EUROPEAN UNION AGENCY
FOR CYBERSECURITY

DEPLOYING PSEUDONYMISATION TECHNIQUES

The case of the Health Sector

MARCH 2022

ABOUT ENISA

The European Union Agency for Cybersecurity, ENISA, is the Union's agency dedicated to achieving a high common level of cybersecurity across Europe. Established in 2004 and strengthened by the EU Cybersecurity Act, the European Union Agency for Cybersecurity contributes to EU cyber policy, enhances the trustworthiness of ICT products, services and processes with cybersecurity certification schemes, cooperates with Member States and EU bodies, and helps Europe prepare for the cyber challenges of tomorrow. Through knowledge sharing, capacity building and awareness raising, the Agency works together with its key stakeholders to strengthen trust in the connected economy, to boost resilience of the Union's infrastructure, and, ultimately, to keep Europe's society and citizens digitally secure. More information about ENISA and its work can be found here: www.enisa.europa.eu.

CONTACT

For contacting the authors please use isdp@enisa.europa.eu
For media enquiries about this paper, please use press@enisa.europa.eu.

CONTRIBUTORS

Fabio Guasconi (Bl4ckswan), Pantelis Angelidis (UOWM & Vidavo), Prokopios Drogkaris (ENISA)

EDITOR

Prokopios Drogkaris (ENISA)

ACKNOWLEDGEMENTS

We would like to thank the ENISA eHealth Security Expert Group[1] members Fotios Gioulekas, Merja Ikäheimonen, Konstantinos Chondropoulos, Ben Kokx, Martha De Cunha Maluf-Burgman, Marta Carbonell Cobo and the ENISA colleagues Athanasios Drougkas and Athena Bourka for their valuable comments and review of the document.

LEGAL NOTICE

This publication represents the views and interpretations of ENISA, unless stated otherwise. It does not endorse a regulatory obligation of ENISA or of ENISA bodies pursuant to the Regulation (EU) No 2019/881.

ENISA has the right to alter, update or remove the publication or any of its contents. It is intended for information purposes only and it must be accessible free of charge. All references to it or its use as a whole or partially must contain ENISA as its source.

Third-party sources are quoted as appropriate. ENISA is not responsible or liable for the content of the external sources including external websites referenced in this publication.

Neither ENISA nor any person acting on its behalf is responsible for the use that might be made of the information contained in this publication.

ENISA maintains its intellectual property rights in relation to this publication.

[1] https://resilience.enisa.europa.eu/ehealth-security

COPYRIGHT NOTICE

ISBN 978-92-9204-576-0, DOI 10.2824/092874

TABLE OF CONTENTS

EXECUTIVE SUMMARY

As the healthcare domain is attempting to make the most of the evolving technical landscape and adapt the provision of services to fulfil the growing needs of patients in a timely manner, additional cybersecurity and data protection challenges come into play. The integration of new technologies in already complex IT infrastructures opens up new challenges regarding data protection and cybersecurity.

This is due to the growing need to exchange and share the health related information of individuals among different stakeholders. It is therefore essential for the entities processing personal data, on the one hand, to collect and further process only data that are necessary for their purposes and, on the other hand, to employ proper organisational and technical measures for the protection of such personal data.

Pseudonymisation is increasingly becoming a key security technique for providing a means that can facilitate personal data processing, while offering strong safeguards for the protection of personal data and thereby safeguarding the rights and freedoms of individuals.

Complementing previous work by ENISA that is relevant, this report demonstrates how pseudonymisation can be deployed in practice to further promote the protection of health data during processing. Obviously, there is not a single solution on how and when to apply it; in fact different solutions might provide equally good results in specific scenarios, depending on the requirements in terms of protection, utility, scalability, etc.

Pseudonymisation can be a ‘simple’ option to adopt but it can also be comprised of a very complex process, both at technical as well as at organisational levels. For this reason, defining the goals and objectives of pseudonymisation in each particular case and processing operation is really important.

This report highlights the added value of pseudonymisation in the healthcare sector and demonstrates its applicability through simple but specific use cases. Complementing relevant ENISA publications in this area, it shows how such techniques can increase the level of protection for personal data being processed in the healthcare domain and will eventually promote and raise awareness on the usability and deployment of such technical measures.

1. INTRODUCTION

Recent decades have witnessed an accelerating pace in the development and adoption of new technologies. This rapid technological change has also affected the healthcare sector which is going through the digitalisation process and has continuously been adopting new technologies to improve patient care, offer new services focusing on patient-at-home care and even preventive schemes.

The integration of new technologies into already complex IT infrastructures opens up new challenges regarding data protection and cybersecurity as there is an increasing need to exchange and share the health related information of individuals among different stakeholders, in some cases across countries, in order to provide better health services. It is therefore essential for the entities processing personal data to collect and further process only data that are necessary for their purposes and, in addition, to employ proper organisational and technical measures for the protection of such data.

The integration of new technologies in healthcare opens up new challenges regarding data protection and cybersecurity.

Pseudonymisation is one well-known measure that can significantly contribute to this end.

Broadly speaking, pseudonymisation aims at protecting personal data by hiding the identities of individuals in a dataset, e.g. by replacing one or more personal identifiers[2] with the so-called pseudonyms (and appropriately protecting the link between the pseudonyms and the initial identifiers).

This process is not at all new in the design of information systems but gained special attention after the adoption of the General Data Protection Regulation (GDPR) [1], where pseudonymisation is explicitly referred as a technique which can both promote data protection by design (Article 25 GDPR), as well as the security of personal data processing (Article 32 GDPR).

