apdcat
Autoritat Catalana de Protecció de Dades

Opinion in relation to the query made by a public company on the communication of anonymized information from the Municipal Register for the development of a predictive tool in the field of social services

A letter from a public company is presented to the Catalan Data Protection Authority in which it asks whether certain municipalities can communicate anonymized information from the municipal register with the aim of developing a predictive tool in the field of social services.

Having analyzed the consultation and given the current applicable regulations, and in accordance with the report of the Legal Advice I issue the following opinion.

I

(...)

II

The entity states in its letter of inquiry that it is constituted as a mercantile society with entirely public capital, attached to an autonomous body with competences in the field of social and health services.

According to its Statutes, the objective is *"the management and administration of transferred services (...) in health and social matters, as well as the management and administration of centers, services and establishments of health protection and health care, socio-health, mental health and social care that determines (...), of the institutional programs in the field of health promotion and protection, disease prevention, social assistance, health and socio-health and rehabilitation, and the services and benefits of the public health system and the social protection and assistance system, (...)" (article* 2).

Within the framework of the public functions attributed to the fulfillment of the aforementioned objective (article 3 Statutes), it states that it seeks to develop an integrated health and social care plan, implementing, in this sense, a predictive tool that helps to size the situations of vulnerability in the population, with the aim of identifying population patterns of demands for social services by people in a situation of vulnerability or at risk of social exclusion, and thus better plan the resources needed for that purpose.

He adds in his query that this predictive tool will be based on a series of *"Machine learning"* algorithms that, based on verifiable data, will allow the system to make a prediction-estimate of service needs for each of the population groups

previously defined. At this point, he says that, among the data necessary to be able to develop the algorithms, they consider the data from the Municipal Register of the municipalities that want to collaborate in the project to be indispensable. It points out, in this regard, that this data would be obtained in an anonymized manner.

In addition to all this, this Authority asks whether the intended communication of anonymized information from the Municipal Register can be legitimate in the exercise of the powers attributed to it.

III

Before examining this issue, it should be noted that, given the description that the public company makes in its consultation on the predictive tool, it seems reasonable to think that its development will entail, beyond the processing of data from the registers municipal authorities, the use of data from other sources of information available to the same entity, for the purpose of carrying out the functions attributed to it in the fields of health and social services, or from other possible sources entities or participating bodies, since the intended objective is to obtain population patterns of the demands for social services of people in a situation of vulnerability or at risk of social exclusion. At the very least, it seems clear that demographic, social and economic data could be collected, without ruling out others that might be needed in this regard.

From the manifestations of the public company, it can be intuited, for the purposes of interest, the intention to cross, combine or correlate this information from multiple databases, using artificial intelligence techniques, in order to generate predictive models that allow know the foreseeable evolution of social needs and the structure of vulnerabilities by territory.

It must be said that the compilation of all this information, prospectively, that is, as an analysis made with the purpose of sizing situations of social vulnerability, would constitute an accumulation of highly intrusive information for the right to the protection of data of the affected persons. And this not only because it may affect data from special categories (in the query mention is made of people with dependency, with disabilities, with mental health problems, etc.) but also because other information necessary to evaluate each of the situations of vulnerability or risk of social exclusion would also constitute sensitive information considered in isolation. Therefore, with more reason if all of it is combined.

To point out that, from the side of data protection, this type of actions can give rise to the elaboration of profiles on the people referred to in said information and, depending on the use made of it, have legal effects or significant effects on these people.

Article 4.4) of Regulation (EU) 2016/679, of the Parliament and of the European Council, of April 27, 2016, General Data Protection (hereafter, RGPD), defines profiling as *"any form of automated processing of personal data consisting of using these data to evaluate certain personal aspects of a natural person; in particular, to analyze or predict aspects related to professional performance, economic situation, health, personal preferences, interests, reliability, behavior,*

the location or movements of that person."

It must be taken into account that the data protection regulations recognize the right of the affected person not to be the subject of an automated decision, including profiling, that produces legal effects on it or that significantly affects it in a similar way (article 22.1 GDPR). The creation of profiles is only allowed, with a certain exceptional character, in the three cases provided for in article 22.2 of the RGPD and with the requirements and guarantees contained therein:

