually incurred in responding to a DU's request. This section outlines the relevant costs associated with data access applications, data requests, and the simplified procedure involving TDHs, and distinguishes between costs that are likely to be compensated for through fees and those that must be excluded.

Excluded costs

Three main categories of costs are excluded from the fees.

- **Data collection and processing realised for medical purposes**: All costs associated with data collection and processing carried out for medical purposes (primary use) are excluded from the fees. These costs cannot be compensated for by DUs. However, specific efforts deployed to support the future secondary use of these data ("by design"), used in the primary care context, could be eligible if they are necessary to enable access for a specific secondary use request and are not related to already funded primary use costs (i.e. no double compensation).
- **Data discovery phase**: All costs related to the creation, updating, and maintenance of the metadata records and catalogue (Article 77) including quality labelling (Article 78), as well as the use of the metadata catalogue by the DU to identify datasets are out of scope (see Annex 3 for list of costs). These costs are not directly related to a data access or request and cannot be compensated by DUs since the dataset descriptions and the catalogue are regulatory obligations, not user-specific services, and therefore not fee-eligible. Although the EHDS Regulation generally excludes costs related to the data discovery phase from recoverable fees, there may be specific cases where project-specific discovery activities required substantial effort and led to key preparation work. While not foreseen in the Regulation, these situations may deserve further reflection during the public consultation
- **Information on significant findings**: All costs related to informing a natural person about significant health findings identified by DUs (Recital 67, Article 58.3) are out of scope for fees. While complying with this obligation under the EHDS regulation may require significant time and resource investment from DHs, these costs cannot be compensated by DUs.

In addition to the excluded categories of costs**,** any eligible costs that are already covered by funding cannot be used to justify fees (i.e. no double compensation). The purpose of the fee is to compensate HDABs and/or DHs for the actual costs incurred in making health data available for secondary use. Therefore, any expenses already financed through other sources, whether public or private, must be excluded from the fee calculation.

Included costs

Eligible costs have been grouped based on activities related to:

- Receipt of a Data Access Application/ Data Request
- Data permit/Data request approval
- Data preparation for the request
- Provision of the data
- Use of the data

Costs may vary depending on the practices and infrastructure in place at the national level

or within individual DHs' organizations. A detailed cost breakdown is provided in Annex 3. This table serves as a reference framework for fee calculation, illustrating the types of costs that are eligible and identifying the parties responsible for them. However, it is intended to be adaptable by each Member State, HDAB or DH, in line with their specific infrastructure and operational practices. For example, costs labelled as "fixed" in the table may be considered "marginal" depending on the chosen method for carrying out the task (see section 3.4.3.2 for example). Additionally, costs associated with HDABs (in Annex 3 or in the text below) may be borne by different stakeholders depending on how responsibilities are distributed among them. Similarly, in cases where the HDAB is a data controller and acts as a DH, the HDAB may claim related costs incurred in the same way as any other DH. Lastly, it should be noted that as data technology evolves, the list of eligible costs presented in this document may also change and it cannot be regarded as exhaustive. In all cases, fees can only be related to costs incurred to reply to the access request, they must be transparent and clearly stated to the DU.

3.4.1 Receipt of a Data Access Application/Data Request

Once submitted by the data applicant, the completeness and the quality of the application are verified to ensure the HDAB has all the necessary information to proceed with the request. Further details on the application check process are provided in [2].

Once checked, the HDAB shall make public the application through electronic means (Article 57.1.j.ii). The HDAB must contact the DHs to establish whether the data can be extracted in the way the applicant has indicated. The DHs will provide to the HDAB a fee estimate associated with the data extraction. The HDAB consolidates the estimated fees from all relevant contributors and communicates a single transparent fee estimate to the applicant. At this stage, the applicant shall be given the possibility not to accept the fees estimation and either review the application to reduce the estimated costs or withdraw the application (Article 62.5). The DU should be provided with a breakdown of the data preparation fees for each dataset to allow for the application to be reviewed and costs reduced. If the application is withdrawn, the applicant shall only be charged for costs incurred so far. For this step, HDAB may charge fees to the data applicant for all actual eligible costs incurred:

- the management of the application (application check process),
- the running and updating of the public information system,
- the regulatory feasibility,
- the assessment of the datasets to be requested to DHs
- the preparation of the quote to respond to the request.
- related administrative overheads directly linked to the request (see Annex 3 for the list of proposed overheads)

DHs may charge fees to the data applicant for all actual costs related to:

- the examination of the protocol and feasibility study to assess the capacity to provide the requested data based on the protocol
- the preparation of the quote to respond to the request.
- related administrative overheads directly linked to the request

Further guidance is provided regarding the examination of the protocol and feasibility study in [1].

3.4.2 Data permit/Data request approval

Once the application is deemed complete and costs estimation approved by the DU, the formal application assessment process begins. Details on the assessment process are provided in [2].

If the data requested include materials subject to intellectual property rights or containing trade secrets, the HDAB must assess whether and/or under which conditions the reuse can be authorised in accordance with Article 52. Further guidance on assessing the intellectual property rights as part of the HDAB's application assessment is provided in [3]. In the specific case of a Data Access Application and if required by the laws of the Member State, the project must be submitted to an ethics committee prior to processing.

Based on the completed assessment, the HDAB shall issue a decision to either grant or deny access to the requested data. All decisions, whether approval or refusal, must be published, including the reasoning in case of refusal, in accordance with Article 58.1.f.

For this step, HDAB may obtain compensation for eligible, incurred costs for the following activities during the assessment phase:

- the assessment of the application against the criteria in EHDS (Article 68.1), including the ethical evaluation where required (Article 68.1.f)
- the risk mitigation analysis for IP and trade secrets (Article 52),
- the permit/data request decision with justification,
- the risk analysis for national defence, security, public security and public order (Article 63 and Article 68)
- the updating of the public information system
- project contracting and monitoring
- related administrative overheads directly linked to the request

3.4.3 Data preparation for the request

Upon issuing a data permit in the case of a data access application or data request approval in the case of a data request application, the HDAB shall request data extraction from the respective DHs (Article 68.7). DHs shall put the requested electronic health data at the disposal of HDAB. Further guidance is provided for DHs in [1].

3.4.3.1 Database constitution and data collection upon request

DHs may adopt different data preparation strategies, depending on their technical architecture, internal capacity, and reuse potential of their datasets.

For instance, the strategy based on the collection of data upon request is likely best suited for DHs with a limited number of data sources and/or relatively simple data, where the costs

and time required to extract and prepare the data are deemed manageable. This strategy involves marginal costs and therefore incurred for each DU request.

The strategy based on the constitution of a database is more adapted to DHs for which the scope of effort for compiling data is high and/or expect a high number of permits concerning their data. This is particularly relevant for hospitals for instance which generate a high volume of diverse data across multiple information systems and with a significant potential for re-use. By identifying, extracting, combining and pre-processing data in advance for secondary use, these DHs can anticipate future requests, accelerate data access and share costs among DUs. This strategy involves fixed costs or structuring costs, which may be compensated for proportionally across multiple data access requests, using a transparent and non-discriminatory cost-sharing methodology.

3.4.3.2 Strategy-dependent costs

In this step, DHs may be compensated for the actual, eligible costs incurred in preparing the requested data from the data applicant for:

- project monitoring
- patient information for the use of their data in the concerned project for which the data are requested
- ∉ data selection, extraction including data minimisation and pseudonymisation/anonymisation (either from operational information systems or from a data warehouse)
- data consolidation
- data export to HDAB

Depending on the DH's strategy, certain additional cost structures may apply. If using a data warehouse model, DHs may be compensated for fixed or structuring costs associated with:

- data extraction from the initial information system
- data quality improvement related to data compiling and preparation activities (see Annex 3 for details)
- data linkage between data from different data systems
- data storage and infrastructure costs for running, maintaining and updating
- regulatory obligation to inform natural persons on the treatment of their data in the data warehouse

If operating on a per-request basis, DHs may be compensated for marginal costs specific to each request, including:

- data quality improvement related to data compiling and preparation activities
- data linkage between data from different data system

When requested data are partly available in a warehouse but require additional data extraction, a combination of fixed and marginal cost recovery is appropriate, using a transparent methodology proportionate to the actual effort incurred.

