PROTECCIÓN DE DATOS: LA «SEUDOANONIMIZACIÓN» INEXISTENTE

Ma Luisa Romana García

Profesora Universidad Pontificia Comillas (Madrid)

Blanca Hernández Pardo

Profesora Universidad Pontificia Comillas (Madrid)

SUMARIO: 1. Protección de datos y despersonalización. 1.1. La protección de la identidad de las personas físicas; 1.2. Procedimientos y nomenclaturas en protección de datos; 1.3. Acuñación de una palabra nueva fuera de la legislación: seudoanonimización; **2. Lenguaje natural y lenguaje especializado.** 2.1. El significado de las palabras en el lenguaje común; 2.2. La creación de la terminología; 2.3. El fenómeno de reanálisis; **3. El caso de la seudoanonimización.** 3.1. La derivación; 3.2. Hibridación de ambas raíces léxicas; 3.3. La anonimización de datos; 3.4. La seudonimización; **4. Resumen y conclusiones; 5. Bibliografías y referencias.**

RESUMEN

Estudiamos la aparición y uso del nuevo término «seudoanonimización» en el ámbito biosanitario; es un vocablo no procedente de los textos legislativos, sino acuñado directa e inconscientemente por los hablantes.

PALABRAS CLAVE

Protección de datos, terminología, seudoanonimización, anonimización, seudonimización.

ABSTRACT

We discuss the emergence of a new term "pseudoanonymization" in the biosanitary field; this word is not to be found in rules or regulations, but has been unconsciously coined by speakers.

KEYWORDS

Data protection, terminology, pseudoanonymization, anonymization, pseudonymization.

1. PROTECCIÓN DE DATOS Y DESPER-SONALIZACIÓN

1.1 La protección de la identidad de las personas físicas

En el ámbito jurídico de la protección de datos, una de las principales preocupaciones del legislador es la posibilidad de reconocer la identidad de las personas cuya información se incluye en bancos o bases de datos; en efecto, de quedar esta circunstancia al margen del ordenamiento jurídico, podrían darse claras vulneraciones de los derechos individuales y, específicamente, el derecho al honor y a la intimidad personal y familiar (artículo 1 de la Ley Orgánica de Protección de Datos, BOE, 1999a).

Por lo tanto, para proteger los derechos fundamentales de las personas físicas, es esencial impedir que los datos que se manejan y almacenan en cualquier tipo de formato puedan llevar a identificar a las personas de las cuales proceden, los «afectados» en la terminología jurídica. Para garantizar esto, se trata de obligar a las entidades que manejan tales datos a que borren o retiren de la masa de datos toda posibilidad de identificación personal.

1.2 Procedimientos y nomenclaturas en protección de datos

El derecho maneja varios conceptos. En primer lugar se habla de disociación, consistente en separar los datos de las personas, para que no haya forma de identificarlas; este procedimiento estaría estrechamente relacionado con la codificación, que oculta tales datos mediante códigos; este procedimiento es una disociación reversible, esto es, la identidad de las personas se puede recuperar. De ahí se pasa a la anonimización, que es una disociación irreversible; los datos que se anonimizan dejan de permitir la identificación de sus titulares, bien porque tal operación sea imposible, bien porque exija unos medios desproporcionados que no es probable se puedan emplear. La última técnica es la seudonimización o pseudonimización, que oculta la identidad pero permite volver a identificarla en caso necesario. El derecho establece salvaguardas; si los datos se anonimizan dejan de quedar comprendidos en el ámbito de la protección de datos, cuyas normas ya no se les aplicarán; en el caso de información seudonimizada, los instrumentos que permitirían volver a personalizar los datos deben mantenerse por separado y estar debidamente protegidos contra malas utilizaciones.

Al margen de las consideraciones jurídicas, se observa la presencia de un neologismo derivado por reanálisis, el verbo «(p)seudoanonimizar», que en nuestra opinión merece un atento estudio al objeto de delimitar con claridad los usos y sentidos que puede tener este procedimiento de protección de los datos.

1.3 Acuñación de una palabra nueva fuera de la legislación: *seudoanonimización*

Como decimos, en un principio las normas hablaban de «datos disociados» (DOUE, 2000a); en unos años se añadió a éste el concepto de «codificación» (BOE, 2007). En el mismo 2007 se define por primera vez la «anonimización» (BOE, 2007),¹

proceso que impide identificar personas concretas a partir de los datos. Como este efecto es muy difícil de conseguir –con la tecnología actual es poco menos que imposible eliminar totalmente esta posibilidad. se introduce un procedimiento menos definitivo, consistente en sustituir los datos por códigos que deshagan la identificación: la legislación europea considera estos códigos como seudónimos, y su aplicación se denomina, naturalmente, «seudonimización» (DOUE, 2007a). Este proceso no va más allá de una mera recodificación, sustituyendo unos códigos identificables por otros que no lo sean: pero el lector experto, posiblemente, no está interpretando este simple efecto, sino que está entendiendo que hablamos de un proceso que no constituye una auténtica anonimización, esto es, una especie de anonimización desleída, que se le acerca pero no llega a serlo. Y esta interpretación le lleva de manera natural (el proceso mental se llama en lingüística «reanálisis», fenómeno sobre el que volveremos algo más adelante), a acuñar la palabra seudo + anonimización, igual que hablamos por ejemplo de «seudociencia», «seudointelectuales» o «seudoliteratura», siempre en el sentido de una versión desvirtuada del original.

Así, para denominar una cosa que no es *anonimizar* pero sí *casi anonimizar*, se ha acuñado naturalmente este neologismo, que por otra parte no es propio solamente de la lengua española.² En lugar de interpretar el procedimiento en el sentido de una recodificación, los especialistas piensan en una versión suave de la anonimización, sin ser conscientes del malentendido.

Entender este fenómeno precisa el manejo de conceptos ajenos al Derecho; tenemos que adentrarnos en la descripción del lenguaje natural y en lo que pasa en la mente del hablante cuando lo utiliza.