1.1 DIGITAL TRANSFORMATION OF THE HEALTH SECTOR

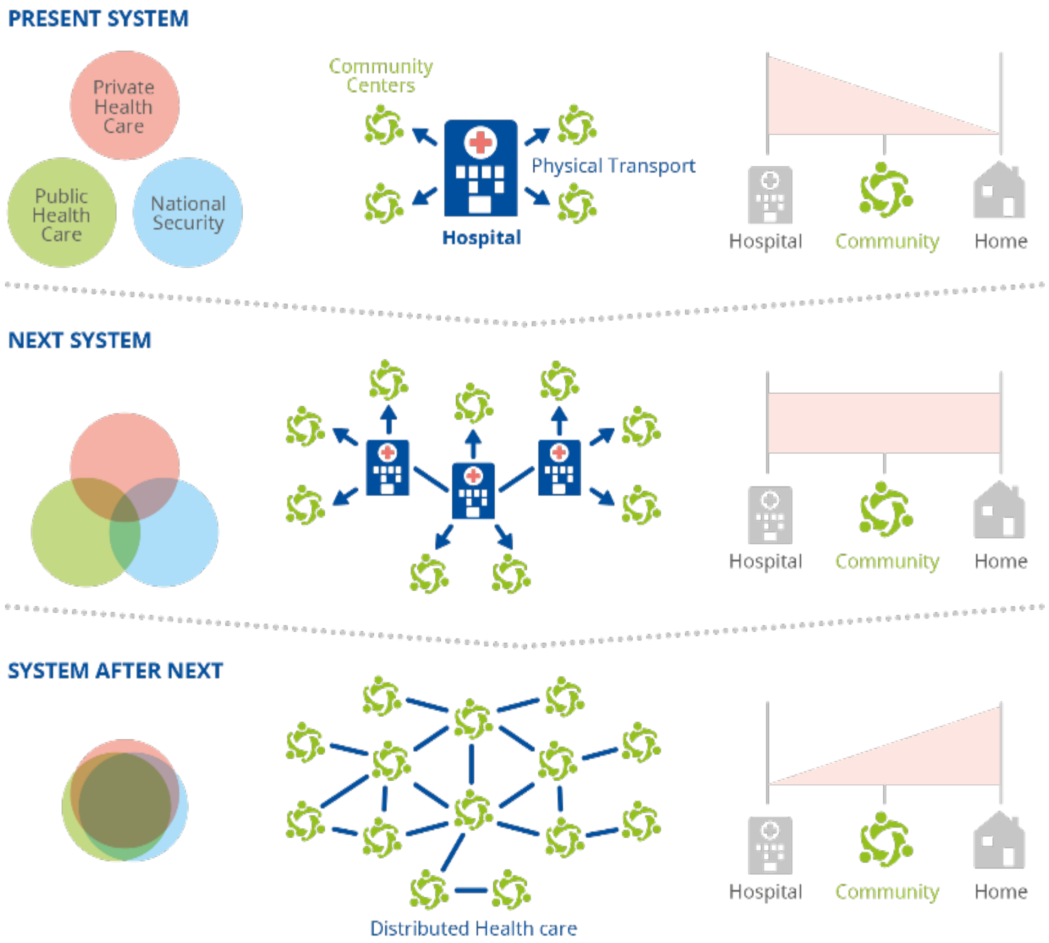
Health data has always been a valuable source of knowledge in healthcare. The healthcare domain has historically generated vast amounts of data, both for the treatment of patients and for research and further analysis. Such processing was mostly performed in paper form but over the last few decades, the accessibility and amount of digitized data has increased massively.

More recently an abundance of new sources of health data occurred as a result of the widespread use of electronic health records, health applications and wearable devices [2]. Furthermore, advances in computational power have enabled the development of novel data analytics and machine learning techniques that improve diagnostics, treatment and administration in healthcare.

The result is a change in assumptions that is increasingly moving the patient away from hospitalization towards a distributed healthcare system provided by a blend of public and private operators while staying closer to home, as depicted in Figure 1 below.

[2] An identifier allows for the identification, directly or indirectly, of an individual. Examples of such identifiers can be name, address, date of birth, national identification number, social security number, etc.

Figure 1: Digital transformation induced shift of value in healthcare [3]



These technological advances are creating a growing demand for big medical data banks to offer solutions in, for example, disease diagnosis and phenotyping, the modelling of clinical outcomes, prediction for early intervention strategies, precision medicine etc. However, the increasing processing of digitised medical data has also increased the risks, in terms of cybersecurity, data protection and the likelihood of data breaches.

Such risks and related threats have already been identified in relevant ENISA publications in the area of eHealth security [4], [5] & [6]. Relevant EU legal instruments such as the NIS Directive [7], the General Data Protection Regulation [1], the Medical Devices Regulation [8], the Directive on the application of patients' rights in cross-border healthcare [9] etc. imposed obligations on healthcare providers and manufactures of medical devices to ensure an adequate and uniform level of protection for medical data and the products and services that use them.

1.2 PROTECTING HEALTH DATA

The protection of health data is considered a high priority due to their sensitive nature and the impact they have on individuals (data subjects). From a cybersecurity point of view, the confidentiality, availability and integrity of medical data and relevant infrastructure is considered essential in order to be able to provide timely, appropriate and uninterrupted medical care.

This is also highlighted by the NIS Directive [7] which categorizes the health sector as an operator of essential service (OES) and calls for minimum security requirements to ensure a level of security appropriate to the level of risks presented. Furthermore, the GDPR distinguishes, in Art. 9, data concerning health as a special category of data (sensitive) and sets out additional requirements and stricter obligations for processing and protecting such data, in order to safeguard the rights and freedoms of individuals (data subjects). Lastly, the Medical Devices Regulation imposes requirements regarding the safety, quality and security of medical devices in order to achieve a high common level for safety.

1.3 SCOPE – TARGET AUDIENCE

The scope of this report is to highlight the added value of pseudonymisation in the healthcare sector and demonstrate its applicability through simple but specific use cases. Complementing relevant ENISA publications in the area [10], [11], [12] the goal is to present how such techniques can increase the level of protection for personal data being processed in the healthcare domain and eventually promote and raise awareness on the usability and deployment of such technical measures.

This document is meant for IT professionals and developers in the healthcare domain as well as Healthcare Authorities that can put forward recommendations on the security of health data processing.