3.4.4 Provision of the data

Once the HDAB receives the data from the DHs, the data are prepared to be made available to the DU.

In this step, HDAB may charge fees to the data applicant for all costs related to

- project monitoring
- data quality, linkage and consolidation
- pseudonymisation or anonymisation
- data treatment to preserve protection of intellectual property and trade secrets
- dataset validation
- related administrative overheads directly linked to the request
- technical resources used (e.g. CPU time, disk space allocation)

In the specific case of a data request, the HDAB shall provide the DU with access to anonymised, aggregated statistical results. The HDAB may also obtain compensation for the actual, eligible costs related to:

- preparation of the analysis plan for data aggregation and implementation
- generation of the aggregated data

In the specific case of a data permit access, the HDAB shall provide the data applicant with access to anonymised or pseudonymised electronic health data within a secure processing environment (SPE). The HDAB may also obtain compensation for the actual, eligible costs related to:

- preparation of the project space in the SPE
- data export to the project space
- adaptation and development of tools
- licences for tool provision
- validation of the project space
- access to the SPE, additional services from SPE providers and environment updates (software and maintenance)
- project and analysis space provision (user training and support)

3.4.5 Use of the data

The DU can perform analyses based on the received data in the SPE for the purpose defined in the permit and for the duration of the data permit (Article 68). An extension of the permit can be asked by the DU (Article 68.12).

In this step, HDAB may obtain compensation for eligible costs related to the data applicant for all costs related to:

- project closure, activities, including final reporting, data archiving, and controlled data destruction
- long-term storage of metadata or audit logs, where required under national law

In the specific case of data access application where project-specific and scalable SPE costs

incurred, HDAB may also obtain compensation for costs for:

- continued allocation of computing resources and storage
- permit extension and related SPE updates linked to the permit extension

More details on the secure processing environment are provided in [4].

3.4.6 Simplified procedure with a TDH

Where an HDAB receives a health data access application or a health data request that only covers electronic health data held by a TDH, a simplified procedure for access shall be applied (Article 72).

Health data access applications and health data requests are forwarded to the relevant TDH who assesses the health data access application or health data request against the same criteria as the HDAB. The HDAB shall not be bound by the proposed assessment submitted by the TDH and makes the decision to issue the data permit or to approve the health data request. Once the permit is granted or the data request approved, the TDH prepares the health data and provides access to the data applicant via a SPE.

In addition to standard DH costs, a TDH may obtain compensation for actual, eligible costs incurred in performing its specific responsibilities under Article 72, including data preparation and SPE configuration (see Article 72 for details on TDH responsibilities). These costs correspond to those borne by the HDAB in the other procedures.

3.4.7 How are the fees computed

The EHDS Regulation does not mandate charging fees; compensation may range from partial to full, depending on the chosen national approach and the funding context of the actors involved. However, any such approach must be non-discriminatory, documented and consistently applied.

Both marginal and fixed eligible costs may be recovered through these fees. In all cases, fees must be transparent, non-discriminatory and not restrict competition (see section 3.3).

> **Marginal costs**: Fees are calculated based on the specific resources used to process the application. For human resources, the cost is determined by the time required to handle and deliver the application, applying an hourly rate derived from staff salaries. For technological resources, the cost is based on the consumption of infrastructure required by the project (e.g., disk space, CPU, GPU).
>
> **Fixed or structuring costs**: These are expenses incurred to build or maintain infrastructure for secondary use independently of any specific application (e.g., database development, data modelling, data quality, standardisation, licences). They may be annually compensated for through proportional allocation among users, using a transparent methodology.

Although the regulation does not mandate a harmonised method for calculating fees, the

general principles of proportionality, fairness, and transparency must still be respected. To uphold these principles, the HDAB and DH should publish the methodology they use for fee calculation. The method may be determined at the national level, according to applicable standards. Over time, efforts should be made to foster convergence across Member States.

As an example, a simplified method is proposed to illustrate a way to calculate extraction and processing fees for data pre-processed in a secondary-use database, aiming to ensure a fair and rational fixed costs distribution for DUs. The method relies on an annual estimation of the costs allocated to health data categories, along with the projected number of projects (i.e. it requires annual recalibration to remain proportional to the evolving volume of applications and projects). Note that the simplified method below is provided as an example to illustrate a possible way to compute fees related to fixed costs for secondary-use database only and is representative of a later phase since three-year costs retrospective is required.

First, data blocks are identified, and the associated eligible costs are isolated. Data block refers to a coherent set of health-related variables, for instance that are extracted and processed together due to their clinical, analytical, or operational relevance. These variables form a meaningful unit of information, as they are typically collected, interpreted, or used in combination to support a specific medical, research, or administrative purpose. For example, in hospitals, data block could include laboratory data, medication administration data, medical questionnaires, intensive care monitoring data, radiology reports, etc.

The annual estimated costs to extract and pre-process each data block from the information system to the databases ($Cost_{DBk}$) can be calculated as the average eligible costs related to the extraction and processing of that block over the last three years. The amount of data for each block can be formalized by the number of patients it concerns (Nb $patients_{DBk}$). For each data block, a weighted average cost (WAC_{DBk}) is calculated as follows:

$$WAC_{DBk} = Cost_{DBk} / \text{Nb } patients_{DBk}$$

For each DU's application, the fees represent the cost allocated per project among the DUs for each requested data block and depending on the number of patients requested by each DU (Nb $patients_{REQ}$). The anticipated number of projects using a specific block of data can be estimated by the average number of projects that used that block of data over the last three years (Nb $project_{DBk}$).

The fees related to the extraction and pre-processing of data in a secondary-use database (i.e. fixed costs) for a given project (F_{REQ}) is calculated by the sum of the costs per data block and weighted by the number of patients requested by the DU:

$$F_{REQ} = \Sigma_{DBk} [(\text{Nb } patients_{REQ} \times WAC_{DBk}) / \text{Nb } project_{DBk}]$$

3.5 To whom the fees are paid

To address this question, three invoicing scenarios were considered. Each model presents distinct advantages and implementation challenges, particularly regarding administrative complexity and legal compatibility with national systems.

3.5.1 Scenario Overview

Scenario 1 (Centralized Model): This scenario, inspired by Recital 70 (non-binding), Article 62 and the enacting terms, involves a centralized invoicing system where all fee estimates are sent to the HDAB. The HDAB then consolidates these costs into a single invoice directed at the DU. Fees are paid to the HDAB who redistributes to stake holders based on incurred costs. This scenario is found to limit the burden for the health data applicants who receive one invoice from all stake holders but to increase the administrative burden for HDAB as it brings many additional invoicing steps and possibly considerable delays by DHs between data delivery and invoicing.

Scenario 2 (Decentralized Model): Under this model, HDAB, TDHs and DHs issue their own invoice directly to the data applicant. Fees are paid to each stake holder based on their invoice. While this reduces the workload for the HDAB, it poses a challenge for smaller DHs, both public and private, who may lack the resources to manage invoicing for activities outside their core activities. It also increases the burden on DUs, who must handle multiple invoices and make payments that are not coordinated with each other, nor aligned with the centralized data delivery managed by the HDAB.

Scenario 3 (Hybrid Model): The hybrid approach allows for flexibility in invoicing. The HDAB offers to either centralize the invoicing or allow DHs to issue their own invoices if they have the maturity to do so. Fees are paid to the invoice sender. This model gives DHs autonomy in managing their fees when relevant, while offering the HDAB the option to streamline the process as necessary. In this scenario, TDHs operate independently in invoicing and fee collection when the request concerns only the electronic health data they hold.

