

# Anonimización de datos personales para la investigación

---

Perspectiva jurídica y práctica

Mallorca 2018



Gonzalo Álvarez Hazas  
Abogado especialista en Derecho de las Tecnologías

### ÍNDICE

#### **1. EL DATO ANÓNIMO**

- I. ¿Qué es un dato anónimo?

#### **2. MARCO NORMATIVO**

- I. Protección de datos
- II. Legislación sanitaria
- III. Seguridad jurídica

#### **3. CONSENTIMIENTO**

- I. ¿Es necesario?
- II. Impacto del RGPD y Doctrina

#### **4. PERSPECTIVA PRÁCTICA**

- I. Criterios generales
- II. Recomendaciones

#### **5. CONCLUSIONES**

### OBJETIVOS DE LA JORNADA

- 1 •Obtener una definición de anonimizar
- 2 •Conocer el marco normativo básico y distinguir otros conceptos
- 3 •Aportar pautas que ayuden a fundamentar jurídicamente los procesos de anonimización.  
•Reconocer riesgos específicos.



**SEGURIDAD JURÍDICA**

### 1. EL DATO ANÓNIMO

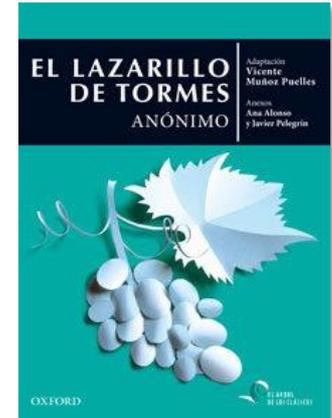
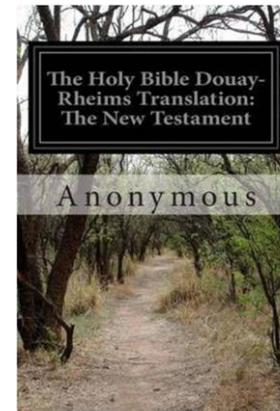
#### ¿Qué es un dato anónimo?



¿Se trata de no ser conocido de ningún modo, de manera irreversible y definitiva?

¿O existen posibilidades de identificar a una persona?

¿Quién escribió la Biblia o el Lazarillo de Tormes?



### 1. EL DATO ANÓNIMO

¿Qué es un dato anónimo?



¿Esta fotografía es un dato anónimo?

¿Y esta fotografía?



### 1. EL DATO ANÓNIMO

#### ¿Qué es un dato anónimo?

Definición en el diccionario de la Real Academia de la Lengua

**anonimizar**

Conjugar

*De anónimo e -izar.*

1. **tr.** Expresar un dato relativo a entidades o personas, eliminando la referencia a su identidad.

### 2. MARCO NORMATIVO

#### I. Protección de Datos

##### Legislación en vigor

#### **Ley Orgánica de Protección de Datos y Reglamento de desarrollo**

- Sin definición de anonimizar
- Se define disociar y en el Proyecto de Ley, seudonimizar

#### **Reglamento Europeo de Protección de Datos (RGPD)**

Sin referencias ni definiciones en el articulado, sólo en el considerando 26.

*[...]. Por lo tanto los principios de protección de datos no deben aplicarse a la información anónima, es decir información que no guarda relación con una persona física identificada o identificable, ni a los datos convertidos en anónimos de forma que el interesado no sea identificable, o deje de serlo. En consecuencia, el presente Reglamento no afecta al tratamiento de dicha información anónima, inclusive con fines estadísticos o de investigación.*

## 2. MARCO NORMATIVO

### I. Protección de Datos

#### Doctrina de la Agencia Española de Protección de Datos

Guía: Orientaciones y garantías en los procedimientos de anonimización de datos personales



*La finalidad del proceso de anonimización es eliminar o reducir al mínimo los riesgos de reidentificación de los datos anonimizados manteniendo la veracidad de los resultados del tratamiento de los mismos, es decir, además de evitar la identificación de las personas, los datos anonimizados deben garantizar que cualquier operación o tratamiento que pueda ser realizado con posterioridad a la anonimización no conlleva una distorsión de los datos reales. [...]. En el proceso de anonimización se deberá producir la ruptura de la cadena de identificación de las personas [...]*