1.4 STRUCTURE OF THE DOCUMENT

Section 2 provides an introduction of pseudonymisation and discusses its advantages and the most common techniques being used. Section 3 exhibits the use of pseudonymisation in three use cases which focus on the collection of health data from patients and the processing of such data by healthcare providers and medical research institutions. Lastly, Section 4 concludes this report and provides the main conclusions and relevant recommendations.

2. PSEUDONYMISATION

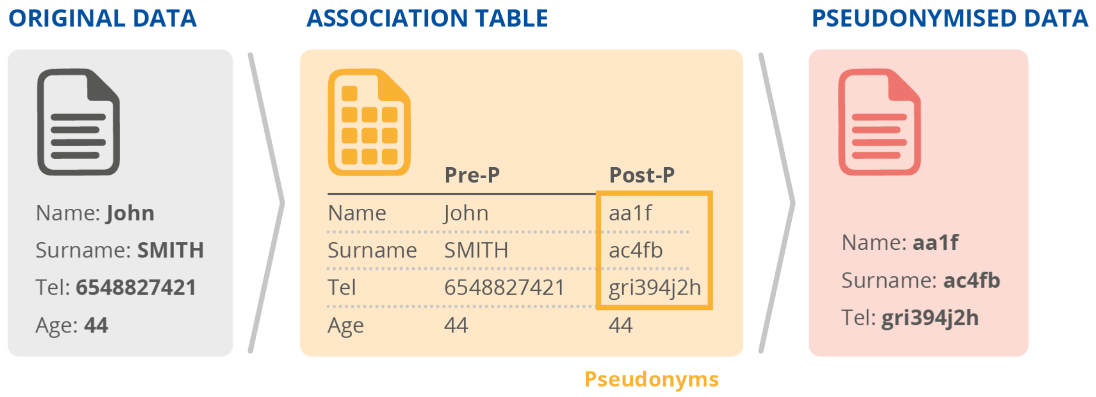
2.1 BACKGROUND

The GDPR defines pseudonymisation in Article 4 (5) as: "*the processing of personal data in such a manner that the personal 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 that the personal data are not attributed to an identified or identifiable natural person*".

Broadly speaking, pseudonymisation aims at protecting personal data by hiding the identity of individuals (data subjects) in a dataset, e.g. by replacing one or more personal data identifiers with the so-called pseudonyms and appropriately protecting the link between the pseudonyms and the initial identifiers. This link, often referred to as the pseudonymisation secret, is usually stored in the association table and can be used re-identify the individual by allowing the pseudonyms and original data to be associated as depicted in Figure 2 below. The association table is meant to be accessible only by the entity that initially performed the pseudonymisation, which is often referred to as the pseudonymisation entity.

Pseudonymisation is a well-established technique that aims to protect personal data by hiding the identity of individuals.

Figure 2: Pseudonymisation example



Pseudonymisation is one of several 'de-identification' techniques (such as aggregation, obfuscation, masking, etc.) intended to remove the association between a set of identifying data and the data principal. Other pseudonymisation definitions such as the ones from ISO, and specifically ISO 25237:2017 Health informatics — Pseudonymization [13], are built upon this assumption and are similar to the GDPR definition discussed above.

2.2 IMPORTANCE OF PSEUDONYMISATION

The main advantage of pseudonymisation, when applied correctly, is to hide the identity of an individual in the context of a specific dataset, so that it is not possible to connect the data with the specific individual. Therefore, it can also reduce the risk of the linkage of personal data for a specific individual across different data processing domains.

In this way, for example, should a breach of personal data occur, pseudonymisation increases the level of difficulty encountered by any third party other than the data controller to correlate the breached data with certain individuals without the use of additional information. Importantly, it can also reduce the level of risk to which the pseudonymised data are exposed.

The difference lies in knowing that our exemplified 'John SMITH' in Figure 2 above suffers from a chronic disease and that a person identified as 'aa1f ac4fb' suffers from a chronic disease.

Even if pseudonymised data is leaked, a third-party company will not be able to make Mr. John SMITH a target of marketing campaigns based on that information, unless the information related to the mechanism for associating pseudonyms to data subjects is also compromised.

The use of pseudonymised data should, for example, give more confidence to data subjects when consenting to the use of their health data for research purposes and even increase the effectiveness of other security controls to be applied to non-pseudonymised data such as, for instance, encryption.

Lastly, it should be noted that pseudonymisation and anonymisation do not refer to the same process and that pseudonymous data are considered by the GDPR as personal data while anonymous data are not[3].

2.3 BASIC PSEUDONYMISATION TECHNIQUES

As already explored in [10], [11] & [12] there are several pseudonymisation techniques. The main differences between them are based on how the pseudonym is generated. For the most common ones, Table 1 below provides a comprehensive summary.

