N/REF: 222535/2016

Se consulta si resulta posible, conforme a lo establecido en la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal, la utilización de datos geolocalizados procedentes de Twitter, con fines de investigación, en el marco de un proyecto de investigación del programa Horizonte 2020. Según señala el consultante, Twitter ofrece la posibilidad de descargar una fracción de todos los tweets que se envían mediante acceso directo al flujo de mensajes a través de APIS, accediendo a los datos públicos de los usuarios. El propósito del consultante es extraer de tales datos patrones agregados de movilidad y actividad de la población (por ejemplo número de personas que viajan de una ciudad a otra) con el objetivo de dar soporte a la planificación y gestión de infraestructuras de transporte. Consulta si es necesario cumplir algún requisito específico para la utilización de los datos y si, en el caso de que la utilización de los datos entrase en conflicto con la legislación aplicable, sería posible resolver dicho conflicto mediante la anonimización de los datos previo a su almacenamiento mediante el procedimiento que describe en la consulta que consiste en borrar determinados datos antes de proceder al almacenamiento y aplicar una función hash a otra parte de ellos, quedando el hash creado como identificador único del usuario, y guardándose el tweet en la base de datos

Los artículos 1 y 2 de la Ley Orgánica 15/1999, extienden su protección a los derechos de los ciudadanos en lo que se refiere al tratamiento de sus datos de carácter personal, siendo definidos éstos en el artículo 3.a) de la citada Ley como "*cualquier información concerniente a personas físicas identificadas o identificables."*

El artículo 5.1 del Reglamento de desarrollo de la Ley Orgánica 15/1999, precisa que constituye un dato de carácter personal *"Cualquier información numérica, alfabética, gráfica, fotográfica, acústica o de cualquier otro tipo concerniente a personas físicas identificadas o identificables."*

La definición contenida en el artículo 3.a de la Ley Orgánica 15/1999 se complementa con la de persona identificable, a la que se refiere el artículo 5.1.o) de su Reglamento de desarrollo, que dispone que lo será "toda persona cuya identidad pueda determinarse, directa o indirectamente, mediante cualquier información referida a su identidad física, fisiológica, psíquica, económica, cultural o social. Una persona física no se considerará identificable si dicha identificación requiere plazos o actividades desproporcionados".

Ambas definiciones tienen su origen en lo establecido en la Directiva 95/46/CE, de 24 de octubre de 1995 del Parlamento Europeo y del Consejo relativa a la protección de las personas físicas en lo que respecta al tratamiento

de daos personales y a la libre circulación de estos datos, cuyo artículo 3 califica como dato personal *"toda información sobre una persona física identificada o identificable (el "interesado"); se considerará identificable toda persona cuya identidad pueda determinarse, directa o indirectamente, en particular mediante un número de identificación o uno o varios elementos específicos, característicos de su identidad física, fisiológica, psíquica, económica, cultural o social".*

Para determinar que se entiende por persona identificada o identificable, cabe acudir a lo señalado, en el Dictamen 4/2007 del Grupo de Trabajo del artículo 29, órgano consultivo independiente de la Unión Europea sobre protección de los datos y la vida privada, creado en virtud de lo previsto en el citado artículo de la Directiva 95/46/CE. Dicho dictamen analiza los términos contenidos en la definición aportada por la Directiva señalando respecto al concepto de persona identificada o identificable lo siguiente:

> "La Directiva requiere que la información se refiera a una persona física «identificada o identificable». Ello suscita las siguientes consideraciones. De modo general, se puede considerar «identificada» a una persona física cuando, dentro de un grupo de personas, se la «distingue» de todos los demás miembros del grupo. Por consiguiente, la persona física es «identificable» cuando, aunque no se la haya identificado todavía, sea posible hacerlo (que es el significado del sufijo «ble»). Así pues, esta segunda alternativa es, en la práctica, la condición suficiente para considerar que la información entra en el ámbito de aplicación del tercer componente. La identificación se logra normalmente a través de datos concretos que podemos llamar «identificadores» y que tienen una relación privilegiada y muy cercana con una determinada persona. Cabe citar como ejemplos su apariencia exterior, es decir su altura, el color del cabello, la ropa, etc. o una cualidad de la persona que no puede percibirse inmediatamente, como su profesión, el cargo que ocupa, su nombre, etc. La Directiva menciona esos «identificadores» en la definición de «datos personales» del artículo 2 cuando establece que «se declarará identificable toda persona cuya identidad pueda determinarse, directa o indirectamente, en particular mediante un número de identificación o uno o varios elementos específicos, característicos de su identidad física, fisiológica, psíquica, económica, cultural o social».