## 2. MARCO NORMATIVO

### I. Protección de Datos

#### Doctrina del Grupo del Artículo 29

Dictamen 05/2014 sobre técnicas de anonimización

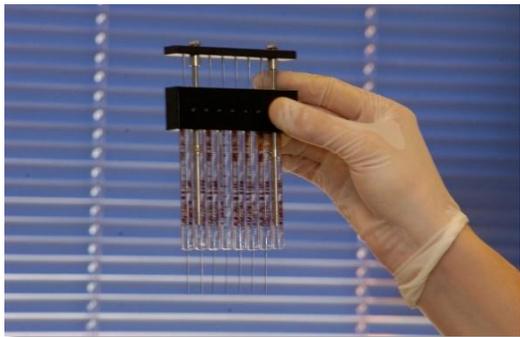


*[...] para anonimizar cualesquiera datos es necesario eliminar de ellos los elementos suficientes para que no pueda identificarse al interesado. Con más precisión, hay que tratarlos de tal manera que no puedan usarse para identificar a una persona física mediante «el conjunto de los medios que puedan ser razonablemente utilizados» por el responsable del tratamiento o por terceros. [...]*

### 2. MARCO NORMATIVO

#### II. Legislación Sanitaria

#### Artículo 3 de la Ley 14/2007 de Investigación biomédica



CC BY-3.0-ES 2012/EJ-GV/Irekia-Gobierno Vasco/Mikel Arrazola

*c) «Anonimización»: proceso por el cual deja de ser posible establecer por medios razonables el nexo entre un dato y el sujeto al que se refiere. Es aplicable también a la muestra biológica.*

*h) «Dato anónimo»: dato registrado sin un nexo con una persona identificada o identificable.*

*i) «Dato anonimizado o irreversiblemente disociado»: dato que no puede asociarse a una persona identificada o identificable por haberse destruido el nexo con toda información que identifique al sujeto, o porque dicha asociación exige un esfuerzo no razonable, entendiéndose por tal el empleo de una cantidad de tiempo, gastos y trabajo desproporcionados.*

### 2. MARCO NORMATIVO

#### III. Seguridad jurídica

#### Definición de anonimizar

REGULACIÓN	DEFINICIÓN
Protección de datos	<ul style="list-style-type: none"><li>• eliminar o reducir al mínimo los riesgos de reidentificación de los datos anonimizados</li><li>• de tal manera que no puedan usarse para identificar a una persona física mediante «el conjunto de los medios que puedan ser razonablemente utilizados» por el responsable del tratamiento o por terceros</li></ul>
Sanitaria	proceso por el cual deja de ser posible establecer por medios razonables el nexo entre un dato y el sujeto al que se refiere

### 2. MARCO NORMATIVO

#### III. Seguridad jurídica

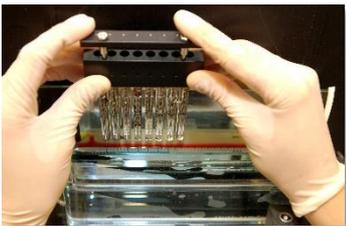
##### Ejemplos



Si dispongo de la foto original, he realizado un proceso de anonimización, dado que, he reducido el riesgo de reidentificación.



La eliminación de datos de identificación directa como nombre y apellidos puede ser anonimización, pero el riesgo de reidentificación es alto utilizando “medios razonables”.



La aplicación de un código a una muestra es un proceso de anonimización y se reduce el riesgo de reidentificación hasta casi eliminarlo.

## 2. MARCO NORMATIVO

### III. Seguridad jurídica

#### Otros conceptos

#### **Seudonimizar** en el RGPD y en el proyecto de LOPD:

*“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”*

#### **Disociar** en la LOPD y su Reglamento de desarrollo

**Dato disociado:** *aquél que no permite la identificación de un afectado o interesado*

**Procedimiento de disociación:** *todo tratamiento de datos personales de modo que la información que se obtenga no pueda asociarse a persona identificada o identificable*

