Pseudonym generation, pseudonym linkage and encrypted pseudonymised data transfer via the European Platform on Rare Disease Registration pseudonymisation tool

1. GENERAL INFORMATION

1.1. Data protection record

Title of the processing operation
Pseudonym generation, pseudonym linkage and encrypted pseudonymised data transfer via the European Platform on Rare Disease Registration pseudonymisation tool

Language of the record
English

Corporate record
no

Last updated on
26/04/2022

Last updated by
Leucio Antonio CUTILLO

Created on
10/08/2021

1.2. Data Protection officer

Data Protection Officer
EC-DPO-INTERNAL@ec.europa.eu

Data Protection Coordinator
JRC-DATA-PROTECTION-COORDINATOR@ec.europa.eu

1.3. Entity of the Operational Controller

Entity of the Operational Controller (DG. UNIT)
Rare disease registries that adhered to the EU RD Platform, located (or operating) in the Member States.

Record editor(s)
Simona MARTIN (Simona.martin@ec.europa.eu)
Leucio Antonio CUTILLO (Leucio-antonio.cutillo@ec.europa.eu)
Andri PAPADOPOULOU (Andri.papadopoulou@ec.europa.eu)
Agnieszka KINSNER-OVASKAINEN (Agnieszka.KINSNER-OVASKAINEN@ec.europa.eu)
Sandra LOURO CALDEIRA (Sandra.Caldeira@ec.europa.eu)
Ciaran NICHOLL (Ciaran.NICHOLL@ec.europa.eu)

1.4. Joint controllership

N/A

1.5. Processors

Processors are involved in the processing
Yes

The processor(s) is/are:
Internal:
Authorised JRC staff from the F1 “Health in Society” Unit
External organisation(s)/entity(ies)
Names and contact details of organisation(s)/entity (ies)
No

1.6. Notes

DPC Notes

> **From the JRC’s perspective**, acting as processor of the different rare disease registries in the Member States adhering to the EU RD Platform, the processing consists in:
> Pseudonym generation, pseudonym linkage and encrypted pseudonymised data transfer via the European Platform on Rare Disease Registration pseudonymisation tool.
>
> **From the Operation data controller’s**, the processing consists in:
> - at registry’s level, compute a Password Based Key Derivation Function PBKDF2 and subsequently encrypt through the web browser a patient’s first name, second name and date of birth, then send the encrypted data to the EU RD Platform’s pseudonymisation tool to get the patient’s pseudonym as a response;
> - at registry’s level, encrypt a pseudonymised patient’s medical data and transfer the encrypted pseudonymised data to a target registry via the EU RD Platform’s pseudonymisation tool.

DPO Notes
-

1.7. Data subjects and keywords

The data subjects that the record concerns
Citizens

2. PURPOSE AND DESCRIPTION OF THE PROCESSING

2.1. Description of the purpose of the processing

The JRC, in support of EU policies in the area of public health set up, maintains and develops the European Platform on Rare Disease Registration (EU RD Platform) to gather at central level scarce and widely fragmented rare disease patient data from across Europe.

In the context of the activities of the EU RD Platform, the JRC facilitates the pseudonymisation, linkage and transfer of encrypted pseudonymised data on individuals affected by rare diseases[1] ('rare disease data') provided by registries in EU Member States and between those registries. This includes rare disease data provided by rare disease registries that already joined or may adhere to EU RD Platform in the future.

JRC aims to facilitate a secured channel to compute a pseudonym, encrypt and transfer rare disease data exclusively among EU RD members.

The EU RD Platform pseudonymisation tool implemented by the JRC provides to the rare disease registries acting as data controllers and processors three main functionalities:

- pseudonym generation,
- pseudonym linkage and
- encrypted pseudonymised data transfer.

The **pseudonym generation functionality** allows a data controller to compute a pseudonym for a data subject without allowing any other party but the same data controller to re-identify the subject.

The **pseudonym linkage functionality** allows a data controller or processor to find other data controllers or processors processing the personal data related to a given local pseudonym.