En lo que respecta a los términos "directa" o "indirectamente" identificable continúa señalando el Grupo de Trabajo del artículo 29 que:

> "Esta idea se aclara aún más en los comentarios a los artículos de la propuesta modificada de la Comisión, en donde se afirma que «una persona puede ser identificada directamente por su nombre y apellidos o indirectamente por un número de teléfono, la matrícula de un coche, un número de seguridad social, un número de pasaporte o por una

combinación de criterios significativos (edad, empleo, domicilio, etc.), que haga posible su identificación al estrecharse el grupo al que pertenece."

Los términos de esta declaración indican claramente que el que determinados identificadores se consideren suficientes para lograr la identificación es algo que depende del contexto de la situación de que se trate. Un apellido muy común no bastará para identificar a una persona - es decir, para aislarla – dentro del conjunto de la población de un país, mientras que es probable que permita la identificación de un alumno dentro de una clase. Incluso una información auxiliar, como, por ejemplo, «el hombre que lleva un traje negro», puede identificar a alguno de los transeúntes que esperan en un semáforo. Así pues, el que se identifique o no a la persona a la que se refiere una información depende de las circunstancias concretas del caso.

Por su parte, cuando hablamos de «indirectamente» identificadas o identificables, nos estamos refiriendo en general al fenómeno de las «combinaciones únicas», sean éstas pequeñas o grandes. En los casos en que, a primera vista, los identificadores disponibles no permiten singularizar a una persona determinada, ésta aún puede ser «identificable», porque esa información combinada con otros datos (tanto si el responsable de su tratamiento tiene conocimiento de ellos como si no) permitirá distinguir a esa persona de otras. Aquí es donde la Directiva se refiere a «uno o varios elementos específicos, característicos de su identidad física, fisiológica, psíquica, económica, cultural o social». Algunas de esas características son tan únicas que permiten identificar a una persona sin esfuerzo (el «actual Presidente del Gobierno de España »), pero una combinación de detalles pertenecientes a distintas categorías (edad, origen regional, etc.) también puede ser lo bastante concluyente en algunas circunstancias, en especial si se tiene acceso a información adicional de determinado tipo. Este fenómeno ha sido estudiado ampliamente por los estadísticos, siempre dispuestos a evitar cualquier quebrantamiento de la confidencialidad."

Además, cabe recordar que el del 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) que será aplicable a partir de 25 de mayo de 2018, conforme a su artículo 99, hace expresa mención a otros tipos de datos personales en la definición que de los mismos efectúa en su artículo 4.1, entendiendo por tales *"toda información sobre una persona física identificada o identificable («el interesado»); se considerará persona física identificable toda persona cuya identidad pueda determinarse, directa o indirectamente, en particular mediante un identificador, como por ejemplo un nombre, un número de identificación, datos de*

localización, un identificador en línea o uno o varios elementos propios de la identidad física, fisiológica, genética, psíquica, económica, cultural o social de dicha persona.

Dicha definición tiene en cuenta la aparición de nuevos datos identificativos de las personas, tal y como expresa el considerando 30 de dicho Reglamento según el cual "*Las personas físicas pueden ser asociadas a identificadores en línea facilitados por sus dispositivos, aplicaciones, herramientas y protocolos, como direcciones de los protocolos de internet, identificadores de sesión en forma de «cookies» u otros identificadores, como etiquetas de identificación por radiofrecuencia. Esto puede dejar huellas que, en particular, al ser combinadas con identificadores únicos y otros datos recibidos por los servidores, pueden ser utilizadas para elaborar perfiles de las personas físicas e identificarlas."*

Por consiguiente, no solamente tendrán la condición de datos de carácter personal los datos relativos a nombre y apellidos, la imagen de usuario que aparezca en su caso en el perfil o los identificadores numéricos, sino también el resto de las informaciones asociadas a la persona, como puede ser precisamente el dato relativo a la geolocalización.