### 2. MARCO NORMATIVO

#### III. Seguridad jurídica

##### Otros conceptos

CONCEPTO	SIGNIFICADO
Anonimizar	Proceso por el cual deja de ser posible establecer por medios razonables el nexo entre un dato y el sujeto al que se refiere
Seudonimizar	Medida de seguridad para añadir o garantizar confidencialidad. Puede utilizarse en el ámbito de la anonimización. Sustitución de un atributo (normalmente un atributo único) por otro en un registro
Disociar	Proceso irreversible, destruimos el nexo de unión entre el dato personal y la información resultante

### 3. EL CONSENTIMIENTO

#### I. ¿Es necesario para la investigación biomédica?



CC BY-3.0-ES 2012/EJ-GV/Irekia-Gobierno Vasco/Mikel Arrazola

¿El consentimiento es necesario en el ámbito de la investigación biomédica?

¿Qué tipo de consentimiento?

¿Y si los datos se encuentran anonimizados?

Recordar que a la información resultante de la anonimización no se aplican los principios de protección de datos

### 3. EL CONSENTIMIENTO

#### I. ¿Es necesario para la investigación biomédica?

#### Ley 41/2002 de autonomía del paciente

##### **Artículo 8 consentimiento informado**

1. *Toda actuación en el ámbito de la salud de un paciente necesita el consentimiento libre y voluntario del afectado, una vez que, recibida la información prevista en el artículo 4, [...].*

2. *El consentimiento será verbal por regla general.*

*Sin embargo, se prestará por escrito en casos de intervención quirúrgica, procedimientos diagnósticos y terapéuticos invasores y, en general, aplicación de procedimientos que suponen riesgos o inconvenientes de notoria y previsible repercusión negativa sobre la salud del paciente. [...]*

4. *Todo paciente o usuario tiene derecho a ser advertido sobre la posibilidad de utilizar los procedimientos de pronóstico, diagnóstico y terapéuticos que se le apliquen en un proyecto docente o de investigación, que en ningún caso podrá comportar riesgo adicional para su salud.*



**Acepto**

### 3. EL CONSENTIMIENTO

#### I. ¿Es necesario para la investigación biomédica?

#### Ley 14/2007 de investigación biomédica

Consentimiento para tratar y ceder a terceros datos de investigación

#### **Artículo 4**

1. *Se respetará la libre autonomía de las personas que puedan participar en una investigación biomédica o que puedan aportar a ella sus muestras biológicas, para lo que será preciso que hayan prestado previamente su consentimiento expreso y escrito una vez recibida la información adecuada.*

#### **Artículo 5**

2. *La cesión de datos de carácter personal a terceros ajenos a la actuación médico-asistencial o a una investigación biomédica, requerirá el consentimiento expreso y escrito del interesado.*

5. *Si no fuera posible publicar los resultados de una investigación sin identificar a la persona que participó en la misma o que aportó muestras biológicas, tales resultados sólo podrán ser publicados cuando haya mediado el consentimiento previo y expreso de aquélla.*



**Acepto**

### 3. EL CONSENTIMIENTO

#### II. Impacto del RGPD y Doctrina de las Agencias

#### Protección de datos

**LOPD** sólo podrán recabarse datos de salud cuando lo disponga una ley (interés general) o la persona afectada consienta expresamente. (art. 7.3 LOPD)

#### **Proyecto de LOPD:**

*2. El tratamiento de datos en la investigación en salud se regirá por los siguientes criterios:*

*a) El interesado o, en su caso, su representante legal podrá otorgar el consentimiento para el uso de sus datos con fines de investigación en salud y, en particular, la biomédica. Tales finalidades podrán abarcar categorías relacionadas con áreas generales vinculadas a una especialidad médica o investigadora.*



CC BY-3.0-ES 2012/EJ-GV/Irekia-Gobierno Vasco/Mikel Arrazola

### 3. EL CONSENTIMIENTO

#### II. Impacto del RGPD y Doctrina de las Agencias

¿Qué tipo de consentimiento recabar? ¿Expreso? ¿Expreso y por escrito?  
¿Previo y expreso? ¿Informado?

#### IDEA DESTACADA

Con el RGPD desaparece la distinción entre consentimiento expreso y tácito

### 3. EL CONSENTIMIENTO

#### II. Impacto del RGPD y Doctrina de las Agencias

#### Informe AEPD 2018

#### **Informe 073667/2018 de la Agencia Española de Protección de Datos**



Objeto: informe acerca de la incidencia que en el ámbito de la investigación biomédica puede producir la plena aplicación del RGPD.

