

Dictamen en relació amb la consulta formulada per un Departament sobre si facilitar les dades sobre el nombre d'alumnes matriculats a primer curs per cada grau i universitat pública catalana, classificades en funció del centre de secundària de procedència (amb indicació del municipi, comarca i nom del centre) quan el nombre d'estudiants sigui inferior a tres persones pot vulnerar el secret estadístic.

Es presenta davant l'Autoritat Catalana de Protecció de Dades un escrit del Departament en el qual, abans de donar compliment a la resolució 720/2021, de 30 de juliol, de la Comissió de Garantia del Dret d'Accés a la Informació Pública (en endavant GAIP) i lliurar al reclamant la informació, es planteja, per tal de garantir el respecte a la protecció de dades dels estudiants, si facilitar les dades sobre el nombre d'alumnes matriculats a primer curs per cada grau i universitat pública catalana, classificades en funció del centre de secundària de procedència (amb indicació del municipi, comarca i nom del centre) quan el nombre d'estudiants sigui inferior a tres persones pot vulnerar el secret estadístic.

Analitzada la consulta, que no s'acompanya de cap document, vista la normativa vigent aplicable i d'acord amb l'informe de l'Assessoria Jurídica, es dictamina el següent:

I

(...)

II

El Departament exposa a la consulta que la Comissió de Garantia del Dret d'Accés a la Informació Pública li demana que lliuri a una persona la informació següent:

“la relació del nombre d'alumnes matriculats a primer curs dels diferents graus universitaris de les universitats públiques catalanes, que hagin accedit via Proves d'Accés a la Universitat (PAU) o Cicles Formatius de Grau Superior (CFGS), en el curs 2020-2021.

La relació ha d'incloure:

- 1) Nombre d'alumnes desagregat per cada grau de cada universitat. No s'ha d'incloure el nom, cognoms o altres dades personals dels alumnes, sinó simplement el nombre d'alumnes per grau de cada universitat.*
- 2) Titularitat del centre d'educació secundària postobligatòria d'origen de l'alumne: si era un centre públic, concertat o privat.*

- 3) *El municipi o comarca d'aquest centre d'origen i, a ser possible, el nom d'aquest centre.*

En cas que no hi hagi dades disponibles del curs 2020-2021, sol·liciten les mateixes dades del darrer curs disponible”.

Amb motiu d'aquesta petició, el Departament, consulta “si facilitar aquestes dades classificades en funció del centre de secundària de procedència amb indicació del municipi, comarca i nom del centre quan el nombre d'estudiants sigui inferior a tres persones pot vulnerar el secret estadístic, en el sentit que aquesta desagregació pot permetre la identificació immediata dels interessats, o condueixin per la seva estructura, contingut o grau de desagregació a la identificació indirecta d'aquests.”

El Reglament (UE) 2016/679 del Parlament Europeu i del Consell de 27 d'abril de 2016, relatiu a la protecció de les persones físiques en lo que respecte al tractament de dades personals i a la lliure circulació d'aquestes dades i per la qual es deroga la Directiva 95/46/CE (en endavant 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 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;*” (article 4.1 RGPD).

En aquest sentit, el considerant 26 de l'RGPD estableix que: *“Els principis de la protecció de dades s'apliquen a tota la informació relativa a una persona física identificada o identificable (...) Per determinar si una persona és identificable cal tenir en compte tots els mitjans que raonablement pot utilitzar el responsable del tractament o qualsevol altra persona per identificar directament o indirectament la persona física, com per exemple la singularització. Per determinar si hi ha una probabilitat raonable que s'utilitzin mitjans per identificar una persona física, cal considerar-ne tots els factors objectius, com els costos i el temps necessaris per a la identificació, tenint en compte tant la tecnologia disponible en el moment del tractament com els avenços tecnològics. (...).”*

