大數據應用下個人資料的法律保護

葉志良*

近年來大數據技術及其應用對資訊經濟、科學發展與商業利益帶來極大貢獻，但一般人經常於未知情下被他人（甚或機器）蒐集與其個人有關之各種資料，滋生隱私侵害疑慮；不過資料保護法律並非將隱私權視為至高無上，在法益衡量上仍須與其他社會價值進行權衡，在資料利用方面或可透過去識別化做法，確保資料處理的合法性。本文強調除非是屬於無法合理期待之隱私資訊，否則法律應賦予當事人控制其資料的權利，在此前提下經去識別化後而可供利用之資訊，應可運用於大數據的探勘與分析，以創造資料經濟之價值，並有利於公共利益與福祉之最大化。

一、大數據浪潮與個資隱憂

時下熱門的大數據技術，透過業者所大量儲存使用者的使用紀錄，將這些資料透過巨量分析與資料探勘等方式，萃取出有用或可供預測的資訊，可幫助業者瞭解使用者行為、進而發展新服務，此大數據技術為資訊經濟開創新的時代，也對科學與商業帶來極大利益，同時也影響國家政策的方向，譬如災防告警系統、天然資源最佳化與資訊基礎建設等。

在大數據巨大商業利益的驅使下，企業以最大程度蒐集、處理、利用個人資料作為提升其產業競爭力的重要手段，使得傳統《個資法》上「知情同意」、「目的限定」、「資料蒐集最小化」等原則遭受嚴峻的挑戰，尤其在資料處理喪失原有的「情境脈絡」（context）[1]下，個人對於其個資遭他人蒐集後進行何種方式之利用往往毫無察覺，且大數據分析所建構、追求的個性化服務，抑或原為合法蒐集之個資遭到不當處理，例如因有缺陷的去識別化（de-identification）使得本

* 元智大學資訊傳播學系助理教授

[1] Helen Nissenbaum. (2010). *Privacy in Context: Technology, Policy, and the Integrity of Social Life.* Stanford Law Books.

不受《個人資料保護法》（下稱《個資法》）所保護的資料，再度恢復為可辨識出特定之個人[2]，這些都對個人的隱私造成莫大威脅。

二、個資保護法律及其界限

大數據應用本質上係追求資料開發的價值最大化，而個人資料保護的目的則在於保障個人對於個資的自主控制，這兩者價值各異，實難以置於同一天平上衡量其輕重；然倘若這兩者價值間產生交錯時，孰輕孰重仍應按法益衡量方式按個案方式處理，或可調整既有框架（如採取 Privacy by Design 概念，或調整《個資法》規範或其範圍等）以尋求利益最大化的解決之道。

由於《個資法》的目的同時保障資料主體之人格權以及資料的合理利用，因此如何兼顧當事人與資料利用者間之平衡，殊值探討。我國《個資法》第 3 條賦予當事人行使權利並近用資料之權利，其規定：「當事人就其個人資料依本法規定行使之下列權利，不得預先拋棄或以特約限制之：一、查詢或請求閱覽。二、請求製給複製本。三、請求補充或更正。四、請求停止蒐集、處理或利用。五、請求刪除。」司法院大法官會議第 603 號解釋也提到：「其中就個人自主控制個人資料之資訊隱私權而言，乃保障人民決定是否揭露其個人資料、及在何種範圍內、於何時、以何種方式、向何人揭露之決定權，並保障人民對其個人資料之使用有知悉與控制權及資料記載錯誤之更正權。」近來歐盟一般資料保護規則（General Data Protection Regulation, GDPR）也相當強調資料處理須經資料當事人的明確同意，因此如何建構與維持資料主體與資料利用者間的信賴關係，誠為資料利用的首要之務。

我國《個資法》對個人資料有其規範意義上的定義：得以直接或間接方式識別該個人之資料，但是否所有與個人有關之資訊皆屬值得受法律保護之個資，實有疑問。1967 年美國聯邦最高法院 *Katz v. US.* 一案，將隱私權保障從原本僅限於私人處所擴大至個人在公共場所中之行為，也從原本僅限於有形之物擴大至無形的資訊，另本案奠定「合理的隱私期待」（reasonable expectation of

[2] 《個資法施行細則》第 17 條有去識別化之規定，「無從識別特定當事人」是指個人資料以代碼、匿名、隱藏部分資料或其他方式無從識別該特定個人者，因此去識別化之資料非屬於個人資料；惟學者 Paul Ohm 認為在大數據時代，資料大量產生後將會被蒐集與連結，因資料來源的廣泛性使得可供比對的數量大量增加，加上資訊技術的發展與資料辨識能力大幅提升，使得資料越來越難以保持匿名化狀態。這種再識別科技的發展可能會摧毀我們對於匿名所欲達到隱私權保護的信心。Ohm, Paul. (2010). Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. *UCLA Law Review*, 57, 1701, 1731.

