DIPUTACIÓ DE VALÈNCIA

*Protecció de Dades i Seguretat de la Informació*

# Quaderns DivalData

Quaderns dirigits a delegats,
responsables i especialistes en protecció
de dades personals

Quadern núm. 18 | Desembre 2021

## LA SEUDONIMIZACIÓ I ANONIMITZACIÓ DE DADES PERSONALS

# Í N D E X

## LA SEUDONIMIZACIÓ I ANONIMITZACIÓ DE DADES PERSONALS


| | Pàgina |
|---|---|
| **INTRODUCCIÓ** | 2 |
| **LA SEUDONIMIZACIÓ NO ÉS EL MATEIX QUE L'ANONIMITZACIÓ** | 4 |
| **ASPECTES A TINDRE EN COMPTE PER A ABORDAR L'ELIMINACIÓ DE LA IDENTIFICACIÓ EN L'ANONIMITZACIÓ** | 5 |
| **TÈCNIQUES D'ANONIMITZACIÓ: CLAUS** | 6 |
| **GARANTIES DE LES TÈCNIQUES D'ANONIMITZACIÓ** | 9 |
| **MATERIAL COMPLEMENTARI I NOTÍCIES** | 10 |


Vos convidem a traslladar-nos les temàtiques que resulten de vostre interés per als pròxims butlletins informatius. Aquestes peticions hauran de dirigir-se a:

Diputació de València

Dpt. de Protecció de Dades i Seguretat de la Informació

Pl. de Manises, 4 46003 València

email: dpdsi@dival.es

SUSCRIPCIONS

Si desitges subscriure's al nostre Butlletí informatiu accedeix al següent

enllaç

*"TRANSFORMAR UN CONJUNT DE DADES PERSONALS EN INFORMACIÓ ANÒNIMA O SEUDONIMIZADA EXIGEIX REALITZAR UN TRACTAMENT SOBRE AQUESTES DADES PERSONALS. EL TRACTAMENT D'ANONIMITZACIÓ GENERA UN ÚNIC I NOU CONJUNT DE DADES, MENTRE QUE EL TRACTAMENT DE *SEUDONIMIZACIÓN GENERA DOS NOUS CONJUNTS DE DADES: LA INFORMACIÓ SEUDONIMIZADA I LA INFORMACIÓ ADDICIONAL QUE PERMET REVERTIR L'ANONIMITZACIÓ."*

## INTRODUCCIÓ

La informació anònima és un conjunt de dades que no guarda relació amb una persona física identificada o identificable (Considerant 26 del RGPD), en tant que la informació seudonimizada és un conjunt de dades que no pot atribuir-se a un interessat sense utilitzar informació addicional, requereix que aquesta informació addicional figure per separat i, a més, estiga subjecta a mesures tècniques i organitzatives destinades a garantir que les dades personals no s'atribuïsquen a una persona física identificada o identificable (Article 4.5).

Transformar un conjunt de dades personals en informació anònima o seudonimizada exigeix realitzar un tractament sobre aquestes dades personals. El tractament d'anonimització genera un únic i nou conjunt de dades, mentre que el tractament de seudonimització genera dos nous conjunts de dades: la informació seudonimizada i la informació addicional que permet revertir l'anonimització.

El conjunt de dades anonimitzades no està sota l'àmbit d'aplicació del Reglament General de Protecció de Dades (RGPD) (Considerant 26) encara que poguera estar sota l'àmbit d'aplicació d'altres normes (p. ex. de seguretat nacional, salut pública, infraestructures crítiques, etc.) En aquest cas ha de tindre's en compte que:

- **El tractament que generen les dades anonimitzades sí que és un tractament de dades personals, que pot considerar-se compatible amb la fi original del tractament de dades personals del qual procedeixen les dades (Dictamen 05/2014 sobre tècniques d'anonimització WP246 apartat 2.2.1. Legitimació del procés d'anonimització).**
- **El conjunt de dades anonimitzades queda fora de l'àmbit d'aplicació del RGPD en la mesura que és possible demostrar objectivament que no existeix capacitat material per a associar les dades anonimitzades a una persona física determinada, directament o indirectament, ja siga mitjançant l'ús d'altres conjunts de dades, informacions o mesures tècniques i materials que pogueren existir a la disposició de tercers.**

És a dir, les dades es consideraran anonimitzats en la mesura que no existisca una probabilitat raonable que qualsevol persona puga identificar a la persona física en el conjunt de dades. Aquesta avaluació ha de tindre en compte els costos, el temps requerit per a dur a terme

la reidentificació o els mitjans tecnològics necessaris per a aconseguir la reversió de l'anonimització, tant els actuals com tenint en compte els avanços tecnològics (Considerant 26).