Table 1: Overview of basic pseudonymisation techniques

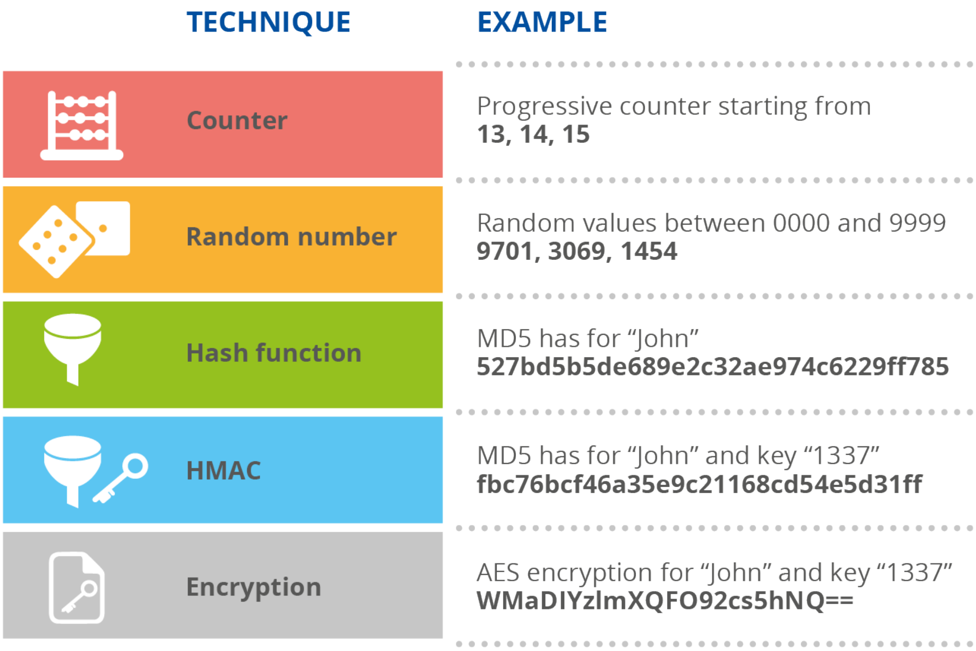
Technique	Pseudonym Generator
Counter	Monotonic counter which starts at a certain value and is increased each time a new pseudonym is necessary
Random number	Random value extracted between a minimum and a maximum boundary each time a new pseudonym is necessary
Hash function	One-way (non-reversible) cryptographic function transforming input personal data in fixed-length values
Hash-based message authentication code (HMAC)	One-way (non-reversible) cryptographic function adding a key that makes it less predictable than a hash function
Encryption	Two-way (reversible) cryptographic function transforming an input personal data in values that can be re-transformed in its original format using a key

As mentioned in [10], while a hash function can significantly contribute towards data integrity, it is generally considered weak as a pseudonymisation technique as it is prone to brute force and dictionary attacks. Similarly counters are also considered as weak pseudonymisation technique as they cannot really scale.

A robust approach to generate pseudonyms can be based on the use of keyed hash functions as depicted in Figure 8 – i.e. hash functions whose output depends not only on the input but also on a secret key (salt). Some practical examples of the application of the aforementioned techniques are presented in Figure 3 below:

[3] Additional information on anonymisation techniques can be found at Working Party 29 Opinion 05/2014 on Anonymisation Techniques

Figure 3: Application of basic pseudonymisation techniques



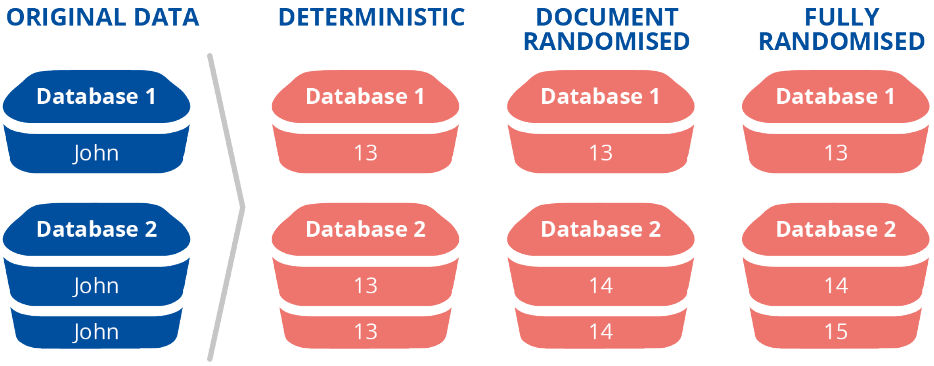
TECHNIQUE	EXAMPLE
Counter	Progressive counter starting from **13, 14, 15**
Random number	Random values between 0000 and 9999 **9701, 3069, 1454**
Hash function	MD5 has for "John" **527bd5b5de689e2c32ae974c6229ff785**
HMAC	MD5 has for "John" and key "1337" **fbc76bcf46a35e9c21168cd54e5d31ff**
Encryption	AES encryption for "John" and key "1337" **WMaDIYzlmXQFO92cs5hNQ==**

The difference between the last three techniques depicted in Table 1 and Figure 3 may not be immediately appreciable but it is nevertheless substantial as regards their implementation, as what is also of primary importance is choosing the extent and the approach, or in other words, the policy, with which those techniques will be applied. Most commonly, the policy could imply either:

1) **Deterministic pseudonymisation** – always using the same pseudonym for the same data;
2) **Document randomised pseudonymisation** – using the same pseudonym for the same data only within a consistent scope;
3) **Fully randomised pseudonymisation** – always using a different pseudonym for the same data.

An application of the policies to the Counter technique, discussed earlier, is exemplified in Figure 4, illustrating three occurrences of the same personal data in two different databases:

Figure 4: Application of pseudonymisation policies



In practice there are more advanced methods and, since it is a topic of wide interest, techniques are also evolving. But, generally speaking, policies tend to combine the most common techniques depicted above or to introduce a number of variations to them.

2.4 PSEUDONYMISATION CONSIDERATIONS

Data controllers and processors can make use of pseudonymisation techniques and related policies either jointly – using the same criteria – or disjointly. The first case is usually driven by the data controller requiring the data processor to adopt the same criteria. Other than the use of criteria, the most relevant difference would be the sharing of the association table depicted in Figure 2**.**

If this table is not shared, the entity receiving pseudonymised data, provided all mechanisms are correctly implemented, would not have any means of retrieving the original personal data. If the table is shared then it should be done with similar care given to exchanging encryption keys, with differences possibly due to the greater volume of data.

A data controller may ask its processors to pseudonymise personal data or even how to do it, especially if pseudonymised data is to be exchanged between the parties. Those specifications should at least include the following elements and should always be based on the results of previously conducted assessments of risk or impact:

- the target **personal data** (e.g. a set of identifiers);
- the **technique** to be used;
- the **parameters** applicable to the technique (e.g. counter rationale, randomness management, employed algorithms, key lengths);
- the **policy** to be used.

Depending on the applicable requirements, most likely including regulations, speed, simplicity, predictability and budget, the technique and related parameters could vary.

3. USE CASES

In an attempt to demonstrate the added value of pseudonymisation in the healthcare domain, this section presents three use cases where the personal medical data being processed are pseudonymised. Even if specific pseudonymisation techniques are not analysed in depth, the different use cases attempt to provide an overview of the possibilities and of the key aspects of their application, in terms of increasing the level of protection of the personal data being processed by removing direct personal identifiers.

