TEHDAS2 Glossary

TEHDAS2 – Second Joint Action Towards the European Health Data Space

Version 1 (08 June 2026)

Co-funded by the European Union

0 Document info

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

0.1 Authors

Author(s)
TEHDAS2 project partners

0.2 Keywords

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

0.3 Document history

Date	Version	Editor	Change	Status
28/04/2025	1.0	Sofia Peltola, Elina Drakvik	Initial document creation	Draft
29/04/2025–19/01/2026	1.1	TEHDAS2 project partners	Adding terms	Draft
27/01/2026	1.2	DG SANTE, C1 Unit	Revision and commenting	Draft
04/06/2026	2.0	Sofia Peltola, Elina Drakvik	Revision and final editing	Final

Contents

1 Introduction 3
2 Glossary 4

1 Introduction

Advancing health data use in the European Health Union

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

TEHDAS2, the second joint action Towards the European Health Data Space, represents a significant step forward in this vision. The project will develop guidelines and technical specifications to facilitate smooth 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.

This document is a collection of the terms used throughout TEHDAS2 deliverables and is intended to serve as assistance to readers of those documents. The aim is to ensure harmonised terminology across all TEHDAS2 deliverables.

Besides this master glossary, each TEHDAS2 deliverable includes a glossary specific to that document, i.e. covering the terms used in that document. The document specific glossary can be found at the end of each document, in the annex section.

The glossary is a joint effort by the TEHDAS coordination team, project partners and document authors who have contributed to its creation. A new, updated version of the glossary will be published towards the end of the project, alongside the final set of deliverables.

This document should be understood as an expert opinion and guidance document developed within the TEHDAS2 framework, reflecting technical and expert input from the

project partners. It is not legally binding and does not constitute a formal guideline or technical specification under the European Health Data Space.

This document does not represent the position of the European Commission.

Legally binding and enforceable requirements under the European Health Data Space are laid down in Regulation (EU) 2025/327 and, where applicable, in Implementing Acts adopted by the European Commission, within the limits of the empowerments provided by the Regulation.

2 Glossary

Disclaimer: This glossary is intended to support consistent terminology within TEHDAS2 deliverables. It does not create legal obligations, does not interpret Union law, and does not prejudge the content of implementing acts, guidelines, or national implementation measures under the EHDS Regulation.

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, version adopted 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).
AI system	A machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual

	environments. AI Act – Regulation (EU) 2024/1689, Article 3(1)
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	Where applicable, 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. It includes details e.g., on applied techniques and transformation logs. This metadata helps assess data protection, track modifications, and ensure compliance with anonymisation criteria.
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. Aggregation is one possible anonymisation technique.
Areas of occupational health	The main disciplines concerned with protecting and promoting works health and safety in the workplace. Areas of occupational health include: - Occupational medicines: Prevention and management of work-related diseases, - Occupational hygiene: Identification and control of workplace hazards - Occupational safety: Prevention of accidents and injuries - Occupational health nursing: Workplace health services - Ergonomics: Adapting work to fit the worker - Occupational psychology: Mental health and well-being at work - Environmental health: Control of environmental risks affecting workers A case of occupational disease is defined as a case recognised by the national authorities responsible for recognition of occupational diseases. The data shall be collected for incident occupational diseases and deaths due to occupational disease.

	Work-related health problems and illnesses are those health problems and illnesses which can be caused, worsened or jointly caused by working conditions. This includes physical and psychosocial health problems. A case of work-related health problem and illness does not necessarily refer to recognition by an authority, and the related data shall be collected from existing population surveys such as the European health interview survey (EHIS) or other social surveys. (Regulation (EU) 1338/2008, Annex V, (b) and WHO2, Article 3(c))
Areas of public health	All elements related to health, namely health status, including morbidity and disability, the determinants having an effect on that health status, health care needs, resources allocated to health care, the provision of, and universal access to, health care as well as health care expenditure and financing, and the causes of mortality. (Regulation (EU) 2021/2282, Article 2(5))
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, version adopted for public consultation)
Authorised user	An authorised natural person or legal person listed in the data permit, giving them the rights to process sensitive data inside a secure processing environment.
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))
Citizen information point (CIP)	A public information system that helps health data access bodies meet their legal obligations laid down in Article 58 of the EHDS regulation. It informs natural persons about the conditions under which electronic health data are made available for secondary use.

