公務機關利用去識別化資料之風險評估及法律責任 上

文／林裕嘉

壹 前言

近年來，拜網路科技便利所賜，民眾生活型態大幅改變，尤其在社交與資訊的接觸管道上，已從實體的人際互動與傳統媒體，轉向網路社交媒體。網路虛擬世界發展備受國際重視，許多先進國家已調整法規制度，因應新興產業活動及生活型態。我國網路科技發展已躋身先進國家之列，網路使用人口亦快速攀升，為使我國相關法制環境與國際接軌，以促進產業發展，行政院特於 103 年成立「虛擬世界發展法規調適工作小組」，透過跨部會的整合協調，致力建構完備的虛擬世界法制環境。此外，為了因應網路時代的來臨，行政院規劃運用「群眾外包」（crowd-sourcing）、「大數據」（big data）與「開放資料」（open data）等 3 項作法，讓政府融入網路世界，除了能直接聽取民眾意見、討論公共事務外，鑑於公部門因執行法定職務，掌握大量的政府資訊，若能透過大數據技術的適當統計分析，加以有效應用，不但可增進公共服務效率、便利人民生活，亦有助於政府整體政策規劃。另一方面，將政府資料開放予人民利用，可結合民間資源及創意，達成施政便民及公開透明之目的。然而，政府資訊若含有個人資料，於轉化成「開放資料」之過程，也衍生個人隱私方面的風險及隱憂，如何兼顧人格權之保護及個人資料之合理利用，乃各界所關切之重要課題。

目前歐盟具有全球最先進之個人資料保護架構，在歐盟規範架構下，政府資訊涉及個人資料時，其開放公眾利用仍須受個人資料保護法之限制，因此除法規另有特別規定外，涉及個人資料之政府資訊大多屬於限制或禁止公開之範圍，但若個人資料已「去識別化」達到無從識別個人資料當事人，而為「匿名」（anonymous）之狀態，歐盟認為此時已非個人資料，自無歐盟個人資料保護指令之適用，原則上應屬政府資訊開放公眾利用之範圍，惟去識別化的資料必須達到「以一切可能合理之方法」（all the means likely reasonably to be used）無從再識別資料當事人之程度，否則仍應受個人資料保護指令規範[1]。

然而，如何界定「個人資料已去識別化」，在大數據時代有其困難性及不確定性，蓋大數據的演算系統有時能從龐大雜亂的資料中挖掘出不容易預見的關聯，即使依當時科技或專業水準被認為已去識別化的資料，日後隨著資料的不斷累積、資訊技術的發展演進、資料儲存成本的下降，仍可能透過與其他資料的對照、組合、連結而被重新識別，在此理論基礎下，嚴格來說，幾乎所有資料均可能是潛在的個人資料。因此，參考歐盟、英國及日本的文獻，均一再強調去識別化不可能絕對毫無風險[2]，故如何依循個人資料保護法之解釋，並採用標準作業流程（standard operation procedure），有效評估並控制去識別化之資料被再識別的風險，為此一保護規範架構的核心。

貳 個人資料之定義與判斷相對性

一、定義—我國法之規定

依我國個人資料保護法（以下簡稱個資法）第 2 條第 1 款規定，個人資料是指姓名、出生年月日、國民身分證統一編號、護照號碼、特徵、指紋、婚姻、家庭、教育、職業、病歷、醫療、基因、性生活、健康檢查、犯罪前科、聯絡方式、財務情況、社會活動及其他得以直接或間接方式識別該個人之資料。簡言之，個資法將個人資料區分成兩類：

（一）直接識別之個人資料：指接觸該資料之人，可直接將該資料與特定人連結。一般而言「姓名」若綜合其他資料，而足以連結特定某一自然人，即會被認為屬直接識別之個人資料，但並不以此為限；只要於具體情形得以直接透過該資料與特定個人建立連結，從而自群體中識別出該特定個人者均屬之。

（二）間接識別之個人資料：由於社會態樣複雜，某些資料雖未直接指名道姓，但一經揭露仍足以識別為某一特定人。所謂「得以間接方式識別」，依個資法施行細則第 3 條規定，係指保有該資料之公務或非公務機關僅以該資料不能直接識別，須與其他資料對照、組合、連結等，始能識別該特定之個人。又所謂經由對照、組合、連結等方式識別特定個人，必須是精準且確定地得以識別出特定個人，而非僅是利用專業知識「猜測」、「推測」可能為某人之資料，即使該「猜測」或「推測」結果是正確的，亦不能因而認為猜測者有取得該個人資料[3]。