It should be noted that these use cases serve the purpose of demonstrating the application of pseudonymisation and do not attempt to cover operational use cases to their full extent. In such scenarios a more in depth analysis of the context, the data processors and processing involved and all the relevant processing operations would have to be taken into account.

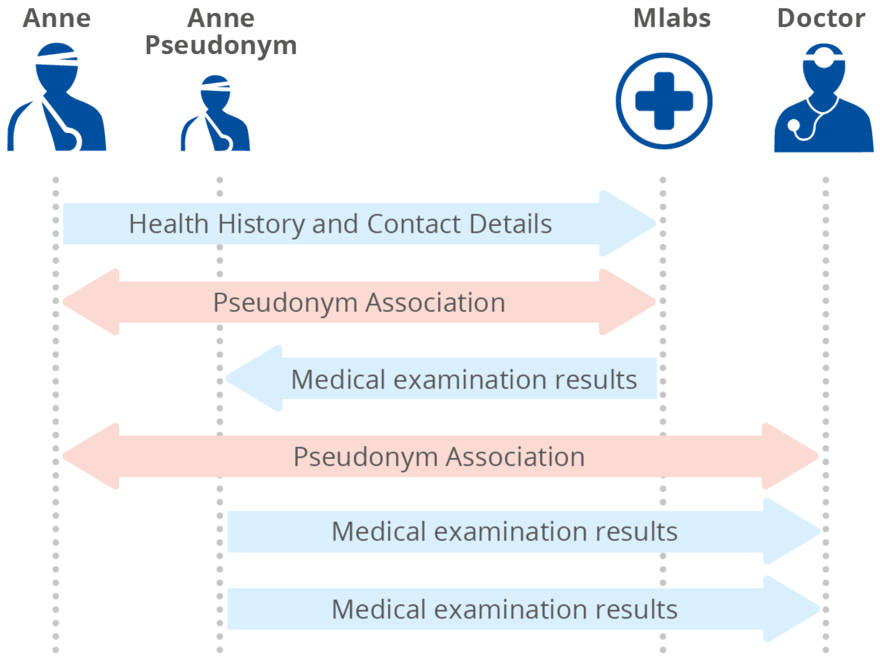
3.1 EXCHANGING PATIENT’S HEALTH DATA

In current medical practice, the exchange of data between organisations is a common tactic, and is mainly used for diagnostic and therapeutic purposes. This includes cases of exchange between different departments within the same entity (e.g. a hospital) and exchange between individuals (e.g. medical professionals, laboratories, etc).

For example, Anne undertakes a set of medical laboratory tests in Mlabs medical laboratory which then need to be reviewed by her doctor. When Anne visits Mlabs for the first time she is asked to provide background health information (medication, medical history etc.) and personal details (name, contact details, social security number) and Mlabs assigns her a Patient_ID. From this point onwards, all data, such as results, etc.) related to a medical examination are being associated with this Patient_ID and not with her personal details, as depicted in Figure 6. The addition of a Patient ID, which can even be a simple progressive registration COUNTER, is used to decouple Anne’s personal details from the results and acts as a simple but effective pseudonym.

The data being transferred from Mlab to the treating doctor do not contain Anne’s personal identifiers but only the pseudonym and the medical test results.

Figure 5: Exchange of patient's health data

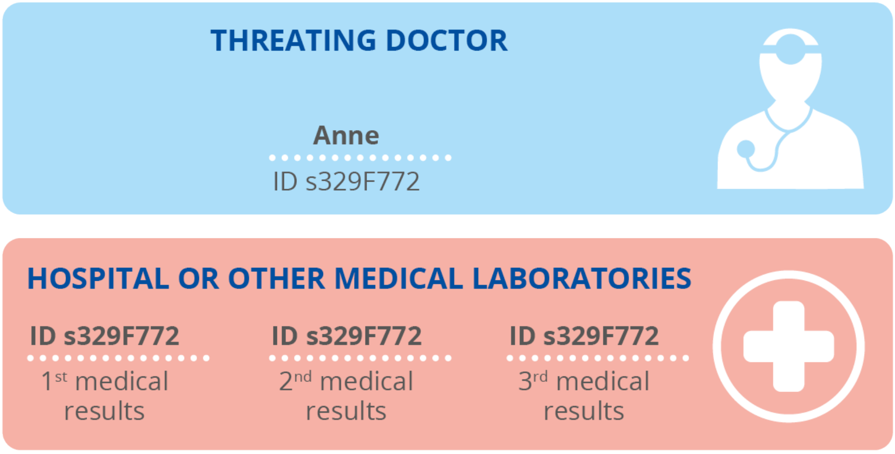

Anne
Anne Pseudonym
Mlabs
Doctor
Health History and Contact Details
Pseudonym Association
Medical examination results
Pseudonym Association
Medical examination results
Medical examination results

Upon completion of the medical examinations, Mlabs will correlate the lab results with her Patient_ID rather than her personal details. When Anne requests a copy of her results, Mlabs will have to look up the Patient_ID she has been assigned and then perform the correlation, as Patient_ID and personal details are stored separately.

Should she ask that the results be sent directly to her treating doctor, who will be responsible for reviewing them and deciding on whether additional tests or any medication is needed, Mlabs can inform the treating doctor of the results for the Patient_ID that Anne has been assigned. The association between Anne and Patient_ID needs to be communicated to the treating doctor only once; afterwards the doctor will already know that the specific Patient_ID refers to Anne. At this point though, the treating doctor gains access to the pseudonymisation secret and can directly identify Anne, which abolishes the level of personal data protection offered by the pseudonymisation process performed earlier. This level of protection would still be in place, if Anne was only sharing the medical examination results and not the pseudonymisation secret.

