Pseudonymous or anonymous, that is the question

Ali Vaziri, Partner at Lewis Silkin LLP, considers the impact of a recent decision of the EU General Court on identifiability

The question of whether data are pseudonymous or anonymous is painfully familiar to those using AI to mine for ever-deeper insights from ever-larger datasets. The answer is not without consequence. If pseudonymous, the GDPR applies with all its restrictions. If anonymous, then it does not.

The EU has traditionally taken an 'absolute' approach to anonymisation, where re-identification must be impossible for everyone. A 2016 decision of the Court of Justice of the EU ('CJEU') (Case C-582/14: *Patrick Breyer v Bundesrepublik Deutschland*), should have heralded a more pragmatic, 'relative' approach, the test being whether re-identification is legally and practically possible (without disproportionate effort). Unfortunately, it did not.

Whilst EU regulators have mostly continued to take a hard line, the UK Information Commissioner's Office ('ICO') has espoused a 'relative', risk-based approach, even acknowledging that "the same information can be personal data to one organisation, but anonymous information in the hands of another organisation."

Recently, a decision of the EU's General Court ('GC'), Case T-557/20, *SRB v EDPS*, reinforced the view that the risk of re-identification has to be considered from the perspective of the holder of the data, and may encourage EU regulators to embrace a more realistic approach to anonymisation.

Background

The EU's Single Resolution Board ('SRB') is the central resolution authority within the EU's banking union. It ensures the restructuring of failing banks to minimise economic harm. The SRB decided to place the Banco Popular Español ('Bank') under resolution. Professional services firm Deloitte was engaged by the SRB to provide a valuation about whether shareholders and creditors would have received better treatment if the Bank had entered into normal insolvency.

The SRB published a preliminary decision on whether compensation needs to be granted to the shareholders and creditors, along with a non-confidential version of the valuation. It then invited the affected shareholders and creditors to express their interest in exercising their right to be heard.

The right to be heard process was in two phases:

- registration — affected shareholders and creditors were invited to express their interest using an online registration form which included a privacy statement, following which the SRB would verify whether they were eligible. They needed to provide proof of identity and proof of ownership of one of the Bank's capital instruments ('Registration Data'); and
- consultation — eligible shareholders and creditors could submit their comments on the preliminary decision to which the valuation was annexed. They were emailed a unique personal link to an online form which contained seven questions with limited space for their comments.

The Registration Data were accessible to a limited number of SRB staff tasked with determining participant eligibility. Different SRB staff were then tasked with processing the comments received in the consultation phase, and the Registration Data were not visible to them. They only received the comments, each of which had been allocated a 33-digit globally unique identifier which was randomly generated at the time of submission.

During an analysis phase, the SRB filtered the comments to remove duplicates and then categorised them into defined themes. In a subsequent review phase, comments relating to the valuation (as opposed to the preliminary decision) were transferred to Deloitte by uploading files to a virtual server to which access was granted to a limited and controlled number of Deloitte staff who were directly involved in the project.

The comments transferred to Deloitte were:

- filtered, categorised and aggregated. Following the de-duplication, individual comments could not be distinguished within a single theme, so Deloitte was unaware

whether a comment had been made by one or more participants; and

- solely those that were received during the consultation phase and that had an alphanumeric code. Only the SRB could use the code to link the comments to the Registration Data. The alphanumeric code was developed for audit purposes to verify, and if necessary to demonstrate subsequently, that each comment had been handled and duly considered. Deloitte had, and still has, no access to the Registration Data.

The complaints

Five data protection complaints were submitted to the European Data Protection Supervisor ('EDPS') by participants of the right to be heard process. (The EDPS supervises processing by EU bodies such as the SRB; and even though the applicable legislation is different from the GDPR, the provisions are equivalent in substance.)

They complained that the SRB had failed to inform them that the data collected through the responses on the forms would be transmitted to third parties, namely Deloitte and the Bank.

The EDPS agreed and issued the SRB with a reprimand. The SRB requested a review and argued that the information transmitted to Deloitte did not constitute personal data. The EDPS found that the information was pseudonymous data (not anonymous data), even though the Registration Data had not been disclosed to Deloitte.

> ***"Breyer should have marked a move away from this zero-tolerance approach to risk towards a 'relative' approach where the risk of re-identification is assessed with reference to what is legally and practically possible. It is unfortunate that the CJEU did not use that opportunity explicitly to express a view on the 'absolute' approach."***

Appeal to the General Court

The key question for the GC was whether the information transmitted to Deloitte constituted personal data within the GDPR's definition of that term, of which there are two key elements (emphasis added): "'personal data' means any information *relating to* an *identified or identifiable* natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly…".

The GC first observed that the EDPS had not considered whether the information transmitted to Deloitte 'related' to a natural person (i.e. whether the information, by reason of its content, purpose or effect, is linked to a particular person (see Case C-434/16 *Peter Nowak v Data Protection Commissioner*).

It then went on to consider the central question of whether information transmitted to Deloitte related to an 'identified or identifiable' natural person. Given the mechanisms put in place by the SRB, the GC quickly determined that information transmitted to Deloitte did not concern 'identified' persons. The more vexed question was whether the information related to an 'identifiable' person.

