DIPUTACIÓ DE VALÈNCIA

Protecció de Dades i Seguretat de la Informació

Cuadernos DivalData

Cuadernos dirigidos a delegados,
responsables y especialistas en protección
de datos personales

Cuaderno nº 30 | Diciembre 2022

ANONIMIZACIÓN DE DATOS PERSONALES

Í N D I C E

ANONIMIZACIÓN DE DATOS PERSONALES

Os invitamos a trasladarnos aquellas temáticas que resulten de vuestro interés para los próximos boletines informativos. Estas peticiones deberán dirigirse a:

Diputación de Valencia

Dpto. de Protección de Datos y Seguridad de la Información

Pl. de Manises, 4 46003 Valencia

email: dpdsi@dival.es

SUSCRIPCIONES

Si deseas suscribirte a nuestra publicación accede al siguiente

enlace

"SIENDO CONSTITUTIVA DE INFRACCIÓN MUY GRAVE LA REVERSIÓN DELIBERADA DE UN PROCEDIMIENTO DE ANONIMIZACIÓN A FIN DE PEREMITIR LA REIDENTIFICACIÓN DE LOS AFECTADOS".

INTRODUCCIÓN

Recientemente la Agencia Española de Protección de Datos (en adelante, "AEPD") ha publicado una traducción de la "Guía básica de anonimización" elaborada por la Autoridad Nacional de Protección de Datos de Singapur (*Personal Data Protection Commission – PDPC)* dada su especial relevancia para responsables, encargados de tratamientos y delegados de protección de datos.

La guía tiene por objetivo proporcionar orientaciones sobre cómo realizar de forma adecuada la anonimización básica y la desidentificación de conjuntos de datos estructurados, textuales y no complejos.

El Reglamento General de Protección de Datos establece, en su artículo 32, como medidas para garantizar un nivel adecuado al riesgo, *la seudonimización y el cifrado de datos personales.* Es por ello que la anonimización y seudonimización de la información de carácter personal permite la protección de los datos personales de los interesados si se realiza de forma adecuada, siendo constitutiva de **infracción muy grave** *"la reversión deliberada de un procedimiento de anonimización a fin de permitir la reidentificación de los afectados"* (artículo 72 Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales)

CONCEPTOS BÁSICOS DE ANONIMIZACIÓN DE DATOS

Para abordarla la citada Guía, es necesario partir de los conceptos centrales:

- **Anonimización**: proceso que consiste en la conversión de datos personales en datos personales que no se pueden utilizar para identificar a ningún individuo. La anonimización se debe considerar como un proceso basado en riesgo, que incluye tanto la aplicación de medidas técnicas de anonimización como salvaguardas para evitar la reidentificación.

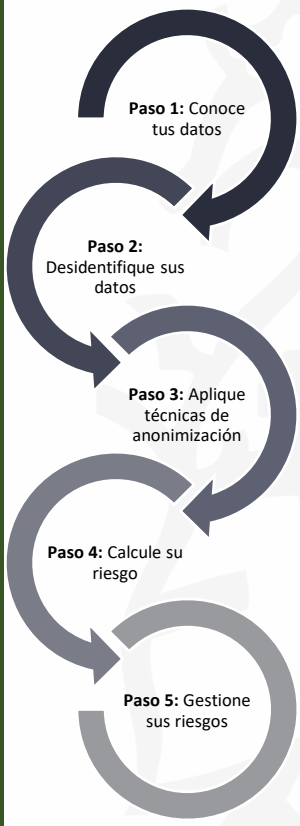
- **Desidentificación:** se basa en la eliminación de identificadores (tales como: nombre, dirección o número de DNI) que identifican directamente a un individuo. En la Guía se incide en que se suele equiparar erróneamente a la anonimización. Sin embargo, es necesario clarificar que la desidentificación es solo el primer paso de la anonimización. Un conjunto de datos desidentificado puede volver a identificarse fácilmente.

- **Reidentificación** se refiere a la identificación de los particulares a partir de un conjunto de datos que previamente fue desidentificado o anonimizado.

Los datos anonimizados no se consideran datos personales y, por lo tanto, no les resulta de aplicación la normativa de protección de datos. En este sentido se establece el Considerando 26 del RGPD.