The same principle can be extended to an entity such as a hospital where a patient undergoes different tests across different departments. Again, the departments performing the tests will process the data using the Patient_ID and the treating doctor will be able to relate the Patient_ID to Anne as depicted below. Access to medical data among doctors of the same medical centre (e.g. hospital) should be based upon authentication and relevant user rights. Depending on the processing operation, they could either have access to the pseudonymisation secret, thus being able to directly identify Anne, or have access only to the pseudonymised identifier of Anne, thus not being able to directly identify Anne.

Figure 6: Pseudonyms decoupling



It should be noted that this use case is not directly applicable to cases where Electronic Health Record (EHR) systems are already deployed or interoperable Electronic Health Records are already deployed[4], as in such cases the means for communication of data and identification of individuals could already be defined. However, in such cases, pseudonymisation could be considered during the early design phases as a technique to increase the level of protection, support compliance and promote broader adoption of the EHRs.

Pseudonymised clinical trial data allow for detection of correlations during the trials and re-identification of specific individuals when deemed necessary.

3.2 CLINICAL TRIALS

Clinical trials study new medical interventions and treatments and evaluate their effects and possible side-effects. They are regarded as a prerequisite for acquiring the necessary approvals from relevant authorities. Typically, they are not performed by the manufacturer (e.g. pharmaceutical company) but by independent organisations called CROs (Clinical Research Organisations).

A typical scenario is the so called double-blinded study where a cohort of individuals with similar characteristics (e.g. patients with the same disease) are split into two sub-groups. The members of one group receive the medication under trial and the members of the other group a placebo medication. In a double-blind study, neither the participants nor the researcher know who is assigned to which group. During the trial the same medical data are accumulated for both cohorts. Once all the data has been obtained, researchers can then compare the results of each group and determine if the new medical intervention (independent variable) had any impact on the treatment (dependent variable).

Beyond the sheer comparison of patient data records, another common use of such medical data is the detection of correlations and patterns among different variables (e.g. age, gender, occupation), in an attempt to identify the upper limit of efficiency and effectiveness as well as additional safety information on, for example, posology. Hence, at these research organisations, the data of all patients must be analysed for common patterns of medical relevance.

For such analyses, the identity of the participants is not directly relevant, and typically researchers do not need to access the real-life identity of any of the research participants whose data gets analysed at the CRO. However, the risk that a patient is identified through indirect information is not minimal. For clinical trials, information such as age, gender,

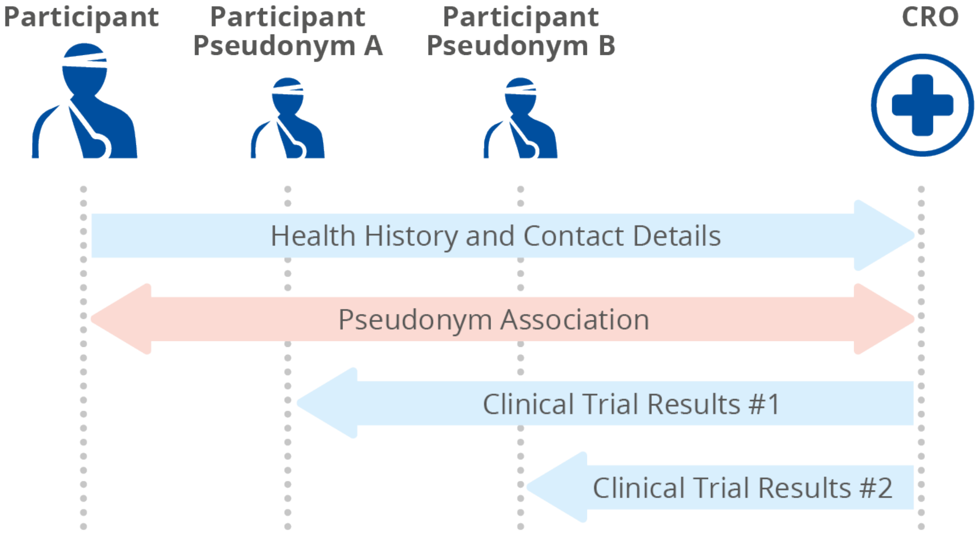
[4] Commission Recommendation of 6.2.2019 on a European Electronic Health Record exchange format, C(2019) 800 final

occupation, place of living may be relevant to the study and under certain conditions may lead to the identification of trial participants.

In this setting, a suitable pseudonymisation scheme can unfold its potential for allowing the tasks of detecting correlations and statistical patterns to be performed without revealing the true value of different data items. In certain trials this can even be extended to correlations and statistical patterns in symptoms and medications. Since several different types of personal data are usually collected during clinical trials, pseudonymisation has to be applied with care so as not to expose patients to unauthorised re-identification. This could be done, for example, by combining two approaches as depicted in Figure 8 below:

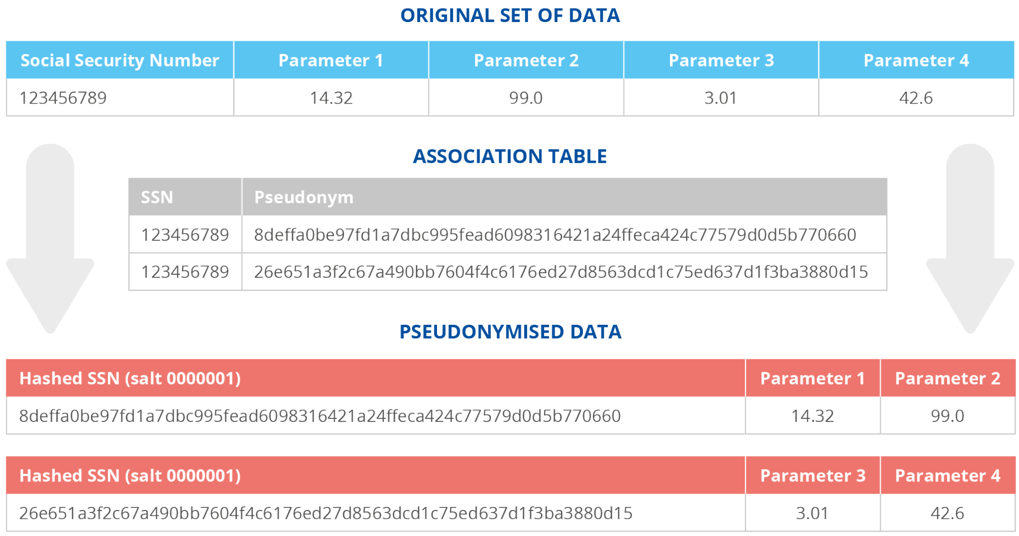
1) employ pseudonymisation on the main identifying data of each participant;
2) use more than one pseudonym for each identifying data for different clinical parameters.

