論 著
© 臺大法學論叢 NTU Law Journal
第 52 卷第 3 期/Vol.52, No.3 (09. 2023)

The Deidentification of Personal Data and its Risk of Reidentification: A Legal Perspective

Ching-Kuen Ueng[*]

Abstract

The concept of "personal data" as the cornerstone for information privacy laws seems workable. Any data relating to an identified or identifiable natural person will trigger the mechanism of personal data protection.

The operation of big data is to derive or infer hidden value from the structured and unstructured raw data through novel reuse. However, the reuse of personal data will be likely beyond the scope of original collection purpose, in violation of the principle of purpose limitation. Furthermore, the ubiquitous use of personal data will lead to privacy risk. As a consequence, one of the solutions is to deidentify personal data in order to use for further purposes or share with third parties.

However, in the age of big data, as the deidentified or anonymized data may be combined with other datasets from various sources, it is not likely to absolutely ensure "a person cannot be identified from a dataset." The reidentification will cause damages to privacy, personality or property, and the chilling effect on freedom of expression.

As there were several famous reidentification cases in the past two decades, the effectiveness of deidentification or anonymization is gradually criticized. However, some scholars insist that the deidentification or anonymization is still effective in protecting privacy because the rate of reidentification is very small. Similarly, the U.S. courts are also divided in their effectiveness.

[*] Associate Professor, School of Law, Fu Jen Catholic University.
E-mail: 071617@mail.fju.edu.tw

In facing the conflict between deidentification and reidentification, there could be some solutions. Firstly, the key point is to adopt a reasonable deidentification standard, thus reducing the risk of reidentification to a not important degree, rather absolutely ruling out its risk. Secondly, data controllers shall evaluate the risk of reidentification and thus adopt the technical, legal, and organizational safeguards subject to the principle of proportionality. Finally, statutes shall include civil and criminal liabilities in order to prohibit improper reidentification.

Keywords: personal data, privacy, big data, identifier, deidentification, anonymization, reidentification, freedom of expression, tolerant of risk, General Data Protection Regulation