3.5.2 Recommendations

The centralised model (scenario 1) is the recommended implementation baseline, as it reflects the approach described in Recital 70 of the regulation. While not legally binding, this recital provides a strong interpretive indication of the legislator's intent to streamline interactions between DUs and the overall system. In this model, the HDAB acts as the sole financial interface: it collects fee estimates from all the relevant actors, issues a single consolidated invoice to the DU, collects payments from DUs and redistributes the relevant shares to the contributors. This approach ensures simplicity and transparency for the DU, while consistent with the centralised procedural logic of the regulation.

However, flexibility must be preserved to account for divergent national legal frameworks, administrative traditions, and operational maturity. Member States may therefore adopt a model adapted to their specific national context, provided that the principles of transparency, consistency, and proportionality are respected, as required under Article 62.6.

In particular, while centralising invoicing at the HDAB level is encouraged, the HDAB should have the option to delegate its invoicing responsibilities to DHs when a single source of data is required by the DU and when this aligns with local practices and with the maturity of those actors. The role of TDHs could be further reinforced in the simplified procedure, where they may operate autonomously in invoicing and fee collection. In this case, the HDAB retains

responsibility for issuing the final permit, but the TDH manages the technical implementation and associated billing. Such flexibility is considered essential to ensure operational efficiency and alignment with local realities.

3.6 When the fees are paid

Two scenarios have been discussed regarding when fees should be invoiced to and paid by DUs during the data access procedure.

3.6.1 Key concepts

To support the scenarios, the following key concepts are introduced:

- **Invoice**: A legally binding commercial document, detailing the complete cost structure with breakdowns by services and DHs. It contains disaggregated cost elements (fixed costs and estimated components), typically at the task level to favour clarity and transparency.
- **Request for Payment:** A formal request submitted to the DU for payment of the actual costs corresponding to work completed during a specific period. It follows the structure defined in the original invoice and refers to the relevant cost components outlined therein.
- **Payment Instalment**: One of several scheduled payments made in response to requests for payment. Each instalment corresponds to a portion of the total cost, aligned with the progress of the procedure or delivery of services.
- **Payment**: The financial transaction by which the user transfers the requested amount to the HDAB, TDH or the DH in response to a request for payment.

3.6.2 Scenario overview

Scenario 1 (Single invoice with conditional staged payments model): a single invoice is issued to the DU at the beginning of the procedure, once the HDAB has received fee estimates from all relevant stakeholders in response to a data request or data access application (as indicated in section 3.4.1) prior to the data permit/data request approval phase. The invoice details all potential cost components that may arise throughout the procedure and clearly indicates which costs are payable upfront and which are conditional upon progression to later phases. While fixed prices should be used whenever possible, any estimated components must be clearly identified as such on the invoice and presented with appropriate ranges or formulas to account for variability (e.g., hourly rates, per gigabyte charges, etc.). Payment instalments are predefined in the invoice, specifying when each conditional payment becomes due. These instalments may be scheduled periodically or based on key milestones, or a combination of both, depending on the phase of the procedure. Requests for payment are submitted to the DU in accordance with the instalment schedule, triggering the corresponding payments of the actual costs as the procedure advances and services delivered.

Scenario 2 (Two-stage milestone-based invoicing model): Unlike Scenario 1, where a single invoice covers the entire procedure from the outset, this model splits invoicing into two

distinct stages aligned with key milestones. The first invoice is issued at the beginning of the procedure, once the HDAB has received all relevant fee estimates in response to a data request or data access application prior to the data permit/data request approval phase. It covers cost components referred in sections Receipt of a Data (3.4.1), Data permit/Data request approval (3.4.2) and Data preparation for the request (3.4.3). The second invoice is sent when the data are ready to be delivered in the SPE. It covers costs components referred in sections Provision of the data (3.4.4), and Use of the data (3.4.5). As in Scenario 1, both invoices include predefined payment instalments, which may be scheduled periodically, by milestone, or both. Requests for payment are issued in line with the instalment plan, triggering the corresponding payments as the procedure progresses.

3.6.3 Recommendations

The single invoice with conditional staged payments model (scenario 1), inspired by Recital 70 of the regulation, is the recommended implementation as it best reflects the legislator's intent. A single invoice is issued at the beginning of the procedure and outlines the full potential cost structure of the procedure, along with a schedule of payment instalments. This approach requires estimating the total cost of the entire procedure from the outset. It enables the HDAB to provide the DU with a clear view of their potential financial exposure throughout the process. If the user proceeds beyond the application stage, the combination of payment instalments and requests for payment ensures that the DU is only charged for costs that have actually been incurred as the procedure progresses.

However, accurately estimating all costs from the beginning can be challenging. As the use cases for projects involving health data continue to evolve and HDABs gain operational maturity, the risk increases that the initial estimates from both HDAB and DH may diverge significantly from the actual costs incurred. This is particularly relevant in data access applications involving data use in a SPE, where infrastructure assumptions and computational needs must be projected long before the project is executed. This is something many HDABs may not yet be equipped to do confidently or that cannot be done as long as data is not prepared.

Therefore, flexibility should be granted to HDAB to apply a two-stage (or more) invoicing model when all costs cannot be reliably estimated at the outset. This alternative allows for more accurate fee estimation in the next invoice and helps avoid disputes or confusion with DUs in case of major discrepancies. Even when later stages are difficult to predict, the first invoice should still provide as much initial cost information as possible over the whole procedure to ensure transparency and user confidence.

To support prospective applicants, it is recommended that DHs optionally provide illustrative examples of past projects, potentially included in dataset descriptions or published on HDAB websites. FAQs and cost estimation tools available on the HDAB website are also encouraged to help users anticipate potential costs prior to applying.

4. Areas for further exploration

Survey feedback, interviews and workshops realised in the context of the WP4.1.1 led to questions that appear to be important to raise while defining rules for fees.

Sanitary crisis: While many countries successfully adapted their practices to facilitate data sharing at the national level during the COVID-19 crisis, it is important to establish clear, EU-wide guidelines for managing similar situations in the future. In particular, guidance is needed on how fees should be handled in times of crisis—whether they should be waived, reduced, or remain in place, and for which types of users or purposes. A coordinated approach would ensure fairness, consistency, and rapid response across Member States during future public health emergencies. Although Article 62.1 allows fees to be charged “where appropriate”, a board-level guidance should clarify this point for crisis settings.

Cost discrepancies: Since the invoice is issued at the start of the procedure, actual costs may differ from the initial fee estimates, potentially leading to disputes between stakeholders. The procedure should therefore clearly define how such cost discrepancies are to be managed. Relying solely on estimated fees may result in overestimations that could discourage or exclude many potential DUs. Conversely, basing charges only on actual costs when discrepancies are significant may jeopardize the successful execution of the DU’s project due to unforeseen financial burdens. Although scenario “Two-stage milestone-based invoicing model” proposed in section 3.6.2 aims to mitigate this point, further guidance should be investigated.

Enriched datasets: A DU can enrich data in datasets provided to him in the SPE. The enriched dataset may have value for future DUs, so the question is how to do this and who bears the costs. The EHDS Regulation leaves the precise mechanisms for the processes associated with data enrichment at national level.
Article 51.3 of the Regulation says: Member States may establish rules for the processing and use of electronic health data containing improvements related to the processing of those data, such as correction, annotation or enrichment, based on a data permit pursuant to Article 68.
There are (at least) two deliverables in TEHDAS2 that define framework for data enrichment (D5.4 Short guide for data enrichment for HDABs, DHs and DUs) and its handling in SPE (M7.4 Draft technical, functional and security specifications of Secure Processing Environments).
The costs associated with data enrichment processes can be tracked for the following entities:

- SPE (possible longer period for dataset retain before it is moved to the destination DH, transfer of the enriched dataset to HDAB)
- HDAB (processing of announcements from DU who managed to enrich data in particular datasets, negotiation of these enrichments with the DH of the original dataset, temporary storing the enriched dataset (in local HDAB SPE), secure transfer of the enriched dataset to the DH, support of personal data protection tasks

- DHs (negotiation and tasks related to acceptance of the new dataset, creation or update of metadata to the new dataset, possible maintenance of two datasets instead of one original dataset)
- TDHs may have a partial combination of the costs indicated for HDAB and DHs

Effectively, there are costs also at the DUs, as they should describe what they enriched in the dataset, and they shall report the enrichment. A framework for allocating and recovering these costs should be thoroughly examined and clearly defined.