Figure 7: Clinical trials pseudonymisation overview


Participant
Participant Pseudonym A
Participant Pseudonym B
CRO
Health History and Contact Details
Pseudonym Association
Clinical Trial Results #1
Clinical Trial Results #2

Such an approach could limit the personal data related to each pseudonym that, paired with a robustness that can be enforced using a solid hashing function like SHA-2 with a random seed value as shown in the following example, would make re-identification even more difficult.

Figure 8: Association of parameters with pseudonyms



ORIGINAL SET OF DATA

Social Security Number	Parameter 1	Parameter 2	Parameter 3	Parameter 4
123456789	14.32	99.0	3.01	42.6

ASSOCIATION TABLE

SSN	Pseudonym
123456789	8deffa0be97fd1a7dbc995fead6098316421a24ffeca424c77579d0d5b770660
123456789	26e651a3f2c67a490bb7604f4c6176ed27d8563dcd1c75ed637d1f3ba3880d15

PSEUDONYMISED DATA

Hashed SSN (salt 0000001)	Parameter 1	Parameter 2
8deffa0be97fd1a7dbc995fead6098316421a24ffeca424c77579d0d5b770660	14.32	99.0

Hashed SSN (salt 0000001)	Parameter 3	Parameter 4
26e651a3f2c67a490bb7604f4c6176ed27d8563dcd1c75ed637d1f3ba3880d15	3.01	42.6

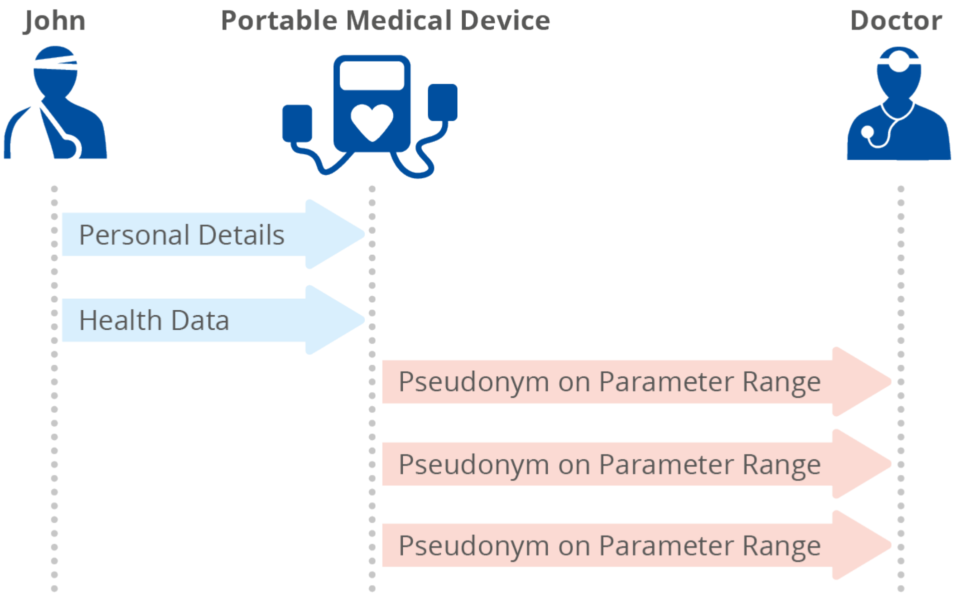
Another exception to the basic assumption occurs when a patient’s data might reveal a risk to her health (e.g. a new diagnosis) and it would be beneficial for the research team to contact the CRO and trigger a patient notification. This is a case very similar to the one discussed in [12] under section 4.2.2.

3.3 PATIENT-SOURCED MONITORING OF HEALTH DATA

Nowadays smart wearable devices are able to monitor vital signs such as heart rate, oxygen saturation and blood pressure level. Regular monitoring of vital signs is a common intervention in patient care, which aims to facilitate the early recognition of abnormal physiological parameters. However, the usual practice is that only the patient him/herself is able to view the measurements and contact the doctor when an abnormal value is observed.

The normal ranges for a person’s vital signs usually vary with age, weight, gender and overall health status, which makes it difficult to set predefined values. This use case envisions the provision of a Health Monitoring System (HMS) where the doctor can access medical information and monitor vital signs and can also receive notification of abnormal values thereof, via properly predefined ranges of values.

Figure 9: Pseudonymisation during patient sourced monitoring of health data

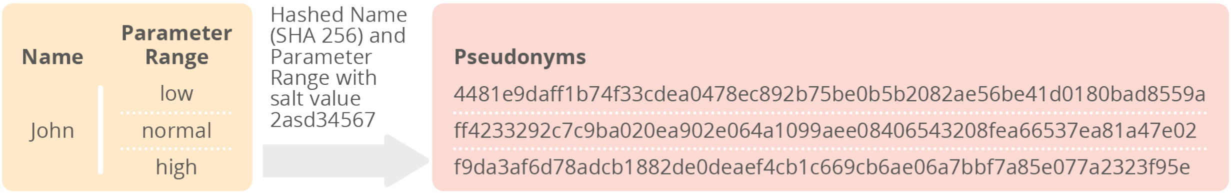

John
Portable Medical Device
Doctor
Personal Details
Health Data
Pseudonym on Parameter Range
Pseudonym on Parameter Range
Pseudonym on Parameter Range

For example, John has been diagnosed with a cardiovascular disease (CVD) and arrhythmia; bradycardia or tachycardia can increase the risk of stroke. John's treating doctor has defined three sets of values with regards to his heart at rest: Low, Normal and High. John's wearable device regularly monitors his heartrate and whether he is moving and can provide a notification to John when his resting heartrate is below the Low or above the High thresholds. At the same time such notification can also be sent to his treating doctor, who can proactively support John or anticipate a call to support him on what to do next.

