Eurotransplant's decision regarding anonymization

Introduction

The General Data Protection Regulation (GDPR) makes a distinction between pseudonymised and anonymised data. With pseudonymised data, the data subject is still relatively traceable and such data fall within the scope of the GDPR. When Eurotransplant uses anonymised data, the data subject can no longer be traced, which means that the data are no longer within the scope of the GDPR.
At present, there are two different opinions on when data should be regarded as personal data. In this decision, we explain the difference between pseudonymised and anonymised data and when we at Eurotransplant should regard data as anonymous. There are four different legal frameworks:

- Recital 26 of the GDPR
- Article 4(1) of the GDPR
- Article 29 Working Party Opinion 05/2014 on anonymisation techniques
- Breyer case: ECLU:EU:C:2016:779

Considerations

There is a legal difference between anonymous data and pseudonymised data. Where data can be regarded as anonymous, they are not viewed as personal data and therefore fall outside the scope of the GDPR, unlike pseudonymised data, which do fall within its scope.

At present, there are two different opinions on when data should be regarded as personal data when anonymisation/pseudonymisation is involved:

1. The opinion based on the text of the GDPR and a European Court of Justice judgment: *'To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments (Recital 26 of the GDPR)'*

2. The opinion based on the opinion of the European privacy regulator, the European Data Protection Board, the successor of the Article 29 Working Party. *'Identification of a natural person not only means the possibility of retrieving a person's name and/or address, but also includes potential identifiability by singling out, linkability (of files/datasets1) and inference based on the information available in the dataset. Generally speaking, therefore, removing directly identifying elements in itself is not enough to ensure that identification of the data subject is no longer possible.'*

[1] A dataset is a collection of data, usually presented in tabular form. Each column represents a particular variable and each row corresponds to a given record of the data set in question.

Following the line taken by the European privacy regulators will mean in practice that virtually no datasets can be considered anonymous because there will almost always be a possibility of their being enriched and matched with additional data, publicly available or otherwise. As a result, even the most non-identifying records will be capable of being traced back to individuals, or to some individuals, and will therefore have to be regarded as personal data within the meaning of the GDPR.

This interpretation of the rules has a restrictive effect on research-based and other practice because strict legal requirements must be met when working with personal data. Those requirements create a great deal of work, many contracts and delays, and can sometimes make collaboration with national or international research partners, the conducting of research or obtaining grants for such research projects impossible.

It is worth noting that there are some striking examples of cases where anonymous data, that were apparently impossible to trace back to individuals, turned out to be traceable after all. However, those cases involved a deliberate exercise where datasets were linked and attempts made to trace individuals. In practice, cases like these could be limited by stipulating in contracts with third parties provided with datasets that such exercises are prohibited.

Recommendation

The Data Protection Officer of Eurotransplant recommends following the line taken in the legislative text of the GDPR and the European Court of Justice judgment (opinion 1). For Eurotransplant, this will mean a higher level of risk acceptance when it comes to assessing whether datasets are anonymous. It should also be noted that the legislative text of the GDPR and the European Court of Justice judgment are of a more recent date (2016) than the opinion of the European regulators (2014).

This means that when assessing whether a dataset is anonymous, Eurotransplant will consider whether in a specific case presented to us it can be expected that it will be possible to identify data subjects in the dataset, instead of the theoretical possibility of an individual being identified. Contracts with third parties to which Eurotransplant makes data available should include a prohibition on the linking of public or non-public datasets with the aim of retrieving the identity of individuals.

In taking the proposed line, Eurotransplant will be accepting the risk that in some cases the regulator (Dutch Data Protection Authority) will not accept this approach. The regulator would then be able to take enforcement action and, for instance, impose a fine for multiple breaches of the GDPR. Examples include cases where there may be no basis and/or ground for exception for the provision of data (Articles 6 and 9 of the GDPR) because no conclusive agreements have been made with cooperation partners (Articles 26 and 28 of the GDPR) and/or because no transfer mechanism is used for international flows (Articles 44 to 49 of the GDPR).

[2] Netflix and taxi drivers' cases, examples from the WP29 opinion, examples from Matthijs Koot's doctoral thesis.