The JRC applies a set of data sharing policies defined by each data controller or processor during the pseudonym generation phase. A data sharing policy states which data controller or processor can derive that the personal data associated with a particular data subject is processed by the data controller or processor that issued the policy.

The **encrypted pseudonymised data transfer functionality** allows a data controller or processor to transfer encrypted pseudonymised data to another target controller or processor (e.g. from a local registry adhering to the EU RD Platform to another local registry adhering to the EU RD Platform) for further processing.

2.2. Processing for further purposes

The purpose(s) for further processing
N/A

Safeguards in place to ensure data minimization
Pseudonymisation

2.3. Modes of processing

The mode of processing

1. Automated processing (Article 24)
    a. Computer/machine
        i. Any other
            A description of 'Any other'

[1] In line with applicable EU policy and scientific standards, a rare disease is defined as a disease or condition affecting no more than 5 in 10.000 persons in the European Union. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on Rare Diseases - Europe's challenges, COM (2008) 679 final.

JRC puts at disposal of registries a web based tool for generating pseudonyms, linking them, and transferring encrypted pseudonymised rare disease patients' data between rare disease registries located in the Member States. Before proceeding to the transfer of the data, the registries join the EU RD Platform, generate pseudonyms of their patients and encrypt data associated with a patient's pseudonym. Rare disease registries users log in into the system via EU Login and, once loaded the cryptographic archive file containing the public and private key pair associated to their registry, as well as the public key certificate issued by the certification authority, in their web browser local storage, they can access the pseudonym generation, pseudonym linkage and encrypted pseudonymised data transfer functionalities. Encryption and decryption operations are performed at the client side through the user's web browser, the personal data never leaves the browser in an unencrypted form. No other party other than the intended recipient can decrypt the data. The JRC does not store nor have access nor generate the private keys used to encrypt or decrypt the rare disease patients' data.

2. Manual processing
3. Any Other

   A description of 'Any other'

Description/additional information regarding the modes of processing

The EU Login username and email address of the authorized registry users that can access the system is stored at JRC and is used just to perform access control operations. Such information is deleted when the user communicates to quit the service.

1. Pseudonym generation

To implement this functionality, a rare disease registry computes locally a Password Based Key Derivation Function PBKDF2 of the data subject's first name, second name and date of birth, encrypts locally the output with an one-time random key provided by the JRC and obtains an encrypted identifier, and finally sends the encrypted identifier to the JRC. The JRC then computes a global and a local pseudonym from such encrypted data, creates an authorisation token containing the global pseudonym encrypted for the JRC itself, and sends it back to the requesting controller together with the local pseudonym without storing any information coming from the received, computed, or sent data.

The JRC cannot derive the data subject's identity neither from the corresponding local or global pseudonym, nor from the encrypted identifier. Even if the random one-time key is provided to the data controller by the JRC, the JRC does not decrypt the encrypted identifier and, moreover, the data subject's identity cannot be derived from the output of the PBKDF2 function, as those functions are non-invertible and designed to make it difficult for a computer to quickly compute them. The JRC can neither link a given local pseudonym to the corresponding global pseudonym without recurring to additional information, such as the relationship between a local pseudonym and an authorization token, which is generated by the JRC and is kept by the controller only. A data controller never communicates this information to the JRC, nor is required to do that. A data controller cannot decrypt the encrypted global pseudonym contained in the authorization token.

A data controller can communicate to the JRC the sharing policy associated to a given data subject by sending the policy, the authorization token, and the centre identifier of the data controller. In this case, the JRC stores the association between the global pseudonym contained in the token, the encrypted controller identifier and the sharing policy until the data controller deletes it.

2. Pseudonym linkage

To implement this functionality, the JRC receives an authorisation token from the data controller, accesses the global pseudonym contained in the token, then retrieves all the encrypted data controller and processor identifiers associated with the same global pseudonym from its local storage, and computes a shared pseudonym for each found identifier. Then, the JRC builds a list containing, in each row, the shared pseudonym and the corresponding data controller or processor identifier and sends it back to the requesting controller or processor.