Instead of communicating John's heartrate values and details, the wearable device can instead perform the pseudonymisation process and communicate a pseudonym of John and the parameter range for his heart rate at rest. The treating doctor will receive the pseudonym and be able to access the patient to which it relates and the state of the patient's heart rate at rest. Similar to the use case discussed in Section 3.1, if the treating doctor gains access to the pseudonymisation secret and is able to directly identify John, the pseudonymisation process should be considered mainly as a security measure to support confidentiality during communication rather than a data protection measure.

Transferring pseudonymised from the medical device to the treating doctor reduces the risk of the data being compromised in transit.

Figure 10: Pseudonyms association table


Hashed Name (SHA 256) and Parameter Range with salt value 2asd34567

Name	Parameter Range	Pseudonyms
John	low	4481e9daff1b74f33cdea0478ec892b75be0b5b2082ae56be41d0180bad8559a
	normal	ff4233292c7c9ba020ea902e064a1099aee08406543208fea66537ea81a47e02
	high	f9da3af6d78adcb1882de0deaef4cb1c669cb6ae06a7bbf7a85e077a2323f95e

Attempting to take this use case one step further, the wearable device could very well be subject to the provisions of the Medical Devices Regulations [8] and the guidance on cybersecurity for medical devices [14] as endorsed by the Medical Device Coordination Group (MDCG). If so, the deployment of pseudonymisation techniques during, for example, the communication and storage of personal medical data could very well contribute towards meeting the principles of both security and data protection by design while also re-enforcing the confidentiality of transferred data.

4. CONCLUSIONS

As the healthcare domain is attempting to make the most out of the evolving technical landscape and adapt the provision of services to fulfil, in a timely manner, the growing needs of patients of all ages and cultures worldwide, additional challenges in cybersecurity and data protection come into play.

Pseudonymisation is increasingly becoming a key security technique by providing a means that can facilitate the processing of personal data, while offering strong safeguards for the protection of personal data and thereby safeguarding the rights and freedoms of individuals.

Complementing previous relevant work done by ENISA, this report demonstrates how pseudonymisation can be deployed in practice to further promote the protection of exchanges of health data.

Obviously, there is not a single solution on how and when to apply pseudonymisation; in fact different solutions might provide equally good results in specific scenarios, depending on the requirements in terms of protection, utility, scalability, etc.

Starting from a plain token, pseudonymisation can be a ‘simple’ option to adopt, but it can also be comprised of a very complex process both at technical and at organisational levels. For this reason, defining the goals and objectives of pseudonymisation in each particular case and each processing operation is really important. To this end, relevant good practices and examples of pseudonymisation in the context of the GDPR can be of great value to healthcare providers and developers of healthcare applications.

Developers and regulators at national and European level should promote the exchange of good practices and provide practical guidance on deploying pseudonymisation in practice.

Advances in technology and in the types of health-related services offered might affect the effectiveness and applicability of a pseudonymisation solution that is already being deployed. This is not only relevant to the choice of the technique itself but also to the overall design of the pseudonymisation process including, especially, the protection of the additional information (i.e. the information that allows the association between pseudonyms and initial identifiers).

Pseudonymisation solutions brought forward are highly dependent on the state-of-the-art and implementation challenges or limitations with regard to each technique may arise over time.

The research community should continue working on data protection and security engineering, including state-of-the-art pseudonymisation techniques and their possible implementations, with the support of the EU institutions in terms of policy guidance and research funding.

5. BIBLIOGRAPHY

1. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Da. (2016)

2. Topol, E.: Individualized Medicine from Prewomb to Tomb. (2014)

3. Angelidis, P.: UOWM Lecture Notes. (2019)

4. ENISA: Security and Resilience in eHealth Infrastructures and Services. (2015)

5. ENISA: Procurement Guidelines for Cybersecurity in Hospitals. (2020)

6. ENISA: Cloud Security for Healthcare Services. (2021)

7. Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union. (2016)

8. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EE. (2017)

9. Directive 2011/24/EU of the European Parliament and of the Council of 9 March 2011 on the application of patients' rights in cross-border healthcare. (2011)

10. ENISA: Recommendations on shaping technology according to GDPR provisions - An overview on data pseudonymisation. (2019)

11. ENISA: Pseudonymisation techniques and best practices. (2019)

12. ENISA: Data Pseudonymisation: Advanced Techniques and Use Cases. (2021)

13. ISO: 25237:2017 Health informatics — Pseudonymization. (2017)

14. Medical Device Coordination Group: 2019-16 - Guidance on Cybersecurity for medical devices. (2019)

TP-01-22-221-EN-N

ABOUT ENISA

The European Union Agency for Cybersecurity, ENISA, is the Union's agency dedicated to achieving a high common level of cybersecurity across Europe. Established in 2004 and strengthened by the EU Cybersecurity Act, the European Union Agency for Cybersecurity contributes to EU cyber policy, enhances the trustworthiness of ICT products, services and processes with cybersecurity certification schemes, cooperates with Member States and EU bodies, and helps Europe prepare for the cyber challenges of tomorrow. Through knowledge sharing, capacity building and awareness raising, the Agency works together with its key stakeholders to strengthen trust in the connected economy, to boost resilience of the Union's infrastructure and, ultimately, to keep Europe's society and citizens digitally secure. More information about ENISA and its work can be found here: www.enisa.europa.eu.

ENISA
European Union Agency for Cybersecurity

Athens Office
Agamemnonos 14, Chalandri 15231, Attiki, Greece

Heraklion Office
95 Nikolaou Plastira
700 13 Vassilika Vouton, Heraklion, Greece

enisa.europa.eu

ISBN 978-92-9204-576-0
DOI 10.2824/092874