apdcat
Autoritat Catalana de Protecció de Dades

Dictamen en relación a la consulta formulada por una empresa pública sobre la comunicación de información anonimizada del Padrón municipal para el desarrollo de una herramienta predictiva en materia de servicios sociales

Se presenta ante la Autoridad Catalana de Protección de Datos un escrito de una empresa pública en la que plantea si determinados municipios pueden comunicarle información anonimizada del Padrón municipal con el objetivo de desarrollar una herramienta predictiva en materia de servicios sociales.

Analizada la consulta y vista la normativa vigente aplicable, y de acuerdo con el informe de la Asesoría Jurídica emito el siguiente dictamen.

(...)

II

La entidad expone en su escrito de consulta que se constituye como una sociedad mercantil con capital íntegramente público, adscrita a un organismo autónomo con competencias en el ámbito de los servicios sociales y de salud.

De acuerdo con sus Estatutos, el objetivo es *"la gestión y administración de los servicios traspasados (...) en materia sanitaria y social, así como la gestión y administración de los centros, servicios y establecimientos de protección de la salud y de atención sanitaria, sociosanitaria, de salud mental y de atención social que determine (...), de los programas institucionales en materia de promoción y protección de la salud, prevención de la enfermedad, asistencia social, sanitaria y sociosanitaria y rehabilitación, y de los servicios y prestaciones del sistema sanitario público y del sistema de protección y asistencia social, (...)" (artículo* 2).

En el marco de las funciones públicas atribuidas para el cumplimiento del citado objetivo (artículo 3 Estatutos), manifiesta que busca desarrollar un plan de atención sanitaria y social integrada, implementando, en este sentido, una herramienta predictiva que ayude a dimensionar las situaciones de vulnerabilidad a la población, con el fin de identificar patrones poblacionales de las demandas de servicios sociales de las personas en situación de vulnerabilidad o con riesgo de exclusión social, y así planificar de mejor forma los recursos necesarios a tal efecto.

Añade en su consulta que esta herramienta predictiva se basará en una serie de algoritmos de *"Machine learning"* que, partiendo de datos constatables, permitirán al sistema realizar una predicción-estimación de necesidades de servicios para cada uno de los colectivos de población

previamente definidos. En este punto, recuerda que, entre los datos necesarios para poder desarrollar los algoritmos, consideran indispensables los datos del Padrón municipal de los municipios que quieran colaborar en el proyecto. Puntualiza, al respecto, que estos datos se obtendrían de forma anonimizada.

En todo ello, plantea a esta Autoridad si la comunicación pretendida de la información anonimizada del Padrón municipal puede legitimarse en el ejercicio de las competencias que tiene atribuidas.

III

Antes de examinar esta cuestión, cabe señalar que, en atención a la descripción que la empresa pública hace en su consulta sobre la herramienta predictiva, parece razonable pensar que su desarrollo comportará, más allá del tratamiento de los datos de los padrones municipales, la utilización de datos de otras fuentes de información de que pueda disponer la propia entidad, con motivo del ejercicio de las funciones que tiene atribuidas en los ámbitos de salud y de servicios sociales, o bien provenientes de otras posibles entidades o entes participantes, dado que el objetivo pretendido es obtener patrones poblacionales de las demandas de servicios sociales de las personas en situación de vulnerabilidad o con riesgo de exclusión social. Al menos, parece claro que se podrían recopilar datos demográficos, sociales y económicos, sin descartar otros que puedan ser necesarios en este sentido.

De las manifestaciones de la empresa pública puede intuirse, a los efectos de su interés, la intención de cruzar, combinar o correlacionar esta información de múltiples bases de datos, utilizando técnicas de inteligencia artificial, para generar modelos predictivos que permitan conocer la evolución previsible de las necesidades sociales y de la estructura de vulnerabilidades por territorios.