2. LENGUAJE NATURAL Y LENGUAJE ESPECIALIZADO

2.1 El significado de las palabras en el lenguaje común

La relación entre palabras y objetos del mundo no es ni predeterminada ni directa: se trata de un acuerdo social, un vínculo que se establece o actualiza en

¹ Encontramos «anonimización» en este sentido en la legislación española ya en el año 1993, en una norma que gobierna la protección de datos en el Centro de Investigaciones Sociológicas (BOE, 1993), pero hasta el año 2007 no recibe una

definición jurídica plena, en la Ley de Investigación Biomédica (BOE, 2007). La legislación europea describe el mismo proceso en 2008 (DOUE, 2008b).

² Un fenómeno equivalente se ha producido también en otras lenguas, como el inglés o el francés.

el cerebro cada vez que hablamos. Contrariamente a lo que suele suponerse, el significado **no** es una propiedad de las palabras: la definición más precisa lo describiría como la actividad sináptica que se produce en el cerebro humano como reacción a la presencia de las palabras.

Ya en 1923 los autores Ogden y Richards establecieron, en el conocido como «triángulo semiótico» (Fig. 1)³, la relación indirecta entre palabras y entidades; la palabra no está vinculada directamente con la realidad –si así fuera no habría lenguas distintas–, sino con un pensamiento, esto es, una o varias pautas sinápticas en el cerebro del hablante.

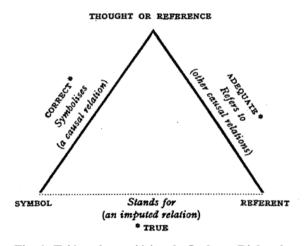


Fig. 1: Triángulo semiótico de Ogden y Richards

Por otra parte, para la escuela cognitiva (el llamado «enfoque cognitivo-prototípico»),⁴ el significado es una acumulación de ideas: a partir de un sentido básico que estaría en el centro de nuestra interpretación, las palabras van sumando nuevas acepciones⁵ mediante relaciones semánticas como la metonimia, generalización, especialización o metáfora. Así, cualquier palabra o expresión permite añadir significados o usos distintos, nuevos, aplicando dichas relaciones de sentido: las piezas léxicas del lenguaje natural son abiertas y no delimitadas, con fronteras semánticas borrosas y nunca definitivas. Así lo representa Geeraerts (Fig. 2):⁶

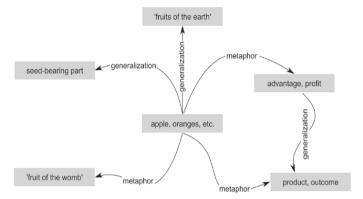


Fig. 2: Red radial del significado, según Geeraerts

Así pues, el significado no puede considerarse una propiedad de las palabras. Con mayor precisión puede definirse como la relación entre palabras e ideas que el cerebro establece activando determinadas redes sinápticas: cuando nos hallamos ante un mensaje, las palabras nos provocan una reacción en el cerebro. Técnicamente, es esta activación neuronal lo que permite que una expresión llame a nuestra mente entidades, sucesos o ideas. Así, el significado lingüístico solo puede describirse como una potencialidad abierta. Los diccionarios no deben dictar -a modo de prescripción- cómo deben utilizarse las palabras ni la interpretación que hay que darles; solo pueden limitarse a registrar sentidos compartidos, comunes a varios hablantes y más generalizados o consolidados cuanto mayor sea el número de los mismos. Veamos un ejemplo. El DRAE⁷ recoge la palabra «mesa» con el siguiente significado en primera acepción:

mesa

Del lat. mensa.

 f. Mueble compuesto de un tablero horizontal liso y sostenido a la altura conveniente, generalmente por una o varias patas, para diferentes usos, como escribir, comer, etc.

Los hablantes utilizan esta misma palabra en muchos otros sentidos, que igualmente recoge el

³ OGDEN, C. K. y RICHARDS, I. A.: The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism, Magdalene College: University of Cambridge, Cambridge 1923, p-11.

⁴ Este concepto puede consultarse en diversas referencias, entre otrosGeeraerts, D. y Cuyckens, H. (eds.): *The Oxford Handbook of Cognitive Linguistics*, Oxford University Press, Oxford 2007.

⁵ Cf. DIRVEN, R. y VERSPOOR, M.: Cognitive Exploration of Language and Linguistics, John Benjamins, Amsterdam/Philadelphia 1998, pp. 32-37.

⁶ GEERAERTS, D.: *Theories of Lexical Semantics*, Oxford University Press, Oxford2009, p.195.

⁷ RAE: Diccionario de la lengua española. [online]. Disponible en la URL: http://dle.rae.es/>. [Con acceso el 21.11.2017].

diccionario: encontramos no menos de 14 acepciones para este sencillo vocablo. Pero, aparte de estos catorce, es muy fácil añadir un nuevo uso; por ejemplo, si en un restaurante oímos la frase «la mesa 7 se ha ido sin pagar», inmediatamente interpretamos que la palabra *mesa* se está utilizando en lugar de «las personas que se han sentado en la mesa»; por abreviar. Vemos, pues, que las palabras no tienen naturalmente un significado fijo, como decíamos.

La Real Academia recoge los significados, junto con datos de uso tales como qué tipo de hablantes utilizan tal vocablo, bien por época (arcaísmos, voces en desuso), procedencia (variantes geográficas), clase social (vulgarismos), situación comunicativa (coloquialismos, cultismos), disciplina (tecnicismos) y otros criterios. La institución recomienda usos cultos y desaconseja otros vulgares o extranjerizantes, pero no puede delimitar con precisión los significados de las palabras, y mucho menos promulgar normas sobre lo que pueden o deben significar; por la sencilla razón de que es imposible aplicar tales acciones al lenguaje natural.

