

CNS 50/2021

Dictamen en relación con la consulta formulada por un Departamento sobre si facilitar los datos sobre el número de alumnos matriculados en primer curso por cada grado y universidad pública catalana, clasificadas en función del centro de secundaria de procedencia (con indicación del municipio, comarca y nombre del centro) cuando el número de estudiantes sea inferior a tres personas puede vulnerar el secreto estadístico.

Se presenta ante la Autoridad Catalana de Protección de Datos un escrito del Departamento en el que, antes de dar cumplimiento a la resolución 720/2021, de 30 de julio, de la Comisión de Garantía del Derecho de Acceso a la Información Pública (en adelante GAIP) y entregar al reclamante la información, se plantea, a fin de garantizar el respeto a la protección de datos de los estudiantes, si facilitar los datos sobre el número de alumnos matriculados en primer curso por cada grado y universidad pública catalana, clasificadas en función del centro de secundaria de procedencia (con indicación del municipio, comarca y nombre del centro) cuando el número de estudiantes sea inferior a tres personas puede vulnerar el secreto estadístico.

Analizada la consulta, que no se acompaña de ningún documento, vista la normativa vigente aplicable y de acuerdo con el informe de la Asesoría Jurídica, se dictamina lo siguiente:

(...)

II

El Departamento expone en la consulta que la Comisión de Garantía del Derecho de Acceso a la Información Pública le pide que entregue a una persona la siguiente información:

“la relación del número de alumnos matriculados en primer curso de los diferentes grados universitarios de las universidades públicas catalanas, que hayan accedido vía Pruebas de Acceso a la Universidad (PAU) o Ciclos Formativos de Grado Superior (CFGS), en el curso 2020 -2021.

La relación debe incluir:

- 1) Número de alumnos desagregado por cada grado de cada universidad. No debe incluirse el nombre, apellidos u otros datos personales de los alumnos, sino simplemente el número de alumnos por grado de cada universidad.**
- 2) Titularidad del centro de educación secundaria posobligatoria de origen del alumno: si era un centro público, concertado o privado.**

3) El municipio o comarca de este centro de origen y, a ser posible, el nombre del mismo centro.

En caso de que no existan datos disponibles del curso 2020-2021, solicitan los mismos datos del último curso disponible”.

Con motivo de esta petición, el Departamento, consulta “si facilitar estos datos clasificados en función del centro de secundaria de procedencia con indicación del municipio, comarca y nombre del centro cuando el número de estudiantes sea inferior a tres personas puede vulnerar el secreto estadístico, en el sentido de que esta desagregación puede permitir la identificación inmediata de los interesados, o conduzcan por su estructura, contenido o grado de desagregación a la identificación indirecta de éstos.”

El 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 la que se deroga la Directiva 95/46/CE (en adelante RGPD) considera “datos personales”: 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 número, 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;” (artículo 4.1 RGPD).

En este sentido, el considerante 26 del RGPD establece que: “Los principios de la protección de datos se aplican a toda la información relativa a una persona física identificada o identificable (...) Para determinar si una persona es identificable es necesario tener en cuenta todos los medios que razonablemente puede utilizar el responsable del tratamiento o cualquier otra persona para identificar directa o indirectamente a la persona física, como por ejemplo la singularización. Para determinar si existe una probabilidad razonable de que se utilicen medios para identificar a una persona física, es necesario considerar 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. (...)”

Ahora bien, el considerante 26 también especifica que “(...) Por tanto, los principios de protección de datos no se aplican a la información anónima, es decir a la información que no tiene relación con una persona física identificada o identificable, ni en los datos convertidos en anónimos de modo que el interesado no sea identificable o deje de serlo. En consecuencia, este Reglamento no afecta al tratamiento de esta información anónima, incluso con fines estadísticos o de investigación.”

