6. CÓMO SEUDONIMIZAR O AGREGAR LOS DATOS ANTES DE COMPARTIRLOS

TERRE DES HOMMES - LAUSANNE

1. Objetivos del documento ..... Error! Bookmark not defined.
2. ¿Por qué tenemos que des-identificar los datos antes de compartirlos? 2
3. ¿Qué tipo de datos tienen que ser des-identificados? 2
4. Técnicas de des-identificación o anonimización 3

1. Resumen ejecutivo

Después de leer este documento, usted sabrá como seudonimizar o agregar los datos sensibles antes de compartirlos.

¿Por qué es importante para Terre des hommes la manera en que se seudonimiza o agrega los datos antes de compartirlos?

Seudonimizar o agregar datos es útil para poder compartir datos con las partes interesadas clave (aliados, donantes) sin poner la información sensible en riesgo.

2. ¿Por qué tenemos que des-identificar los datos antes de compartirlos?

Al compartir datos fuera de su propia organización, necesita mantener un cierto control sobre esta información y des-identificarla cuando tiene que ver con beneficiarios o actividades sensibles.

La agregación de datos es un punto crucial de la evaluación de riesgos: en ocasiones datos independientes considerados seguros se convierten en datos que pueden causar un perjuicio al combinarlos con otros conjuntos de datos, o datos que usted pensaba se habían anonimizado se hacen fácilmente discernibles una vez que se les combina con otros datos, usando técnicas de triangulación.

Los datos pueden tener consecuencias reales para las personas, y a menudo estas consecuencias son tan involuntarias como perjudiciales. Este es por lo general el caso cuando se comparte o publica ***Información de Identificación Personal (IIP)***, o cuando datos aparentemente inocuos se mezclan y compilan con otros conjuntos de datos.

3. ¿Qué tipo de datos tiene que ser des-identificados?

La respuesta corta es, todos y cada uno de los datos que tengan algún potencial de identificar a una persona.

- Edad
- Etnicidad
- Género
- Distrito/país
- Nivel más alto de educación
- Medicamentos recetados
- Geo-localización desde donde se inicia la sesión
- …

Es importante realizar una evaluación de riesgos concienzuda de las posibles consecuencias de la divulgación de datos durante el diseño del proyecto y actualizarla antes de su publicación (vea la "Herramienta de autoevaluación de protección de los datos" que le ayudará en las primeras etapas de una evaluación de riesgos). Una vez que los datos se han divulgado, es muy tarde. Tiene que considerar lo siguiente:

- ¿Se puede identificar a una persona a partir de los datos, o, a partir de los datos y otra información relativamente accesible?
- ¿Se 'refieren' los datos a la persona identificable, ya sea en su vida personal o familiar, en la empresa o profesión?
- ¿Son los datos 'obviamente sobre' una persona en particular?
- ¿Están los datos `vinculados a' una persona de manera que proporcionan información particular sobre esta?
- ¿Se utilizan o se utilizarán los datos para informar o influir en acciones o decisiones que afecten a una persona identificable?
- ¿Tienen los datos algún significado biográfico en relación con la persona?
- ¿Se centran o se concentran los datos en la persona como su tema central en lugar de en algún objeto, transacción o evento?

- ¿Impactan los datos o tienen el potencial de impactar en una persona, ya sea en el ámbito personal, familiar, empresarial o profesional?

Si la respuesta a cualquiera de estas preguntas es "**sí**" o incluso "**tal vez**", **inecesita anonimizar sus datos!**

La información que a primera vista parece no ser identificable podría hacerse identificable, especialmente cuando se le combina con otros datos y se le somete a un análisis algorítmico poderoso. Esta lista da algunos ejemplos, pero no es en absoluto exhaustiva.

⚠ Tenga presente que algo tan básico como tener la categoría de edad de un menor vinculada al estado de desplazamiento de sus padres a veces puede ser suficiente para identificarlo, idependiendo del perímetro de los datos!

4. Técnicas de des-identificación o anonimización

En caso sea necesario hacer una anonimización, existen algunas técnicas que se pueden utilizar:

- ***Enmascaramiento de datos****: Esta técnica consiste en proporcionar solamente una parte de un conjunto de datos (por ejemplo, extracción de columnas de una hoja de cálculo) o la eliminación completa de estas partes del conjunto de datos (por supuesto, iocultar columnas no es suficiente!). Asegúrese de guardar de forma segura una copia de los datos sensibles que extrajo de su conjunto de datos antes de empezar a trabajar en este.*

- ***Seudonimización****: Esta técnica consiste en el intercambio de valores por códigos, de esta manera, por ejemplo, un nombre puede ser reemplazado por un número, pero el mismo número aparecerá en todos los casos en los que estaba el nombre (por ejemplo, se utiliza el código "XYZ" en lugar de "John Bates"). Asegúrese sin embargo de que su código no sea demasiado obvio para que no se pueda encontrar demasiado fácilmente (por ejemplo: crear códigos a partir de fechas de nacimiento).*
- ***Agregación****: En lugar de proporcionar los datos brutos, los datos compartidos agruparán la información sobre la base de un criterio determinado, mediante el intercambio de sumas de datos que responden a determinados criterios, por ejemplo ("23 niños que corren el riesgo de sufrir abuso físico" en lugar de una lista con los nombres de los niños afectados), o proporcionando solo pequeñas muestras de conjuntos de datos más amplios.*
- ***Datos derivados****: Esta técnica consiste en un proceso en el que se mantiene la intención original, pero se cambia la salida. Un ejemplo sería proporcionar la edad de una persona en lugar de la fecha exacta de nacimiento.*