Lo que debes saber sobre la seudonimización de los datos

Post del blog

Julio 2020

eCityclic by semic

Índice

1. ¿Qué es la seudonimización?

El *Reglamento (UE) 2016/679, del Parlamento Europeo y del Consejo, de 27 de abril de 2016,* relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos, art. 4.5, establece que "**la seudonimización es el tratamiento de datos personales de manera tal que ya no puedan atribuirse a un interesado sin utilizar información adicional,** siempre que dicha información adicional figure por separado y esté sujeta a medidas técnicas y organizativas destinadas a garantizar que los datos personales no se atribuyan a una persona física identificada o identificable".

¡Sigue leyendo! Te contamos más sobre la seudonimización de los datos.

2. Anonimización y seudonimización de datos personales

Según la AEPD (la Agencia Española de Protección de Datos):

> **Los procesos de anonimización y seudonimización son una herramienta válida para garantizar la privacidad de los datos personales** y sus limitaciones son inherentes al avance de la tecnología.

Existe una proporcionalidad manifiesta en lo que respecta a la capacidad tecnológica de anonimizar y la posibilidad de la reidentificación de las personas cuyos datos han sido anonimizados, es decir, la misma capacidad de la tecnología para anonimizar datos personales puede ser utilitzada para la reidentificación de las personas. Además, se debe tener en cuenta el riesgo que la propia sociedad de la información añade a los datos anonimizados, riesgo que por otra parte evoluciona a lo largo del tiempo, por lo que habrá que contemplar el riesgo de los procesos de anonimización como una contingencia latente a lo largo de la vida de la información y no en un momento concreto, y en consecuencia, las medidas encaminadas a valorar y gestionar los riesgos deben tener carácter periódico.

No es posible considerar que los procesos de anonimización garanticen al 100% la no reidentificación de las personas, por lo que será necesario sustentar la fortaleza de la anonimización en medidas de evaluación de impacto (EIPD), organizativas, de seguridad de la información, tecnológicas y, en definitiva, cualquier medida que sirva tanto para atenuar los riesgos de reidentificación de las personas como para paliar las consecuencias de que éstos se materialicen.

2.1. Diferencias entre seudonimización y anonimización

Cabe destacar que **existe diferencia entre los conceptos de seudonimización y anonimización**. Como señala la AEPD en el documento "La K-Anonimidad como medida de la privacidad" los datos seudonimizados constituyen información sobre una persona física a partir de la cual es posible realizar su identificación dentro de una probabilidad razonable teniendo en cuenta medios y factores objetivos, así como los costes, el tiempo y la tecnología necesarios para materializar su identificación.

La diferencia del término en ambas normas ha evolucionado desde una limitado anonimización a una materialización de ésta en el término seudonimización del RGPD donde se pone de manifiesto la dificultad de conseguir, en la actualidad, una anonimización perfecta o que garantice, en términos absolutos el enmascaramiento de la identidad de las personas.

> En conclusión, **con un dato anonimizado en ningún caso es posible la vinculación de los datos con la persona a la que hubiese identificado**. **La anonimización es** además un procedimiento **irreversible**. En cambio, **la seudonimización se reduce a limitar la trazabilidad entre el conjunto de datos tratados y la persona física cuya identidad queda asociada a estos,** por tanto, **es un procedimiento reversible** en la mayoría de sus técnicas.

3. Técnicas de seudonimización

Para la correcta seudonimización es muy importante la custodia de información adicional que permite vincular el dato seudonimizado con el titular del mismo.

Tal y como aparece en el Dictamen 05/2014 del Grupo de Trabajo sobre Protección de Datos de Carácter del artículo 29, de 10 de abril de 2014, las cinco técnicas de seudonimización más relevantes son:

- **Cifrado con clave secreta:** el poseedor de la clave puede reidentificar al interesado fácilmente. Con esta técnica de seudonimización solamente es necesario descifrar el conjunto de datos, debido a que este contiene los datos personales, aunque sea en forma cifrada. Si se aplican los sistemas de cifrado más avanzados, tan solo es posible descifrar los datos si se conoce la clave.

- **Función hash:** se trata de una función de seudonimización que devuelve un resultado de tamaño fijo a partir de un valor de entrada de cualquier tamaño (esta entrada puede estar formada por un solo atributo o por un conjunto de atributos). Esta función no es

reversible, es decir, no existe el riesgo de revertir el resultado, como en el caso del cifrado. Sin embargo, si se conoce el rango de los valores de entrada de la función hash, se pueden pasar estos valores por la función a fin de obtener el valor real de un registro determinado. Habitualmente, las funciones hash se diseñan para poder ejecutarse de manera relativamente rápida, por lo que están sujetas a ataques de fuerza bruta que consisten en probar todas las posibles entradas para crear tablas de correspondencia . También se pueden crear tablas precalculadas para lograr una reversión masiva de un gran número de valores hash.

- **Función con clave almacenada:** es un tipo de función hash que usa una clave secreta como valor de entrada suplementario. El responsable del tratamiento puede reproducir la ejecución de la función con el atributo y la clave secreta.

- **Cifrado determinista o función hash con clave de borrado de clave:** es una técnica de seudonimización que equivale a generar un número aleatorio a modo de seudónimo para cada atributo de la base de datos y, porteriormente, borrar la tabla de correspondencia. Este tipo de seudonimización reduce el riesgo de vinculabilidad entre los datos personales del conjunto de datos y los datos personales relativos a la misma persona contenidos en otro conjunto de datos en el que se usa un seudónimo distinto.

- **Descomposición en tokens:** se trata de una técnica de seudonimización que se usa normalmente en el sector financiero para reemplazar los números de identificación de tarjetas por valores que son de poca utilidad para los atacantes. Tiene su origen en las técnicas mencionadas en los puntos anteriores, y suele basarse en la aplicación de mecanismos de cifrado unidireccionales, o bien en la asignación, mediante una función de índice, de un número de secuencia o un número generado aleatoriamente que no derive matemáticamente de los datos originales.

4. Errores frecuentes sobre seudonimización

Hay algunos errores sobre seudonimización que se repiten frecuentemente:

1. **Pensar que un conjunto de datos seudonimizado es anónimo.**
2. **Usar la misma clave en bases de datos diferentes**: es de vital importancia usar claves distintas para reducir la vinculabilidad.
3. **Usar claves distintas o claves rotatorias para cada usuario:** se debe evitar usar claves distintas para diferentes conjuntos de usuarios y cambiar la clave según el uso.
4. **Conservar la clave:** si la clave secreta se almacena junto con los datos seudonimizados, un atacante podría llegar a vincularlos con el atributo original.

En conclusión, **la seudonimización consiste en sustituir un atributo por otro en un registro**. Así, a pesar de que siga existiendo la posibilidad de vincular a la persona física de manera indirecta con el conjunto de datos origen se dificulta la acción.

Esperamos haber resuelto tus dudas acerca del concepto de seudonimización, si crees que nos hemos dejado información relevante o deseas que escribamos sobre otros temas, envíanos un correo a hola@ecityclic.com e intentaremos hacerlo ;)

¡Gracias por leernos!