En este sentido, es necesario tener en cuenta el Dictamen 4/2007 “sobre el concepto de datos personales”, del antiguo Grupo de Trabajo 29 (en adelante GT 29) sobre la definición de datos personales, el cual determina:

“A los efectos de la Directiva, los datos anónimos pueden definirse como cualquier información relativa a una persona física que no permita su identificación por el responsable del tratamiento de las datos o por cualquier otra persona, teniendo en cuenta el conjunto de medios que puedan razonablemente ser utilizados por el responsable del tratamiento o por cualquier otra persona, para identificar a dicha persona. «Datos anonimizados» serán, por el

tanto, los datos anónimos que con anterioridad se referían a una persona identificable, cuya identificación ya no es posible. El considerando 26 también hace referencia a este concepto cuando dice que «los principios de protección no se aplicarán a aquellas datos hechos anónimos de modo tal que ya no sea posible identificar al interesado». Una vez más, la decisión sobre si los datos permiten la identificación de una persona y sobre si la información puede considerarse como anónima o no depende de las circunstancias concretas de cada caso, por lo que debe realizarse un análisis caso por caso, en lo que habrá que prestar especial atención a hasta qué punto cabe considerar razonablemente posible que se utilicen los medios para identificar a dicha persona, tal y como se describe considerando 26. Esto es particularmente importante en el caso de la información estadística, puesto que a pesar de que ésta pueda facilitarse en forma de datos agregados, la muestra original puede no ser suficientemente amplia y otros datos pueden permitir la identificación de

Por otra parte, el artículo 25 de la Ley 23/1998, de 30 de diciembre, de estadística de Cataluña, de aplicación a las actividades estadísticas reguladas por esta Ley, establece lo siguiente en relación con el secreto estadístico:

“1. El secreto estadístico ampara todos los datos individualizados de carácter privado personal, familiar, económico o financiero, utilizados para elaborar la estadística obtenidos tanto directamente de la persona informante como de fuentes administrativas, salvo los datos que se han hecho públicos mediante registros públicos o publicaciones no declaradas ilegales o contra las que no haya abierto procedimiento judicial alguno.

2. En virtud de la obligación a que se refiere el apartado 1, los datos individuales de comunicación obligatoria no pueden ser hechos públicos ni comunicados a ninguna persona o entidad, ni siquiera a las administraciones públicas, salvo las instituciones o las entidades que también están vinculadas por la obligación del secreto estadístico y exclusivamente con el fin de ser usadas para operaciones estadísticas.”

Y el artículo 26 de esta Ley dispone que el secreto estadístico se vulnera “por cuya comunicación de datos se pueda deducir razonablemente una información individual”.

En el mismo sentido, el artículo 13.2 de la Ley 12/1989, de 9 de mayo, de la función estadística pública dispone que “se entiende que son datos personales los referentes a personas físicas o jurídicas que, o bien permitan la identificación inmediata de los interesados, o bien conduzcan por su estructura, contenido o grado de desagregación a la identificación indirecta de estos últimos.”

La decisión sobre si los datos permiten la identificación de una persona y sobre si la información puede considerarse como anónima o no, depende de las circunstancias concretas de cada caso (en especial del número de miembros del grupo, del número de casos dentro de cada grupo, y de la información que puede relacionarse y del contexto en el que nos encontramos) y se engloba en el análisis de riesgos que el responsable del tratamiento debe llevar a cabo con anterioridad al tratamiento de los datos. Por tanto, será el Departamento, a la hora de facilitar esta información, como responsable del tratamiento, en el proceso de anonimización de los datos, quien debe evaluar los riesgos y la posibilidad de reidentificación a posteriori de los titulares de los datos teniendo en cuenta las circunstancias concretas de cada caso.

En cualquier caso, en el supuesto que se analiza, en que los conjuntos de datos, cuando se trata de la agrupación por centros, puede ser relativamente bajo en algunos casos, no puede descartarse que pueda acabar conociéndose un determinado atributo de la mayor parte de miembros del grupo. Así, si en un determinado centro sólo un alumno se ha matriculado en un Grado en una universidad pública catalana, una vez conocido este alumno (y esto aparte del propio alumno también pueden hacerlo fácilmente, por ejemplo, personas de su entorno más o menos cercano) puede inferirse fácilmente que el resto de integrantes del grupo no reúnen este atributo. Esta información no permitirá saber en qué Grado se han matriculado el resto de alumnos, pero sí que no se han matriculado en universidades públicas catalanas. Esto podría ser aplicable también a supuestos en los que el número de personas conocidas sea más de una pero sea reducido. Especialmente en un entorno como un grupo/clase o centro educativo, donde puede ser relativamente fácil que sus integrantes conozcan directamente por otras vías la identidad de un determinado número de alumnos inferior a tres que se ha matriculado en algún grado de una universidad pública de Cataluña.