Debe decirse que la recopilación de toda esta información, de forma prospectiva, es decir, como un análisis realizado con el fin de dimensionar las situaciones de vulnerabilidad social, constituiría una acumulación de información altamente intrusiva por el derecho a la protección de datos de las personas afectadas. Y esto no sólo porque pueda afectar a datos de categorías especiales (en la consulta se hace mención a personas con dependencia, con discapacidad, con problemas de salud mental, etc.) sino también porque otra información necesaria para evaluar cada una de las situaciones de vulnerabilidad o riesgo de exclusión social constituiría también una información sensible considerada aisladamente. Por tanto, con más motivo si se combina toda ella.

Hacer notar que, desde la vertiente de la protección de datos, este tipo de actuaciones puede dar lugar a la elaboración de perfiles sobre las personas a las que hace referencia dicha información y, en función de cuál sea su utilización, tener efectos jurídicos o efectos significativos sobre estas personas.

El artículo 4.4) del Reglamento (UE) 2016/679, del Parlamento y del Consejo Europeo, de 27 de abril de 2016, General de Protección de Datos (en adelante, RGPD), define la elaboración de perfiles como *"cualquiera forma de tratamiento automatizado de datos personales consistente en utilizar estos datos para evaluar determinados aspectos personales de una persona física; en especial, para analizar o predecir aspectos relativos al rendimiento profesional, la situación económica, la salud, las preferencias personales, los intereses, la fiabilidad, el comportamiento,*

la ubicación o los movimientos de esa persona."

Hay que tener en consideración que la normativa de protección de datos reconoce el derecho de la persona afectada a no ser objeto de una decisión automatizada, incluida la elaboración de perfiles, que produzca efectos jurídicos en ella o que le afecte significativamente de forma similar (artículo 22.1 RGPD). La elaboración de perfiles se admite sólo, con cierto carácter excepcional, en los tres supuestos que prevé el artículo 22.2 del RGPD y con los requisitos y garantías que se recogen:

> *"1. Todo interesado tendrá derecho a no ser objeto de una decisión basada únicamente en el tratamiento automatizado, incluida la elaboración de perfiles, que produzca efectos jurídicos en él o le afecte significativamente de modo similar.*
> *2. El apartado 1 no se aplicará si la decisión:*
> *a) es necesaria para la celebración o la ejecución de un contrato entre el interesado y un responsable del tratamiento;*
> *b) está autorizada por el Derecho de la Unión o de los Estados miembros que se aplique al responsable del tratamiento y que establezca asimismo medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado, o*
> *c) se basa en el consentimiento explícito del interesado.*
> *3. En los casos a que se refiere el apartado 2, letras a) yc), el responsable del tratamiento adoptará las medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado, como mínimo el derecho a obtener intervención humana por parte del responsable, a expresar su punto de vista ya impugnar la decisión.*
> *4. Las decisiones a que se refiere el apartado 2 no se basarán en las categorías especiales de datos personales contempladas en el artículo 9, apartat 1, salvo que se aplique el artículo 9, apartat 2, letra a) og), y se hayan tomado medidas adecuadas para salvaguardar los derechos y libertades y los intereses legítimos del interesado."*

En el presente caso, la finalidad de la elaboración de los perfiles, según manifiesta la empresa pública, sería ayudar a la administración pública en la toma de decisiones en materia de políticas sociales y concretamente en la planificación eficiente de los recursos públicos. Por tanto, no parece en principio que la elaboración de estos perfiles o patrones de los colectivos de personas en situación de vulnerabilidad o riesgo de exclusión social debiera llevar a la adopción de decisiones automatizadas sobre individuos concretos, es decir , no parece que debiera comportar efectos jurídicos o consecuencias negativas en las posibles personas afectadas.