Por otra parte, debe hacerse referencia al concepto de tratamiento de datos personales, toda vez que de lo expuesto en la consulta se desprende la realización de diversos posibles tratamientos. A este respecto, constituye un tratamiento de datos, conforme a lo establecido en el artículo 3.c de la Ley Orgánica 15/1999, las *"operaciones y procedimientos técnicos de carácter automatizado o no, que permitan la recogida, grabación, conservación, elaboración, modificación, bloqueo y cancelación, así como las cesiones de datos que resulten de comunicaciones, consultas, interconexiones y transferencias."* La realización de alguna o varias de estas operaciones respecto de datos personales, esto es, de cualquier información asociada a una persona física identificada o identificable, dará lugar a un tratamiento de datos personales. En consecuencia, en lo que al presente supuesto se refiere, el tratamiento consistente en la recogida de los datos se encontrará sujeto a lo previsto en la Ley Orgánica 15/1999.

En lo que respecta al proceso que el consultante efectúa respecto de los datos ya recogidos, cabe hacer recordar que la Ley Orgánica 15/1999 describe el proceso de disociación en su artículo 3 f) como *"Todo tratamiento de datos personales de modo que la información que se obtenga no pueda asociarse a persona identificada o identificable"*

En el presente supuesto se utiliza una función hash, sirviendo el hash creado como identificador único del usuario. A este respecto cabe recordar que la utilización de seudónimos no constituye un proceso de disociación. El aludido Grupo de Trabajo del Artículo 29 define, en su Dictamen 5/2014 sobre

técnicas de anonimización, la seudonimización como “la sustitución de un atributo (normalmente un atributo único) por otro en un registro” señalando que “por consiguiente, sigue existiendo una alta probabilidad de identificar a la persona física de manera indirecta; en otras palabras, el uso exclusivo de la seudonimización no garantiza un conjunto de datos anónimo. La seudonimización reduce la vinculabilidad de un conjunto de datos con la identidad del interesado; se trata, por tanto, de una medida de seguridad útil, pero no es un método de anonimización.”

El concepto de seudonimización aparece expresamente recogido en el artículo 4.5) del Reglamento europeo que define la «seudonimización» como “*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”*

El considerando 26 de dicho Reglamento recuerda que *“*Los principios de la protección de datos deben aplicarse a toda la información relativa a una persona física identificada o identificable. Los datos personales seudonimizados, que cabría atribuir a una persona física mediante la utilización de información adicional, deben considerarse información sobre una persona física identificable. Para determinar si una persona física es identificable, deben tenerse en cuenta todos los medios, como la singularización, que razonablemente pueda utilizar el responsable del tratamiento o cualquier otra persona para identificar directa o indirectamente a la persona física. Para determinar si existe una probabilidad razonable de que se utilicen medios para identificar a una persona física, deben tenerse en cuenta todos los factores objetivos, como los costes y el tiempo necesarios para la identificación, teniendo en cuenta tanto la tecnología disponible en el momento del tratamiento como los avances tecnológicos.”

Y en el considerando 28, viene a poner de manifiesto la diferencia entre anonimización y seudonimización al señalar que “La *aplicación de la seudonimización a los datos personales puede reducir los riesgos para los interesados afectados y ayudar a los responsables y a los encargados del tratamiento a cumplir sus obligaciones de protección de los datos. Así pues, la introducción explícita de la «seudonimización» en el presente Reglamento no pretende excluir ninguna otra medida relativa a la protección de los datos.”*

Por parte de esta Agencia se ha venido señalando que para que un procedimiento de disociación pueda ser considerado suficiente a los efectos de la Ley Orgánica 15/1999, es necesario que de la aplicación de dicho procedimiento resulte imposible asociar el dato o datos de que se disponga a un sujeto determinado. Para ello será preciso que no exista la posibilidad, incluso remota, de que, mediante la utilización, con carácter previo, coetáneo o

posterior de cualquier medio, la información concerniente a los afectados por el tratamiento de datos, que obre en poder del consultante, pueda revelar su identidad.