Research costs for data access: optimally, researchers should be aware of the costs related to a data access application or data request in advance. They probably need this information when applying for a research grant. However, many problems arise with this:

- The researcher already requires some funding for the initial evaluation of the request and cannot include that part of the costs in the anticipated budget in the research grant proposal
- Research grants proposals are usually written at a very early stage and sometimes, 1 or 2 years pass until the research grant is actually granted so the estimation of the costs might not be accurate anymore

Public-facing cost estimation tools or standardised case examples could be provided by HDABs to address this point, the guidance that DH could provide to DU before they apply for the data should be further investigated and described.

5. References

[1] TEHDAS2 document M6.1 “Guideline for health data holders on making personal and non-personal electronic health data available for reuse”.

[2] TEHDAS2 document M6.3 “Guideline for Health Data Access Bodies on the Procedures and Formats for Data Access”.

[3] TEHDAS2 document M5.2 “Guideline for Health Data Access Bodies on allowed purposes and prohibited secondary use according to EHDS”.

[4] TEHDAS2 document M7.1 “Guideline on how to use data in a secure processing environment”

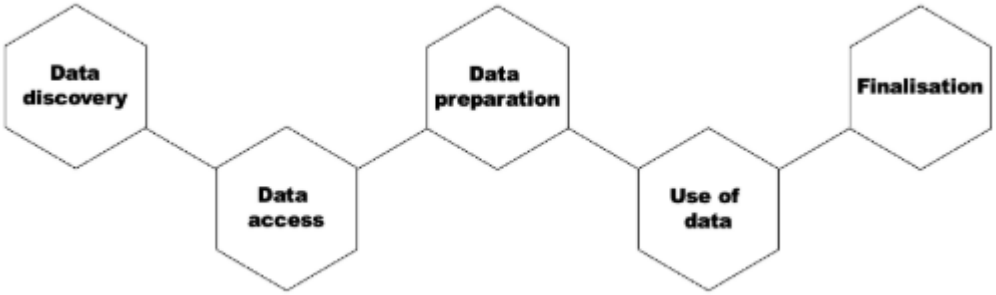
6. Annexes

6.1 Annex 1: User journey

When a data user[i] 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[iv]. 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.

6.2 Annex 2 : Methodology

A survey designed to gather insights on health data fees practices and benchmarking current fee structures and policies across EU Member States, was sent in October 2024 to the Competent Authorities of TEHDAS2 covering 29 countries. Respondents completed it remotely via a dedicated tool, and the survey was closed at the end of January 2025.

A total of 24 responses were received : 20 from member states (Belgium, Croatia, Cyprus, Denmark, Finland, France, Hungary, Ireland, Italy, Iceland, Lithuania, Malta, Norway, the Netherlands, Poland, Portugal, Romania, Spain-Aragones, Spain and Sweden) 2 from European institutions (EIT Health, ECDC) and 2 from industrial federations (EFPIA, INFARMA). A comparative analysis of the responses was conducted, along with targeted interviews with 4 of the respondents (Sweden, Poland, Norway, Denmark). Based on this preparatory work, scenarios have been developed and discussed with the WP4.1.1 group members during workshops to converge towards consensual proposals, which form the foundation of the recommendations presented in this guideline.