Esto hace que la lengua, como código natural, no sirva para establecer conceptos fijos e inmutables: las palabras tenderán siempre a ir modificando sus significados en la medida en que vayan interviniendo en su uso y comprensión nuevas personas. Esta flexibilidad, que supone una ventaja innegable para la comunicación humana, representa un obstáculo de gran calado para el ejercicio del Derecho, que trata de establecer normas comunes a todas las personas. Por ejemplo, es necesario que las leyes que castigan la violencia definan con la máxima claridad qué actos deben considerarse violentos y cuáles no. Además el esfuerzo es constante, no termina, puesto que a la indeterminación semántica se añade el hecho de que las definiciones irán perdiendo validez con el paso del tiempo y con la transformación de la sociedad. Lo que hace cincuenta años se consideraba en España violencia sexual contra una mujer tiene muy poco, poquísimo, que ver con nuestra concepción presente. Así, el Derecho está abocado a debatir, consensuar, redefinir y reconsensuar siempre los significados: la terminología jurídica sí será entonces un conjunto de significados cerrados e inamovibles, que confieran seguridad jurídica a la formulación de las normas de convivencia.

En este sentido, el hablante debería haber tomado el término «seudonimización» directamente de las leyes, como un término jurídico fijo, puesto que la norma lo define con ese nombre; pero esto no ha ocurrido. Lo que ha tenido lugar en la comunidad científica es un proceso mental—que es el mismo que gobierna el lenguaje natural— que ha generado una palabra híbrida, «seudoanonimización», que realmente no está en las leyes y que, sin embargo, los especialistas han creído leer en ellas, cuando en realidad se ha formado únicamente en su mente.

2.2 La creación de la terminología

Como antes decíamos, las palabras están a la disposición de los hablantes: su significado no es fijo de por sí, sino que quien lo consolida en cada caso es la conversación, el contexto o la propia cultura. Por supuesto, se trata de una situación muy poco satisfactoria para las disciplinas humanas, que necesitan definiciones precisas y bien delimitadas; por tanto, los expertos de las diversas ramas del saber tratan de limitar las posibilidades de ambigüedad. Nace así la terminología.

Los ámbitos especializados y, en lo que a nosotros respecta, concretamente el Derecho, necesitan utilizar lo que Harari (Fig. 3) denomina «código parcial» (partial script):⁸ un sistema de signos que **no** tenga la ductilidad semántica propia de la lengua natural. Para explicar la noción de código parcial, Harari utiliza el ejemplo de las matemáticas: pueden servir para lo mismo que las palabras en algunos casos (más bien pocos), como en la frase «uno más uno, dos»; pero las unidades lingüísticas, tan útiles para otras comunicaciones, se revelan completamente inefectivas para codificar, por ejemplo, el sistema tributario de un país.



Fig. 3: Código parcial (partial script), esquema de Harari.

⁸ Harari, Y. N.: Sapiens: a brief history of humankind, Harper, New York2015, p. 122.

Pues bien, hemos de considerar el lenguaje del derecho como un código parcial; en algunos casos servirá para lo mismo que el natural, pero en su mayor parte necesita una codificación exacta y bien delimitada de palabras y expresiones, cualidad totalmente ajena a la naturaleza misma del código lingüístico humano.

A esta necesidad atiende la terminología. Los significados o acepciones terminológicas solo comparten con los naturales, quizá, el punto de partida; los profesionales del Derecho han de tratar de acordar la comprensión común de los vocablos para conseguir, siquiera sea imperfectamente, una mínima seguridad jurídica. Con mucha frecuencia, las normas escritas incluyen un apartado de definiciones que cierra el paso, por lo menos relativamente, a una interpretación demasiado abierta de las situaciones reguladas. Por naturaleza, la tipificación jurídica de un concepto lleva aparejada la terminologización de las palabras o expresiones que lo denominan.

La terminología no es otra cosa que el resultado de este intento de controlar el significado de las palabras, cercenando la posibilidad de ampliar usos por vías de metáfora, generalización, metonimia y otras relaciones de sentido. Así, definiremos un «término» o «unidad terminológica» como un conjunto **cerrado** de significados, convenido artificial y conscientemente por los especialistas de una disciplina.

2.3 El fenómeno de reanálisis

En la legislación, cada procedimiento de protección de datos recibe su nombre. Una información se anonimiza cuando se hace imposible identificar a la persona, y se seudonimiza cuando sus datos se sustituyen por seudónimos. Pero, en el caso de la comprensión del término «(p)seudonimización», parece existir cierta ambigüedad. En puridad debería entenderse como la asignación de seudónimos como método de despersonalización de datos; pero también parece evocar en el hablante de castellano un proceso que no llega a ser anonimización, pero se le parece mucho. Para que este fuera el caso debería haberse empleado como forma primitiva el verbo «anonimizar» y no el nombre «seudónimo»:los expertos han corregido inconscientemente esta decisión léxica, asumiendo como raíz anonimizar para nombrar un proceso que no lo es, pero casi; han puesto así remedio a la imprecisión entre el sentido de poner nombre falso (seudonimizar) y el de quitar datos, despersonalizar (casi anonimizar o seudoanonimizar).

El proceso que estos hablantes llevan a cabo de manera inconsciente se denomina «reanálisis», y es muy común en la historia de cualquier lengua: la forma de una palabra nos lleva a pensar que puede dividirse en partes con significado independiente -es decir, las interpretamos como morfemas-, y esto puede llevar por ejemplo a acuñar neologismos como «monokini» o «trikini» a partir del extranjerismo «bikini»⁹, cuya primera sílaba se reanaliza como prefijo. Pues bien, en el caso de la palabra «seudonimizar», determinados hablantes han corregido esta forma y han conformado, de nuevo cuño, la palabra «seudoanonimizar», que como hemos visto respondería a una técnica que no llega a ser anonimizar del todo. Así puede verificarse en la práctica: cualquier búsqueda en Internet arroja una enorme profusión del uso de este término entre los expertos.