De este modo, si la información relativa a las personas se encuentra asociada a un seudónimo, esta operación de seudonimización, por sí misma, no constituye un procedimiento de disociación puesto que a partir de la información de que se tiene conocimiento será posible realizar la operación inversa a la seudonimización, lo que da lugar a que la aplicación de la normativa de protección de datos no quede excluida. A ello deben añadirse las posibilidades de reidentificación de la persona a partir de otros datos que pudieran estar contenidos en el tweet.

Por consiguiente, no cabe considerar que el consultante trata datos anónimos, en primer lugar, porque no lo son en el momento de su recogida, operación que por sí misma constituye un tratamiento de datos lo que ya determina la aplicación de la normativa de protección de datos. Por otra parte, el proceso descrito en la consulta no puede calificarse como un procedimiento de disociación sino de seudonimización, en tanto no convierte los datos en anónimos por lo que igualmente se aplicará la normativa de protección de datos a los tratamientos de datos realizados con los datos que han sido sometidos a dicho proceso. No obstante, debe valorarse positivamente el proceso de seudonimización que se aplique a los datos recogidos, en tanto constituye una medida de seguridad añadida a las que deben implementarse conforme a la Ley Orgánica 15/1999 y responde a las nuevas exigencias en materia de seguridad que impondrá el Reglamento general de protección de datos que, como se ha señalado, será aplicable a partir del 25 de mayo de 2018.

II

Teniendo en cuenta que los datos que va a tratar el consultante se encuentran sujetos a lo previsto en la normativa de protección de datos, debe señalarse, en cuanto a la consideración hecha por el consultante de que los datos en redes sociales tienen el carácter de públicos, que la normativa de protección de datos no efectúa una distinción entre datos públicos y privados, y que dicha normativa no permite, sin más, el uso de datos que los afectados hayan hecho públicos, sino que otorga con carácter general una protección a los datos personales determinando aquéllos supuestos en que dicho tratamiento resulta conforme a la misma.

Así en lo que a datos personales tratados en el marco de las redes sociales se refiere el Grupo de Trabajo del Artículo 29, órgano consultivo independiente de la UE sobre protección de los datos y la vida privada, creado en virtud de lo previsto en el citado artículo de la Directiva 95/46/CE 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, en su Dictamen 5/2009 relativo a las redes sociales en línea señala que cuando una persona se convierte en usuaria de una red social, tiene "*una expectativa legítima de que los datos personales que revelan sean tratados de acuerdo con la legislación europea y nacional relativa a la protección de datos y de la intimidad. "*

En este sentido ya el Memorandum de Roma, adoptado en marzo de 2008 por el Grupo de Trabajo internacional de Berlín sobre protección de datos en las telecomunicaciones pone de manifiesto que los datos personales contenidos en las redes sociales no constituyen datos de libre uso al señalar que "Social Networks sites are not –while the term "social" may suggest otherwise- public utilities."

Por consiguiente, no existiendo en el marco regulador de la protección de datos tal distinción entre datos públicos y privados, el tratamiento de los datos suministrados voluntariamente a una red social que resulten accesibles a todos los miembros de la misma con una finalidad distinta de aquellas para las que el usuario ha prestado su consentimiento informado debe encontrar un fundamento en la Ley Orgánica 15/1999, siendo contraria a la misma en otro caso.

Con carácter general el artículo 6 de dicha Ley dispone respecto de la legitimación para el tratamiento de datos "*1. El tratamiento de los datos de carácter personal requerirá el consentimiento inequívoco del afectado, salvo que la ley disponga otra cosa.*

2. No será preciso el consentimiento cuando los datos de carácter personal se recojan para el ejercicio de las funciones propias de las Administraciones públicas en el ámbito de sus competencias; cuando se refieran a las partes de un contrato o precontrato de una relación negocial, laboral o administrativa y sean necesarios para su mantenimiento o cumplimiento; cuando el tratamiento de los datos tenga por finalidad proteger un interés vital del interesado en los términos del artículo 7, apartado 6, de la presente Ley, o cuando los datos figuren en fuentes accesibles al público y su tratamiento sea necesario para la satisfacción del interés legítimo perseguido por el responsable del fichero o por el del tercero a quien se comuniquen los datos, siempre que no se vulneren los derechos y libertades fundamentales del interesado."