6.3 Annex 3 : List of costs

Data application/ data request/ both	Eligible to be considered for fees?	TOPICS CATEGORY	Marginal/ Fixed	TOPICS	ACTIVITY	Task related to cost	Type of costs	HDAB	TDH	DH
Both	NO	Data discovery	Fixed	National Metadata catalogue	Quality and utility label	Implement provisions on data quality and utility label	HR			X
Both	NO	Data discovery	Fixed	National Metadata catalogue	Quality and utility label	Supervise DH to ensure implementation of the provisions on data quality a	HR	X		
Both	NO	Data discovery	Fixed	National Metadata catalogue	Implementation	Produce of database documentation	HR			X
Both	YES	Data discovery	Fixed	National Metadata catalogue	Implementation	Database update due to project-specific requirements	HR			X
Both	NO	Data discovery	Fixed	National Metadata catalogue	Administrative overheads	Communicate to HDAB a description of the database	HR			X
Both	NO	Data discovery	Fixed	National Metadata catalogue	Database accuracy	Check accuracy of database in the national dataset catalogue on annual ba	HR			X
Both	NO	Data discovery	Fixed	National Metadata catalogue	Implementation	Make and maintain a national dataset catalogue (internet site)	HR	X		
Both	NO	Data discovery	Fixed	National Metadata catalogue	Implementation	Develop tools to consolidate catalogs from DH	HR	X		
Both	NO	Data discovery	Fixed	National Metadata catalogue	Implementation	Integrate data and update in national catalogue	HR	X		
Both	NO	Data discovery	Fixed	National Metadata catalogue	Implementation	Server sizing and security costs	Infrastructure	X		
Both	NO	Data discovery	Fixed	European Metadata catalogue	Implementation	Synchronize at european level	HR	X		
Both	YES	Submission of Data Access application/request	marginal	Application/request management	Examination of the data application/re	Receive and analyse the file (details in M6.3)	HR	X		
Both	YES	Submission of Data Access application/request	marginal	Application/request management	Public Information	Make publicaly available data application and request	HR	X		
Both	YES	Submission of Data Access application/request	Fixed	Application/request management	Public Information	Run a management system to record and process health data access applic	Infrastructure	X		
Both	YES	Submission of Data Access application/request	marginal	Application/request management	Examination of the data application/re	Assessment of the datasets to be requested to data holders	HR	X		
Both	YES	Submission of Data Access application/request	marginal	Application/request management	Administrative overheads	Transmit the file to concerned data holders	HR	X		
Both	YES	Submission of Data Access application/request	marginal	Application/request management	Administrative overheads	Coordinate between HDAB at EU and national level	HR	X		
Both	YES	Submission of Data Access application/request	marginal	Project feasibility	Examination of the data application/re	Validate method and protocol	HR			X
Both	YES	Submission of Data Access application/request	marginal	Project feasibility	Feasibility study	Validate regulatory feasibility	HR	X		
Both	YES	Submission of Data Access application/request	marginal	Project feasibility	Feasibility study	Perform feasability study	HR			X
Both	YES	Submission of Data Access application/request	marginal	Fee estimate	Quoting	Draft the quote	HR	X		X
Both	YES	Submission of Data Access application/request	marginal	Fee estimate	Administrative overheads	Transmit the quote to HDAB	HR			X
Both	YES	Submission of Data Access application/request	marginal	Fee estimate	Administrative overheads	Compile and transmit the quote to data user	HR	X		
Data application	YES	Data permit/Data request approval	marginal	Data permit assessment	Permit management	Assess data application	HR	X	X	
Data application	YES	Data permit/Data request approval	marginal	Ethical assessment	Submission to ethical committee	Prepare file for ethical committee	HR	X		
Data application	YES	Data permit/Data request approval	marginal	Ethical assessment	Submission to ethical committee	Ethical committee assessment costs	HR	X		
Data application	YES	Data permit/Data request approval	marginal	Data permit	Permit management	Grant (or refuse) permit and justify	HR	X		
Data application	YES	Data permit/Data request approval	marginal	Data permit	Permit publication	Make publicaly available granted permit and refusal	HR	X		
Data request	YES	Data permit/Data request approval	marginal	Data request approval	Data request management	Assess data request (details in M6.3)	HR	X	X	
Data request	YES	Data permit/Data request approval	marginal	Data request approval	Data request management	Approve (or refuse) data request and justify	HR	X		
Data request	YES	Data permit/Data request approval	marginal	Data request approval	Data request publication	Make publicaly available data request approval/refusal	HR	X		
Both	YES	Data permit/Data request approval	marginal	Data request approval	Risk mitigation	Analyse risks for IP rights and trade secrets	HR	X		X
Both	YES	Data permit/Data request approval	marginal	Data request approval	Risk mitigation	Analyse risk for data protection GDPR	HR	X		
Both	YES	Data permit/Data request approval	marginal	Data request approval	Risk mitigation	Analyse risks for national defence, security, public security and public order	HR	X		
Both	YES	Data permit/Data request approval	marginal	Contracting	Contracting management	Prepare contract and signature	HR	X	X	
Both	YES	Database constitution, infrastructure and quality	Fixed	Compiling	Data extraction	Perform data collection to feed datawarehouse for secondary use	HR			X
Both	YES	Database constitution, infrastructure and quality	Fixed	Compiling	Data extraction	Create of new data pipeline for datawarehouse	HR			X
Both	YES	Database constitution, infrastructure and quality	Fixed	Compiling	Data quality	Monitor data (quality control on data collection)	HR			X
Both	YES	Database constitution, infrastructure and quality	Fixed	Compiling	Data quality	Align terminology (protocol + mapping)	HR			X
Both	YES	Database constitution, infrastructure and quality	Fixed	Compiling	Data quality	Improve data quality (accuracy, completeness, format,...)	HR			X
Both	YES	Database constitution, infrastructure and quality	Fixed	Compiling	Data linkage	Perform internal data linkage	HR			X
Both	YES	Database constitution, infrastructure and quality	Fixed	Compiling	Infrastructure	Depreciation cost of investment for secondary use infrastructure	Infrastructure			X
Both	YES	Database constitution, infrastructure and quality	Fixed	Compiling	Data storage	Disk space and infrastructure costs (running+maintaining+update)	Infrastructure			X
Both	YES	Database constitution, infrastructure and quality	Fixed	Regulatory obligation	Patient information	Drafting and validation of patient information	HR			X
Both	YES	Database constitution, infrastructure and quality	Fixed	Regulatory obligation	Patient information	Dissemination of patient information	HR			X
Both	YES	Database constitution, infrastructure and quality	Fixed	Regulatory obligation	Patient information	Maintain a public information system to comply regulatory obligations	HR	X		
Both	YES	Database constitution, infrastructure and quality	Fixed	Regulatory obligation	Patient information	Run a public information system to comply regulatory obligations	Infrastructure	X		
Both	YES	Data preparation for the request	marginal	Regulatory obligation	Patient information	Draft and validate of patient information	HR			X
Both	YES	Data preparation for the request	marginal	Regulatory obligation	Patient information	Disseminate patient information	HR			X
Both	YES	Data preparation for the request	marginal	Data preparation DH	Data preparation follow up	Monitor project	HR			X
Both	YES	Data preparation for the request	marginal	Data preparation DH	Data selection	Check inclusion/exclusion criteria	HR			X
Both	YES	Data preparation for the request	marginal	Data preparation DH	Data extraction	Develop file/data extraction protocole	HR			X
Both	YES	Data preparation for the request	marginal	Data preparation DH	Data extraction	Extract file/data	HR			X
Both	YES	Data preparation for the request	marginal	Data preparation DH	Data treatment and consolidation	Perform pseudonymisation/ anonymisation	HR			X
Both	YES	Data preparation for the request	marginal	Data preparation DH	Data treatment and consolidation	Data minimisation	HR			X
Both	YES	Data preparation for the request	marginal	Data preparation DH	Data treatment and consolidation	Consolidate table	HR			X
Both	YES	Data preparation for the request	marginal	Data preparation DH	Data storage	Storage and computing resources preparation space	Infrastructure			X
Both	YES	Data preparation for the request	marginal	Data preparation DH	Data transmission	Export data to HDAB	HR			X
Both	YES	Provision of the data	marginal	Data preparation in SPE	Project space preparation	Monitor project	HR	X	X	
Both	YES	Provision of the data	marginal	Data preparation in SPE	Data transmission	Receive data and perform quality control	HR	X	X	
Both	YES	Provision of the data	marginal	Data preparation in SPE	Data quality	Align terminology (protocole + mapping)	HR	X	X	
Both	YES	Provision of the data	marginal	Data preparation in SPE	Data linkage	Define linkage strategy and implement	HR	X	X	
Both	YES	Provision of the data	marginal	Data preparation in SPE	Data linkage	Make linkage report	HR	X	X	
Data Application	YES	Provision of the data	marginal	Data preparation in SPE	Data treatment and consolidation	Perform pseudonymisation/ anonymisation	HR	X	X	
Both	YES	Provision of the data	marginal	Data preparation in SPE	Data treatment and consolidation	Take measures necessary to preserve the confidentiality of intellectual proj	HR	X	X	
Both	YES	Provision of the data	marginal	Data preparation in SPE	Dataset validation	Validate data set	HR	X	X	
Both	YES	Provision of the data	marginal	Data preparation in SPE	Run	Storage and computing resources preparation space	Infrastructure	X	X	
Data application	YES	Provision of the data	marginal	Project environment SPE	Project space preparation follow up	Monitor project	HR	X	X	
Data application	YES	Provision of the data	marginal	Project environment SPE	Project space preparation	Prepare project space SPE	HR	X	X	
Data application	YES	Provision of the data	marginal	Project environment SPE	Project space preparation	Adapt existing tools/create new tools for project space	HR	X	X	
Data application	YES	Provision of the data	marginal	Project environment SPE	Project space preparation	Validate project space	HR	X	X	
Data application	YES	Provision of the data	marginal	Project environment SPE	Project space preparation	Licence for tool provision in project space	Licences	X	X	
Data application	YES	Provision of the data	marginal	Project environment SPE	Project space preparation	Additional services from SPE providers and environment updates	HR	X	X	
Data application	YES	Provision of the data	marginal	Project environment SPE	Project space preparation	Approve project space on security level	HR	X	X	
Data application	YES	Provision of the data	marginal	Project environment SPE	Project space preparation	Approve data environment in project space SPE	HR	X	X	
Data application	YES	Provision of the data	marginal	Project environment SPE	Project space preparation	Data export from preparation space to project space	HR	X	X	
Data application	YES	Provision of the data	marginal	Project environment SPE	Project space preparation	Project space access	Infrastructure	X	X	
Data application	YES	Provision of the data	marginal	Project environment SPE	SPE access	Project and analysis space provision (user training and support)	HR	X	X	
Data application	YES	Provision of the data	marginal	Project environment SPE	SPE access	Initial disk space in project space	Infrastructure	X	X	
Data request	YES	Provision of the data	marginal	Data request implementation	Data aggregation	Prepare and validate analysis plan	HR	X	X	
Data request	YES	Provision of the data	marginal	Data request implementation	Data aggregation	Develop analysis method	HR	X	X	
Data request	YES	Provision of the data	marginal	Data request implementation	Data aggregation	Generate aggregated statistical report	HR	X	X	
Data request	YES	Provision of the data	marginal	Data request implementation	Data aggregation	CPU for aggregation	Infrastructure	X	X	
Both	YES	Provision of the data	marginal	Invoice	Administrative overheads	Consolidate fee from contributors and send invoice	HR	X		
Both	YES	Provision of the data	marginal	Invoice	Administrative overheads	Reception of fees and redistribution to contributors	HR	X		
Both	YES	Use of the data	marginal	Regulatory obligation	Patient information	Maintain an information system to make public the results or output of sec	HR	X		
Both	YES	Use of the data	Fixed	Regulatory obligation	Patient information	Run an information system to make public the results or output of seconda	Infrastructure	X		
Data application	YES	Use of the data	marginal	Project running	Computation capacity	Computing resources in project space	Infrastructure	X	X	
Data application	YES	Use of the data	marginal	Project running	Disk space capacity	Disk space in project space	Infrastructure	X	X	
Data application	YES	Use of the data	marginal	Project running	Project extension	Infrastructure costs related to an extension of the data permit	Infrastructure	X	X	
Data application	YES	Use of the data	marginal	Project running	Project extension	HR costs related to an extension of the data permit	HR	X	X	
Both	YES	Use of the data	marginal	Project Closure	Project closure and archiving	Monitor project closure and database archiving	HR	X	X	
Both	YES	Use of the data	marginal	Project Closure	Project closure and archiving	Provide technical support for project closure and database archiving	HR	X	X	
Both	YES	Use of the data	marginal	Project Closure	Project closure and archiving	Perform all task for data archiving	HR	X	X	
Both	YES	Use of the data	marginal	Project Closure	Project closure and archiving	Data storage and archiving	Infrastructure	X	X	
Both	YES	Use of the data	marginal	Project Closure	Project closure and archiving	Close the project space	HR	X	X	
Both	YES	Use of the data	marginal	Project Closure	Project closure and archiving	Perform data destruction	HR	X	X	
Both	YES		Fixed	Overhead		Administration		X		X
Both	YES		Fixed	Overhead		Electricty bill		X		
Both	YES		Fixed	Overhead		Location costs		X		
Both	YES		Fixed	Overhead		Employer's social insurance contributions		X		
Both	YES		Fixed	Overhead		Equipment (PC, Telephone,...)		X		
Both	YES		Fixed	Overhead		Infrastructure		X		
Both	YES		Fixed	Overhead		Back office activities		X		

