GDPR series: personal data — an expanding concept?

Davinia Brennan, Associate at A&L Goodbody, examines changes to the definition of personal data under the GDPR and the impact of a recent European Court ruling

Understanding whether information is 'personal data' is essential in order to determine whether or not data protection law applies, and yet much uncertainty surrounds the concept.

Whilst the Article 29 Working Party Opinion 4/2007 and guidance from the Information Commissioner's Office ('ICO') helps us to understand the scope of the concept, uncertainty prevails, particularly in regard to the status of online identifiers (such as IP addresses and cookies) and pseudonymous data (such as key-coded data). Therefore the recent decision of the Court of Justice of the European Union ('CJEU') in Breyer (Case C-582/14) provides some welcome clarity on the concept.

This article considers the expanding concept of personal data under the existing Data Protection Directive (95/46/EC) and the General Data Protection Regulation ('GDPR') coming into force in May 2018.

What is personal data?

The Directive defines 'personal data' as 'any information relating to an identified or identifiable living individual.' It provides that an identifiable person is 'one who can be identified directly or indirectly, in particular by reference to an identification number, or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity'.

The Data Protection Act 1998 (the 'DPA') repeats the substance of the Directive's definition, defining personal data as data that 'relate to a living individual who can be identified— (a) from those data, or (b) from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller, and includes any expression of opinion about the individual and any indication of the intentions of the data controller or any other person in respect of the individual'.

A person's name may be 'personal data', where it directly identifies that individual. A person may also be indirectly identifiable by their car registration number, social security number or a combination of criteria such as age, occupation, and place of residence.

Whilst the concept appears relatively straightforward, its interpretation has proved difficult in practice. For example, a common name may not be sufficient to identify someone, i.e. single someone out, from the whole of a country's population, but it may be sufficient to identify them in their workplace. Therefore, the extent to which certain identifiers are sufficient to achieve identification is very much dependent on the context of a particular situation.

Opinion 4/2007 helpfully broke the definition down into four component parts: 'any information'; 'relating to'; 'an identified or identifiable', and 'natural person'. It considers that these elements together determine whether a piece of information should be considered 'personal data'.

The first element, 'any information', calls for a wide interpretation, and includes objective and subjective information, such as statements and opinions about a person, whether true or false, and irrespective of the technical medium on which it is contained.

As regards the second element, information can be considered to 'relate' to an individual when it is 'about' that individual. The third element requires that the information relate to a natural person that is 'identified or identifiable'. The Working Party considers that a person is identifiable when, although the person has not been identified yet, it is 'possible' to do so. The question of when a person is 'identifiable' has proved to be a particularly difficult one, and therefore this article focuses on this element.

Finally, the fourth element requires the personal data to be about a 'living individual'.

When is a person 'identifiable'?

The Directive contains a broad test for determining whether an individual is 'identifiable'. Rectal 26 provides that to determine when a person is 'identifiable', account should be taken of 'all the means likely reasonably to be used either by the controller or by any other person to identify the said person'. Therefore to determine whether a person is 'identifiable', you must examine what means and available datasets might be used to identify a data subject.

The Working Party considers that a mere hypothetical possibility to single out an individual is not enough to consider the person as 'identifiable' if that possibility is negligible. It suggests that the criterion, 'all the means likely reasonably to be used', to identify a person, requires a range of factors to be taken into account, such as the cost of conducting identification, the purpose and advantage pursued by the controller in the data processing. The Working Party highlights that 'to argue that individuals are not identifiable where the purpose of the processing is precisely to identify them, would be a sheer contradiction in terms'.

In its guidance on personal data, the ICO similarly states that 'when considering identifiability, it should be assumed that you are not looking just at the means reasonably likely to be used by the ordinary man in the street, but also the means that are likely to be used by a determined person with a particular reason to want to identify individuals. Examples would include investigative journalists, estranged partners, stalkers, or industrial spies'.

> *"Breyer v Bundersrepublik Deutschland (Case C-582/14) has been lauded as clarifying that a dynamic IP address may be 'personal data' in the hands of a person, such a website provider, even though additional information has to be sought from a third party ISP to identify the data subject."*

Are IP addresses identifiable personal data?

The status of IP addresses has caused much controversy.

The Working Party discusses the processing of IP addresses by copyright owners at Example 15 of its Opinion 4/2007. This notes that where such processing is carried out by copyright owners for the purpose of identifying and prosecuting copyright infringers, the copyright owner anticipates that the 'means likely reasonably to be used' to identify the persons will be available, such as through the courts appealed to, 'otherwise the collection of the information makes no sense'. Therefore, the Working Party considers that IP information should be treated as 'personal data'.

The Working Party further looked at the scenario where an IP address does not allow identification of the user, such as where an IP address is attributed to a computer in an internet café, where no identification of the customers is requested. It points out that the ISP will probably not know whether the IP address in question is one allowing identification or not, and so the ISP 'will have to treat all IP information as personal data, to be on the safe side'.

The decision in Breyer

Breyer v Bundersrepublik Deutschland (Case C-582/14) has been lauded as clarifying that a dynamic IP address may be 'personal data' in the hands of a person, such a website provider, even though additional information has to be sought from a third party ISP to identify the data subject. The CJEU held that the key question in determining whether information is 'personal data', is whether there is a legal means, reasonably likely to be used, to identify the person to whom the data belongs. In determining whether those means are likely to be used, the Court will take into consideration the effort involved, in terms of time, cost and manpower. Accordingly, whether information is 'personal data' must be determined on a cases-by-case basis. The decision is in line with Opinion 4/2007 discussed above.