The computed shared pseudonyms may be cached by the JRC servers to speed up the response time. It is not possible to derive the data subject's identity from a shared pseudonym.

While computing a shared pseudonym and filling the lists, the JRC applies a set of data sharing policies defined by each data controller or processor after the pseudonym generation phase. A data sharing policy states which data controller or processor can derive that the personal data associated with a particular data subject is processed by the data controller or processor that issued the policy. By default, no third controller or processor can link its local pseudonyms to those of another controller or processor.

3. Encrypted pseudonymised data transfer

The pseudonymised data transfer takes place through an end-to-end encrypted channel: the data is encrypted locally by the sender, sent to the JRC servers and stored here until it is forwarded to the intended recipient.

Once this data is delivered, the recipient can try to link the pseudonymised data with other datasets across different sources or to transfer it again. The JRC prevents the recipient from performing such actions if not authorized by the original controller through the data sharing policies specified during the pseudonym generation phase.

When an encrypted pseudonymised data transfer takes place to share information about a mutual data subject, a shared pseudonym is used as pseudonym value in the transfer. In all the other cases, the hash of the local pseudonym is used as pseudonym value in the transfer. In those cases, the authorization token is also sent to allow the recipient to find information on additional data controllers and processors processing personal data belonging to the same subject, according to the defined data sharing policies. In order to guarantee data integrity, data confidentiality and sender authentication, the JRC implements a Public Key Infrastructure and provides each data controller or processor with a public key certificate.

Therefore, each controller or processor can generate their own private and public cryptographic key pair, obtain a public key certificate signed by the EU RD Platform pseudonymisation tool Certification Authority, and use such keys to sign, encrypt and decrypt the transferred pseudonymised data.

As a consequence, the signed and encrypted pseudonymised data transiting the JRC cannot be tampered by any party in the communication path, nor be decrypted by anyone other than the intended recipient, as the private key is generated and known by such recipient and nobody else.

The JRC can also act as an encrypted data storage and allow a data controller for working on the same dataset from multiple devices. In this case, a device can download the encrypted data from the JRC storage, decrypt it locally, apply modifications, sign and encrypt it again and store it safely on the same storage.

Such storage is also used to temporarily store encrypted signed pseudonymised data when a data controller or processor recipient is not online. In this case, the data will be delivered when the recipient connects to the system and deleted immediately after.

When using the pseudonym generation, pseudonym linkage and encrypted pseudonymised data transfer functionalities, a data controller must ensure the lawfulness of the processing as stated in Art.6 of the Regulation (EU) 2016/679. The JRC processes the data on behalf of a data controller or on behalf of a data processor with prior specific or general written authorization of the controller.

2.4. Storage medium

The medium of storage (one or more)
Description/additional information regarding the storage medium
The encrypted pseudonymised data for which the transfer is facilitated via the EU RD Platform is stored in the JRC servers until it is forwarded to the intended recipient, and is then automatically deleted.

The encrypted data that is kept in the JRC encrypted data storage service to facilitate the user access from multiple users belonging to the same registry and the synchronization of data from multiple devices is stored in the JRC servers until all users belonging to such registry communicate to quit the service.

A global pseudonym generated by the system is stored in the JRC servers until all the users that have defined a data sharing policy linked to that global pseudonym delete such policy.

A shared pseudonym generated by the system is cached in the JRC servers until a user modifies the data sharing policy that allowed to generate such shared pseudonym.

The JRC servers are implemented as Docker containers and store the above mentioned data in their local H2 database. Since data is already encrypted, the database does not apply further encryption measures.

An automated full backup of the databases is performed every day as a standard service by the JRC I.5. Unit.

2.5. Comments

Comments/additional information on the data processing
The data processing is always carried out by authorized JRC staff in the F1 Unit on behalf of the rare disease registries that act as data controllers or data processors (in accordance with Art.29 of Regulation (EU) 2018/1725).

Pseudonymisation and Encryption

The data processed by the JRC is always encrypted in such a way that the JRC cannot have access to the rare disease patients' identities nor their data.