3. EL CASO DE LA SEUDOANONIMIZA-CIÓN

3.1 La derivación

Los tres términos que nos ocupan, dos utilizados en las leyes (anonimización y seudonimización) y un tercero acuñado por reanálisis del hablante (seudoanonimización) son, morfológicamente, derivaciones: la derivación es uno de los mecanismos que tienen las lenguas naturales para forjar nuevos vocablos a partir de palabras ya existentes. Con la raíz «anónimo» se ha creado «anonimizar» (convertir en anónimo) y de ahí «anonimización» (acción y efecto convertir en anónimo); con la raíz «seudónimo», lo mismo.

 $an\'{o}nimo\ (sust.) \rightarrow anonimizar\ (verbo) \rightarrow anonimizaci\'{o}n\ (sust.)$

(p)seudónimo (sust.) \rightarrow (p)seudonimizar (verbo) \rightarrow (p)seudonimización (sust.)

Esta derivación no ofrece problemas con respecto a los procedimientos habituales de generación de nuevas palabras: el doble derivado «anonimización», que no se encuentra actualmente recogido en el DRAE ni tampoco en el DEJ, es una palabra perfectamente construida y admisible en el castellano genérico con arreglo al procedimiento morfológico descrito. Lo mismo vale para el término «pseudonimización», introducido por el Derecho de la UE e incorporado posteriormente al ordenamiento jurídico español.

⁹ Originariamente, «Bikini» es el nombre de un atolón de las islas Marshall, en el océano Pacífico.

3.2 Hibridación de ambas raíces léxicas

Como estamos viendo, la legislación hace uso de dos raíces distintas, pero el instinto lingüístico del hablante experto ha mezclado ambas creando el vocablo «seudoanonimizar» a partir de la raíz «anónimo». Se mezclan pues usos naturales con usos terminológicos del lenguaje, con la consiguiente confusión de nomenclaturas.

Las propias instituciones europeas han publicado dos textos en el que se simultanea la «seudonimización» con la «seudoanonimización» como traducciones indistintas del inglés pseudonymization, mezclando las dos denominaciones: esto ocurre en el Dictamen C192 de 2009 (2009/C 192/02), Dictamen del Supervisor Europeo de Protección de Datos relativo a la propuesta de Directiva del Parlamento Europeo v del Consejo sobre normas de calidad y seguridad de los órganos humanos destinados a trasplantes (DOUE, 2009b). Las personas responsables de la traducción vuelven a acuñar la palabra «seudoanonimizar» en una Comunicación de la Comisión (DOUE, 2016d, p. 15). Pero, como es de esperar, no existe una definición jurídica propia del concepto de «seudoanonimizar» en ninguna norma, ni española ni europea; por consiguiente, tampoco aparece definido en el DEJ.

Además, el uso de la palabra que sí definen las leyes, «seudonimizar», plantea un problema añadido. Se trata del procedimiento que sustituye por seudónimos algunos de los datos personales, impidiendo así la identificación del afectado; pero, si tenemos en cuenta que uno de tales datos puede a su vez ser un seudónimo, esto plantea la dificil disyuntiva de tener que distinguir, en última instancia, entre un seudónimo (un alias que ya tenía el interesado) y un seudónimo (un nuevo nombre falso que ha sustituido un dato anterior), con la consiguiente posibilidad de que las personas responsables de la protección de datos se vean en el caso de sustituir un seudónimo (propio) por otro seudónimo (acuñado para la ocasión). En nuestra opinión, se trataría de una situación más propia de los trabalenguas que del Derecho; no obstante, serán los letrados quienes quedarán encargados de resolver las posibles ambigüedades o imprecisiones que se deriven de estas decisiones léxicas del legislador.

3.3 La anonimización de datos

Como queda indicado, la legislación española utiliza el concepto de anonimización desde 1993, en una norma que gobierna la protección de datos en el Centro de Investigaciones Sociológicas (BOE, 1993); en 2007 se formula una definición jurídica plena en la Ley de Investigación Biomédica (BOE, 2007):

«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.

Por su parte, la Unión Europea se limita, en un principio, a hablar de datos «hechos anónimos» o «convertidos en anónimos», únicamente para excluir este tipo de datos del ámbito de aplicación de la primera Directiva de protección de datos, publicada en 1995 (DOUE, 1995);¹⁰ el mismo uso se observa en la propuesta de Reglamento sobre protección de datos (DOUE, 1999), que ya sí habla explícitamente del procedimiento de anonimización.

Uno de los actos legislativos principales en este ámbito es el Reglamento 45/2001 sobre protección de datos (DOUE, 2001); aún no se emplea aquí expresamente el término «anonimización» (tampoco anonymization en la versión inglesa), pero sí se da una formulación que parecen recoger los posteriores actos jurídicos de la UE (el resaltado es nuestro):

- » [...] Los principios de la protección no deben aplicarse a los datos convertidos en anónimos de forma que la persona a quien se refieren ya no resulte identificable.
- » [...] La institución o el organismo comunitario establecerá para los datos personales que deban ser archivados por un período más largo del mencionado para fines históricos, estadísticos o científicos, que dichos datos se archivenbien únicamente **en forma anónima**, o, cuando ello no sea posible, sólo con la identidad codificada del interesado.
- » [...] Estos datos se deberán **suprimir o convertir en anónimos** en los plazos más breves [...]

Pese a un uso bastante frecuente del concepto, la «anonimización» propiamente dicha no recibe definición explícita hasta el año 2007, en una propuesta de reglamento que ya no está vigente (DOUE, 2007b, p. 13), donde se expone el sentido de las palabras «microdatos anonimizados»: «los ficheros estadísticos individuales que hayan sido modificados al objeto de minimizar, con arreglo a las mejores prácticas actuales, el riesgo de identificación de las unidades estadísticas a las que hacen referencia».

¹⁰ Es exactamente lo mismo que hace el Reglamento 679 en 2016, que deroga y sustituye la Directiva.

La Comisión entra en detalles de procedimiento en 2008, en un Informe sobre regulación de estadísticas (DOUE, 2008):

La anonimización de ficheros individuales se obtiene mediante la eliminación o recodificación de las variables que podrían utilizarse indirectamente para identificar individuos u hogares.