6.4 Annex 4 : Glossary

Project partners have added key terms and their definitions used in the milestones and deliverables to this glossary. The aim is to ensure harmonised terminology in all the TEHDAS2 deliverables.

This is a copy from a living document to be updated throughout the joint action. This version is from 5 September 2025.

Term	Definition
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
Access point	A component of the HealthData@EU infrastructure that ensures secure, point-to-point message exchange between National Contact Points and the central platform. Access Points exist at both the national and EU levels and enable the technical interconnection required by Articles 36(3d) and 75 of the Regulation.
Additional information (related to pseudonymisation)	Additional information is information whose use enables the attribution of **pseudonymised data** to identified or identifiable persons (EDPB Guideline 01/2025, Glossary, adopted version for public consultation). This term is specific to **pseudonymisation** and related to the "additional information" referred to in Regulation (EU) 2016/679 Article 4(5) (GDPR).
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)
Anonymisation metadata	Anonymisation metadata refers to a structured set of detailed information describing (a) the methods and parameters used to anonymise a dataset, and (b) the resulting **quality metrics** used to anonymise a dataset or data processing result, or to assess their anonymisation. This metadata helps assess data protection, track modifications, and ensure compliance with anonymisation criteria.

Term	Definition
Anonymisation result	The output of anonymisation, which can be an anonymised dataset or a data processing result including **anonymisation metadata**.
Anonymised statistical format	An anonymised statistical format refers to aggregated data that does not include information on individual data subjects or entities.
Anonymous data	Anonymous data is data which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. (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)
Attribution of pseudonymised data to data subjects	Process that establishes that **pseudonymised data** relate to an already identified person, or links the data to other information with reference to which the data subjects could be identified. (EDPB Guideline 01/2025, Glossary, adopted version for public consultation)
Benefits (of data use)	Refers broadly to positive outcomes of data use. It can encompass social, health and environmental aspects, among others.
Central Platform	An interoperability platform established by the European Commission, providing services to support and facilitate the exchange of information between National Contact Points and authorised participants in HealthData@EU for secondary use of electronic health data. (EHDS Regulation, Article 75(8))
Consistent pseudonymisation	Two sets of data are considered to be pseudonymised consistently if data contained in those sets and relating to the same person can be linked on the basis of the **pseudonyms** they contain (EDPB Guideline 01/2025, Glossary, adopted version for public consultation). Consistency is context-specific and may be limited to a **pseudonymisation domain**.
Cross-border gateway	Handles the transmission and reception of communications between one National Contact Point and Central Services in a secure and technically standardised manner. It

Term	Definition
	supports the eDelivery protocol (HD@EU Pilot WP5 - Architecture Definition).
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 aggregation	A process by which information is collected, manipulated and expressed in summary form (ISO/TR 12300:2014(en), 2.1.4)
Data anonymisation framework	A set of processes and practices designed to ensure data privacy through anonymisation and **privacy risk assessment**.
Data consolidation	A process of combining data from multiple sources, cleaning and verifying them, removing errors so that they can be prepared for provision. Data consolidation may include creation of data subsets, data extraction, duplicates elimination, quality control and data linkage aspects.
Data controller	A data controller is a person or organisation that determines the purposes and essential means of the processing of personal data. The role of the data controller can be shared by several people or organisations. In that case, they are defined as joint controllers. The controller is accountable and responsible for establishing a lawful data processing workflow and observing the rights of data subjects. (GDPR Article 4(1)(7)).
Data extraction	Data extraction is the process of retrieving data from its source dataset. Structured data extraction involves extracting data from datasets that are already organised in predefined formats. Unstructured data extraction pertains to extracting data from databases handling unstructured formats such as PDFs, images, or free text.

Term	Definition
	There may be one or more different data sources from which data extraction may be required.
Data holder application (a software linked to the Secure Processing Environment)	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 linkage	The process of combining **datasets** "from several sources on one topic or data subject" (ISO 5127:2017, 3.1.11.12,). This can be done using unique identifiers, probabilistic methods, or a combination of techniques.
Data minimisation	A principle mandating to only collect, store and process personal data in a manner that is adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed. (GDPR Article 5(1)(c)) Access is only provided to electronic health data that is "adequate, relevant and limited to what is necessary in relation to the purpose of processing indicated in the health data access application by the health data user and in line with the data permit issues pursuant to Article 68." (EHDS Regulation, Article 66(1)) Data minimisation applies to all stages of the data lifecycle.
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 EHDS Regulation. (EHDS Regulation, Article 2(2v))
Data preparation	Data preparation is the process in which an organisation (in this case the data holder) transforms and organises raw personal or non-personal health data into one or more datasets (either in individual-based or aggregated form), to comply with a data permit or a data request

Term	Definition
	approval issued by a data user and approved by the competent Health Data Access Body.
Data processing	Any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction. (GDPR Article 4(2))
Data processing result	Data processing result refers to outputs from data processing activities carried out by the health data user. It may be generated from statistical analysis or machine learning algorithms, including descriptive statistics, model coefficients, performance indicators, visualisations.
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 their behalf. (GDPR, Article 4(1)(8))
Data provenance	Data provenance means a **description of the source** of the data, including context, purpose, method and technology of data generation, documenting agents involved in the provenance of data, data validation routines, source data verification, traceability of changes, and quality control of data.
Data provision	The stage in the data user journey where prepared health data is made accessible to authorised users for secondary purposes.
Data quality	Data quality means the degree to which the elements of electronic health data are suitable for their intended primary use and secondary use; (EHDS Article 2 (2z))
Data quality and utility label	Data quality and utility label means a graphic diagram, including a scale, describing the data quality and conditions of use of a dataset. (EHDS Article 2 (2aa))