Where a registry has information which would allow the JRC to access a rare disease patient's data and/or to attribute such data to a specific data subject, such information shall not be made accessible to the JRC. The JRC shall not request the rare disease registry to provide such information.

Professional secrecy

JRC staff having access to the JRC servers performing the data processing shall be subject to a duty of professional secrecy.

In accordance with Art. 13 of Regulation (EU) 2018/1725, the JRC shall ensure that the following measures are observed to safeguard the rights and freedoms of data subjects:

Access limitation and staff awareness

Access to encrypted pseudonymised rare disease data within JRC shall be limited to selected staff having been formally authorized by the JRC management. JRC shall put in place regular awareness-raising measures among authorized staff on their duties with regard to the level of security to be ensured.

Other technical and organizational measures

The processing of rare disease data by JRC shall be subject to the requirements set out in Commission Decision COM (2017)/46[2] and its implementing rules[3] on the security of communication and information systems in the European Commission. Compliance with these rules shall provide appropriate safeguards in regard to personal data in full compliance with the data protection rules applicable to the Commission.

In particular, the following measures shall be adopted:

- standard mandatory measures for access control and authentication, IT asset management, backups, business continuity, compliance, control against malicious code;
- information security and risk management;
- encryption, logging and monitoring of access, management of vulnerabilities from removable media, physical and environmental security, secure systems development, IT vulnerability and remediation management, web applications security standards;
- the system compliance shall be checked by a risk management process followed by the Local Informatics Security Officer (LISO).

3. DATA SUBJECTS AND DATA CATEGORIES

3.1. Data subjects' categories

Data subject(s) are:
External to the organization

A description of the data subjects (external to the organization)
Citizens suffering from rare diseases.

3.2. Data categories/fields

3.2.1. General description of the data categories

Description of the categories of data that will be processed

Encrypted rare disease patients' identity data (first name, second name, date of birth) for the purposes of computing a pseudonym for such subjects.
Encrypted pseudonymised personal data (including encrypted pseudonymised special categories of personal data) for the purpose of delivering such data to the intended registry.

[2] Commission Decision (EU, Euratom) 2017/46 of 10 January 2017 on the security of communication and information systems in the European Commission, C/2016/8998, OJ L 6, 11.1.2017, p. 40–51

[3] Commission Decision (EU, Euratom) 2018/559 of 6 April 2018 laying down implementing rules for Article 6 of Decision (EU, Euratom) 2017/46 on the security of communication and information systems in the European Commission, C/2018/1726, OJ L 93, 11.4.2018, p. 4–10

Encrypted personal data (including encrypted special categories of personal data) for the purpose of storing it to make it available to the same sender when connecting from another device or to another authorized user belonging to the same data controller or processor organization.

For further information on the privacy policy of the European Platform on Rare Disease Platform, please refer to record “DPR-EC-01972 - European Platform on Rare Disease Registration”.
EU Login usernames belonging to the users that are allowed to access the system for the purposes of performing access control. For further information on EU Login please refer to the dedicated processing operation “DPR-EC-03187: Identity & Access Management Service (IAMS)”.

3.2.2. Special categories of personal data

The processing operation concerns any 'special categories of data' which fall(s) under Article 10(1), which shall be prohibited unless any of the reasons under Article 10(2) applies:
Yes, the processing concerns special categories

The processing operation concerns one or more of the following special categories of data:

- Data concerning health

A general description of the special categories of data being processed
The EU Platform on Rare Disease Registration includes data concerning health, a special category of data that falls under Article 10 of Regulation (EU)2018/1725. The health data include medical descriptions, medical diagnosis and tests results, medical classifications, and socio-demographic data.
Such data is processed in an encrypted form on behalf of the rare disease registries that have joined the EU RD Platform, and therefore cannot be interpreted by the JRC.