En el año siguiente la autoridad europea de protección de datos emite un Dictamen (DOUE, 2009b, p. 9) que analiza y compara pormenorizadamente los conceptos de anonimización y seudonimización. En este texto se expone que el «término "anonimización", según el contexto al que se aplique, se emplea a veces para aludir a datos indirectamente identificables, como en el caso de las estadísticas. Este empleo, sin embargo, no es correcto en el ámbito de la protección de datos». Vemos que el propio SEPD está ya señalando un peligro de falsas interpretaciones en este campo; vuelve sobre ello en otro Dictamen emitido el año siguiente (DOUE, 2010, p. 2), dentro de una nota al pie: «La anonimización debe entenderse como la imposibilidad de identificar a la persona. Para algunos tipos de información, como las grabaciones de voz, la anonimización completa no será posible, lo que justifica la necesidad de adoptar garantías más estrictas para evitar usos indebidos».

Por tanto, la situación que se describe hace comprensible la nueva acuñación: la anonimización hace que no se pueda identificar al dueño de los datos, pero no es demasiado realista contar con ello, y será necesario rebajar el nivel de exigencia en la protección de los datos. Así, el concepto de anonimización se somete a matizaciones de tipo instrumental y aspectual, introduciendo en su definición cualificaciones tales como «esfuerzo razonable» o «de forma directa o indirecta»: en otro dictamen -esta vez en materia de estadísticas sobre delincuencia (DOUE, 2011)- el SEPD equipara datos anónimos con anonimizados, y maneja indistintamente las expresiones «anonimizar», «convertir en anónimos» y «hacer anónimos» los datos. Ya en 2014 hallamos la primera definición completa de la «anonimización» (DOUE, 2014, p. 25):

«anonimización»: la eliminación, en las notificaciones de sucesos, de todos los datos personales referentes al notificante y a las personas mencionadas en relación con el suceso, y de todos aquellos datos, como el nombre de la organización o las organizaciones implicadas en el suceso, que permitan identificar al notificante o a terceros o que den lugar a que esa identidad se deduzca de dicha información;

En el año 2016, en una propuesta de Reglamento, la Comisión vuelve a considerar conveniente aclarar explícitamente el sentido del término «anonimizar», que se utiliza en los actos normativos de manera habitual, sin mayores explicaciones (DOUE, 2016b) (el resaltado es nuestro):

Los datos personales reales adoptados a efectos de ensayo se anonimizarán de forma que el sujeto de los datos deje de ser identificable.

De este mismo año data el Reglamento General de Protección de Datos o RGPD (DOUE, 2016c), norma que viene a sistematizar todas las cuestiones relacionadas con este campo y entra en vigor en mayo de 2018. Con respecto al concepto de «anonimizar»—término que no utiliza en absoluto—, el Reglamento se limita a excluir de su ámbito de aplicación, en el Considerando 26, 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». Así pues, las normas europeas siguen hablando hasta el presente de «anonimización» como un procedimiento de salvaguardia adicional de los datos personales.

3.4 La seudonimización

Un seudónimo es un nombre falso, luego -por la misma regla de tres- seudonimizar debería utilizarse con el significado de poner a algo un nombre falso, y no de asignar códigos como en efecto ha sido el caso en la formulación del legislador. Como ya veíamos, la legislación española sí emplea el término «seudónimo» en numerosos actos normativos, como mínimo desde el año 1966. Se utiliza en contextos de autoría -por ejemplo en leyes de prensa e imprenta, instituto bibliográfico, leyes del libro, de premios nacionales (a veces con los sinónimos «lema» y «sobrenombre»)- y denominación de objetos o entidades -cooperación aduanera, propiedad intelectual (en este caso se habla de «obras seudónimas o anónimas»), Registro Mercantil («denominaciones subjetivas») – y otros ámbitos jurídicos.

La legislación también ha utilizado tradicionalmente esta palabra en el propio campo de la protección de datos personales: así ocurre en la regulación de la firma electrónica tanto en el Real Decreto 14 de 1999 (BOE, 1999b) como en la Directiva 1999/93/CE de 2000

(DOUE, 2000b)11; en el Código Civil de Cataluña (BOE, 2008)¹², en la regulación de la gestión pública de ficheros con datos de carácter personal (BOE, 2012¹³ y BOE, 2014¹⁴) en el Ministerio de Sanidad y Consumo. Señalamos especialmente estos casos porque vienen a equiparar «nombre» y «seudónimo», empleando naturalmente esta última voz en el sentido de nombre falso, que muchas veces se incluye en la lista de los propios datos objeto de intercambio, junto con los apellidos, nombres, ocupación, domicilio, nacionalidad, etc. En coherencia con estos usos, la tipificación jurídica del derivado «seudonimizar» debería tener en cuenta el significado académico de la voz «seudónimo», anteriormente recogida. Vemos que se trata de legislación tanto española como europea. Esto, como ya observábamos antes, pone de manifiesto cierto grado de incoherencia neológica en la acuñación de la voz «seudonimizar» que tratamos con el sentido de despersonalizar los datos.

El Diccionario del Español Jurídico de la RAE sí recoge la palabra «seudonimización», aunque no «seudonimizar» ni, como veíamos, «anonimización» (Fig.4).

En los textos de la UE, la seudonimización es introducida por el Dictamen (DOUE, 2007a) de un grupo consultivo, que especifica que «la utilización de seudónimos («seudonimización») se utiliza para ocultar identidades» (p. 19), y pone el cifrado de datos como ejemplo de seudonimización (p. 20). Por su parte, el ya citado Dictamen del SEPD sobre trasplantes (DOUE, 2009b) emplea asimismo el término «seudonimización», pero sin especificar su sentido: este texto presenta el problema de hablar indistintamente de «seudonimización» y «seudoanonimización» para denominar el mismo procedimiento, junto con la expresión «datos seudoanonimizados» (cf. más adelante). En ese mismo año, el SEPD emite otro dictamen (DOUE, 2009a, p. 23) en el que, en la nota 23 a pie de página, se especifica la naturaleza de

Diccionario del español jurídico

dej.rae.es

seudonimización

Gral. 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.