No obstante, no puede descartarse que, en función de cuál sea la información finalmente tratada, el modo o las condiciones en que se trate y el resto de circunstancias concurrentes, un tratamiento de este tipo pudiera tener efectos significativos en las personas afectadas . Por ejemplo, la herramienta predictiva podría ser un mecanismo adecuado para avanzar en la personalización de los servicios sociales para cada ciudadano. De ser así, entrarían en juego las previsiones del artículo 22 del RGPD, citado, es decir, que la elaboración de estos perfiles requeriría del consentimiento explícito de las personas afectadas o bien que estuviera autorizada por una norma con rango de ley que se aplique al responsable del tratamiento y que establezca asimismo medidas adecuadas para salvaguardar los derechos e intereses legítimos de las personas interesadas.

A todo esto, si nos atenemos a las manifestaciones de la empresa pública en las que afirma que *"en ningún caso se identificará a las personas, familias o hogares de forma individual"* parece que en el presente

caso nos encontraríamos ante un tratamiento de datos anonimizados, por lo que de ser así no habría que tener en cuenta las previsiones del artículo 22 del RGPD, dado que, como veremos a continuación, la normativa de protección de datos no resultaría de aplicación. Esto sin perjuicio de las consideraciones hechas en el apartado V de este dictamen.

IV

Centrándonos en la consulta que se formula, la empresa pública plantea si la comunicación de información anonimizada del Padrón municipal por parte de los municipios que quieran colaborar en el proyecto podría legitimarse en el ejercicio de sus competencias la entidad atribuidas.

El RGPD establece que todo tratamiento de datos personales, como es el caso de las comunicaciones de datos (artículo 4.b)), debe ser lícito, leal y transparente (artículo 5.1.a)) y, en este sentido, establece un sistema de legitimación del tratamiento de datos que se fundamenta en la necesidad de que concurra alguna de las bases jurídicas establecidas en su artículo 6.1.

Ahora bien, es necesario tener presente que cuando este tratamiento comprende información anónima, esto es información que ha perdido toda vinculación directa o indirecta con la persona física -o que ya no la ha tenido desde su obtención-, por lo que la persona afectada deja de ser identificable sin esfuerzos desproporcionados, ya no es necesario contar con una base jurídica que legitime el tratamiento (como podría ser el caso de la relativa al cumplimiento de una misión en interés público (artículo 6.1.e) RGPD), a que menciona la consulta), dado que en este supuesto no resultan de aplicación los principios y garantías de la protección de datos.

Así se desprende claramente del considerante 26 del RGPD, que dispone lo siguiente:

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

Por tanto, en el presente caso, en el que se plantea una comunicación de información anonimizada de varios padrones municipales en la empresa pública, no habría inconvenientes desde la vertiente de la protección de datos para llevar a cabo esta comunicación, al no resultar de aplicación la legislación de protección de datos.

Hay que tener en cuenta que en la consulta no se indica cómo se llevará a cabo el proceso de anonimización de los datos del Padrón municipal en su origen, es decir, por parte de los ayuntamientos que son responsables ni, por tanto , tampoco qué información en concreto de este registro municipal, ni en qué términos, se facilitará finalmente a la empresa pública para llevar a cabo su proyecto (herramienta predictiva).

Por eso, es importante recordar, en este punto, que cualquier proceso de anonimización, aplicado a datos personales, debe tener por finalidad destruir el vínculo o nexo entre el dato personal y la persona física afectada, a quien se refiere la información. El objetivo es que la persona afectada no resulte identificable por terceros sin esfuerzos desproporcionados.

Mientras este nexo entre el dato y la persona física a que se refiere pueda ser reconstruido de forma relativamente sencilla –en este sentido, hay que tener 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 también, como veremos más adelante, la información con la que se puede cruzar, no se puede considerar que la información ha sido objeto de un procedimiento de anonimización adecuado y seguirá sujeta a los principios y obligaciones derivados de la normativa de protección de datos.