二、個人資料判斷之相對性

上述有關個人資料之定義包含得以直接或間接方式識別該個人之資料，對於個人資料或隱私之保護固然周全，但因得為比對組合之人及資料範圍難有一致性、明確性之認定標準。如同個資法施行細則第 3 條修正理由說明所示：「是否得以直接或間接方式識別者，需從蒐集者本身個別加以判斷，此宜於個案中加以審認，……」換言之，欲判斷一份資料是否為個人資料，對於不同蒐集者而言，即可能會產生不同之判斷結果，蓋蒐集者原先所持有或可能得以獲取之資料範圍有所不同，其與新蒐集之資料，是否得以透過對照、組合、連結等方式識別特定之個人，亦會有不同之結果，自難以一概而論，此即所謂「個人資料判斷之相對性」。

三、英國墮胎統計資料公開案

英國資訊保護官辦公室（Information Commissioner's Office）於 2012 年發布「匿名化：資料保護風險管理實踐準則（Anonymisation: Managing Data Protection Risk Code of Practice）」，為歐洲首部由個人資料保護機關所發布之官方匿名化實踐準則。而觸動英國政府開始重視此議題之原因，係源於英國墮胎統計資料公開之爭訟案件，該案涉及政府資料開放所帶來之公共利益及個人隱私保護間之權衡問題，而法院調查之重點包含原告請求政府公布之系爭資料是否仍為個人資料？系爭資料是否已有效匿名化，轉變為單純的統計數據而無從識別特定之個人？值得參考。

（一）事實背景：2003 年以前，英國健康部（Department of Health）所公布的醫療統計數據非常詳細。惟 2003 年以後，英國健康部大幅減少釋出詳細的統計數據，包含移除任何少於 100 人的數據資料，也不提供晚期墮胎人數。The ProLife Alliance 團體於 2005 年 2 月 21 日要求英國健康部公布某地區懷孕晚期墮胎的完整統計資料，英國健康部拒絕公布，其認為墮胎統計資料將促使有心人士（例如記者、媒體）有充分動機去識別相關當事人（包含醫生及婦女），且統計資料與在地知識見聞對照、組合、連結後，容易使人臆測或誤信其已成功追蹤識別特定當事人，此一風險並將隨著社交網站之普及、資訊網路科技之發展而日益增加，並主張該統計資料仍屬個人資料，若予以公布將違反英國資訊自由法（Freedom of Information Act 2000, 以下簡稱 FOIA）之禁止開放條款以及個人資料保護法（Data Protection Act of 1998, 以下簡稱 DPA），因而拒絕公布。英國資訊保護官（the Information Commissioner）則認為系爭統計資料並非個人資料，因此並無 FOIA 禁止開放條款之適用，亦無 DPA 之適用。

（二）法院判決：

1. 英國資訊法庭（the Information Tribunal）認為系爭統計資料在英國健康部保管之下，因英國健康部仍得對照、組合資料來源識別特定個人，依據DPA規定確實構成個人資料；然而，當其對外揭露系爭統計資料被第三人識別之可能性已微乎其微，則該揭露並不違反DPA，不符合FOIA之禁止開放事由，且經權衡公開系爭資料的公共利益，可認為符合比例原則及正當性。因此，英國健康部拒絕揭露系爭統計資料違反FOIA之規定，判決英國健康部應公布系爭統計資料[4]。

2. 英國健康部上訴高等法院（the High Court），高等法院參考上議院（the House of Lords）有關公布白血病孩童統計資料的判決先例[5]，認為法院有權針對系爭資料是否已完全匿