- 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.
 - tratamiento de datos personales.

Fig.4 Diccionario del Español Jurídico de la RAE «seudonimización»

este procedimiento: «La seudonimización es un proceso que puede utilizarse para ocultar la identidad de la persona a que se refieren los datos, manteniendo al mismo tiempo la rastreabilidad de los datos».

Estas indicaciones, que no llegan a constituir una tipificación jurídica propia, sirven de base para posteriores actos: en 2015 se vuelve sobre ello en un nuevo dictamen de la SEPD, que dice (DOUE, 2015):

Las personas deben ejercer de un modo más eficaz sus derechos en lo tocante a cualquier información que pueda servir para identificarlas o señalarlas, incluso si la información se considera «seudonimizada» (14). (p. 3)

A menos que exista una definición clara y vinculante de «datos seudonimizados» distinta de la de «datos personales», este tipo de datos debe mantenerse dentro del ámbito de protección de las normas de protección de datos. (Nota 14, p. 7)

Como vemos, el supervisor reclama (pero no ofrece) una definición «clara y vinculante» que

¹¹ Artículo 8: «Sin perjuicio de los efectos jurídicos concedidos a los seudónimos con arreglo al Derecho nacional, los Estados miembros no impedirán al proveedor de servicios de certificación que consigne en el certificado un seudónimo del firmante en lugar de su verdadero nombre»; en parecidos términos se expresa el Considerando 25.

¹² Artículo 311-4: «No puede incluirse en la denominación de una persona jurídica el nombre o el pseudónimo de una persona física sin su consentimiento expreso o sin la autorización de las personas legitimadas, después de su muerte, para ejercer acciones de protección de su honor, intimidad o imagen».

¹³ Anexo II, punto 1.c.1.

¹⁴ En este caso se enumera el «seudónimo» junto con sus sinónimos «nick» y «alias».

permita identificar el concepto de «datos seudonimizados». Señal de que existe un problema terminológico también con respecto a esta voz. El RGPD (DOUE, 2016c), fija ya una definición formal:

«seudonimización»: 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;

En 2016 el SEPD, hablando de la protección de datos en el sistema ECRIS¹⁵, establece la siguiente recomendación (DOUE, 2016a, p. 8):

Las referencias a datos anónimos deberían eliminarse de la propuesta y sustituirse por referencias precisas al proceso de seudonimización.

Observamos aquí la preferencia por la seudonimización sobre la anonimización, debido a que «la información tratada a efectos del ECRIS sobre los NTP son datos personales que se han sometido a un proceso de seudonimización, y que no constituyen datos anónimos» (p. 7); por tanto, se pide el esclarecimiento terminológico de las normas, aunque solamente en este caso. La definición de «seudonimización» del RGPD ha quedado fijada jurídicamente en los términos expuestos.

Por lo tanto, no es de extrañar que el usuario de la norma no tenga una idea clara de qué significa anonimizar y seudonimizar y que, por lo tanto, recurra a su instinto lingüístico para crear un nuevo término que, sin estar en la legislación, le parece bastante más claro que los dos anteriores y, por eso mismo, más propio de la terminología jurídica.

4. RESUMEN Y CONCLUSIONES

En la bibliografía española de protección de datos en el ámbito biosanitario, puede observarse que el término «seudoanonimización» está totalmente generalizado: los expertos lo consideran un procedimiento de protección de datos tipificado en las leyes, ya sea las europeas (RGPD) o en las españolas. La simple búsqueda en Internet da testimonio de este hecho: en un ejercicio muy sencillo, sin ánimo descriptivo ni ningún rigor científico, de 4680 documentos que trataban este tema encontramos la palabra seudoanonimización en 404, esto es, aproximadamente un 10%. No parece necesario establecer una medición exacta y rigurosa de la prevalencia de este uso en la literatura para colegir que, en efecto, estamos hablando de un efecto extendidísimo, que sería conveniente corregir.

Encontramos incluso, en numerosas ocasiones, que en textos especializados sobre la aplicación de las nuevas técnicas de protección de datos con arreglo al Reglamento que entra en vigor en 2018, un experto que está comentando las definiciones y tipificaciones del RGPD recoge en una cita el término «seudonimización» para pasar a referirse a este procedimiento, a renglón seguido, como «seudoanonimización», como si efectivamente estuviera citando el mismo nombre que incluyen y definen las normas. La situación es muy poco satisfactoria, puesto que da a entender que existe un procedimiento, la anonimización, que sería lo ideal; pero, como la tecnología avanza inexorablemente, es poco realista pensar que se van a conseguir anonimizar datos de verdad, sin posibilidad de recuperar las claves que permitan su identificación, y que por tanto la norma se contenta con una seudo anonimización, con un procedimiento que podría ser, por llamarlo de alguna manera, un sucedáneo. No es así. Lo que la norma contempla no es un procedimiento que casi anonimiza los datos pero no del todo, sino que está estableciendo como protección el cambio de unos datos identificables a unos datos no identificables, que llama seudónimos; además, como hemos visto, también tiene en cuenta que desde un punto de vista técnico es muy poco realista esperar que ese cambio no pueda revertirse y, con unos medios técnicos muy determinados (que deben ser desproporcionados o «poco razonables»), recuperar los datos originales a partir de los seudónimos.

Pensamos que, a la luz de todo lo expuesto, puede afirmarse que no queda suficientemente claro lo que significa exactamente *seudonimizar* desde el punto de vista jurídico; si no fuera así, el usuario no experimentaría tantos problemas en la lectura y escritura de estos términos, hasta el punto de llegar a acuñar un neologismo por reanálisis para poder hacerse una idea cabal del panorama léxico.