> *"1. All interested parties will have the right not to be subject to a decision based solely on automated processing, including profiling, that produces legal effects on them or significantly affects them in a similar way.*
> *2. Section 1 will not apply if the decision:*
> *a) is necessary for the celebration or execution of a contract between the interested party and a data controller;*
> *b) is authorized by the Law of the Union or of the Member States that applies to the person responsible for the treatment and that also establishes adequate measures to safeguard the rights and freedoms and the legitimate interests of the interested party, or*
> *c) is based on the explicit consent of the interested party.*
> *3. In the cases referred to in section 2, letters a) and c), the person in charge of treatment will adopt the appropriate measures to safeguard the rights and liberties and legitimate interests of the interested party, at least the right to obtain human intervention by the person in charge, to express his point of view now contest the decision.*
> *4. The decisions referred to in section 2 will not be based on the categories especiales de datos personales contemplated in article 9, section 1, except that article 9, section 2, letter a) og) applies, and appropriate measures have been taken to safeguard the rights and freedoms and the legitimate interests of the interested party."*

In the present case, the purpose of the elaboration of the profiles, as stated by the public company, would be to help the public administration in making decisions regarding social policies and specifically in the efficient planning of public resources. Therefore, it does not seem in principle that the elaboration of these profiles or patterns of groups of people in a situation of vulnerability or risk of social exclusion should lead to the adoption of automated decisions about specific individuals, that is , it does not seem that it should entail legal effects or negative consequences for the possible people affected.

However, it cannot be ruled out that, depending on what information is finally processed, the way or conditions in which it is processed and the rest of the concurrent circumstances, a treatment of this type could have significant effects on the people affected . For example, the predictive tool could be a suitable mechanism to advance the personalization of social services for each citizen. If so, the provisions of article 22 of the RGPD would come into play, cited, that is to say, that the preparation of these profiles would require the explicit consent of the persons affected or that it was authorized by a rule with range of law that applies to the person responsible for the treatment and that also establishes appropriate measures to safeguard the rights and legitimate interests of the persons concerned.

In addition to all this, if we look at the statements of the public company in which it states that *"in no case will individuals, families or households be identified individually"* it seems that in the present

case we would be dealing with anonymized data, so if this were the case it would not be necessary to take into account the provisions of article 22 of the RGPD, given that, as we will see below, the data protection regulations do not would be applicable. This notwithstanding the considerations made in section V of this opinion.

IV

Focusing on the query that is formulated, the public company considers whether the communication of anonymized information from the Municipal Register by the municipalities that wish to collaborate in the project could be legitimate in the exercise of the powers it has the attributed entity.

The RGPD establishes that all processing of personal data, as is the case with data communications (article 4.b)), must be lawful, loyal and transparent (article 5.1.a)) and, in this sense, establishes a system of legitimizing the processing of data that is based on the need for one of the legal bases established in its article 6.1 to apply.

However, it must be borne in mind that when this treatment includes anonymous information, this is information that has lost all direct or indirect connection with the natural person -or that has no longer had it since it was obtained-, so that the person affected is no longer identifiable without disproportionate efforts, it is no longer necessary to have a legal basis that legitimizes the treatment (as could be the case of the one relating to the fulfillment of a mission in the public interest (Article 6.1.e) RGPD), to what does the query mention), since in this case the principles and guarantees of data protection do not apply.

This is clear from recital 26 of the RGPD, which provides the following:

> *"The principles of data protection must be applied to all relevant information to an identified or identifiable natural person. Pseudonymized personal data, which could be attributed to a natural person through the use of additional information, must be considered information about an identifiable natural person. To determine whether a natural person is identifiable, all means, such as identification, that can reasonably be used by the data controller or any other person to directly or indirectly identify the natural person must be taken into account. To determine whether there is a reasonable probability that means will be used to identify a natural person, all objective factors must be taken into account, such as the costs and time required for identification, taking into account both the technology available at the time of the treatment as technological advances.* ***Therefore, the principles of data protection should not be applied to anonymous information, that is, information that is not related to an identified or identifiable natural person, nor to data converted into anonymous data in such a way that the interested party is not identifiable, or to be*** *Consequently, this* ***Regulation does not affect the treatment of said anonymous information,*** *including for statistical or research purposes."*

Therefore, in the present case, in which a communication of anonymized information from several municipal registers in the public company, there would be no inconvenience from the data protection side to carry out this communication, as the data protection legislation does not apply.

However, it should be borne in mind that the consultation does not indicate how the process of anonymizing the municipal register data will be carried out at its origin, that is to say, by the municipalities that are responsible or, therefore , nor what specific information from this municipal register, nor under what terms, will finally be provided to the public company to carry out its project (the predictive tool).

For this reason, it is important to remember, at this point, that any anonymization process, applied to personal data, must aim to destroy the link or nexus between the personal data and the affected natural person, to whom the information refers. The aim is that the affected person cannot be identified by third parties without disproportionate efforts.