El conjunt de dades seudonimizadEs, i la informació addicional vinculada amb aquest conjunt de dades, estan sota l'àmbit d'aplicació del RGPD, així com el tractament que els genera.

Els drets i llibertats dels interessats han d'estar igualment protegides tant en els tractaments d'anonimització com en els processos de seudonimització. Tenint en compte que sobre el conjunt de dades anonimitzades no caldrà atendre els requisits establits pel RGPD quant a la limitació del tractament, la conservació de les dades, les comunicacions i les transferències internacionals, o les mesures per a protegir la confidencialitat, s'han de dissenyar i validar els tractaments d'anonimització pensant en la protecció dels drets anteriorment assenyalats. Això exigeix poder demostrar un nivell objectiu de qualitat en el tractament d'anonimització i aconsella determinar com evoluciona el risc de reidentificació al llarg del temps. En qualsevol cas, la reversió de l'anonimització suposa la plena aplicació del RGPD als subjectes obligats que tracten les dades personals.

*“LES DADES ANÒNIMES, NO PODEN ASSOCIAR-SE AMB UN INDIVIDU EN PARTICULAR. UNA VEGADA QUE LES DADES SÓN REALMENT ANÒNIMS I ELS INDIVIDUS DEIXEN DE SER IDENTIFICABLES, DEIXEN D'ESTAR INCLOSOS EN L'ÀMBIT D'APLICACIÓ DEL RGPD”*

## LA SEUDONIMIZACIÓ NO ÉS EL MATEIX QUE L'ANONIMITZACIÓ

La seudonimització no és el mateix que la anonimització». El RGPD defineix «seudonimització» com «el tractament de dades personals de manera que no puguen atribuir-se a un interessat sense utilitzar informació addicional, sempre que aquesta informació addicional figure per separat i estiga subjecta a mesures tècniques i organitzatives destinades a garantir que les dades personals no s'atribuïsquen a una persona física identificada o identificable». Això significa que l'ús de «informació addicional» pot suposar la identificació dels individus; per aqueix motiu les dades personals seudonimizats també són dades personals.

Per contra, les dades anònimes, no poden associar-se amb un individu en particular. Una vegada que les dades són realment anònims i els individus deixen de ser identificables, deixen d'estar inclosos en l'àmbit d'aplicació del RGPD.

La seudonimització consisteix en la substitució d'un atribut (normalment un atribut únic) per un altre en un registre. Per consegüent, continua existint una alta probabilitat d'identificar a la persona física de manera indirecta; en altres paraules, l'ús exclusiu de la seudonimització no garanteix un conjunt de dades anònim. No obstant això, el present dictamen examina aquest mètode degut a les nombroses idees falses i errors existents sobre ell. La seudonimització redueix la vinculabilitat d'un conjunt de dades amb la identitat de l'interessat; es tracta, per tant, d'una mesura de seguretat útil, però no és un mètode d'anonimització.

El resultat de la seudonimització pot ser independent del valor inicial (tal seria el cas d'un número aleatori generat pel responsable del tractament o d'un cognom triat per l'interessat) o bé derivar-se dels valors originals d'un atribut o conjunt d'atributs, com per exemple en el cas de funcions hash o sistemes de xifratge.

*"EL FET DE REDUIR LA INFORMACIÓ EXISTENT ALS MÍNIMS NECESSARIS PER A SATISFER ELS OBJECTIUS QUE HA DE COMPLIR LA INFORMACIÓ ANONIMITZADA IMPLICA DE MANERA DIRECTA UNA REDUCCIÓ DEL RISC DE REIDENTIFICACIÓ, DONCS A MENOR QUANTITAT DE DADES PERSONALS MENOR SERÀ EL RISC INHERENT QUE RESULTE DEL TRACTAMENT REALITZAT DURANT EL PROCÉS D'ANONIMITZACIÓ"*

## ASPECTES A TINDRE EN COMPTE PER A ABORDAR L'ELIMINACIÓ DE LA IDENTIFICACIÓ EN L'ANONIMITZACIÓ

L'objecte és reduir al mínim necessari la quantitat de variables que permeten la identificació de les persones, restringint l'accés a la informació confidencial a l'equip de treball implicat en el procés i optimitzant el cost computacional de les operacions amb dades anonimitzades. El fet de reduir la informació existent als mínims necessaris per a satisfer els objectius que ha de complir la informació anonimitzada implica de manera directa una reducció del risc de reidentificació, doncs a menor quantitat de dades personals menor serà el risc inherent que resulte del tractament realitzat durant el procés d'anonimització: vulneració del deure de secret, pèrdua d'informació, bretxes de seguretat, robatori de claus, etc.