"LA ANONIMIZACIÓN SE DEBE CONSIDERAR COMO UN PROCESO BASADO EN RIESGO, QUE INCLUYE TANTO LA APLICACIÓN DE MEDIDAS TÉCNICAS DE ANONIMIZACIÓN COMO SALVAGUARDAS PARA EVITAR LA REIDENTIFICACIÓN"

EL PROCESO DE ANONIMIZACIÓN



En la Guía se presentan cinco pasos para anonimizar los conjuntos de datos cuando sea apropiado.

Un registro de datos personales se compone de atributos de datos que tienen diversos grados de identificabilidad y sensibilidad a un individuo. Por ejemplo: los "identificadores directos" son atributos que son exclusivos de un individuo y se pueden usar como atributos de datos clave para volver a identificar al individuo.

Para desidentificar los datos, es necesario eliminar todos los identificadores directos (por ejemplo, los nombres).

Opcionalmente, se puede asignar un seudónimo a cada registro si es necesario vincular el registro a un individuo único.

En el siguiente paso, se aplicarán técnicas de anonimización de los identificadores indirectos para que no se puedan combinar fácilmente con otros conjuntos de datos que puedan contener información adicional para volver a identificar a las personas.

Al abordar el paso 4, en la Guía se menciona el método k-anonimidad para calcular el nivel de riesgo de reidentificación de un conjunto de datos.

"PARA DESIDENTIFICAR LOS DATOS ES NECESARIO ELIMINAR TODOS LOS IDENTIFICADORES DIRECTOS".

"LA SEUDONIMIZACIÓN SE UTILIZA CUANDO LOS VALORES DE LOS DATOS DEBEN DISTINGUIRSE DE FORMA ÚNICA Y NO SE CONSERVA NINGÚN CARÁCTER O CUALQUIER OTRA INFORMACIÓN IMPLÍCITA SOBRE LOS IDENTIFICADORES DIRECTOS DEL ATRIBUTO ORIGINAL".

TÉCNICAS BÁSICAS DE ANONIMIZACIÓN

En el Anexo A de la citada Guía se incluye un catálogo de técnicas básicas de anonimización de datos:

- **Supresión de registros:** se refiere a la eliminación de un registro completo en un conjunto de datos. Esta técnica afecta a múltiples atributos al mismo tiempo. La eliminación debe ser permanente y no solo una función de "ocultar fila".
- **Enmascaramiento de caracteres:** se refiere al cambio de los caracteres de un valor de datos. Puede realizarse mediante el uso de un símbolo (por ejemplo "*" o "x". Habitualmente el enmascaramiento se aplica solamente a algunos caracteres de un atributo. Esta técnica se suele utilizar cuando el valor de los datos es una cadena de caracteres y ocultar parte de ella es suficiente para proporcionar el grado de anonimato requerido.
- **Seudonimización (o "codificación"):** consiste en la sustitución de datos de identificación por valores inventados. Los seudónimos puede ser irreversibles cuando los valores originales se eliminan correctamente y la seudonimización se realiza de una manera no repetible. Esta técnica se utiliza cuando los valores de los datos deben distinguirse de forma única y no se conserva ningún carácter o cualquier otra información implícita sobre los identificadores directos del atributo original.
- **Generalización:** se basa en la reducción deliberada de la precisión de los datos. Un ejemplo de esta técnica consistiría en convertir la edad de una persona en un rango de edad. Normalmente se utiliza esta técnica para supuestos en los que la generalización permite que la información resultante siga siendo útil para el propósito previsto.
- **Intercambio:** esta técnica – que también se conoce como barajado y permutación – consiste en reorganizar los datos en el conjunto de información de forma que los valores de los atributos individuales sigan representados en el

conjunto de datos, pero generalmente no responden a los registros originales.

- **Perturbación de datos:** mediante esta técnica los valores del conjunto de datos original se modifican para que sean ligeramente diferentes. Se suele utilizar para identificadores indirectos (normalmente números y fechas), que pueden ser potencialmente identificables cuando se combinan con otras fuentes de datos, pero los cambios leves en el valor son aceptables para el atributo. Esta técnica no debe utilizarse cuando la precisión de los datos sea crucial.
- **Agregación de datos:** consiste en la conversión de un conjunto de datos de una lista de registros a valores resumidos. Se suele utilizar cuando no se requieren registros individuales y los datos agregados son suficientes para la finalidad.

"MEDIANTE LA TÉCNICA DE PERTURBACIÓN DE DATOS LOS VALORES DEL CONJUNTO DE DATOS ORIGINAL SE MODIFICAN PARA QUE SEAN LIGERAMENTE DIFERENTES".