privacy）判斷標準：個人在主觀上必須先有「隱私期待」，再進一步判斷此主觀的期待於客觀上是否可被認為是合理的，其中主觀要件判斷關鍵在於反面角度觀察個人是否已「自願」放棄隱私利益，藉以判斷其主觀上是否仍具有隱私期待，通常這種判斷是建立在隱私主體的「同意」與否。至於主觀的隱私期待在客觀上是否「合理」之判斷，則須綜合考量時間與空間等相對因素，在不同時代、地域、社會、科技發展與社會認知等，會有不同的判斷結果。

在目前個資定義「二分法」判斷下，對於能否識別特定個人的「關聯性」或「連結性」的片段資料，並非足以區分資料主體的識別要素，而是作為充實其既有之個人剖繪（enrich existing profiles of individuals）之要素，藉此增強資料的累積效應[3]，也因此能相互連結的資料量越多，可識別出特定個人的可能性也越高，即知曉其屬於何人並建構、充實其個人剖繪的機率就越高。故判斷一筆資料是否構成個人資料，須在具體情境脈絡中進行個案權衡而無法跳脫情境作抽象、靜態的考量[4]。

傳統上對個資進行去識別化被視為是允許企業得以利用資料分析技術並同時保護個人隱私，例如 Google 透過移除連結（de-link）將個資中可直接辨識之資料，盡可能移除指向某一特定主體之資料集片段的「連接性」，讓資料盡量分散，使其變為足夠小的片段，讓搜尋到完整資料的可能性大幅降低，然而這種操作在實務上難度頗高，經常讓一般人誤認資料去識別化後是永久性，或認為匿名資料被利用可安心無憂等，而忽略了個資風險的存在。由於匿名資料經常會被再識別而將該資料指向特定個人，因此去識別化資料其實只是一種暫時的狀態，且經常引起個人隱私的顧慮。

三、近來司法個案認定對個資利用的影響

2012 年某行動電話業者提供一項 App 應用服務名為「M+ Messenger」，允許使用者得以知悉其手機通訊錄中親友電話號碼所屬之行動電話業者別，藉以確認通訊錄中的親友是否屬於網內（intra-network）以享有較便宜的通話費率；但該 App 遭到消費者提告，認為該 App 違法蒐集並使用其個人資料，請求因隱私

[3] Article 29 Data Protection Working Party, *Opinion 05/2014 on Anonymisation Techniques*, 0829/14/EN, WP 216 (Apr. 10, 2014), at 4.

[4] Paul M. Schwartz & Daniel J. Solove. (2011). The PII Problem: Privacy and a New Concept of Personally Identifiable Information, 86 NYU L. REV. 1814, 1848.

權受侵害之損害賠償。然而法院對於「電話號碼所屬之電信業者別」這項資料是否屬於個資，見解各異。

採肯定說（臺灣臺北地方法院 103 年度小北字第 1360 號、103 年度小上字第 155 號民事判決）者肯認電話號碼是一種個人資料，而「電話號碼所屬之電信業者別」乃個人電話號碼之附屬資料，可以間接識別特定個人，應屬於個人資料。採否定說（臺灣臺北地方法院 103 年度訴字第 255 號民事判決）者則認為行動電話號碼僅為數字之組合，無法直接識別特定個人，且與其他資料對照、組合、連結方式而識別特定個人，亦不具間接識別性，而「電信業者別」並無法直接識別特定個人，縱與電話號碼相結合，依前述電話號碼已難直接或與其他資料連結間接識別特定人之情況，仍極難識別特定個人，非屬《個資法》所定之個人資料。

作為電信事業主管機關之國家通訊傳播委員會，於 2017 年 6 月 28 日第 754 次委員會議討論此案是否涉及違反《個資法》而予裁罰，最後決議不予裁罰，其理由認為：經由 App 提供用戶檢視通訊錄中聯絡電話號碼之所屬電信業者，或變更檢視結果為「網內／網外」，均不足以直接或間接識別該等電話號碼之個人身分，自無人格權受侵害之事實，亦非《個資法》所欲保護之標的及適用。

從以上個案卻出現法院見解分歧的情形，顯見國內企業對於資料利用仍趨於保守，而包括政府在內的多數人仍誤認《個資法》之目的僅在避免個人之人格權受侵害，殊不知「促進個人資料之合理利用」亦為本法立法目的之一。實則本文贊同否定說法院見解，認為間接識別之個資並非毫無限制，倘未予適當限制，則舉目所見之資料盡屬《個資法》所保護之個資[5]。本文認為當間接識別資料對於識別特定個人太過於困難，諸如須經反覆多層遞次對照、組合、連結而耗費鉅大時間、費用、人力後，或可識別某一特定人；抑或該資料根本無法協助識別以至於不構成合理的隱私期待，此非《個資法》為促進個人資料之合理利用之立法目的，且恐不當壓縮言論自由、資訊自由及公共利益。