Term	Definition
Data user application (a software linked to the Secure Processing Environment)	A software application that provides the data user with secure, computerised access to their workspace within the Secure Processing Environment. Its primary functions include facilitating the upload and download of data while ensuring robust authentication and authorisation mechanisms to prevent unauthorised access.
Dataset	A structured collection of electronic health data. (EHDS Article 2(2)(w))
Dataset Catalogue	A collection of dataset descriptions, arranged in a systematic manner and including a user-oriented public part, in which information concerning individual dataset parameters is accessible by electronic means through an online portal. (EHDS Article 2(2y))
Dataset record	A dataset record is a single, structured unit of data within a dataset, analogous to a row in a table or a record in a database. It contains specific information about a single entity or instance within the broader dataset.
Dataset subset	Dataset subset contains only selected records, variables or elements from a larger dataset while maintaining its key characteristics and relationships.
Dataset description	Health data access bodies shall, through a publicly available and standardised machine-readable dataset catalogue, provide a description in the form of metadata of the available datasets and their characteristics (EHDS Article (77(1))
Direct identifier	A data element (or set thereof) that has been assigned or is being used to distinguish the data subject it refers to from all others in the given context without requiring the use of **additional information**. Examples are passport or social security number, or the set consisting of first and last name as well as date of birth. (EDPB Guideline 01/2025, Glossary, adopted version for public consultation)
Disclosure control	Disclosure control refers to techniques and procedures applied to datasets to reduce the privacy risks for individuals when the data is disclosed to data users.
Dispatcher	A component of the HealthData@EU infrastructure that enables the secure transmission, routing and delivery of

Term	Definition
	structured electronic messages (such as dataset records and access requests) between national and central systems.
Electronic health data	Personal or non-personal electronic health data (EHDS Article 2(2c)).
EU dataset catalogue	A dataset catalogue means a collection of dataset descriptions, arranged in a systematic manner and including a user-oriented public part, in which information concerning individual dataset parameters is accessible by electronic means through an online portal. (EHDS Regulation, Article 2 (2y)) The EU dataset catalogue, the national dataset catalogues and the dataset catalogues of authorised participants in HealthData@EU shall be made publicly available. (EHDS Regulation, Article 79 (1–2))
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 learning (“federated

Term	Definition
	learning"), analytics ("federated analysis"), and secure data collaborations across multiple organisations.
Fidelity	Fidelity (or resemblance**)** refers to the extent to which processed data—such as anonymised data—retains the statistical properties, relationships, and structural characteristics of the **original data**. High fidelity means that distributions, correlations, and key patterns remain intact.
Health data access application	An application seeking to access personal-level electronic health data for secondary use in an anonymised or a pseudonymised format (EHDS Article 67).
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 applicant	A natural or legal person submitting a health data access application or a data request to a Health Data Access Body for the purposes referred to in Article 53 of EHDS Regulation.
Health data holder	Any person, organisation or public body involved in healthcare, care services, health-related products, wellness apps or health(care) research, that has the right to process data for health care provision or for public health purposes, reimbursement, research, policy making, official statistics or patient safety. This includes, for example, hospitals, insurers, research institutes and EU institutions. For a more detailed definition: EHDS Regulation, Article 2(2t))

Term	Definition
Health data request	A request to access data in an anonymised statistical format for the purposes referred to in EHDS Article 53. (EHDS Regulation, Article 69)
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 Regulation, 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.
Intellectual Property (IP)	(a) a trade mark; (b) a design; (c) a copyright or any related right as provided for by national or Union law; (d) a geographical indication; (e) a patent as provided for by national or Union law; (f) a supplementary protection certificate for medicinal products as provided for in Regulation (EC) No 469/2009 of the European Parliament and of the Council of 6 May 2009 concerning the supplementary protection certificate for medicinal products (1); (g) a supplementary protection certificate for plant protection products as provided for in Regulation (EC) No 1610/96 of the European Parliament and of the Council of 23 July 1996 concerning the creation of a supplementary protection certificate for plant protection products (2); (h) a Community plant variety right as provided for in Council Regulation (EC) No 2100/94 of 27 July 1994 on Community plant variety rights (3); (i) a plant variety right as provided for by national law; (j) a topography of semiconductor product as provided for by national or Union law; (k) a utility model in so far as it is protected as an intel lectual property right by national or Union law; (l) a trade name in so far as it is protected as an exclusive intellectual property right by national or Union law. (Regulation concerning customs enforcement of intellectual property rights and repealing, Article 2(1))

Term	Definition
Intermediation entity	A legal person that may be established by national law for the purpose of fulfilling the obligations of certain categories of health data holders and that is able to process, make available, register, provide, restrict access to and exchange electronic health data for secondary use provided by health data holders. (EHDS Regulation, Article 50 (3) and Recital 59)
Interoperability	Ability of organisations, as well as of software applications or devices from the same manufacturer or different manufacturers, to interact through the processes they support, involving the exchange of information and knowledge, without changing the content of the data, between those organisations, software applications or devices. (EHDS Regulation, Article 2(2f))
Irreversible pseudonymisation	A pseudonymisation method where the **pseudonymising transformation** cannot be reversed. The information necessary to re-establish the link between the **pseudonym** and the **original data** has been permanently destroyed or is otherwise unavailable.
Legal basis of data processing	The conditions under which personal data processing is considered lawful (GDPR, Article 6). Purposes for which the electronic health data can be processed for secondary use are laid down in EHDS Regulation, Article 53.
Metadata	A structured description of the contents or the use of data facilitating the discovery or use of that data. (Data Act, Article 2)
National contact point (NCP)	A National Contact Point for secondary use is the organisational and technical gateway for making electronic health data available for secondary purposes, including research, innovation, policy-making, and public health. It plays a crucial role in connecting national data infrastructures to the HealthData@EU Central Platform, enabling secure and efficient data sharing across borders. (EHDS Regulation, Article 75(1))
Non-compliance	Any failure to comply with any requirement under the Union harmonisation legislation

Term	Definition
Non-personal electronic health data	Electronic health data other than personal electronic health data, including both data that have been anonymised so that they no longer relate to an identified or identifiable natural person (the 'data subject') and data that have never related to a data subject. (EHDS Regulation, Article 2(2b))
Observational Medical Outcomes Partnership (OMOP) common data model (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.
Open data	Data in an open format that can be freely used, re-used and shared by anyone for any purpose. Open format means a file format that is platform-independent and made available to the public without any restriction that impedes the re-use of documents. (EU Open Data Directive)
Open (data) database	Publicly accessible digital data that anyone can freely use, reuse, and redistribute for any purpose.
Original data	Individual-level health data prior to any application of ***pseudonymisation, anonymisation***, or ***synthetic data generation***. It consists of raw data that directly represent real-world individuals.
Personal electronic health data	Data concerning health and genetic data, relating to an identified or identifiable natural person, processed in an electronic form. (EHDS Regulation, Article 2(2a))
Privacy (of synthetic or anonymised data)	Privacy measures the extent to which anonymised or synthetic data protects individuals from re-identification, membership inference, or sensitive information leakage. High privacy ensures that no single individual can be traced back to the real dataset, nor can their participation in the ***dataset*** be inferred.