Alguns aspectes que poden ser tinguts en compte pel responsable del tractament per a abordar l'eliminació o emmascarar les variables d'identificació poden ser els següents:

- Determinar la finalitat de les dades anonimitzades: terminis de conservació, ús estadístic, científic, o qualsevol ús posterior.
- Establir les variables confidencials necessàries per al tractament de les dades anonimitzades i identificar les variables de confidencialitat que no vagen a ser necessàries en el tractament de les dades anonimitzades. Com a variable de confidencialitat s'entendrà qualsevol informació existent sobre una persona, tant si permet la seua identificació com si a priori la seua identificació no és possible.
- Eliminació de dades identificatives directes o indirectes no necessaris: noms, data de naixement, telèfon, DNI, email, adreça postal, nombre de comptes bancaris, matrícules de vehicles, identificador dispositiu mòbil, número de sèrie, adreça IP, identificadors biomètrics, fotografia o imatge, etc.
- Control segregat d'usuaris amb accés a les dades personals i usuaris amb accés a les dades anonimitzades: la informació, anonimitzada o no, pot estar estructurada i cada usuari té accés a les dades que li són necessaris per a fer el seu treball, de manera que

*"L'ALGORISME DE HASH GENERA UNA EMPREMTA DIGITAL I FA IMPOSSIBLE RECONSTRUIR LA DADA ORIGINAL PARTINT DE LA PETJADA I D'ALTRA BANDA QUALSEVOL VARIACIÓ EN LA DADA ORIGINAL DONARÀ LLOC A UNA EMPREMTA DIGITAL DIFERENT, LA QUAL COSA EXPRESSAT EN TERMES COMPUTACIONALS PODRIA DIR-SE QUE LA MODIFICACIÓ D'UN SOL BIT EN LA INFORMACIÓ ORIGINAL EMMAGATZEMADA EN UN ORDINADOR DONARIA LLOC A UNA CLAU DIFERENT O UNA EMPREMTA DIGITAL DIFERENT."*

la resta d'informació o variables que pogueren permetre la reidentificació dels subjectes i no siguen necessàries per a l'acompliment de les funcions d'un treballador no li siguen accessibles.

- Utilització de rangs per a emmascarar a les persones quan existeixen microdades concretes que permeten la identificació directa de persones o col·lectius específics. Per exemple, en el cas de col·lectius de persones extremadament reduïts s'ha de diluir la informació d'aquest xicotet grup de persones en un col·lectiu de major rang numèric afegint, si és necessari, una referència a un percentatge en el qual es pose de manifest l'existència del col·lectiu menor com a part d'un conjunt major. – Disposar d'una política d'ús de claus per a ocultar la identificació de les persones serà de gran utilitat en aquesta fase de l'anonimització. La política de claus establirà els nivells i el nombre de claus mínim a tindre en compte en funció de l'objecte de la informació anonimitzada.

## TÈCNIQUES D'ANONIMITZACIÓ: CLAUS

ALGORISMES DE HASH: és inqüestionable la utilitat que tenen els algorismes de xifratge quan necessitem anonimitzar microdades, resultant especialment útils els algorismes de "hash". Un algorisme de hash és un mecanisme que, aplicat a una dada concreta, genera una clau única o quasi única que pot utilitzar-se per a representar una dada. Per exemple, disposem d'una dada que volem ocultar o anonimitzar i per a això utilitzem un algorisme de hash, com per exemple SHA1 o MD5. De l'aplicació de l'algorisme a una determinada dada obtenim una clau o empremta digital que pot utilitzar-se per a reemplaçar la dada real. L'algorisme de hash genera una empremta digital i fa impossible reconstruir la dada original partint de la petjada i d'altra banda qualsevol variació en la dada original donarà lloc a una empremta digital diferent, la qual cosa expressat en termes computacionals podria dir-se que la modificació d'un sol bit en la informació original emmagatzemada en un ordinador donaria lloc a una clau diferent o una empremta digital diferent.