Se realiza contraste entre las reglas del consentimiento contempladas en la legislación de investigación biomédica (Ley 14/2007) y el RGPD

### 3. EL CONSENTIMIENTO

#### II. Impacto del RGPD y Doctrina de las Agencias

##### Informe AEPD 2018

**Consentimiento:** Recabar de modo explícito (libre, voluntario, inequívoco, informado...) manifestación de la voluntad positiva. Carga de la prueba.



*Artículo 17 Ley 41/2002 Conservación de la documentación clínica*

*2. La documentación clínica también se conservará a efectos judiciales de conformidad con la legislación vigente. Se conservará, asimismo, cuando existan razones epidemiológicas, de investigación o de organización y funcionamiento del Sistema Nacional de Salud. Su tratamiento se hará de forma que se evite en lo posible la identificación de las personas afectadas.*

### 3. EL CONSENTIMIENTO

#### II. Impacto del RGPD y Doctrina de las Agencias

##### Informe AEPD 2018

**Finalidad:** Si la base jurídica es la investigación biomédica y la propia legislación sanitaria, no interpretar de modo restrictivo, limitado a una concreta investigación.

*“será suficientemente inequívoco y específico el consentimiento prestado en relación con una rama amplia de investigación, como por ejemplo, la investigación oncológica, o incluso para ámbitos más extensos”*



CC BY-3.0-ES 2012/EJ-GV/Irekia-Gobierno Vasco/Mikel Arrazola



### REFLEXIÓN

**Demostrar haber obtenido el consentimiento informado es un problema de prueba**

**¿cómo conservar traza y garantías del otorgamiento?**

¿Es posible digitalizar un consentimiento en papel?



¿Y recabarlo mediante tableta?



### 4. PERSPECTIVA PRÁCTICA

#### I. Criterios generales



### 4. PERSPECTIVA PRÁCTICA

#### I. Criterios generales

##### Referencias doctrinales

- Guía de la Agencia de Protección de Datos: Orientaciones y garantías en los procedimientos de anonimización de datos personales
- Dictamen del Grupo del Artículo 29: Dictamen 05/2014 sobre técnicas de anonimización

### 4. PERSPECTIVA PRÁCTICA

#### I. Criterios generales

##### Principios generales



CC BY-3.0-ES 2012/EJ-GV/Irekia-Gobierno Vasco/Mikel Arrazola

1. Principio proactivo. *A priori, no reactivo. Ej. Clasificación de datos (microdatos, identificación indirecta, sensibles) y valoración cuantitativa.*
2. Privacidad por defecto: *en el proceso*
3. Privacidad objetiva: *Índice residual de reidentificación*
4. Principio de plena funcionalidad: *garantía de no distorsión de los datos (Ej, distorsión geográfica por enfermedades raras)*
5. Privacidad en el ciclo de vida de la información: *auditorías, destrucción, etc.*
6. Principio de información y formación: *personal involucrado*

### 4. PERSPECTIVA PRÁCTICA

#### I. Criterios generales

#### Fases de la anonimización

1. Equipo de trabajo e independencia de funciones
2. **Evaluación de riesgos de reidentificación**
3. Definición de objetivos y finalidad de la información anonimizada
4. Viabilidad del proceso
5. Preanonimización: definición de variables de identificación
6. Eliminación/reducción de variables
7. Selección de técnicas de anonimización
8. Segregación de la información: *entornos separados*
9. *Proyecto piloto*
10. **Anonimización**