Term	Definition
Privacy risk assessment	Overall process of identifying, analysing, evaluating, consulting, communicating and planning the treatment of potential ***privacy*** impacts with regard to the processing of personally identifiable information (3.7), framed within an organisation's broader risk management framework (ISO/IEC 29100:2024(en), 3.18). ***Re-identification risk assessment*** falls under privacy risk assessment, together with attribute inference and group membership, for example.
Pseudonym	Identifier that is added to data during the ***pseudonymising transformation*** and set in such a way that it can be attributed to data subjects only using ***additional information***. (EDPB Guideline 01/2025, Glossary, version adopted for public consultation)
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))
Pseudonymisation domain	Environment in which the controller or processor wishes to preclude attribution of data to specific data subjects. May incorporate persons acting under the authority of the controller or processor, respectively, other natural or legal persons, public authorities, agencies or other bodies, and their respective technological and informational resources. Does not include persons authorised to process additional data allowing the attribution of the ***pseudonymised data*** to data subjects. (EDPB Guideline 01/2025, Glossary, version adopted for public consultation)
Pseudonymisation entity	The entity responsible of processing identifiers into pseudonyms using the pseudonymisation function. It can be a data controller, a data processor (performing pseudonymisation on behalf of a controller), a ***trusted third party*** or a data subject, depending on the pseudonymisation scenario. It should be stressed that, following this definition, the role of the pseudonymisation entity is strictly relevant to the practical implementation of

Term	Definition
	pseudonymisation under a specific scenario. (ENISA, Pseudonymisation techniques and best practices, p. 10)
Pseudonymisation secrets	Data that is used in the application of the ***pseudonymising transformation*** or is created during that process, for example cryptographic keys or salts, and allows the computation of pseudonyms from certain identifying attributes. Part of ***additional information***. (EDPB Guideline 01/2025, Glossary, version adopted for public consultation)
Pseudonymised data	Result of applying the ***pseudonymising transformation*** to some personal data. Cannot be attributed to a specific data subject without ***additional information***. (EDPB Guideline 01/2025, Glossary, version adopted for public consultation)
Pseudonymising controller or processor	Controller or processor that uses pseudonymisation as a safeguard and modifies ***original data*** according to Regulation (EU) 2016/679 (GDPR) Article 4(5). (EDPB Guideline 01/2025, Glossary, version adopted for public consultation)
Pseudonymising transformation	Procedure that modifies ***original data*** in a way that the result cannot be attributed to a specific data subject without ***additional information***. (EDPB Guideline 01/2025, Glossary, version adopted for public consultation)
Public use file	A ***dataset*** made available to the public, typically containing anonymised, synthetic or aggregated data to protect individual privacy. These files can be released to data users for information and testing purposes before they apply for a data permit. It is based on **original data**.
Public value (of data use)	Public value means a weighted composite of risks and benefits of the data use taking into account the sustainability of benefits, addressing future societal needs, distributing benefits fairly, evaluating potential harm, ensuring stable safeguards through risk assessment, and correcting any harms that may occur.
Purpose limitation	Personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner

Term	Definition
	that is incompatible with those purposes. (GDPR, Article 5(1b).
Quality metrics	Quality metrics refer to qualitative and quantitative indicators used to assess the fitness for purpose of a dataset. In the context of synthetic and anonymised data, quality metrics are particularly relevant to evaluate how transformations affect the data's ***utility***, ***fidelity***, and ***privacy***. Quality metrics may also be used to assess pseudonymised or original datasets, particularly when serving as a benchmark or when evaluating fitness for specific secondary use purposes. (Adapted from ISO and EHDS principles; EHDS Regulation, Article 66 and Recital 58)
Quality metrics evaluation	Quality metrics evaluation refers to the calculation or derivation of the ***quality metrics***.
Quality metrics tool	Quality metrics tool (or "metrics tool") refers to a software, an algorithm, a processing pipeline, a documented manual process, or a combination of these, designed to perform ***quality metrics evaluation***.
Quasi-identifier	A ***dataset*** attribute that, when considered in conjunction with other attributes are sufficient to attribute at least part of the pseudonymised data to data subjects. (EDPB Guideline 01/2025, Glossary, version adopted for public consultation)
Re-identification	The process of associating data in a de-identified ***dataset*** with the original data principal (i.e., data subject) (ISO/IEC 20889:2018(en), 3.31).
Re-identification risk	The risk of a successful re-identification attack (ISO/IEC 20889:2018(en), 3.33), which describes an action performed on de-identified data by an attacker with the purpose of ***re-identification*** (ISO/IEC 20889:2018(en), 3.32).
Representational State Transfer Application Programming Interface (RESTful API)	An application programming interface used for building scalable and interoperable web services. RESTful API follows the principles of Representational State Transfer (REST), using standard HTTP methods 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.

Term	Definition
Reversible pseudonymisation	The ***pseudonymisation entity*** uses a ***pseudonymising transformation*** process that allows the pseudonymisation entity to reverse the ***pseudonym***, if necessary. For example, by using separately kept matching tables of pseudonyms and identifying data, or computable secrets allowing for calculating back to the original input.
Secondary use	Processing of electronic health data for the purposes set out in Chapter IV of EHDS Regulation, other than the initial purposes for which they were collected or produced. (EHDS Regulation, Article 2(2e))
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 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)
Sensitive data	Data with potentially harmful effects in the event of disclosure (i.e., providing access to data to a third party) or misuse (ISO 5127:2017(en), 3.1.10.16)).
Synthetic data	Data that is artificially generated. The concept of synthetic data generation is to take an original data source (dataset) and create new, artificial data, with similar statistical properties from it.
Synthetic data documentation	Documentation of a synthetic dataset generated automatically or semi-automatically by the ***synthetic data generator***. The documentation shall be anonymised so that it can be accompanied with the synthetic data set when released for the data user or for public use.
Synthetic data generator	A synthetic data generator is a software application, model or algorithm designed to generate ***synthetic data***. It uses real-world data as input and generates a synthetic dataset. It is also possible to use parameters derived from the

Term	Definition
	original data as input and/or modify additional parameters entered by the user.
Tabular data	Data organised in a structured format of rows and columns, where each row represents a single record or entity, and each column represents a specific attribute or variable. This structure is commonly found in spreadsheets or relational databases, making it easy to store, query, and analyse. Tabular data is often used for structured datasets where relationships between variables are well-defined.
Trade secret(s)	Information which meets all of the following requirements: (a) it is secret in the sense that it is not, as a body or in the precise configuration and assembly of its components, generally known among or readily accessible to persons within the circles that normally deal with the kind of information in question; (b) it has commercial value because it is secret; (c) it has been subject to reasonable steps under the circumstances, by the person lawfully in control of the information, to keep it secret. (Trade Secret Directive (2016/943), Article 2(1))
Transfer of data outside the EU/EEA	General principles, adequacy decisions, appropriate safeguards and specific derogations for transferring personal data to third countries or international organisations (GDPR, Chapter 5, Articles 44–50). The European Data Protection Board (EDPB) identifies three cumulative criteria to identify a transfer outside the EEA: • ”a controller or a processor is subject to the GDPR for the given processing; • this controller or processor discloses by transmission or otherwise makes personal data available to another organisation (controller or processor); • this other organisation is in a country outside EEA or is an international organisation.”
Trusted health data holder	Member State designated health data holder for whom a simplified procedure can be followed for the issuance of data permits. Trusted health data holders leverage their expertise on the data they hold to assist the Health Data Access Body by providing assessments of data requests or access applications. Once data permits are authorised, these

Term	Definition
	trusted data holders provide the data within a Secure Processing Environment that they manage. (EHDS Regulation, Article 72 and Recital 76)
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. TREs are not the same as Secure Processing Environments, which are legally defined in the EHDS Regulation.
Trusted third party (TTP)	A **pseudonymisation entity** which is independent from the data user and data holder that processes identifiers into pseudonyms. (ENISA, Pseudonymisation techniques and best practices). The TTP needs only to know the identifiers of the data subjects on the basis of which it will compute the **pseudonyms**, and no other data. (EDPB Guideline 01/2025, version adopted for public consultation)
Invoice	A legally binding commercial document, detailing the complete cost structure with breakdowns by services and data holders. It contains disaggregated cost elements, typically at the task level to favour clarity and transparency.
Request for payment	A formal request submitted to the data user for payment of the actual costs corresponding to work completed during a specific period. It follows the structure defined in the original invoice and refers to the relevant cost components outlined therein.
Payment instalment	One of several scheduled payments made in response to requests for payment. Each instalment corresponds to a portion of the total cost, aligned with the progress of the procedure or delivery of services
Payment	The financial transaction by which the user transfers the requested amount to the Health Data Access Body, Trusted Data Holder or the Data Holder in response to a request for payment.
Utility	Utility refers to how well the data supports its intended use, such as syntactical testing, analytical tasks, decision-making, or machine learning model performance. In the

Term	Definition
	context of anonymised and synthetic data high utility means that insights, predictions, or outcomes derived from the data closely match those obtained using the **original data**.