IP addresses are also likely to be considered 'personal data' under the GDPR. The GDPR defines personal data as including information such as 'an online identifier' where it can lead to identification of individuals, when combined with other information. Recital 30 comments that online identifiers, such as an IP addresses and cookies, 'may leave traces which, in particular when combined with unique identifiers and other information received by the servers, may be used to create profiles of individuals and identify them'.

Is pseudonymous data 'personal data'?

Pseudonymisation is a method of replacing one attribute in a record, such as a name, with another, such as a unique number. It is therefore still possible for an individual to be identified indirectly from such data. Although the Breyer decision does not expressly refer to pseudonymous data, it is recognised as supporting the view that pseudonymous data may be 'personal data' and fall within the scope of the Directive, where there is a legal means by which the data can be retraced to an individual, which does not involve disproportionate effort in terms of time, costs and manpower.

Whilst the ICO guidance suggests that a person who puts in place appropriate technical, organisational and legal measures to prevent individuals being identifiable from the data held may prevent such data falling within the scope of the Directive, it would be prudent, due to the risk of re-identification, for organisations to treat pseudonymised data as personal data. This approach is consistent with the Working Party Opinions 4/2007 and 5/2014 (the latter concerned 'Anonymisation

(Continued on page 14)

(Continued from page 13)

Techniques').

Is anonymised data 'personal data'?

The Working Party is of the view that 'irreversibly' anonymised data does not constitute 'personal data'. However, the threshold put forward by the ICO is lower.

The ICO's Code of Practice on Anonymisation states that organisations need not guarantee that data are 100% anonymised in order for them to be outside of the scope of the Act. Instead, the ICO has said that provided that there is no more than a 'remote' chance that data subjected to anonymisation measures can be traced back to individuals, then those data would be treated as having been anonymised and no longer 'personal data'. Organisations should be aware though that the current state of technology and the information that is available online for re-identification purposes increase the risk of re-identification, and make the threshold for truly anonymised data extremely high.

Pseudonymous data and the Privacy Shield

Interestingly, the issue of key-coded data in relation to pharmaceutical research was addressed in the European Commission's FAQs to the Safe Harbor framework, and in Principle 14(g) of the Privacy Shield replacing that framework. Both indicate that key-coded data transferred from the EU to the US for pharmaceutical research purposes does not constitute a transfer of 'personal data' under the Directive, and hence is not subject to the Principle.

In Opinion 4/2007, the Working Party considered the Safe Harbor FAQs as not being inconsistent with its view that key-coded data are personal data for all parties that might be involved in the possible identification of an individual. The Working Party states that its conclusion is based on the premise that the recipient in the US (i.e. the pharmaceutical company) receives only the key-coded data and will never be aware of the identity of the patients, which is known only to the medical professional/researcher in the EU. It does not consider the issue of whether the US pharmaceutical company has a legal means reasonably likely to be used, to identify a particular patient.

The Working Party's conclusion on this issue may need to be reconsidered in light of the broad construction of 'personal data' taken by the CJEU in Breyer, and the new categories of 'personal data' in the GDPR. Accordingly, organisations would be wise not to view Principle 14(g) of the Shield as providing a carte blanche for transferring key-coded data for pharmaceutical research purposes. Instead, organisations would be safer treating such key-coded data as 'personal data' within the remit of EU data protection law, and subject to the Shield's principles.

Pseudonymous and anonymous data under the GDPR

The GDPR explicitly recognises the concept of pseudonymised data, distinguishing it from anonymised data for the first time. The GDPR encourages pseudonymisation of personal data as a privacy-enhancing technique. Recital 28 of the GDPR provides that pseudonymisation 'can reduce the risks to the data subjects' and 'help controllers and processors to meet their data-protection obligations'. However, pseudonymised data remains within the remit of EU data protection law, where it can lead to identification of an individual.

The ICO's guidance on the GDPR highlights that pseudonymised data 'can fall within the scope of the GDPR depending on how difficult it is to attribute the pseudonym to a particular individual'. Recital 26 of the GDPR also states 'personal data which have undergone pseudonymisation, which could be attributed to a natural person by the use of additional information, should be considered to be information on an identifiable natural person' (i.e.. 'personal data'). On the other hand, Recital 26 states that anonymised data, namely 'data rendered anonymous in such a way that the data subject is not or no longer identifiable', does not fall within the remit of the GDPR. Accordingly, whilst pseudonymisation is a useful security measure, it is not a method of anonymisation, and pseudonymised data remains 'personal data' where it can attributed to an individual.

Is the concept of 'personal data' broader under the GDPR?

Like the Directive, the GDPR defines personal data as 'any information relating to an identified or identifiable natural person'. However it specifies some new identifiers, including 'location data' and 'online identifiers'. Although location data and online identifiers are not expressly included in the definition of personal data in the Directive, it is clear from the Working Party's Opinions, that they are both regarded as a means of indirectly identifying individuals. Accordingly, the GDPR should not be a game-changer for those organisations that have been following such guidance.

Conclusion

In light of the decision in Breyer and the new categories of personal data in the GDPR, along with the hefty fines for non-compliance, it would be prudent for organisations to take steps now to review the data they collect, and assess whether such data falls within the definition of personal data in the GDPR. The express inclusion of location data and online identifiers will, in particular, affect network providers, app developers, device manufacturers, and those involved in data analytics, behavioural advertising and social media, and such organisations will need to amend their policies and procedures to ensure compliance with the new rules.

Davinia Brennan
A&L Goodbody
dbrennan@algoodbody.com