De acuerdo con el artículo 16.2 de la Ley 7/1985, de 2 de abril, reguladora de las bases de régimen local (LRBRL), la inscripción en el Padrón municipal contendrá como obligatorios sólo los siguientes datos:

a) Número y cogidos.
b) Sexo.
c) Domicilio habitual.
d) Nacionalidad.
e) Lugar y fecha de nacimiento.
f) Número de documento nacional de identidad o, tratándose de extranjeros:
– Número de la tarjeta de residencia en vigor, expedida por las autoridades españolas, o en su defecto, número del documento acreditativo de la identidad o del pasaporte en vigor expedido por las autoridades del país de procedencia, tratándose de ciudadanos nacionales de Estados Miembros de la Unión Europea, de otros Estados parte en el Acuerdo sobre el Espacio Económico Europeo o de Estados a los que, en virtud de un convenio internacional se extienda el régimen jurídico previsto para los ciudadanos de los Estados mencionados.
– Número de identificación de extranjero que conste en documento, en vigor, expedido por las autoridades españolas o, en su defecto, por no ser titulares de éstos, el número del pasaporte en vigor expedido por las autoridades del país de procedencia, tratándose de ciudadanos nacionales de Estados no comprendidos en el inciso anterior de este párrafo, salvo que, por virtud de Tratado o Acuerdo Internacional, gozan de un régimen específico de exención de visado en materia de pequeño tráfico fronterizo con el municipio en el que se pretenda el empadronamiento, en cuyo caso, se exigirá el correspondiente visado.

g) Certificado o título escolar o académico que se posea.
h) Cuantas otras datos puedan ser necesarias para la elaboración del Censo Electoral, siempre que se garantice el respeto a los derechos fundamentales reconocidos en la Constitución."

Aparte de estos datos, el artículo 57.2 del Reglamento de Población y Demarcación Territorial de las Entidades Locales, aprobado por Real Decreto 1690/1986, de 11 de julio (RPDTEL), dispone que en el Padrón Municipal se pueden recoger con carácter voluntario los siguientes datos:

"a) Designación de las personas que pueden representar a cada vecino ante la Administración municipal a efectos padronales.
b) Número de teléfono."

En atención al conjunto de información que puede constar en el Padrón municipal, hacer notar que un primer paso necesario para la anonimización de la información sería eliminar los datos que permiten identificar de forma directa y fácilmente a una persona física, como las relativas al nombre y apellidos, número de DNI o documento equivalente. Asimismo, sería necesario eliminar aquellos datos que permitan la identificación indirecta de una persona física, tales como domicilio y número de teléfono.

Más allá de esto, también podría ser necesario, en atención a las características y necesidades del proyecto, generalizar ciertos atributos, por ejemplo, por períodos de tiempo o bien agruparlos por meses o años (caso de las fechas de nacimiento), por intervalos de valores o por áreas de población suficientemente amplias.

En cualquier caso, sobre cómo anonimizar la información del Padrón municipal resulta de especial interés el Dictamen 5/2014 sobre técnicas de anonimización, elaborado por el Grupo de Trabajo del Artículo 29 (GTA29), disponible en la web de la Comisión Europea. En este dictamen, al que nos remitimos, se analiza la eficacia y limitaciones de las diferentes técnicas de anonimización existentes, atendiendo al marco legal sobre protección de datos, y se formulan recomendaciones para la adecuada gestión de estas técnicas por parte de los responsables del tratamiento.

Así pues, siempre que el proceso de anonimización aplicado a la información del Padrón municipal y al resto de la información con la que se combine, garantice que las personas físicas a las que hace referencia esta información no podrán ser identificadas sin esfuerzos desproporcionados, no sería necesario disponer de una base jurídica que legitimara su comunicación a la empresa pública.

V

A todo ello, a raíz del previsible cruce de información obtenida de orígenes diversos para el desarrollo de la herramienta predictiva, no está de más poner de manifiesto la necesidad de evaluar los riesgos de una eventual reidentificación posterior de las personas físicas afectadas.

Como se ha visto, el objetivo del proyecto es desarrollar una herramienta predictiva que ayude a dimensionar las situaciones de vulnerabilidad en la población. La finalidad es identificar patrones poblacionales de las demandas de servicios sociales de las personas en situación de vulnerabilidad o con riesgo de exclusión social por territorios, y así planificar y distribuir mejor los recursos que resulten necesarios.

