Comment on legislation and anonymisation of data in the commercial sphere

On 10 April 2014, the European Union's ("**EU**") independent data protection advisory body, the Article 29 Working Party ("**A29WP**"), adopted an opinion on the use of anonymisation to reap the benefits of Big Data and 'open data' whilst ensuring the protection of personal data ("**Opinion**"). Although not legally binding, the Opinion provides a clear indicator to businesses on best practice for achieving compliance in this complex yet commercially vital area of data protection law.

The need for anonymization

Collecting and processing personal data is heavily regulated under EU data protection law, however if the personal data in question is rendered anonymous so as to prevent identification then it falls outside the scope of regulation as personal data (Recital 26, Data Protection Directive 95/46/EC).

Anonymisation therefore offers businesses a viable way to commercially exploit the value of data relating to individuals (e.g. shopping preferences, medical conditions) without having to surmount the numerous hurdles posed by EU data protection law, such as the requirement to only use data for the purposes consented to at the time of collection and to delete the data once it is no longer required for such purposes.

Determining effective anonymisation

One of the key obstacles for businesses wishing to use anonymised data is the current lack of a prescribed standard for the type and extent of anonymisation required in order to exclude the application of EU data protection law. In the Opinion, the A29WP clarifies that whilst appropriate measures will always depend on the circumstances, ultimately, anonymization should irreversibly prevent identification and that in order to determine the robustness of each technique the following three criteria should be assessed in each case:

a. Singling Out: is it still possible to single out an individual?
b. Linkability: is it still possible to link records relating to an individual?
c. Inference: can information be inferred concerning an individual?

To provide businesses with greater insight into the effectiveness of the myriad techniques for anonymising personal data, the A29WP examines the following techniques:

- **Randomisation**: removing the link between the data and the individual by altering the veracity of certain data.
- **Noise addition**: modifying data to make it broader and less accurate by modifying it with randomised values. For example, a measurement accurate to 1cm could become accurate to 10cm. The amount of ‘noise’ added will depend on the impact of disclosure and the information involved.
- **Permutation**: swapping values within a data set (as opposed to noise addition, above, which adds artificial values). Permutation may be limited to a subset of a larger dataset, for example, within a medical dataset the information of persons with certain

symptoms or conditions may be swapped with other persons in that subset, so as to preserve the correlation of the wider dataset.

- **Differential privacy**: this is a form of data randomisation which involves
- anonymisation upon release of the data (instead of altering the data at an early stage as with noise addition), for example generating an anonymised dataset for a particular audience.
- **Generalisation**: involves generalising or diluting the data (e.g. a region rather than a city).
- **Aggregation and K-anonymity**: generalising and aggregating values. For example, generalising location to a region rather than a city, then aggregating all persons who live in that region.
- **L-diversity/T-closeness**: these are variations on data aggregation which require (respectively) at least L number of variables within a dataset value, or T number of equivalent classes. The use of such formulae aims to prevent identification through inference where there few other variables in the aggregated data. For example, if there are only two cities in a particular region, the effectiveness of aggregating that data will be considerably less than if the region incorporated 10 cities.
- **Pseudonymisation**: pseudonymisation is commonly used to disguise an identity by replacing an identifier (for example, a name) with an artificial alternative. In large datasets this is often achieved automatically via a cryptographic tool which encrypts the name or generates an alternative. However, the A29WP highlights the common misconception that pseudonymising data is

equivalent to anonymising it (which it is not) and emphasises that pseudonymisation is simply a useful security measure.

The A29WP concludes that each current anonymisation technique is limited and that none provides an effective anonymisation solution in every circumstance. To establish what is appropriate in each case a thorough risk assessment should be carried out to determine the most effective protection against the three main risks of singling out, linking, and inference. Clarifying that anonymisation is not a one-off exercise, the A29WP also emphasises the need to regularly re-evaluate the risks associated with anonymisation (together with the measures taken to protect against them) and identify any new risks.

The Opinion also clarifies the important question of whether the act of anonymisation in itself constitutes processing of personal data for the purposes of EU data protection legislation. In the A29WP's view, anonymisation does constitute further processing but this will be compatible with the original purposes of the processing (e.g. the personal data may be anonymised without obtaining additional consent from the relevant individuals) provided that the anonymisation process reliably produces anonymised information in the sense described in the Opinion.

WAB comment

The Opinion makes it clear, however, that current anonymisation techniques are an imperfect solution which must be carefully tailored to the circumstances and kept under ongoing review.

Link to Guidance: http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

SourceWhite & Black Limited - Mathys & Neil

http://www.lexology.com/library/detail.aspx?g=de76c2b3-7202-4222-9c39-cb98f001da7f