（文轉三版）

註釋

1. 參照 1995 年歐盟個人資料保護指令前言第 26 點 (Recital 26 of Directive 95/46/EC)：「26. Whereas the principles of protection must apply to any information concerning an identified or identifiable person; whereas, to determine whether 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; whereas the principles of protection shall not apply to data rendered anonymous in such a way that the data subject is no longer identifiable; whereas codes of conduct within the meaning of Article 27 may be a useful instrument for providing guidance as to the ways in which data may be rendered anonymous and retained in a form in which identification of the data subject is no longer possible.」。
2. 歐　盟A29WP，Opinion 6/2013 on open data and public sector information ('PSI') reuse (Wp207), p.13。另可參考Kieron O'Hara（南安普敦大學）於 2011 年為英國內閣辦公室所撰寫的報告 "Transparent Government, Not Transparent Citizens: A Report on Privacy and Transparency for the Cabinet Office"，作者於該文中警告從匿名化資料中結合其他資料進而識別特定個人之「拼圖識別（"jigsaw identification"）」可能性，並認為並無完美周全的技術方法可徹底解決「去匿名化（de-anonymisation）」的問題，資料來源：http://www.cabinetoffice.gov.uk/sites/default/files/resources/transparency-and-privacy-review-annex-b.pdf
3. Information Commissioner's Office, Anonymisation: managing data protection risk code of practice, pp26-27.
4. Department of Health v. Information Commissioner, Case Ref EA/2008/0074.

JUDICIAL 司法 從稿 院網站。 一律不退件，請自留底稿。 經刊登，酌送稿酬，並寄贈該期刊物。 訂閱辦法 本刊每 全年 費用新臺 如蒙 請郵政 帳 5005 帳 司法院

（文接二版）

名為事實認定，亦即判斷該資料是否已無法單獨或與其他資料結合後識別特定自然人，並認為系爭統計資料並非個人資料（即認為資訊法庭認定其屬個人資料之見解有誤，採認英國資訊保護官認為其非屬個人資料之見解），縱使資料控制者（即英國健康部）擁有足以辨識相關當事人之資料，但並不代表其對外公開釋出已充分匿名之資料仍然是個人資料，重點並不在於資料控制者手中所持有資料之狀態，而是在於資料控制者對外釋出揭露之資料內涵。固然資料控制者擁有辨識相關當事人之鑰匙（key），故資料控制者所保有的整體資料仍然屬於個人資料，但其對外公布釋出完全匿名的資料並無違反DPA的問題，因為當該資料對外公布釋出之時，其已不再是個人資料。

3. 上議院於前揭有關公布白血病孩童統計資料的判決先例中提出一項測試方法：「其他資料」與系爭統計資料對照、組合、連結後，是否將得以識別出特定當事人？若非，則該統計資料並非個人資料；若「其他資料」本身即可單獨識別出特定當事人，而非與系爭統計資料對照、組合、連結後才得以識別，則亦應認為系爭統計資料已經匿名化，而非個人資料。在本案中，高等法院透過上述測試方法，認定英國健康部墮胎統計資料並非個人資料[6]。英國健康部決定不再上訴，並依法院判決公布系爭晚期墮胎統計資料。

本案為關於個人資料匿名化的重要案例，高等法院認為英國健康部在將系爭資料轉化為統計資料呈現時，已經有效地將該資料匿名化，考量一切可能合理之方法，任何接收該統計資料者均已無從識別特定個人，因此英國健康部所公布的詳細墮胎統計資料並非個人資料。綜上可知，高等法院於本案亦採行了「個人資料判斷相對性」之見解，只要公布之資料依其呈現方式無從使資料接收者直接或間接識別特定個人，則該公布之資料即非屬個人資料。

參 個人資料之去識別化

一、意義及該行為之定性

依我國個資法第 2 條第 1 款有關個人資料之定義反面解釋可知，所謂「去識別化」，即指透過一定技術與程序的加工處理，使個人資料不再具有直接或間接識別性；故個人資料去識別化之行為應定性為個資法第 2 條第 4 款所稱之「處理」[7]。而觀察個資法之規範體系架構，係將「蒐集」及「處理」行為規範於相同法定要件之下，蓋個人資料之「蒐集」大多緊密伴隨著「處理」行為，故個資法並未特別區隔兩者之行為要件；且因去識別化資料須達到無從直接或間接識別特定人之程度，故去識別化之加工處理若確實有效執行，其並不會增加對當事人權益之額外侵害，應可認為去識別化之處理並未逾越原先蒐集之特定目的。因此，如公務機關或非公務機關原先蒐集個人資料之行為符合個資法第 6 條、第 15 條或第 19 條之規定，其在進行去識別化之加工處理前，可不必再另外取得當事人之同意[8]。