四、結論：邁向雙贏的個資保護

資通訊科技（ICT）改變人與人間的互動方式，透過關鍵字搜尋、cookies、串流資料、tweets、社群媒體、位置資訊服務、近場通訊（Near Field Communication）行動交易等方式所獲得新的資料來源，不但可提供給處理資料的產業進行資料

[5] 按臺灣臺北地方法院 103 年度訴字第 255 號民事判決，其指出：「……解釋《個資法》第 2 條第 1 款所定間接識別之個人資料，應以『合理、可能、容易之方法』為限。」

的加值利用，同時也可進一步成為資料管理產業開發創新應用的核心所在，並用以協助企業與政府增進其業務與政務的管理、服務的優化，進而創造消費者與人民的滿意與信賴；此外，藉由個人與機構間的資訊分享，將可促成資料加值與支出減少的可能性。

在大數據時代下，企業不應將重心僅集中在如何使去識別化做到「完美無缺」，而是在承認去識別化本身並不完善且具有一定風險的基礎上，重視如何將資料隱私風險掌控在合理且可接受的範圍內。判斷去識別化的有效性，實際上是對風險本身的評估，更正確來說，是將個資的生命週期（從隱私權政策、風險評估、去識別化操作、重新識別評鑑等）都納進來，才是真正的去識別化。

學者 Tene 與 Polonetsky 倡議，可比對資料再識別本身在統計學上的可能性，並搭配企業對於不進行再識別的合法承諾以及契約義務[6]，以達到真正的防制效果。本文認同兩位學者看法，認為應將基於資料安全與課責性原則所採取的去識別化方法視為一種保護措施（protective measure），而非作為大數據謎題的解方。企業藉由盡可能將個資去識別化，同時不放棄對資料進行有益利用，或許是值得讓大數據對於資料進行蒐集與利用較有智慧的做法[7]。

在鑑別去識別化的有效性上，首應考慮匿名資料之使用目的，或資料接收者有無利用個資之動機。由於企業對匿名資料的二次利用通常不同於資料初始之處理目的，因此從目的角度切入去識別化是否充分以及是否合乎「善意不知情」，可成為較為有效的判斷方式[8]。

大數據應用對於法規範所產生的巨變，例如告知同意原則之適用，本身就是相當棘手的問題。本文於分析 M+ Messenger 判決後嘗試檢討可能、合法的作為，例如業者可考量先取得使用者同意，若有不同意者，則業者可以屏蔽或不顯示電信業者別之方式作為解套。但此《個資法》下「告知同意原則」極可能摧

---

[6] 國內學者范姜真媺教授認為，因匿名化處理而留存轉換程式或對照表，應將轉換程式或對照表毀棄，或至少應將原始資料與匿名化處理資料分由兩個不同機關保管，讓被提供利用之資料與原始資料形成「連結不可能匿名化」狀態，對當事人之保護始得稱為完全。參見范姜真媺（2013）。〈個人資料保護法關於「個人資料」保護範圍之檢討〉，《法學研究》，第 41 期，頁 100。

[7] Tene, O. & Polonetsky, J. (2013). Big Data for All: Privacy and User Control in the Age of Analytics. *Northwestern Journal of Technology and Intellectual Property*, 11, 239, 259. 去識別化處理方法有很多種，但必須考量去識別化之目的、資訊欄位、所要串接的資料是否可能連結導致重新識別等因素才能決定。對於個資去識別化，在技術上或可建立一套驗證標準規範，以供企業遵循，雖目前我國已公告 CNS29100 與 CNS29191 兩項國家標準，但因國家標準並無強制性，在目前檢驗成本並不低廉且業界普遍採納度不高的情況下，企業要能明確且合法的利用資料仍有一大段路要走。

[8] 劉定基（2012）。〈個人資料的定義、保護原則與個人資料保護法適用的例外 ——以監控錄影為例（上）〉，《月旦法學教室》，第 115 期，頁 50-51。

毀軟體設計本身所欲達到之效果，究竟資料合理利用（創新）與個資隱私保護（規範）該如何兼備，值得有識者持續研究。

參考文獻

范姜真媺（2013）。〈個人資料保護法關於「個人資料」保護範圍之檢討〉，《法學研究》，第 41 期，頁 91-123。

劉定基（2012）。〈個人資料的定義、保護原則與個人資料保護法適用的例外——以監控錄影為例（上）〉，《月旦法學教室》，第 115 期，頁 42-54。

Helen Nissenbaum. (2010). *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford Law Books.

Ohm, Paul. (2010). Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization. *UCLA Law Review*, *57*, 1701-1777.

Schwartz, P. & Solove, D. (2011). The PII Problem: Privacy and a New Concept of Personally Identifiable Information. *New York University Law Review*, *86*, 1814-1894.

Tene, O. & Polonetsky, J. (2013). Big Data for All: Privacy and User Control in the Age of Analytics. *Northwestern Journal of Technology and Intellectual Property*, *11*, 239-273.