Ara bé, el considerant 26 també especifica que *“(...) Per tant, els principis de protecció de dades no s'apliquen a la informació anònima, és a dir a la informació que no té relació amb una persona física identificada o identificable, ni a les dades convertides en anònimes de manera que l'interessat no sigui identificable o deixi de ser-ho. En conseqüència, aquest Reglament no afecta el tractament d'aquesta informació anònima, fins i tot amb finalitats estadístiques o de recerca.”*

En aquest sentit, cal tenir en compte el Dictamen 4/2007 “sobre el concepte de dades personals”, de l'antic Grup de Treball 29 (en endavant GT 29) sobre la definició de dades personals, el qual 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 los 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 lo

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 aquellos datos hechos anónimos de manera 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 el 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 como se describe en el 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 personas físicas.

Per altra banda, l'article 25 de la Llei 23/1998, de 30 de desembre, d'estadística de Catalunya, d'aplicació a les activitats estadístiques regulades per aquesta Llei, estableix el següent en relació amb el secret estadístic:

“1. El secret estadístic empara totes les dades individualitzades de caràcter privat personal, familiar, econòmic o financer, utilitzades per a elaborar l'estadística obtingudes tant directament de la persona informant com de fonts administratives, llevat de les dades que s'han fet públiques mitjançant registres públics o publicacions no declarades il·legals o contra les quals no hi ha obert cap procediment judicial.

2. En virtut de l'obligació a què fa referència l'apartat 1, les dades individuals de comunicació obligatòria no poden ésser fetes públiques ni comunicades a cap persona o entitat, ni tan sols a les administracions públiques, llevat de les institucions o les entitats que també estan vinculades per l'obligació del secret estadístic i exclusivament amb la finalitat d'ésser usades per a operacions estadístiques.”

I l'article 26 d'aquesta Llei disposa que el secret estadístic es vulnera “*per la comunicació de dades de les quals es pugui deduir raonablement una informació individual*”.

En el mateix sentit, l'article 13.2 de la Llei 12/1989, de 9 de maig, de la funció estadística pública disposa que “*s'entén que són dades personals les referents a persones físiques o jurídiques que, o bé permetin la identificació immediata dels interessats, o bé condueixin per la seva estructura, contingut o grau de desagregació a la identificació indirecta d'aquests últims.*”

La decisió sobre si les dades permeten la identificació d'una persona i sobre si la informació pot considerar-se com anònima o no, depèn de les circumstàncies concretes de cada cas (en especial del nombre de membres del grup, del nombre de casos dins de cada grup, i de la informació que es pot relacionar i del context en el qual ens trobem) i s'engloba en l'anàlisi de riscos que el responsable del tractament ha de portar a terme amb anterioritat al tractament de les dades. Per tant, serà el Departament, a l'hora de facilitar aquesta informació, com a responsable del tractament, en el procés d'anonimització de les dades, qui ha d'avaluar els riscos i la possibilitat de reidentificació a posteriori dels titulars de les dades tenint en compte les circumstàncies concretes de cada cas.

En qualsevol cas, en el supòsit que s'analitza, en que els conjunts de dades, quan es tracta de l'agrupació per centres, pot ser relativament baix en alguns casos, no es pot descartar que pugui acabar coneixent-se un determinat atribut de la major part de membres del grup. Així, si en un determinat centre només un alumne s'ha matriculat en un Grau en una universitat pública catalana, un cop conegut aquest alumne (i això a banda del propi alumne també ho poden fer fàcilment, per exemple, persones del seu entorn més o menys proper) es pot inferir fàcilment que la resta d'integrants del grup no reuneixen aquest atribut. Aquesta informació no permetrà saber en quin Grau s'han matriculat la resta d'alumnes, però sí que no s'han matriculat a universitats públiques catalanes. Això podria ser aplicable també a supòsits en què el nombre de persones conegudes sigui més d'una però sigui reduït. Especialment en un entorn com un grup/classe o centre educatiu, on pot ser relativament fàcil que els seus integrants coneguin directament per altres vies la identitat d'un determinat nombre d'alumnes inferior a tres que s'ha matriculat en algun Grau d'una universitat pública de Catalunya.