The SRB argued that even if the information allowing re-identification is not permanently deleted, data are rendered anonymous for a third party as long as re-identification is not reasonably likely. In support of its assessment of the risk of re-identification, it relied on the fact that Deloitte:

- cannot re-identify the participants from the alphanumeric code assigned to comments — additional information in the form of the decoding database would be needed, to which only the SRB has access; and
- has no lawful means of gaining access to the additional, identifying information.

The EDPS argued that the distinction between pseudonymous and anonymous data came down to whether there was any 'additional information' that could be used to attribute the data to a specific data subject. If there was, then the data were not pseudonymous; if there was not, then they were anonymous. In its view, the Registration Data together with the alphanumeric code constituted a perfect example of 'additional information', because it could be used by the SRB to attribute the data to a specific data subject.

Breyer

The GC referenced the *Breyer* case (full citation above) in its reasoning. That decision concerned the question of whether a dynamic IP address was personal data in the hands of the online media services provider which had registered it even though:

- that services provider was unable to identify the user from the IP address alone; and
- the necessary additional information which, if combined with the IP address would enable the user to be identified, was held by the internet service provider ('ISP').

After drawing a parallel between the alphanumeric code and an IP address, the GC gave a nod to the following elements of the CJEU's reasoning in *Breyer*:

- although the relevant Recital states that "to determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used...either by the controller

(Continued on page 12)

(Continued from page 11)

or by another person" it is not required that all the information enabling the identification of the data subject must be in the hands of one person;

- the fact that the additional information necessary to identify the user of a website is not held by the online media services provider, but by that user's ISP, does not exclude the IP addresses registered by the online media services provider from being personal data;
- it must nevertheless be determined whether the possibility to combine the IP address with the additional information held by the ISP is a means likely reasonably to be used to identify the data subject; and
- that would not have been the case if the identification of the data subject had been prohibited by law or had been practically impossible on account of the fact that it would have required a disproportionate effort in terms of time, cost and person-power, so that the risk of identification would have appeared in reality to be insignificant.

The decision

The GC first noted that it was not disputed that:

- the alphanumeric code appearing on the information transmitted to Deloitte did not in itself allow participants to be identified; and
- Deloitte did not have access to the Registration Data from which participants could be linked to their comments.

The GC then went on to observe that the EDPS had concluded that the information transmitted to Deloitte was personal data simply because the SRB held additional information from which participants could be re-identified (i.e. the Registration Data), even though the EDPS had acknowledged that the Registration Data had not been communicated to Deloitte. The EDPS had therefore merely examined whether it was possible to re-identify the participants from the SRB's perspective and not from Deloitte's.

Instead, the EDPS should have considered whether the possibility of combining the information that had been transmitted to Deloitte with the additional information held by the SRB (i.e. the Registration Data) constituted a means likely reasonably to be used by Deloitte to identify the participants. Since the EDPS had not investigated whether Deloitte had legal means available to it which could in practice enable it to access the Registration Data, the EDPS should not have concluded that the information transmitted to Deloitte constituted information relating to an 'identifiable natural person'.

Comment

The Article 29 Working Party's Opinion 05/2014 on Anonymisation Techniques encapsulated the EU's 'absolute' approach, stating (for example) that "it is critical to understand that when a data controller does not delete the original (identifiable) data at event-level, and the data controller hands over part of this dataset (for example after removal or masking of identifiable data), the resulting dataset is still personal data." The Opinion has yet to be updated by the European Data Protection Board.

Breyer should have marked a move away from this zero-tolerance approach to risk towards a 'relative' approach where the risk of re-identification is assessed with reference to what is legally and practically possible. It is unfortunate that the CJEU did not use that opportunity explicitly to express a view on the 'absolute' approach.

It is equally unfortunate that the facts of *Breyer* were such that it was possible to obtain through legal channels the information needed to identify individuals from their ISPs in the event of a cyber attack. So despite applying a 'relative' approach, the outcome in that case was that the dynamic IP address did comprise personal data: an outcome that has doubtless contributed to the uncertainty surrounding this issue. Whilst this GC decision is expected to be appealed to the CJEU, it is nonetheless welcome as a more realistic approach to anonymisation from an EU court. If it is appealed, it will join at least one other case on the CJEU's cause list on the issue of identifiability: Case C-319/22 *Gesamtverband Autoteile-Handel.* In that case, the Advocate General recently delivered an opinion on whether a Vehicle Identification Number ('VIN') is personal data by considering the issue from the perspective of the particular recipient: does an independent garage reasonably have at its disposal the means to link a VIN to a vehicle owner, such as through a register of registration certificates (to which a public administrative authority will have access)?

Meanwhile, the UK awaits finalisation of the ICO's draft Anonymisation, pseudonymisation and privacy enhancing technologies guidance, and passage of the Data Protection and Digital Information (No.2) Bill that narrows the definition of personal data to limit the assessment of identifiability.

Ali Vaziri
Lewis Silkin LLP
ali.vaziri@lewissilkin.com