"SIEMPRE QUE SEA POSIBLE, DEBE ESTABLECER UN VALOR DE K-ANONIMIDAD MÁS ALTO (POR EJEMPLO, 5 O MÁS) PARA EL INTERCAMBIO DE DATOS EXTERNOS"

K-ANONIMIDAD

k- anonimidad es un modelo utilizado para garantizar que no se haya superado el umbral de riesgo, como parte de la metodología de anonimización.

Se utiliza para confirmar que las medidas de anonimización implementadas alcanzan el umbral deseado contra los ataques de enlace.

A estos efectos, se entiende como "umbral de riesgo" el nivel de riesgo que la entidad está dispuesta a tolerar.

El modelo k-anonimidad se utiliza como guía antes de que se hayan aplicado técnicas de anonimización (por ejemplo, generalización), y para la verificación posterior también, para garantizar que los identificadores indirectos de cualquier registro sean compartidos por al menos k-1 otros registros. Por lo tanto, no es posible vincular o señalar el registro de un individuo, ya que siempre hay k atributos idénticos.

Además, en la Guía se añade la siguiente nota respecto a la k-anonimidad:

"Siempre que sea posible, debe establecer un valor de k-anonimidad más alto (por ejemplo, 5 o más) para el intercambio de datos externos, mientras que se puede establecer un valor más bajo (por ejemplo, 3) para el intercambio de datos internos o la retención de datos a largo plazo.

Sin embargo, si no puede anonimizar sus datos para lograrlo, debe implementar medidas de seguridad más estrictas para garantizar que los datos anonimizados no se divulguen a partes no autorizadas y se mitiguen los riesgos de reidentificación.

Alternativamente, puede contratar a expertos para que proporcionen métodos de evaluación alternativos para lograr riesgos de reidentificación equivalentes."

MATERIAL COMPLEMENTARIO

- Guía básica de anonimización. Elaborada por Autoridad Nacional de Protección de Datos de Singapur. Consulta este enlace.
- Directrices de asesoramiento del PDPC sobre la Ley de Protección de Datos Personales para Temas Seleccionados. Consulta este enlace.
- Nota técnica de la Agencia Española de Protección de Datos sobre la k-anonimidad como medida de la privacidad. Consulta este enlace.
- 10 Malentendidos relacionados con la anonimización. Consulta este enlace.
- Introducción al hash como técnica de seudonimización de datos personales. Consulta este enlace.
- Dictamen 05/2014 sobre técnicas de anonimización. Consulta este enlace.

NOTICIAS

- **La AEPD impone sanción de apercibimiento por infracción del artículo 13 RGPD relativo al deber de informar a los interesados sobre el tratamiento de sus datos personales.** La reclamación se formuló en relación con el protocolo por la Consellería para la realización de PCR al alumnado, que se llevan por empresas externas contratadas por aquella entidad. Solicita que se elimine la información y se corrija el protocolo "de manera que ningún laboratorio pueda cruzar el ADN con datos del propietario sin un consentimiento expreso e informado". La Consellería alega que la conservación de las muestras se realiza en condiciones de anonimización. Sin embargo, la AEPD considera que cuando los datos personales se recaben directamente del interesado, la información deberá facilitarse en el momento mismo en que tiene lugar esa recogida de datos.
  Consulta la resolución en este enlace.

- **La AEPD impone una multa de 5.000€ por una infracción del artículo 5.1.c) "minimización de datos".** Se interpone una reclamación ante la AEPD principalmente por la publicación de escritos en internet con su contenido íntegro, figurando en la mayoría de ellos los nombres y apellidos de los reclamantes. Concretamente, el número del DNI de un reclamante, como parte de la firma digital, sin anonimizar. En los fundamentos de derecho, la AEPD establece: *"Respecto del DNI, corresponda éste a una persona de carácter público o una persona de carácter privado, o profesional, se entiende que el conocimiento de este dato no es relevante a los efectos de la finalidad del tratamiento de B.B.B., toda vez que sin el mismo se cumple la finalidad del mismo. En este caso, se impone la anonimización de dicho dato. La firma de un documento como el que es objeto de difusión, un escrito privado de reclamante no precisa de asociar el DNI/NIF de su titular por mas que vaya asociado o incluido en la firma electrónica de este, ya que suficiente es la información e identificación por el nombre y apellidos que en otros muchos escritos ya figuran del mismo"*. Consulta la resolución en este enlace.