As long as this link between the data and the physical person it refers to can be reconstructed in a relatively simple way - in this sense, it is necessary to take into account all the objective factors, such as the costs and time required for identification, taking into account both technology available at the time of treatment such as technological advances and also, as we will see later, the information with which it can cross-, it cannot be considered that the information has been subject to an appropriate anonymization procedure and will remain subject to the principles and obligations derived from data protection regulations.

In accordance with article 16.2 of Law 7/1985, of April 2, regulating the bases of the local regime (LRBRL), the registration in the Municipal Register must contain only the following data as mandatory:

"a) Number and surnames.
b) Sex.
c) Usual address.
d) Nationality.
e) Place and date of birth.
f) National ID number or, in the case of foreigners:
– Number of the valid residence card, issued by the Spanish authorities, or failing that, number of the valid identity document or passport issued by the authorities of the country of origin, in the case of citizens
nationals of Member States of the European Union, of other States part of it Agreement on the European Economic Area or of States to which, by virtue of an international agreement, the legal regime provided for the citizens of the aforementioned States is extended.
- Foreigner's identification number contained in a valid document issued by the Spanish authorities or, failing that, because they do not hold these, the valid passport number issued by the authorities of the country of origin, in the case of citizens nationals of States not included in the previous paragraph of this paragraph, unless, by virtue of an International Treaty or Agreement, they enjoy a specific regime of visa exemption in the matter of small border traffic with the municipality in which registration is sought, in which case, the corresponding visa will be required.

g) School or academic certificate or title that is presented.
h) How many other data may be necessary for the preparation of the Electoral Census, provided that respect for the fundamental rights recognized in the Constitution."

Apart from these data, article 57.2 of the Regulation of Population and Territorial Demarcation of Local Entities, approved by Royal Decree 1690/1986, of July 11 (RPDTEL), states that in the Municipal Register they can be collected with the following data are voluntary:

"a) Designation of the persons who can represent each neighbor before the Municipal administration for land use purposes.
b) Telephone number."

With regard to the set of information that can be contained in the Municipal Register, note that a necessary first step for the anonymization of the information would be to eliminate the data that allow a natural person to be directly and easily identified, such as relating to the first and last name, ID number or equivalent document. Also, it would be necessary to remove those data that allow the indirect identification of a natural person, such as address and telephone number.

Beyond this, it could also be necessary, in attention to the characteristics and needs of the project, to generalize certain attributes, for example, by periods of time or group them by months or years (in the case of dates of birth), by intervals of values or by sufficiently large population areas.

In any case, on how to anonymize information from the Municipal Register, Opinion 5/2014 on anonymization techniques is of particular interest, prepared by the Article 29 Working Group (GTA29), available on the European Commission's website. In this opinion, to which we refer, the effectiveness and limitations of the different existing anonymization techniques are analyzed, taking into account the legal framework on data protection, and recommendations are formulated for the appropriate management of these techniques by those responsible for treatment.

So, as long as the anonymization process applied to the information from the Municipal Register and the rest of the information with which it is combined, guarantees that the natural persons to whom this information refers cannot be identified without disproportionate efforts, it would not be necessary to have a legal basis that legitimizes its communication to the public company.

V

In addition to all this, due to the predictable cross-checking of information obtained from various sources for the development of the predictive tool, it is not superfluous to also consider the need to assess the risks of any subsequent re-identification of the affected physical persons.

As we have seen, the aim of the project is to develop a predictive tool that helps to measure the situations of vulnerability in the population. The purpose is to identify population patterns of demands for social services by people in a situation of vulnerability or at risk of social exclusion by territory, and thus plan and distribute the resources that are necessary in a better way.

The public company in its consultation mentions the need to have the data of the Municipal register, in the terms set out above, to be able to develop the algorithms on which this predictive tool will be based, although, in attention to the stated purpose, it can

assume that this will not be the only information necessary for this purpose. In this regard, it seems that other sets of information, such as data, could also be used demographic, social, economic, etc.

As this Authority has agreed in previous opinions (for example, CNS 26/2021, CNS 12/2021, CNS 10/2016 or CNS 52/2015, available on the Authority's website), in the environment of *big data* and with the possibilities offered by techniques such as data mining and artificial intelligence - concepts to which the public company refers - the crossing of information obtained from various sources, until and even if it has been anonymized, it can end up making a natural person identifiable.

That is to say, depending on the volume and nature of the data that, in the context of the development of the predictive tool, are made available to the public company -or other bodies that participate -, and according to the no matter how they are offered, the possibility that the combination of this information obtained from various sources may end up making specific people identifiable should not be ruled out (based on general characteristics, the number of individuals at the intersection of all of them decreasing until specific individuals are identified).