El legislador podría haber tenido más en cuenta la naturaleza del lenguaje natural del ser humano, como fenómeno genéticamente programado en la mente del hombre. En este punto, hay que hacer

¹⁵ ECRIS: Sistema Europeo de Información de Antecedentes Penales.

una reflexión sobre lo que significa definir y tipificar conceptos, y las consecuencias que en efecto tiene bautizar nuevos procedimientos. Podríamos pensar que los usuarios no han manejado los textos con la suficiente atención; pero también cabe esperar una actuación más reflexiva de los responsables de legislar las vidas y haciendas de los ciudadanos.

No obstante, se trata de interrogantes que ya no corresponde resolver al lingüista, sino al experto en Derecho. Dejamos, pues, el debate abierto: doctores tiene la Iglesia.

5. BIBLIOGRAFÍA Y REFERENCIAS

5.1 Fuentes primarias

5.1.1 Boletín Oficial del Estado (BOE)

- Boe (2014). Orden SSI/1843/2014, de 29 de septiembre, por la que se modifica la Orden de 21 de julio de 1994, por la que se regulan los ficheros con datos de carácter personal gestionados por el Ministerio de Sanidad y Consumo, BOE núm 245, de 9 de octubre de 2014, pp. 82522-28, dir. electr. https://www.boe.es/boe/dias/2014/10/09/pdfs/BOE-A-2014-10296.pdf, consultado 31/1/18.
 - (2012). Orden SSI/2301/2012, de 16 de octubre, por la que se modifica la Orden de 21 de julio de 1994, por la que se regulan los ficheros con datos de carácter personal gestionados por el Ministerio de Sanidad y Consumo, BOE núm 259, de 27 de octubre de 2012, pp. 75758-60, dirección electrónica https://www.boe.es/boe/dias/2012/10/27/pdfs/BOE-A-2012-13338.pdf, consultado 31/1/18.
 - (2008). Ley 4/2008, de 24 de abril, del Libro Tercero del Código Civil de Cataluña, relativo a las personas jurídicas, BOE núm 131, de 30 de mayo de 2008, pp. 25149-74, dirección electrónica https://www.boe.es/boe/dias/2008/05/30/pdfs/A25149-25174. pdf>, consultado 31/1/18.
 - (2007). Ley 14/2007, de 3 de julio, de Investigación biomédica, BOE núm 159, de 4 de julio de 2007, pp. 28826-48, dirección electrónica https://www.boe.es/boe/dias/2007/07/04/pdfs/A28826-28848.pdf, consultado 2/2/18.

- (1999a). LEY ORGÁNICA 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal, BOE núm 298, de 14 de diciembre de 1999, pp. 43088-99, dirección electrónica https://www.boe.es/boe/dias/1999/12/14/pdfs/A43088-43099.pdf, consultado 7/2/18.
- (1999b). REAL DECRETO-LEY 14/1999, de 17 de septiembre, sobre firma electrónica, BOE núm 253, de 21 de octubre de 2000, pp. 36290-91, dirección electrónica https://www.boe.es/boe/dias/2000/10/21/pdfs/A36290-36301.pdf, consultado 31/1/18.
- (1993). Orden de 31 de marzo de 1993 por la que se regula el acceso al Banco de Datos del Centro de Investigaciones Sociológicas., BOE núm 78, de 1 de abril de 1993, pp. 9600-1, https://www.boe.es/boe/dias/1993/04/01/pdfs/A09600-09601.pdf, consultado 2/2/18.

5.1.2 Diario Oficial de la Unión Europea (DOUE)

- DOUE (2016a). Resumen ejecutivo del dictamen del Supervisor Europeo de Protección de Datos sobre el intercambio de información relativa a nacionales de terceros países en lo que respecta al Sistema Europeo de Información de Antecedentes Penales (ECRIS), DOUE C 186 de 25/5/2016, pp. 7–9, http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52016 XX0525(02)&from=ES>, consulta 2/2/18.
 - (2016b). Propuesta de REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSE-JO relativo a la creación del sistema «Eurodac» para la comparación de las impresiones dactilares para la aplicación efectiva del [Reglamento (UE) n.º 604/2013, por el que se establecen los criterios y mecanismos de determinación del Estado miembro responsable del examen de una solicitud de protección internacional presentada en uno de los Estados miembros por un nacional de un tercer país o un apátrida] y de la identificación de un nacional de un tercer país o un apátrida en situación ilegal, y a las solicitudes de comparación con los datos de Eurodac presentadas por los servicios de seguridad de los Estados miembros y Europol a efectos de aplicación de la ley (refundición), Documento COM(2016) 272 final, 4/5/2016, http://eur-lex.europa.eu/ legal-content/ES/TXT/PDF/?uri=CELEX:520 16PC0272&from=ES>, consulta 2/2/18.

- (2016c). 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 y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos), DOUE L 119 de 4/5/2016, pp. 1-88 http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?urieCELEX:32016R0679&qid=1517921366521 &from=ES>, consultado 6/2/18.
- (2016d). Comunicación de la Comisión al Parlamento Europeo, al Consejo, al Comité Económico y Social europeo y al Comité de las Regiones: Digitalización de la industria europea: Aprovechar todas las ventajas de un mercado único digital, Documento COM(2016) 180 final de 19.4.2016 http://ec.europa.eu/transparency/regdoc/rep/1/2016/ES/1-2016-180-ES-F1-1.PDF>, consultado 16/4/18.
- (2015). Recomendaciones del SEPD sobre las opciones de la UE en cuanto a la reforma de la protección de datos, DOUE C 301 de 12/9/2015, pp. 1-8, http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52015XX0912(01)&from=ES, consultado 6/2/18.
- (2014). REGLAMENTO (UE) No 376/2014 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 3 de abril de 2014 relativo a la notificación de sucesos en la aviación civil, que modifica el Reglamento (UE) no 996/2010 del Parlamento Europeo y del Consejo, y por el que se derogan la Directiva 2003/42/CE del Parlamento Europeo y del Consejo y los Reglamentos (CE) no 1321/2007 y (CE) no 1330/2007 de la Comisión, DOUE L 122 de 24/4/2014, pp. 18-43, http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32014R0376&from=ES, consultado 6/2/18.
- (2011). Dictamen del Supervisor Europeo de Protección de Datos sobre la Propuesta de Reglamento del Parlamento Europeo y del Consejo relativo a las estadísticas europeas sobre seguridad frente a la delincuencia, DOUE C 343 de 23/11/2011, pp. 1-9, http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52011XX1123(01)&from=ES>, consultado 5/2/18.