En aquest sentit, el Dictamen 05/2014 sobre tècniques d'anonimització del GT 29 que formula recomanacions per a la gestió de les tècniques d'anonimització tenint en compte el risc residual de la identificació inherent a les mateixes, posa de manifest que el concepte d'identificació no comporta únicament la possibilitat de recuperar el nom o l'adreça d'una persona, sinó que inclou també la *"identificabilitat potencial per singularització, vinculabilitat o inferència"*, es a dir, no es tracta només de la possibilitat d'identificació directa de la persona, sinó també indirecta.

Per tant, no es pot descartar que existeixi un cert risc que, un cop identificada una persona com a alumne d'un centre de secundària, es pugui acabar inferint respecte d'ella l'atribut que no s'ha matriculat a una universitat pública catalana.

En qualsevol cas, cal fer notar també que es tracta d'un risc baix no només des del punt de vista de la seva probabilitat (es produiria només en centres amb un baix percentatge d'alumnes matriculats en un Grau d'universitat pública catalana), sinó també en la seva gravetat, atès que la informació que es pot inferir no seria una informació que formi part de categories especials de dades, que sigui especialment sensible, o que pugui donar lloc a discriminació, tenint en compte, especialment, que la informació que es podria inferir no seria ni tan sols que les persones objecte d'inferència no s'han inscrit a una universitat, sinó només que no s'han matriculat en un Grau d'una universitat pública catalana.

III

D'acord amb el que s'ha exposat, en determinats supòsits amb un nombre baix de d'alumnes d'un centre que s'hagin matriculat en un Grau d'universitats públiques catalanes, no es pot garantir que lliurant la informació a nivell de centre es mantingui la confidencialitat sobre aquest atribut a nivell individual.

Per intentar reduir els riscos de reidentificació es poden aplicar diferents tècniques:

En primer lloc, cal tenir en compte que la resolució de la GAIP estableix que la informació s'ha de lliurar a nivell de municipi i comarca. La inclusió del nom del centre s'indica que només s'ha d'incloure "a ser possible". És a dir, des del punt de vista de la normativa de protecció de dades, si la difusió d'informació a nivell de centre permet la reidentificació per donar-se un número de

casos molt baix, caldria lliurar la informació a nivell de comarca i municipi. Això reduiria sensiblement les possibilitats de reidentificació, que es circumscriurien en gran mesura a aquells municipis que només comptin amb un únic centre de secundària.

Més enllà d'aquesta tècnica de generalització o agregació dels resultats, es poden aplicar també altres tècniques per reduir els riscos.

D'entrada una possibilitat seria facilitar la informació per rangs. Així, en lloc de facilitar el número exacte d'alumnes de cada centre, es pot facilitar la informació per rangs que siguin rellevants. Per exemple, es pot oferir la informació indicant el rang al qual pertany cada centre (centre amb entre 1 i 5 alumnes que s'ha matriculat en un Grau d'una universitat pública catalana, centre amb entre 5 i 10 alumnes, entre 10 i 20, etc.).

Una altra possibilitat seria oferir la informació en percentatges. En aquest cas, és cert que si es coneix la informació exacta sobre el nombre d'alumnes de cada centre es pot acabar obtenint el nombre absolut d'alumnes inclosos en el percentatge, però el nombre exacte d'alumnes de cada centre no està fàcilment a l'abast de qualsevol persona.

Conclusió

La normativa de protecció de dades personals no impedeixen que es lliuri la informació sol·licitada si en els supòsits en que existeix risc que es puguin atribuir la informació sol·licitada a persones identificables s'apliquen les tècniques exposades en aquest informe, en especial facilitar la informació per rangs.

Barcelona, 10 de novembre de 2021