En este sentido, el Dictamen 05/2014 sobre técnicas de anonimización del GT 29 que formula recomendaciones para la gestión de las técnicas de anonimización teniendo en cuenta el riesgo residual de la identificación inherente a las mismas, pone de manifiesto que el concepto de identificación no comporta únicamente la posibilidad de recuperar el nombre o dirección de una persona, sino que incluye también la “identificabilidad potencial por singularización, vinculabilidad o inferencia”, es decir, no se trata sólo de la posibilidad de identificación directa de la persona, sino también in

Por tanto, no se puede descartar que exista un cierto riesgo de que, una vez identificada una persona como alumno de un centro de secundaria, se pueda acabar infiriendo respecto a ella el atributo que no se ha matriculado en una universidad pública catalana.

En cualquier caso, cabe destacar también que se trata de un riesgo bajo no sólo desde el punto de vista de su probabilidad (se produciría sólo en centros con un bajo porcentaje de alumnos matriculados en un Grado de universidad pública catalana), sino también en su gravedad, dado que la información que se puede inferir no sería una información que forme parte de categorías especiales de datos, que sea especialmente sensible, o que pueda dar lugar a discriminación, teniendo en cuenta, especialmente, que la información que podría inferirse no sería ni siquiera que las personas objeto de inferencia no se han inscrito en una universidad, sino sólo que no se han matriculado en un Grado de una universidad pública catalana.

III

De acuerdo con lo expuesto, en determinados supuestos con un número bajo de alumnos de un centro que se hayan matriculado en un Grado de universidades públicas catalanas, no se puede garantizar que entregando la información a nivel de centro se mantenga la confidencialidad sobre este atributo a nivel individual.

Para intentar reducir los riesgos de reidentificación pueden aplicarse diferentes técnicas:

En primer lugar, hay que tener en cuenta que la resolución de la GAIP establece que la información debe entregarse a nivel de municipio y comarca. La inclusión del nombre del centro se indica que sólo debe incluirse “a ser posible”. Es decir, desde el punto de vista de la normativa de protección de datos, si la difusión de información a nivel de centro permite la reidentificación para darse un número de

casos muy bajo, debería entregarse la información a nivel de comarca y municipio. Esto reduciría sensiblemente las posibilidades de reidentificación, que se circunscribirían en gran medida a aquellos municipios que sólo cuenten con un único centro de secundaria.

Más allá de esta técnica de generalización o agregación de los resultados, pueden aplicarse también otras técnicas para reducir los riesgos.

De entrada una posibilidad sería facilitar la información por rangos. Así, en lugar de facilitar el número exacto de alumnos de cada centro, puede facilitarse la información por rangos que sean relevantes. Por ejemplo, se puede ofrecer la información indicando el rango al que pertenece cada centro (centro con entre 1 y 5 alumnos que se ha matriculado en un Grado de una universidad pública catalana, centro con entre 5 y 10 alumnos, entre 10 y 20, etc.).

Otra posibilidad sería ofrecer la información en porcentajes. En este caso, es cierto que si se conoce la información exacta sobre el número de alumnos de cada centro puede acabarse obteniendo el número absoluto de alumnos incluidos en el porcentaje, pero el número exacto de alumnos de cada centro no está fácilmente al alcance de cualquier persona.

Conclusión

La normativa de protección de datos personales no impiden que se entregue la información solicitada si en los supuestos en que existe riesgo de que se puedan atribuir la información solicitada a personas identificables se aplican las técnicas expuestas en este informe, en especial facilitar la información por rangos.

Barcelona, 10 de noviembre de 2021