CC BY-3.0-ES 2012/EJ-GV/Ireka-Gobierno Vasco/Mikel Arrazola

### 4. PERSPECTIVA PRÁCTICA

#### II. Recomendaciones

#### Evaluación de riesgos de reidentificación

Variable de identificación	Probabilidad	Impacto Gravedad	Valoración del riesgo
<ul style="list-style-type: none"><li>• Microdato</li><li>• Identificación indirecta</li><li>• Datos sensibles</li><li>• Sin datos</li><li>• Otros</li></ul>	Identificación de riesgo valor determinado de una escala cuantitativa o cualitativa. Ej, vulneración del deber de secreto; riesgos de revelación de claves de anonimización; existencia de un atacante/adversari o potencial, equipo de trabajo	Riesgos de reidentificación <ul style="list-style-type: none"><li>• Ej, bajo, medio, alto</li><li>• Sin consecuencias, sin consecuencias reseñables, consecuencias graves, etc</li></ul>	Resultado del valor y nivel de riesgo <ul style="list-style-type: none"><li>• Ej, muy bajo a crítico</li><li>• De 1 a 10 u otras escalas</li></ul>

Informe de análisis de riesgos  
**matriz**



Determinación del umbral de riesgo aceptable  
**riesgo residual**



Gestión del riesgo en caso de alto riesgo **EIPD**

### 4. PERSPECTIVA PRÁCTICA

#### II. Recomendaciones

##### Riesgos clave

1. Singularización  
*posibilidad de extraer de un conjunto de datos algunos registros (o todos los registros) que identifican a una persona*
2. Vinculabilidad  
*capacidad de vincular como mínimo dos registros de un único interesado o de un grupo de interesados*
3. Inferencia  
*posibilidad de deducir con una probabilidad significativa el valor de un atributo a partir de los valores de un conjunto de otros atributos*

##### Técnicas de anonimización

- Algoritmos de hash [*Generar claves para los datos*]
- Algoritmos de cifrado [*clave de descifrado*]
- Sello de tiempo
- Capas de anonimización [*Por organización, por tipos o variables de datos, etc*]
- Perturbación de datos [*Permutación, redondeo, ruido aleatorio, microagregación, etc.*]
- Reducción de datos [*Eliminación de variables, reducción de registros, supresión de registros, etc.*]
- Seudonimización [*sustitución de un atributo (normalmente un atributo único) por otro en un registro*]

### 4. PERSPECTIVA PRÁCTICA

#### II. Recomendaciones



#### Ejemplos

- Estudio de hábitos de vida saludable de la población juvenil de 14 a 18 años

*Si tenemos datos de menores de edad y conocemos que algunos consumen alcohol y además enriquecemos los datos con ubicación geográfica/sexo/talla/peso/IMC/ejercicio físico/comidas al día/Organización de servicios sanitarios/ejercicio físico, etc*

- Perfiles genéticos

*En riesgo de ser identificados si solo se utiliza la técnica de eliminación de la identidad del donante. Diversos estudios científicos han demostrado que, al combinar los recursos genéticos disponibles para el público (p. ej., registros genealógicos, obituarios y resultados de consultas en motores de búsqueda) y los metadatos sobre donantes de ADN (fecha de donación, edad o lugar de residencia), se puede revelar la identidad de determinadas personas aunque el ADN se haya donado de forma «anónima»*

### 4. PERSPECTIVA PRÁCTICA

#### II. Recomendaciones

##### IDEA DESTACADA

##### **Protección de datos**

Obligación de resultado, no de medios

No hay ninguna resolución sancionadora de las Agencias por aplicar erróneamente técnicas de anonimización

### 4. PERSPECTIVA PRÁCTICA

#### II. Recomendaciones

#### Ejemplos

- Resoluciones sancionadoras AGPD

#### **Error de anonimización, Word transformado a pdf**

*no utilizar editores que simplemente tapan u ocultan datos mediante etiquetas sobre impresas.*

*“el mero ocultamiento, tachado o sombreado en negro de esos datos no conlleva la anonimización de los mismos, máxime cuando dicha información de carácter personal puede resultar accesible a través de los motores de búsqueda”*

#### **Radiografía con datos personales publicada en internet**

*Vulneración del principio de seguridad, por revelar información personal a terceros sin autorización previa.*

#### RESOLUCIÓN: R/02202/2015

En el procedimiento sancionador PS/00368/2015, instruido por la Agencia Española de Protección de Datos a la entidad SOCIEDAD ESPAÑOLA DE CIRUGIA Y TRAUMATOLOGIA, vista la denuncia presentada por ██████████ en virtud de los siguientes,

#### ANTECEDENTES

**PRIMERO:** Con fecha de 25 de julio de 2014 tiene entrada en esta Agencia un escrito del ██████████ el que declara que con fecha 22 de julio de 2014, ha realizado una búsqueda en Google sobre una enfermedad de la cual ha sido diagnosticado, y en uno de los resultados de la búsqueda ha verificado que se ha hecho público su historial clínico con el diagnóstico y sus datos personales en una imagen de una Resonancia Magnética.