二、國外法制之分類

參考歐盟、英國及德國的文獻，去識別化依其加工程度不同，有以下「匿名化資料」及「擬匿名化資料」之類型:

（一）匿名化資料（anonymised data）[9]：

匿名化資料必須對任何人（包含原資料保有者）而言，均無法採取任何合理可能之方法識別特定個人，亦即資料經加工後，毫無保留連結之可能性；至於判斷資料是否已達匿名化之程度，仍須評估各種情況而視個案而定；例如：即使在呈現聚集化統計或彙總資料時，樣本數是否超過足以識別特定個人之門檻。歐盟第 29 條資料保護工作小組（Article 29 Data Protection Working Party，以下簡稱歐盟A29WP）甚至對此採取非常嚴格之門檻標準，認為如果資料保有者在事件層次（event-level）並沒有刪除原始（可識別）資料（例如：在移除或摘蓋可識別資料之後），即提供資料集的一部分，則所產生的資料集，仍然是個人資料[10]。

（二）擬匿名化資料[11]（pseudonymised data，另有譯為「假名化資料」）：

擬匿名化資料乃是以編碼或別名取代識別符（例如姓名、國民身分證統一編號等），使研究或統計人員得以針對個體資訊進行分析而無須識別個體身分。可再細分為 2 種態樣：

1. 不可逆（non-retraceable/irreversible）：

此態樣欲使重新識別不具可能性，以非專屬代碼、單向加密（one-way cryptography）或其他技術處理後，使任何人（包含原資料保有者）均無法透過資料比對或其他方式再直接或間接辨識出特定個人。

2. 可逆（retraceable/reversible）：

以專屬代碼、雙向加密（two-way cryptography）或其他技術處理後，處理後之編碼資料雖無從識別特定個人，惟原資料保有者仍得透過代碼與原始識別資料對照表或解密工具（鑰匙）還原為識別資料，則屬可逆之擬匿名化資料（例如：進行醫療實驗研究時，為能適時回溯追蹤調整對受試病患之醫療處置）。對於原資料保有者而言，因其仍得透過對照表或解密工具等其他資料之對照、組合、連結而識別特定個人，該資料仍屬個人資料；惟對於資料接收者而言，則須區分下列情形[12]：

(1) 倘若經加工降低識別可能性後，其所提供之擬匿名化資料，因去連結之程度低，資料接收者運用一般知識予以連結之再識別可能性仍大，此時對於原資料保有者及資料接收者應仍屬間接可識別之個人資料，而有個資法之適用。

(2) 倘若經加工降低識別可能性後，其所提供之擬匿名化資料，因去連結之程度高，使資料接收者運用一般知識予以連結之再識別可能性甚小，對於原資料保有者而言雖仍屬間接可識別之個人資料，因而仍有個資法之適用，但對於資料接收者而言，則可認為已非個人資料，而無個資法之適用，此為德國見解[13]。另前述英國資訊保護官及英國法院認為，即使原資料保有者仍保有足以辨識特定個人之資料，只要原資料保有者並未將對照表或解密方法等連結工具提供給資料使用者，且採取適當安全措施，則該釋出的資料亦屬匿名化資料[14]。英國資訊保護官對此議題之指導方針係認為，若權衡其可能性後，已無法透過該資料與其他公眾可得之資料對照、組合、連結而識別出特定個人，則可認為該資料確實已匿名化，並非DPA所規範之個人資料，因而無DPA之適用。保有原始資料之公務機關，原本即可透過該原始資料本身識別特定個人，故不應認為該公務機關對外所釋出之匿名化資料，會因公務機關有能力識別而轉變成個人資料。英國資訊委員會此一見解亦被前述上議院有關公布白血病孩童統計資料的判決先例及高等法院墮胎統計公布案判決所支持採認。至於個人是否有可能透過該匿名化資料被重新識別乃是事實問題，必須視個別案例之具體情形而定。