No obstant això, un algorisme de hash per si sol no és suficient per a fer irreversible l'anonimització, ja que xicotetes cadenes de text com, per exemple, els

microdades corresponents al codi postal d'una persona, un número de telèfon, etc., poden ser fàcilment reidentificables amb un programa informàtic que genere xifres consecutives i les seues corresponents empremtes digitals. Si el que volem és garantir l'anonimització d'una microdada cal utilitzar un mecanisme criptogràfic que ens garantisca el secret de l'empremta digital que hem generat. Una bona opció és l'algorisme HMAC basat en RFC20147 HMAC pot utilitzar-se en combinació amb diversos algorismes de hash com, per exemple, amb MD5 i sobre l'empremta digital o clau resultant de l'algorisme de hash aplica un algorisme criptogràfic que genera una nova empremta digital o clau en funció d'una clau secreta.

ALGORISMES DE XIFRATGE: els algorismes de xifratge amb propietats homomòrfiques obrin noves possibilitats per al tractament de dades anonimitzades. Un algorisme de xifratge homomòrfic permet realitzar operacions amb dades xifrades de tal manera que el resultat de les operacions és el mateix que si les operacions s'hagueren realitzat amb les dades sense xifrar. Els resultats de les operacions amb dades xifrades donen per resultat valors igualment xifrats que poden ser desxifrats posteriorment per l'usuari que dispose de la clau per a desxifrar. L'esquema de xifratge homomòrfic obri la possibilitat del tractament de dades personals anonimitzades garantint la privacitat del tractament i que els resultats dels tractaments seran accessibles únicament al posseïdor de la clau de desxifrat.

La implantació d'esquemes de xifratge homomòrfic pot aportar un alt grau de confidencialitat als tractaments de dades en cloud, tractaments de dades obtingudes dels wearables, sistemes de telemedicina, etc.

SEGELL DE TEMPS: també cal tindre en compte la possibilitat d'utilitzar en el procés d'anonimització algorismes de segell de temps amb la finalitat de garantir la data i hora en la qual l'anonimització ha sigut realitzada, o fins i tot algorismes de signatura electrònica que permeten garantir la identitat electrònica de qui ha realitzat l'anonimització.

***"ELS CRITERIS PER A L'ANONIMITZACIÓ PER CAPES PODEN FIXAR-SE EN RESPOSTA A REQUISITS DE LA INFORMACIÓ, DEL TRACTAMENT DE LES DADES ANONIMITZADES O DE LA PRÒPIA POLÍTICA D'ANONIMITZACIÓ DEL RESPONSABLE DEL TRACTAMENT DE LES DADES ANONIMITZADES, ATENENT DIFERENTS CRITERIS ES PODRAN ABORDAR DIFERENTS TÈCNIQUES."***

CAPES D'ANONIMITZACIÓ: juntament amb aquests processos d'emmascarament i anonimització podem utilitzar el que podria denominar-se l'anonimització per capes. Per exemple, el responsable del tractament ha anonimitzat totes les dades que puguen servir per a reidentificar a les persones i remitent la informació al seu legítim destinatari qui, a fi d'evitar que poguera produir-se la reidentificació, decideix realitzar una segona anonimització de les dades ja anonimitzades. D'aquesta manera, el destinatari de la informació anonimitzada assegura que els seus processos utilitzen els seus propis recursos d'anonimització, evitant que en cas de fragilitat dels processos d'anonimització del responsable del tractament la identitat de les persones poguera veure's afectada.

En definitiva, el destinatari de la informació assegura la privacitat de les persones en el tractament de dades anonimitzades amb les seues pròpies garanties de qualitat. Aquest procés de reanonimització pot utilitzar-se de manera interdepartamental de manera que, a mesura que les dades vagen passant d'un departament a un altre, es realitzen diferents processos d'anonimització fent que per a una reidentificació real siga necessària la concurrència de tots dels actors involucrats en les diferents capes d'anonimització.

Pot dir-se, per tant, que el procés d'anonimització pot ser monocapa o multicapa. Parlem d'un procés d'anonimització monocapa quan l'anonimització de les variables es realitza una única vegada i es dona per finalitzat el procés però, a vegades, la reanonimització de variables o anonimització multicapa pot proporcionar garanties addicionals per a evitar la reidentificació de les persones.