Aporta copia de su historia clínica publicada en la página web [www.secot.es](http://www.secot.es)..... en el epígrafe “Visor web de casos clínicos de residentes en Cirugía Ortopédica y Traumatología”. La historia del denunciante se publica con el ██████████ como autores varios médicos del Complejo H. Universitario de Badajoz. En la historia no constan datos identificativos del paciente, no obstante, se publica una imagen de una prueba diagnóstica, donde figuran el nombre y apellidos del paciente (denunciante), y su fecha de nacimiento.



CC BY-3.0-ES 2012/EJ-GV/Ireki-Gobierno Vasco/Jon Bernardes

### 5. CONCLUSIONES

#### 1. Qué es un dato anónimo

Aquella información resultante de un proceso por el cual no puede identificarse a una persona física de manera directa o indirecta.

El dato personal se mantiene en la fuente original, puede ser una muestra, una imagen, un episodio, la propia historia clínica y evidentemente, el consentimiento informado.



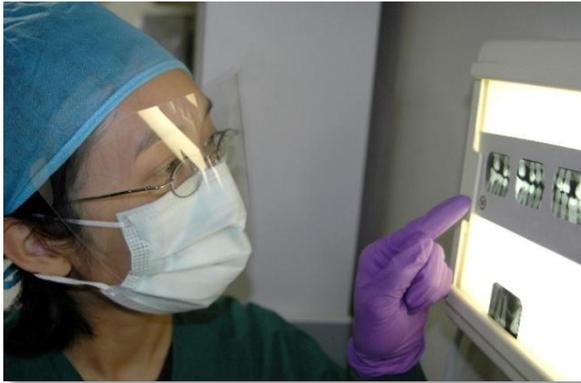
*CC BY-3.0-ES 2012/EJ-GV/Irekia-Gobierno Vasco/Mikel Arrazola*

#### Otros conceptos:

- seudonimización (medida de seguridad)
- disociación (proceso irreversible)

### 5. CONCLUSIONES

#### 2. Consentimiento informado



- Recabar de manera explícita (manifestación inequívoca de la voluntad libre, específica e informada). Ponderar el riesgo por el beneficio para ser garantistas. Carga de la prueba.
- Finalidades de investigación de modo genérico y no limitativo/restrictivo. Porque prima el interés general/público. Interpretación flexible.

### 5. CONCLUSIONES

#### 3. Proyecto de anonimización

- No existe el riesgo cero
- Documentar: incluir **informe de análisis de riesgos** (Evaluación de impacto por probabilidad), tiene su encaje en protección de datos en coherencia con el principio de responsabilidad proactiva del RGPD.
- Nivel alto de riesgo: exige EIPD, gestión del riesgo para minimizar, reducir o eliminar.
- Obligación de resultado, no de medios: NO hay resoluciones sancionadoras sobre procesos de anonimización. Ni siquiera relativa medidas de seguridad técnicas. (Ej, porque no se hayan cifrado correctamente los datos)
- Sector público



CC BY-3.0-ES 2012/EJ-GV/Irekia-Gobierno Vasco/Mikel Arrazola

### Curiosidad



**en la nueva Ley Orgánica de Protección de Datos se tipifica como muy grave la reversión de la anonimización?**

El artículo 72 apartado p) regula como muy grave “La reversión deliberada de un procedimiento de anonimización a fin de permitir la reidentificación de los afectados”

**Para la reflexión ¿Es posible cometer una infracción sobre anonimización cuando no existe una definición en la Ley?**

### **Sugerencia de lectura**

Disposición adicional decimoséptima de la futura LOPD. Tratamientos de datos de salud.

**MUCHAS GRACIAS POR SU ATENCIÓN**

Contacto: [galvarez@globalfactory.es](mailto:galvarez@globalfactory.es)

Más información en [www.gahazas.com](http://www.gahazas.com)