（下期待續）

（作者為法務部法律事務司科長）

註釋

5. Common Services Agency v. Scottish Information Commissioner [2008] UKHL 47 (CSA).
6. Department of Health v. Information Commissioner, [2011] EWHC 1430 (Admin).
7. 德國聯邦個人資料保護法所規定之「處理」行為之內涵包含變更（Verändern），而「匿名（Anonymisieren）」定性為個人資料之變更，「假名（Pseudonymisieren）」乃是以符號替代可識別資訊，亦為變更之一種內涵，故「假名」之定義增列於第 3 條第 6 款「匿名」定義之後（即第 3 條第 6a 款）。
8. 英國 ICO 認為，英國 1998 年資料保護法（The Data Protection Act 1998）規範了各種合法處理個人資料之情形，而取得當事人同意僅是其中一種情形，只要匿名化能被確實有效的執行，其不具有造成當事人不當損害或傷害之可能性，則匿名化之處理無須事先得到當事人的同意，並認為取得當事人同意將可能是一項非常繁重的過程，例如涉及大量個人資料之情形，甚至是不可能的，例如個人資料是老舊的，因而欠缺可靠方法與該當事人進行聯繫，進而取得其同意（按：英國未特別區分蒐集、處理及利用行為，一律均稱之為process，在此均譯為「處理」）。可參考Information Commissioner's Office, 同註 3，pp.28-29。
9. 可參考CNS29100（X6112）2.2 及 4.4.4 之定義。
10. 國家發展委員會委託研究報告（受委託單位：電子治理研究中心；研究主持人：項靖），開放資料及其對政府治理與個人隱私影響之研究，第 68 頁；歐盟A29WP，Opinion 5/2014 on Anonymisation Techniques(WP216), p.9.
11. 同註9。
12. 「可逆之擬逆名化資料」並非全然排除個資法之適用，此時對於原資料保有者（即掌握對照表或解密工具者）而言，仍屬個人資料，而有個資法之適用；然對於資料接收者而言，尚須視加工程度高低、連結之可能性大小等因素個案判斷。簡言之，倘若該假名化資料之利用目的及程序得排除後續重新連結識別之機會或可能性（例如：資料接收者利用該編碼資料進行研究所得之結果，不會透過對照表重新識別出個人資料當事人後，將該研究結果回溯套用至該當事人，從而對當事人產生影響），且在制度上嚴格排除資料接收者再回溯識別的能力（例如：未掌握對照表或解密工具、加工程度高故再連結可能性甚小），並在組織或技術上採取適當安全措施，避免資料使用者以其他方式重新識別相關資料（例如：使用高強度的加密方式、嚴格管控防堵對照表或解密工具之外洩可能、訂立契約禁止資料使用者還原資料等），此時就資料接收者而言，如已無從採取任何合理可能之方法直接或間接識別特定個人，則對該資料接收者而言，該資料即非個人資料，從而不受個資法的規範。參考歐盟A29WP，Opinion 4/2007 on the concept of personal data(Wp136)，pp18-20；劉定基，雲端運算與個人資料保護—以台灣個人資料保護法與歐盟個人資料保護指令的比較為中心，東海大學法學研究第 43 期，2014 年 8 月，第 71 頁；W. Kuan Hon, Christopher Millard, and lan Walden, The problem of "personal data" in cloud computing: what information is regulated?—the cloud of unknowing, International Data Privacy Law, 2011, Vol.1, No.4, pp216-217。
13. 參考Directorate-General Justice, Freedom and Security, Comparative Study on Different Approaches to New Privacy Challenges, in Particular in the Light of Technological Developments, Country Studies A.4–Germany (Douwe Korff, Editor), p.4.
14. Information Commissioner's Office,同註3，p.7；Information Commissioner's Office, Determining what is personal data, pp27-28；另可參考 W. Kuan Hon, Christopher Millard, and lan Walden,同註 12, pp213-214.

WEEKLY 周刊 徵稿 高論 啓事
一、本刊歡迎兼顧理論與實務、探討司法制度、法學新知的作品。文長宜在八千字以內，以利一次刊出。文後請加註本人、事及法院的故事，以利參考。
二、本刊並歡迎一次刊出文章並將置司法周刊出版。
三、來稿有關人、事、時、地、物的敘述，請加註說明；引用他人著作，請加註出處，以免侵害著作權。
四、不願接受刪改者，請於來稿時敘明。
WEEKLY 周刊
每週出刊
訂閱
優惠價400元
訂閱
劃撥
號
4377
戶
司法周刊