La empresa pública en su consulta hace mención a la necesidad de disponer de los datos del Padrón municipal, en los términos antes expuestos, para poder desarrollar los algoritmos en los que se basará esta herramienta predictiva, si bien, en atención a la finalidad apuntada, puede

presuponerse que ésta no será la única información necesaria a tal efecto. En este sentido, parece que también se podrían utilizar otros conjuntos de información, como datos demográficas, sociales, económicas, etc.

Tal y como ha recuerdo esta Autoridad en dictámenes anteriores (a modo de ejemplo, CNS 26/2021, CNS 12/2021, CNS 10/2016 o CNS 52/2015, disponibles en la web de la Autoridad), en el entorno del *big data* y con las posibilidades que ofrecen técnicas como la minería de datos y la inteligencia artificial -conceptos a los que se refiere la empresa pública-, el cruce de información obtenida de orígenes diversos, fines y todo si ha sido anonimizada, puede acabar haciendo identificable a una persona física.

Es decir, en función del volumen y la naturaleza de los datos que, en el contexto del desarrollo de la herramienta predictiva, se pongan a disposición de la empresa pública -u otros entes que participen-, y según la forma como se ofrezcan, la posibilidad de que la combinación de estas informaciones obtenidas de diversas fuentes pueda acabar haciendo identificables personas concretas no debe descartarse (a partir de rasgos generales, el número de individuos en la intersección de todos ellos va disminuyendo hasta que se identifiquen personas concretas).

En este mismo sentido se ha pronunciado el GTA29, en su Dictamen 5/2014, citado antes: *"(...) los responsables del tratamiento deben ser conscientes de que un conjunto de datos anonimizado puede entrañar aún riesgos residuales para los interesados . Efectivamente, por una parte, la anonimización y la reidentificación son campos de investigación activos en los que se publican con regularidad nuevos descubrimientos y, por otra, incluso las datos anonimizados, como las estadísticas, pueden usarse para enriquecer los perfiles existentes de personas, la consiguiente creación de nuevos problemas de protección de datos. En suma, la anonimización no debe contemplarse como un procedimiento esporádico, y los responsables del tratamiento de datos deben evaluar regularmente los riesgos existentes."*

Es necesario, por tanto, tener presente que el riesgo de reidentificación es inherente a cualquier técnica de anonimización, por lo que la intimidad y el derecho a la protección de datos de la persona titular de los datos podría verse comprometida, aunque las datos hayan sido anonimizados.

Por este motivo, en estos casos es necesario realizar siempre un análisis inicial y periódico de posibles riesgos de reidentificación y, a la vista del resultado obtenido, articular las medidas necesarias para atenuar la probabilidad de que se materialicen, previendo, incluso , medidas reactivas para atenuar el posible daño que pudiera derivarse hacia una persona física si dicha reidentificación ocurriera. Estas medidas o garantías tendrán que ser superiores en aquellos casos en que se traten categorías especiales de datos u otra información merecedora de una especial reserva o confidencialidad (como parece que ocurriría en el presente caso), dado que el riesgo es mayor en atención al mayor impacto que representaría esta reidentificación, de materializarse, sobre los derechos y libertades de las personas afectadas.

Esta identificación y análisis del riesgo de reidentificación debería entenderse en un caso como el planteado como una actividad enmarcada dentro de la evaluación de impacto en la protección de datos (AIPD) a que se refiere el artículo 35 del RGPD.

El RGPD requiere realizar una evaluación de impacto sobre la privacidad *"cuando sea probable que un tipo de tratamiento, en particular si utiliza nuevas tecnologías, por su naturaleza, alcance, contexto o finas, entrañe un alto riesgo para los derechos y libertades de las personas físicas"*

(artículo 35.1). Y menciona expresamente como un supuesto en el que habrá que realizar una evaluación de impacto, la evaluación sistemática y exhaustiva que permita la elaboración de perfiles (artículo 35.2.a)) o el tratamiento a gran escala de categorías especiales de datos (artículo 35.2.b)).