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, version adopted 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 supports the eDelivery protocol (HD@EU Pilot WP5 – Architecture Definition).
Data access	A phase in the EHDS user journey during which the user fills in and submits a dedicated and standardised data access application form or a data request to a health data access body (HDAB). The user must complete the information required in the form, upload necessary documents, and provide justifications as needed.
Data 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 combination	The process of bringing together data from multiple **datasets** that can be processed pursuant to one or multiple data permit(s) or data request(s) (Regulation (EU) 2015/327 (EHDS) Articles 57, 68, 69) or other legal basis (such as consent or permits based on other legislation than EHDS). **Data linkage** can be part of this process.
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. There may be one or more different data sources from which data extraction may be required.
Data holder application (a software linked to a 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(2) point (v))
Data preparation	Data preparation is the process in which an organisation (in this case the data holder or the

	health data access body) 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.
Data processing	Any operation or set of operations which is performed on personal/non-personal data or on sets of personal/non-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. (Modified from the 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 protection	Processing data respecting the principles laid down in GDPR Article 5(1). The “implementation of appropriate administrative, technical or physical means to guard against unauthorised intentional or accidental disclosure, modification, or destruction of data (ISO/IEC 20944-1:2013(en), 3.6.5.1).
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 **EHDS 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(2) point (z))
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(2) point (aa))
Dataset	A structured collection of electronic health data. (EHDS Article 2(2) point (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(2) point (y)
Dataset description	A description in the form of metadata of the available datasets and their characteristics (EHDS Article (77(1))
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.
Data user application (a software linked to a 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.
Development activities	The concept 'Development activities' is not clearly defined in any legal act. However, there is a definition of the notion of research and development in Directive 2009/81/EC, Article 1(27): 'Research and development' mean all activities comprising fundamental research, applied research and experimental development, where the latter may include the realisation of technological demonstrators, i.e. devices that demonstrate the performance of a new concept or a new technology in a relevant or representative environment." (Directive 2009/81/EC, Article 1(27))
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, version adopted 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 structured electronic messages (such as dataset records and access requests) between national and central systems.
European Health Data Space (EHDS) user journey	The path of a data user applying for electronic health data for secondary use purposes within the European Health Data Space (EHDS). A simplified version of a EHDS user journey is included in the annexes of TEHDAS2 Deliverables. It consists of five main phases: data discovery, data access, data preparation, use of data and finalisation.
Electronic health data	Personal or non-personal electronic health data (EHDS Article 2(2) point (c)).
Entity	Is “something capable of being uniquely identified. Note 1 to entry: *Entities* include material objects, electronic representations of content, abstract items (such as times, places), parties (human and corporate), as well as anything else that can be identified uniquely. Note 2 to entry: A defined fragment of an *entity* is itself an entity.” (ISO 5127:2017(en), 3.1.13.27
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(2) point (y))
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 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/source data**. High fidelity means that distributions, correlations, and key patterns remain unchanged.
Healthcare	'Healthcare' means health services provided by health professionals to patients to assess, maintain or restore their state of health, including the prescription, dispensation and provision of medicinal products and medical devices. Directive 2011/24/EU, Article 3(a)
Health data access application	An application form used to seek access for 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. (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(2) point (t))
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(2) point (u))
Health record	A “data repository regarding the health and care” (ISO/TS 16551:2025(en), 3.9, shortened) of a data subject.
Health technology assessment (HTA)	A multidisciplinary process that summarises information about the medical, patient and social aspects and the economic and ethical issues related to the use of a health technology in a systematic, transparent, unbiased and robust manner. (Regulation (EU) 2021/2282 on health technology assessment and amending Directive 2011/24/EU, Article 2(5))
High performance computing (HPC)	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 (EU) No 608/2013 concerning customs enforcement of intellectual property rights and repealing, Article 2(1))
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(2) point (f))
Invoice	A legally binding commercial document, detailing the complete cost structure with breakdowns by services and data holders. It contains disaggregated cost elements.
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. If the pseudonymising transformation is truly irreversible and re-identification is no longer reasonably possible, the resulting data qualify as anonymised data rather than pseudonymised data under the GDPR.
Legal basis of data processing	The criteria defined in EHDS Regulation Article 68 for health data access bodies to assess whether an applicant can be given a permit to process electronic health data. The conditions under which personal data processing is considered lawful are laid down in GDPR, Article 6. Purposes for which the electronic health data can be processed for secondary use are laid down in EHDS Regulation, Article 53.
Link	The derived or assumed classification between ***health records.***
Linking organisation	The organisation performing data linkage (HDAB, data holder, trusted data holder).
Match	A true relationship between ***health records***.
Match quality	Based on ISO/IEC 19794-14:2022(en), 3.4.12, but with broader scope: The “level of ***agreement*** between different ***health records***.
Medicinal product	Any substance or combination of substances presented for treating or preventing disease in human beings. Any substance or combination of substances which may be administered to human beings with a view to making a medical diagnosis or to restoring, correcting or modifying physiological functions in human beings is likewise considered a medicinal product. Directive 2011/24/EU referring to Directive 2001/83/EC, Article 1(2)
Medical device	Any instrument, apparatus, appliance, software, implant, reagent, material or other article intended by the manufacturer to be used, alone or in combination, for human beings for one or more of the following specific medical purposes: • diagnosis, prevention, monitoring, prediction, prognosis, treatment or alleviation of disease • diagnosis, monitoring, treatment, alleviation of, or compensation for, an injury or disability • investigation, replacement or modification of the anatomy or of a physiological or pathological process or state