The reasons under Article 10(2) allowing the processing of the special categories of data
N/A

Description/additional information regarding special categories of personal data
In accordance with Art. 10.2(j) of Regulation (EU) 2018/1725 (henceforth ‘EUDPR’), [the prohibition in principle to process special categories of data in] paragraph 1 shall not apply if […] the processing is necessary for […] scientific […] research purposes […] based on Union law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.

3.3. Data related to ‘criminal convictions and offences’

The data being processed contains sensitive data which fall(s) under Article 11 'criminal convictions and offences'
N/A

4. RETENTION PERIOD

4.1. The administrative time limit(s) for keeping the personal data per data category

1. **Data category:**
   All data categories

Retention period:

The JRC will act as processor of the different rare disease registries in the Member States and will follow their specific instruction on the retention policy of their data. Nevertheless a general and common retention will be adopted as follows:

- the encrypted pseudonymised data transferred from a registry to another is kept until it is forwarded to the intended recipient, and is then automatically deleted;
- the encrypted data that is stored to provide the data synchronization service is kept until all users belonging to the same registry communicate to quit the service;
- a global pseudonym is kept until all the users that have issued data sharing policies associated to that global pseudonym delete such policies;
- a shared pseudonym is cached until a user modifies the data sharing policy that allowed to generate such shared pseudonym;
- an EU Login username and the corresponding email address are kept until the user associated to such username and email address communicates to quit the service.

5. RECIPIENTS

5.1. Origin of the recipients of the data

The origin of the data recipients
Within the EU organization
A description of the indicated recipients of the data
Authorized JRC staff in the F.1 “Health in Society Unit”, who is performing EU RD Platform ICT services maintenance, as well as JRC I.5. “Advanced Computing & ICT Support” staff who is performing JRC ICT maintenance can have access to the encrypted pseudonymised personal data as well as to the global and shared pseudonyms. It is not possible for such recipients to retrieve the data subject identity nor their personal data from such encrypted information.
Authorized EU RD Platform users having access to the EU RD Platform pseudonymisation tool can receive encrypted pseudonymised personal data in case they belong to the intended recipient registry. They can decrypt this encrypted data locally and derive the corresponding pseudonymised data.
In case the pseudonymised data is associated to a shared pseudonym, and in case this shared pseudonym is linked to a local pseudonym related to a data subject whose identity is known by the recipient, then the retrieved pseudonymised data can be referred to a specific data subject by the recipient. In all the other cases it is not possible for a recipient to retrieve the data subject identity from the pseudonyms associated to the pseudonymised data.

Outside the EU organisation
A description of the indicated recipients of the data

The JRC will act as processor of the different rare disease registries in the Member States and will follow their specific instructions provided in the data sharing policy.

5.2. Categories of the data recipients

The categories (one or more) of the data recipients

- A natural or legal person
- Public authority
- Agency

Description of the indicated category (ies) of data recipients

Within the JRC: authorized staff in the F1 Unit have access to, but cannot decrypt, the encrypted pseudonymised data that is stored and forwarded to the intended recipient, as well as the encrypted data that is stored for data synchronization purposes.

Recipients outside the JRC: Investigators involved in a research study receive pseudonymised data that contain only selected variables necessary to perform the study, as well as the associated shared or hash of local pseudonym value.

Who has access to which parts of the data

Within the JRC, the encrypted pseudonymised personal data can be accessed by the JRC authorized staff in the F1 "Health on Society Unit" only, but cannot be decrypted.

Outside the JRC, all registries adhering to the EU Platform on Rare Disease Registration that decide to use the pseudonymisation tool have access to pseudonymised data exclusively for scientific research purposes. JRC may transfer, on behalf of such registries, rare disease data to researchers belonging to other registries adhering to the EU RD Platform using the same tool for the purpose of scientific studies, subject to the provisions of Art. 9 of Regulation (EU) 2018/1725. In addition to such provisions, the approval of the rare disease registry at the origin of the rare disease data to be disclosed shall be required. The data shall be used by the recipient exclusively for scientific research purposes. The JRC enforces this constraint in the legal act governing the processing that is signed in accordance with Art.29 n.3 of Regulation (EU) 2018/1725 and Art.28 n.3 of Regulation (EU) 2016/679.