Els criteris per a l'anonimització per capes poden fixar-se en resposta a requisits de la informació, del tractament de les dades anonimitzades o de la pròpia política d'anonimització del responsable del tractament de les dades anonimitzades, atenent diferents criteris es podran abordar diferents tècniques.

## GARANTIES DE LES TÈCNIQUES D'ANONIMITZACIÓ

En relació a les tècniques d'anonimització, el Grup de Treball de l'article 29 va emetre el seu Dictamen 05/2014 en el qual es realitza una anàlisi tècnica sobre la robustesa, feblesa i garanties de les tècniques d'anonimització. En aquest Dictamen es mostren alguns dels límits, riscos i errors que poden tindre lloc com a resultat de les tècniques d'anonimització utilitzades.

Cal tindre en compte que l'anonimització de la informació sempre generarà, independentment de les bones pràctiques emprades, un cert grau de distorsió entre la informació anonimitzada i la informació no anonimitzada, i que a vegades el procés d'anonimització no pot assegurar la impossibilitat de reidentificació de les persones en termes absoluts, per la qual cosa s'han de tindre en compte les garanties jurídiques necessàries per a preservar els drets dels interessats.

En aquest sentit alguns dels aspectes que han de ser tinguts en compte, segons el parer de l'AEPD, són:

- **Acords de confidencialitat** que impliquen els següents actors:
- Responsable del tractament.
- Responsable del procés d'anonimització.
- Responsable del tractament de dades anonimitzades.
- Personal amb accés a la informació anonimitzada.
- Obtindre el **compromís del destinatari** de la informació per a mantindre l'anonimització i l'obligació d'informar el responsable del tractament davant qualsevol sospita de reidentificació.
- Realització de **auditories d'ús de la informació anonimitzada** per part del responsable del tractament al responsable del tractament de les dades anonimitzades.
- Les garanties estaran **incloses en el contracte subscrit entre el responsable del tractament i el destinatari** de la informació anonimitzada.

## MATERIAL COMPLEMENTARI

- Dictamen 05/2014 del GRUP DE TREBALL SOBRE PROTECCIÓ DE DADES DE L'ARTICLE 29. Consulta aquest enllaç.
- "Orientacions i garanties en els procediments de *anonimizacion de dades personals" GUIA (AEPD). Consulta aquest enllaç.
- "10 Malentesos relacionats amb l'anonimització" (AEPD). Consulta aquest enllaç.
- Blog de l'AEPD "Anonimització i seudonimització". Consulta aquest enllaç.
- Blog de l'AEPD "Anonimització i seudonimització (II): la privacitat diferencial". Consulta aquest enllaç.
- Resolució Procediment Núm.: PS/00052/2020 en relació amb un aplicatiu que mostra dades que haurien de ser anonimitzats (AEPD). Consulta aquest enllaç.

## NOTÍCIES

- Entra en vigor el Reial decret llei 24/2021, de 2 de novembre, que transposa directives de la Unió Europea en determinades matèries com la relativa a les dades obertes i la reutilització de la informació del sector públic. El Llibre tercer de l'incorpora a l'ordenament jurídic espanyol les novetats de la Directiva (UE) 2019/1024 del Parlament Europeu i del Consell, de 20 de juny de 2019, relativa a les dades obertes i la reutilització de la informació del sector públic. Consulta la norma en aquest enllaç.
- L'AEPD sanciona a la Conselleria d'Educació del Principat d'Astúries per no utilitzar protocol segur en el seu web i per incorporar un apartat de "Treballa amb nosaltres" sense donar compliment a la Llei de Serveis de la Societat de la Informació. En el procediment instruït per l'AEPD, se sanciona a la citada Conselleria per entendre que la pàgina no és segura, perquè utilitza el protocol http://; així com per considerar que no compleix amb la Llei de Serveis de la Societat de la Informació, en considerar-la "prestador de serveis de la societat de la informació", perquè gestionava la borsa de treball de l'Institut, Consulta la Resolució en aquest enllaç.
- L'AEPD sanciona a la Conselleria d'Educació i Joventut de la Comunitat de Madrid per l'enregistrament telemàtic de les sessions d'avaluació d'un IES. Els docents que no van poder assistir a la sessió presencial van participar de forma telemàticament. Tals sessions van ser gravades per decisió de la direcció del centre educatiu. En el present cas, l'AEPD considera acreditat l'incompliment del principi de transparència (per la falta d'informació als professors participants), així com el principi de licitud del tractament. Consulta la Resolució en aquest enllaç.