	• providing information by means of in vitro examination of specimens derived from the human body, including organ, blood and tissue donations, and which does not achieve its principal intended action by pharmacological, immunological or metabolic means, in or on the human body, but which may be assisted in its function by such means. The following products shall also be deemed to be medical devices: • devices for the control or support of conception • products specifically intended for the cleaning, disinfection or sterilisation of devices as referred to in Article 1(4) and of those referred to in the first paragraph of this point. Regulation (EU) 2017/745 and (EU) 2017/746, Article 2(1)
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.
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. (Directive (EU) 2019/1024 on open data, "Open Data Directive")
Open (data) database	Publicly accessible digital data that anyone can freely use, reuse, and redistribute for any purpose.
Original/source 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.
Overlinkage (false positives)	When one ***health record*** is incorrectly ***linked*** to multiple data subjects.
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.
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.
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.
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 pseudonymisation under a specific scenario. (ENISA, Pseudonymisation techniques and best practices, p. 10, version adopted for public consultation)
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 sector body	'Public sector body' means the state, regional or local authorities, bodies governed by public law, or associations formed by one or several such authorities or one or several such bodies governed by public law." Regulation (EU) 2022/868, Data Governance Act, Article 2(17).
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**. (Eurostat CROS definition)
Public value (of data use)	For analytical or policy discussion purposes, public value could be understood as 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 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).
Re-pseudonymisation	The processing of ***pseudonymised data***, where project ***pseudonyms*** are generated using a pseudonymisation algorithm, replacing previously generated pseudonyms. Re-pseudonymisation should not be confused with attempts to reverse the pseudonymisation, which is not meant here.
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.
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.

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(2) point (e))
Secure processing environment (SPE)	An environment in which access to electronic health data can be provided in following a data permit. A secure processing environment 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)).
Serious cross-border threats	The following categories are considered serious cross-border threats to health: (a) threats of biological origin, consisting of: (i) communicable diseases, including those of zoonotic origin; (ii) antimicrobial resistance and healthcare-associated infections related to communicable diseases ('related special health issues'); (iii) biotoxins or other harmful biological agents not related to communicable diseases; (b) threats of chemical origin; (c) threats of environmental origin, including those due to the climate; (d) threats of unknown origin; and

	(e) events which may constitute public health emergencies of international concern under the International Health Regulations (IHR) ('public health emergencies of international concern'), provided that they fall under one of the categories of threats set out in (a–d) (Regulation (EU) 2022/2371, Article 2(1))
Statistics	Quantitative and qualitative, aggregated and representative information characterising a collective phenomenon in a considered population. (Regulation (EU) 223/2009, Article 3(1))
Statistical disclosure control	Statistical disclosure control can be defined as the set of "methods to reduce the risk of disclosing information on the statistical units (natural persons, households, economic operators and other undertakings, referred to by the data), usually based on restricting the amount of, or modifying, the data released" (Eurostat CROS).
Synthetic data	Artificially generated data created from an original dataset to reproduce its statistical properties, while not directly corresponding to real individuals. Synthetic data may constitute personal data where individuals remain identifiable, in accordance with Regulation (EU) 2016/679 (GDPR).
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 **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.

Topic	An "***entity*** (3.1.13.27) used as a subject of a work (3.2.1.07)." (ISO 5127:2017(en), 3.2.1.17)
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	Transfer of data outside of the European Union or European Economic Area according to the three cumulative criteria identified by the European Data Protection Board (EDPB): • "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 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)	A research environment that aims 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. TREs emerged from the UK health research sector, shaped by community-led principles and structured around flexible, function-based zones.
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 Glossary, version adopted for public consultation)
Underlinkage (false negatives)	When ***matching*** health records belonging to the same data subjects are not ***linked.***
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 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**.