A dichas causas legitimadoras del tratamiento de datos debe añadirse la prevista en el artículo 7.f de la 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, cuyo efecto directo ha sido declarado expresamente por la Sentencia del Tribunal de Justicia de la Unión Europea de 24 de noviembre de 2011, según el cual "Los Estados miembros dispondrán que el tratamiento de datos personales sólo pueda efectuarse si (...) es necesario para la satisfacción del interés legítimo perseguido por el responsable del tratamiento o por el tercero o terceros a los que se comuniquen los datos, siempre que no prevalezca el interés o los derechos y libertades fundamentales del interesado que requieran protección con arreglo al apartado 1 del artículo 1 de la presente Directiva".

No obstante, en el presente supuesto, cabe analizar si resulta de aplicación lo previsto en el artículo 4.2 de la Ley Orgánica 15/1999 que en su último inciso dispone que "*No se considerará incompatible el tratamiento posterior de éstos con fines históricos, estadísticos o científicos."*

Ahora bien, como tiene declarado esta Agencia en sus informes, no todo proyecto que se intitule como histórico, estadístico o científico tiene tal consideración, debiendo analizarse detenidamente y de forma individualizada si efectivamente a la vista del ente que desarrolla el estudio o investigación, el fin del mismo y la proporcionalidad de la intromisión o limitación del derecho a la protección de datos que el estudio conlleve, el tratamiento de datos resultante se encuentra amparado por lo previsto en los preceptos transcritos.

El Reglamento de desarrollo de la Ley Orgánica 15/1999, aprobado por Real Decreto 1720/2007 concreta en el segundo párrafo de su artículo 9.1 el criterio que debe adoptarse para calificar la finalidad de un tratamiento como histórica, estadística o científica estableciendo que "Para la determinación de los fines a los que se refiere el párrafo anterior se estará a la legislación que en cada caso resulte aplicable y, en particular, a lo dispuesto en la Ley 12/1989, de 9 de mayo, Reguladora de la función estadística pública, la Ley 16/1985, de 25 junio, del Patrimonio histórico español y la Ley 13/1986, de 14 de abril de Fomento y coordinación general de la investigación científica y técnica, y sus respectivas disposiciones de desarrollo, así como a la normativa autonómica en estas materias."

En cuanto a la legislación aplicable al presente caso, en que el consultante señala que el tratamiento de datos se llevaría a cabo en el marco de un proyecto del programa Horizonte 2020, vendría constituida por el

Reglamento UE (nº 1291/2013 del Parlamento Europeo y del Consejo de 11 de diciembre de 2013 por el que se establece Horizonte 2020, Programa Marco de Investigación e Innovación (2014-2020 y por el que se deroga la Decisión n o 1982/2006/CE. Por consiguiente, tratándose de un proyecto de I+D financiado por la Comisión Europea a través del aludido Programa Marco de Investigación e Innovación, cabe considerar que el tratamiento de los datos personales obtenidos de la red social con el objetivo señalado por el consultante en el marco de dicho proyecto, puede ser calificado como científico a los efectos de encontrarse legitimado en lo previsto en la Ley Orgánica 15/1999.

En cuanto a los requisitos específicos exigibles para utilizar los datos, debe señalarse que deberá darse cumplimiento a todos los principios, obligaciones y garantías exigidas por la normativa de protección de datos. Debe asimismo tenerse en cuenta que la legitimación para el tratamiento de datos fundada en el carácter científico del mismo, alcanzará únicamente al tratamiento de datos que se efectúe en el marco de dicho proyecto, sin que pueda extenderse a usos posteriores de los datos obtenidos, usos que requerirán igualmente encontrarse legitimados en lo previsto en la Ley Orgánica 15/1999.