- (2010). Dictamen del Supervisor Europeo de Protección de Datos sobre la propuesta de Reglamento del Parlamento Europeo y del Consejo sobre la investigación y prevención de accidentes e incidentes en la aviación civil, DOUE C 132 de 21/5/2010, pp. 1-5, http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52010XX0521(01) & from=ES>, consultado 31/1/18.
- (2009a). Dictamen del Supervisor Europeo de Protección de Datos sobre la propuesta de Reglamento del Parlamento Europeo y del Consejo que modifica, en lo que respecta a la farmacovigilancia de los medicamentos de uso humano, el Reglamento (CE) nº 726/2004, por el que se establecen procedimientos comunitarios para la autorización y el control de los medicamentos de uso humano v veterinario v por el que se crea la Agencia Europea de Medicamentos, y sobre la propuesta de Directiva del Parlamento Europeo v del Consejo que modifica, en lo que respecta a la farmacovigilancia, la Directiva 2001/83/CE, por la que se establece un código comunitario sobre medicamentos para uso humano, DOUE C 229 de 23/9/2009, pp. 19-http://eur-lex.europa.eu/legal-content/ ES/TXT/PDF/?uri=CELEX:52009XX0923(04)&from=ES>, consultado 6/2/18.
- (2009b). Dictamen del Supervisor Europeo de Protección de Datos relativo a la propuesta de Directiva del Parlamento Europeo y del Consejo sobre normas de calidad y seguridad de los órganos humanos destinados a trasplantes, DOUE C 192 de 15/8/2009, pp. 6-13, https://edps.europa.eu/sites/edp/files/publication/09-03-05_organ_transplantation_es.pdf>, consultado 31/1/18.
- (2008). Informe de la Comisión al Consejo y al Parlamento Europeo sobre la aplicación del Reglamento (CE) nº 1177/2003 del Parlamento Europeo y del Consejo, de 16 de junio de 2003, COM (2008) 160 final, 31/3/2008, http://eur-lex.europa.eu/legalcontent/ES/TXT/PDF/?uri=CELEX:52008D C0160&from=ES>, consultado 5/2/18.
- (2007a). Dictamen 4/2007 sobre el concepto de datos personales, Grupo de trabajo del artículo 29, 01248/07/ES WP 136, 20/6/2007, http://www.redipd.es/actividades/encuentros/VI/common/wp136_es.pdf>, consulta 6/2/18.

- (2007b). Propuesta de reglamento del Parlamento Europeo y del Consejo relativo a los censos de población y vivienda, COM (2007) 69 final, 23/2/2007, http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:52007PC0069&from=ES, consulta 5/2/18.
- (2001). Reglamento (CE) n° 45/2001 del Parlamento Europeo y del Consejo, de 18 de diciembre de 2000, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales por las instituciones y los organismos comunitarios y a la libre circulación de estos datos, DOUE L 008 de 12/1/2001, pp. 1-22, dirección electrónica http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32001 R0045&qid=1517919914918&from=ES>, consultado 6/2/18.
- (2000a). 2000/520/CE:Decisión de la Comisión, de 26 de julio de 2000, con arreglo a la Directiva 95/46/CE del Parlamento Europeo y del Consejo, sobre la adecuación de la protección conferida por los principios de puerto seguro para la protección de la vida privada y las correspondientes preguntas más frecuentes, publicadas por el Departamento de Comercio de Estados Unidos de América, DOUE L 215 de 25/8/2000, pp. 7-47, dirección electrónica http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:32000D0520&from=ES, consulta 2/2/18.
- (2000b). Directiva 1999/93/CE del Parlamento Europeo y del Consejo, de 13 de diciembre de 1999, por la que se establece un marco comunitario para la firma electrónica, DOUE L 13 de 19/1/2000, pp. 12-20, dirección electrónica https://www.boe.es/doue/2000/013/L00012-00020.pdf, consultado 31/1/18.
- (1999). Propuesta de Reglamento del Parlamento Europeo y del Consejo sobre la protección de las personas físicas en lo que respecta al tratamiento de datos personales por las instituciones y los organismos de la Comunidad y sobre la libre circulación de estos datos, DOUE C 376 E de 28/12/1999, pp. 24-37, http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:51999P C0337&from=ES>, consultado 2/2/18.

- (1995). Directiva 95/46/CE del Parlamento Europeo y del Consejo, de 24 de octubre de 1995, relativa a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos, DOUE L 281 de 23/11/95, pp. 31-50, http://eur-lex.europa.eu/legal-content/ES/TXT/PDF/?uri=CELEX:31995 L0046&qid=1517922278822&from=ES> consultado 6/2/18.
- RAEa (s/f). *Diccionario de la lengua española*, en línea < http://dle.rae.es/> [consulta: 21/11/2017].
- RAEb (s/f). *Diccionario del español jurídico*, en línea < http://dej.rae.es/> [consulta: 21/11/2017].

5.2 Bibliografía

- DIRVEN, R. y VERSPOOR, M.:Cognitive Exploration of Language and Linguistics, John Benjamins, Amsterdam/Philadelphia 1998.
- GEERAERTS, D. y CUYCKENS, H. (eds.): *The Oxford Handbook of Cognitive Linguistics*, Oxford University Press, Oxford 2007.
- GEERAERTS, D.: *Theories of Lexical Semantics*, Oxford University Press, Oxford 2009.
- HARARI, Y. N.: Sapiens: a brief history of humankind, Harper, New York 2015.
- OGDEN, C. K. y RICHARDS, I. A.:The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism, Magdalene College: University of Cambridge, Cambridge 1923, p-11