En relación con esta evaluación de impacto, la Ley orgánica 3/2018, de 5 de diciembre, de protección de datos personales y garantía de los derechos digitales (LOPDGDD), enumera, en su artículo 28.2, algunos supuestos en los que se entiende probable la existencia de un alto riesgo para los derechos y libertades de las personas, entre los que *"cuando se produjese el tratamiento no meramente incidental o accesorio de las categorías especiales de datos a las que se refieren los artículos 9 y 10 del Reglamento (UE) 2016/679 y 9 y 10 de esta ley orgánica (...)" (letra* c); *"cuando el tratamiento implicase una evaluación de aspectos personales de los afectados con el fin de crear o utilizar perfiles personales de los mismos, en particular mediante la análisis o predicción de aspectos referidos al su rendimiento en el trabajo, su situación económica, su salud, sus preferencias o intereses personales, su fiabilidad o comportamiento, su solvencia financiera, su localización o sus movimientos" (letra* d); o *"cuando se lleve a cabo el tratamiento de datos de grupos de afectados en situación de especial vulnerabilidad y, en particular, de menores de edad y personas con discapacidad" (letra* e)); o *"cuando se produzca un tratamiento masivo que implique a un gran número de afectados o conlleve la recogida de una gran cantidad de datos personales" (letra* f)).

Además, para facilitar a los responsables de los tratamientos la identificación de aquellos tratamientos que requieren una AIPD, el RGPD dispone que las autoridades de control deben publicar una lista con los tratamientos que requieran de una AIPD. Esta Autoridad considera que es necesario realizar una AIPD en los tratamientos incluidos en la siguiente lista, disponible en la web de la Autoridad.

En el presente caso, y partiendo de la información de que se dispone, debe tenerse en cuenta que concurrirían las circunstancias a las que se ha mencionado:

- Tratamiento que implicaría el perfilado o valoración de personas físicas;
- Tratamiento que implicaría el uso de categorías especiales de datos y también de información merecedora de especial reserva o confidencialidad;
- Tratamiento que haría referencia a datos de sujetos vulnerables o en riesgo de exclusión social, incluidos, niños, personas dependientes, con discapacidad reconocida y con problemas de salud mental;
- Tratamiento que implicaría un nuevo uso de tecnologías emergentes;
- Tratamiento que podría acarrear un gran número de afectados o la recogida de una gran cantidad de datos.

Aunque, como se ha dicho, la normativa de protección de datos no resulta de aplicación al tratamiento de datos anónimos y por tanto a priori la realización de una PIA no resultaría exigible, dado que se trata de un procedimiento que busca identificar y controlar los riesgos para los derechos y libertades de las personas asociados a un tratamiento de datos y que , como se ha visto, el riesgo de reidentificación es inherente a cualquier técnica de anonimización, el hecho de que en el proyecto planteado por la empresa pública concurran las circunstancias mencionadas pone de manifiesto la conveniencia de la realización de una evaluación de impacto relativa a la protección de datos que, al menos, permita medir, evaluar y gestionar el riesgo de reidentificación.

A estos efectos, puede resultar de interés consultar la Guía sobre la evaluación de impacto relativa a la protección de datos en el RGPD, disponible en la web de la Autoridad.

Conclusión

La comunicación de información anonimizada del Padrón municipal a la empresa pública no requeriría de una base jurídica que la legitimara, al no resultar de aplicación en estos casos la legislación de protección de datos.

En todo caso, es necesario velar por que el proceso de anonimización aplicado a los datos del Padrón municipal garantice que las personas físicas afectadas no podrán ser identificadas por terceros sin esfuerzos desproporcionados, así como evaluar los riesgos de una eventual reidentificación posterior de estas personas y, en su caso, adoptar las medidas adecuadas para mitigarlo.

Barcelona, 16 de junio de 2022