In this same sense, the GTA29 has pronounced, in its Opinion 5/2014, cited above: *"(...) those responsible for the treatment must be aware that a set of anonymized data can still carry residual risks for the interested parties . Effectively, on the one hand, anonymization and re-identification are active research fields in which new discoveries are regularly published and, on the other hand, even anonymized data, such as statistics, can be used to enrich existing profiles of people, with the consequent creation of new data protection problems. In short, anonymization should not be seen as a sporadic procedure, and those responsible for data processing must regularly evaluate existing risks."*

It is therefore necessary to bear in mind that the risk of re-identification is inherent in any anonymization technique, so the privacy and the right to data protection of the person holding the data could be compromised, even though the data have been anonymized.

For this reason, in these cases it is necessary to always carry out an initial and periodic analysis of possible risks of re-identification and, in view of the result obtained, articulate the necessary measures to reduce the probability of them materializing, anticipating, even , reactive measures to mitigate the possible damage that could be caused to a natural person if said re-identification were to take place. These measures or guarantees will have to be higher in those cases in which special categories of data or other information deserving of a special reservation or confidentiality are treated (as it seems would happen in the present case), given that the risk is greater in attention to greater impact that this re-identification would represent, if materialized, on the rights and freedoms of the people affected.

This identification and analysis of the risk of re-identification should be understood in a case like the one proposed as an activity framed within the data protection impact assessment (AIPD) referred to in article 35 of the RGPD.

The RGPD requires an impact assessment on privacy *"when it is likely that a type of treatment, in particular if it uses new technologies, por su naturela, alcance, context or purposes, entails a high risk for the rights and freedoms of natural persons"*

(article 35.1). And it is expressly mentioned as a case in which an assessment will need to be made of impact, the systematic and comprehensive evaluation that allows the elaboration of profiles (article 35.2.a)) or the large-scale processing of special categories of data (article 35.2.b)).

In relation to this impact assessment, Organic Law 3/2018, of December 5, on the protection of personal data and guarantee of digital rights (LOPDGDD), lists, in its article 28.2, some cases in which likely the existence of a high risk for the rights and freedoms of people, among which *"when the treatment occurs not merely incidental or accessory of the special categories of data to which articles 9 and 10 of the Regulation (EU) 2016/679 and 9 and 10 of this organic law (...)" (letter* c); *"when the*

treatment involves an evaluation of personal aspects of those affected in order to create or use personal profiles of them, in particular through the analysis or prediction of aspects related to their performance at work, their economic situation, their health, their preferences or personal interests, his reliability or behavior, his financial solvency, his location or his movements" (letter d); or *"when data processing is carried out for groups of affected persons in a situation of special vulnerability and, in particular, minors and persons with disabilities" (letter* e)); or *"when a massive treatment is produced that involves a large number of affected or entails the collection of a large amount of personal data" (letter* f)).

In addition, to make it easier for data controllers to identify those treatments that require an AIPD, the RGPD provides that the control authorities must publish a list of the treatments that require an AIPD. This Authority considers that it is necessary to carry out an AIPD for the treatments included in the following list, available on the Authority's website.

In the present case, and based on the information available, it must be taken into account that the circumstances mentioned would occur:

- Treatment that would involve the profiling or assessment of natural persons;
- Treatment that would involve the use of special categories of data and also information deserving of special reservation or confidentiality;
- Treatment that would refer to data of vulnerable subjects or at risk of exclusion social, including children, dependents, recognized disabilities and mental health problems;

- Treatment that would involve a new use of emerging technologies;
- Treatment that could involve a large number of affected or the collection of a large amount of data.

Although, as has been said, the data protection regulations do not apply to the treatment of anonymous data and therefore a priori the performance of a PIA would not be required, given that it is a procedure that seeks to identify and control the risks to the rights and freedoms of people associated with data processing and that , as seen, the risk of re-identification is inherent in any anonymization technique, the fact that in the project raised by the public company the aforementioned circumstances concur highlights the convenience of carrying out an impact assessment relating to data protection that, at least, allows to measure, evaluate and manage the risk of re-identification.

For these purposes, it may be of interest to consult the Guide on relative impact assessment to data protection in the RGPD, available on the Authority's website.

conclusion

The communication of anonymized information from the Municipal Register to the public company would not require a legal basis to legitimize it, as data protection legislation does not apply in these cases.

In any case, it is necessary to ensure that the anonymization process applied to the municipal register data guarantees that the natural persons affected cannot be identified by third parties without disproportionate efforts, as well as to assess the risks of possible subsequent re-identification of these people and, where appropriate, adopt the appropriate measures to mitigate it.

Barcelona, June 16, 2022