6. INTERNATIONAL DATA TRANSFERS

6.1. Transfer outside of the EU or EEA

Data is transferred to countries outside the EU or EEA

Yes

Countries to which data is transferred

Following data controller's instruction, the data can be transferred to United Kingdom, other non EU/EEA countries. Several members of the EU Platform on Rare Diseases Registration are from the UK, so it is foreseen that in the future the EU RD Platform will receive data requests for new studies from these researchers. Following Commission implementing decision of 28.06.2021, UK has an adequate level of protection of personal data. In addition, researchers from other non EU/EEA countries might also adhere to the platform and transfer data through the pseudonymisation tool.

6.2. Transfer to international organisation(s)

Data is transferred to international organisation(s)

N/A, transfers do not occur and are not planned to occur

6.3. Legal base for the data transfer

The legal base for the data transfer

1. Transfer on the basis of the European Commission's adequacy decision (Article 47)
2. Transfer subject to appropriate safeguards (Article 48.2 and .3)
    a. Standard data protection clauses adopted by
        i. The Commission
3. Subject to the authorization from the European Data Protection Supervisor
    a. Contractual clauses between the entity of the operational controller or processor and the entity of the operational controller, processor or the recipient of the personal data in the third country or international organization
    b. Administrative arrangements between public authorities or bodies which include enforceable and effective data subject rights

6.4. Derogations for specific situations (Article 50.1 (a) - (g))

Are there derogations for specific situations (Article 50.1(a)-(g))
N/A

6.5. Comments

Comments/additional information on international data transfers
All the encrypted pseudonymised personal data transfers, including the international ones, are performed by the users of the rare disease registries that adhered to the EU Platform on Rare Diseases Registration and decided to use the pseudonymisation tool provided by the EU RD Platform. The JRC, which processes the personal data on behalf of such registries, cannot perform a data transfer without an explicit instruction of those registries which is communicated via the authorised registry users.

7. INFORMATION TO DATA SUBJECTS ON THEIR RIGHTS

7.1. Privacy statement

Rights of the data subjects

The processing performed by the JRC on behalf of the registries shall ensure the respect of the following rights of data subjects:

- Article 17 - Right of access by the data subject
- Article 18 - Right to rectification
- Article 19 - Right to erasure (right to be forgotten)
- Article 20 - Right to restriction of processing
- Article 21 - Notification obligation regarding rectification or erasure of personal data or restriction of processing
- Article 22 - Right to data portability
- Article 23 - Right to object
- Article 24 – Rights related to automated individual decision making, including profiling

The data subjects are informed about their rights and how to exercise them in the form of a privacy statement attached to this record

N/A, the information about the data subject’s rights must be provided by the data controller.

While providing the functionalities of the EU RD Platform pseudonymisation tool, the JRC processes the personal data on behalf of the registries adhering to the EU Platform on Rare Diseases Registration, and therefore acts as a data processor.

Publication of the privacy statement

N/A

Guidance for Data subjects which explains how and where to consult the privacy statement is available and will be provided at the beginning of the processing operation.

N/A

8. SECURITY MEASURES

8.1. Short summary of overall Technical and Organisational measures implemented to ensure Information Security

All processing operations are carried out pursuant to the Commission Decision (EU, Euratom) 2017/46 of 10 January 2017 on the security of communication and information systems in the European Commission, and follow the Information Security Policy and Internal Rules for handling ICT Information Security Incidents and the Commission Information Systems Security Policy C(2006)3602. Standard mandatory measures are applied for access control and authentication, IT asset management, backups, business continuity, compliance, control against malicious code, information security and risk management, cryptography, logging and monitoring, removable media, physical and environmental security, secure systems development, IT vulnerability and remediation management, web applications security standards. Individual records are stored in an encrypted form in the Docker containers running in the JRC Continuous Integration/Continuous Delivery infrastructure. The system compliance was checked by a risk management process followed by the Local Informatics Security Officer (LISO).