《安全與情報研究》 第五卷第一期 中華民國 111 年 1 月
Security and Intelligence Studies, Vol. 5, No. 1. January 2022

資料去識別化到聯盟式學習：跨組織資料分析之隱私議題探討

李思壯
淡江大學資訊傳播學系助理教授

黃彥男
中央研究院資訊科技創新研究中心特聘研究員

陳意文
淡江大學資訊傳播學系副教授

摘 要

近年來各式資料探勘技術進展迅速，大數據分析不論在實務上或學術上都開始廣泛受到採納。有效整合不同來源的資料能使大數據分析的效益倍增，但若資料中涉及個人資料，多項資料來源的串連將帶來更大的隱私威脅。個人資料法規的進展使得過去數年來資料去識別化的議題廣受重視，而聯盟式學習則提供了另一種毋須先聚合資料，即可共同訓練機器學習模型的可能。本文介紹各國個人資料保護法規中對個人資料的定義如何影響資料共享，並探討資料去識別化及聯盟式學習是否足以協助資料處理者符合現行規範。研究發現，若以合乎目前各國相關規範為目的，去識別化及聯盟式學習均為可行之技術選項，資料處理者可依使用情境挑選兩者中較為適合之機制使用之。

關鍵詞：個人資料保護、資料隱私、去識別化、聯盟式學習

From Data Anonymization to Federated Learning: Dealing with Privacy Issues in Interorganzational Data Analysis

Szu-Chuang Li

Assistant Professor at Department of Information and Communication, Tamkang

Yennun Huang

Research Center for Information Technology Innovation, Academia Sinica.

Yi-Wen Chen

Associate Professor at Department of Information and Communication, Tamkang

Abstract

With the rapid development of data mining techniques recently, the analysis of "big data" has been adopted widely for academic and practical purposes. The aggregation of various data sources could significantly improve the effectiveness of big data analysis. But it could also be a threat to data privacy, as data linkage enables an adversary to identify a specific person with more confidence. Several countries updated their personal data protection regulations to cope with the challenge of data privacy, putting data anonymization in the spotlight during the last few years. Federated learning

provides yet another way to enable organizations to train machine learning cooperatively without data aggregation. In this study, we explore data privacy's legislative and technical facets to clarify the compatibility of data de-identification and federated learning with current regulations and found that both techniques are compatible if performed correctly. Therefore, data processors should choose the most suitable technology on a case-by-case basis, depending on the data's actual usage.

Keywords: personal data protection; data privacy; data de-identification; federated learning

壹、研究背景與目的

近年來各式資料探勘技術進展迅速，尤其在深度學習技術興起後，不論是影片、影像、聲音、文字等各類型資料，都能夠有效地被運用於建立自動化預測模型。輔以電腦硬體在運算能力以及儲存容量上的持續進展，大數據分析不論在實務上或學術上，都開始廣泛收到採納。在實務上，大數據可望改變目前企業高層的角色，由過去依賴多年經驗累積的直覺進行決策，轉為在由人工智慧模型所推論得出的各項可行方案中選擇最適合公司當前情境者[1]。在學術研究方面，各領域之研究者也開始探索利用大數據進行各項研究之可行性，並已肯定其在探索性研究的實用性[2]。也因此，持有巨量資料的政府單位以及大型企業，無不全力探索各種運用大數據的可能性，以求發揮資料的最大效益，創造價值。

有效整合不同來源的資料能使大數據分析的效益倍增。近年來政府大力推行的政府開放資料平台，除了希望盡可能釋出政府所擁有的資料資源供一般民眾擷取、分析以探索新的應用方向之外，也有利於民間資料擁有者將相關政府資料與自行持有資料相連結，以求整合出資料屬性欄位更為豐富、資料筆數更大的資料集，以進行更深入的統計分析作業，並訓練出準確率更高的預測模型。以空氣品質監測為例，不同的民間組織或政府單位可能在都市中的不同地點佈建了針對不同類型氣體微粒設置的感測器。

---

1 Andrew McAfee and Erik Brynjolfsson , “Big Data: The Management Revolution,” *Harvard Business Review*, Vol. 90, No. 10, pp. 60-68.

2 Rob Kitchin, “Big Data, New Epistemologies and Paradigm Shifts,” *Big Data & Society*, Vol. 1, No. 1, pp. 1-12.

如果各種不同來源的資料能夠透過某些技術方案進行整合分析，則可以對都市中的各式污染源的相關性進行比單一資料來源更深入的研究。

然而，不論是政府開放資料平台公開資料的作為，或是機構間進行資料分享的合作，如果資料內容包含個人資料，則在現有國內外的隱私規範下，除非經過使用者明示同意，或將資料進行預處理，使之符合法規規範，機構才能進行資料的釋出或分享。資料隱私是現代人的重要權益，而過去層出不窮的資料外洩事件，以及資料在未經使用者同意之下被不當運用的種種案例，也使得相關法規日趨嚴格。包括台灣之個人資料保護法、歐盟的一般資料保護指令 (General Data Protection Act，GDPR)，均不允許機構在未經使用者同意的情況下，將使用者提供的原始資料分享予其他機構。以 GDPR 為例，其第八十三條即規範如有觸犯的企業將會面臨嚴厲的處罰，最高可達全球總年營業額的百分之四或兩千萬歐元。

種種保護使用者隱私的設計，使涉及個人資料的綜合運用十分困難。依據現行國內外資訊隱私法規的規範，資料如能去除個別個人被直接識別或間接識別之資訊，使其中的個人資料不再有被識別之虞，即可進行資料交換。唯在過去的研究[3]中我們可以得知，單純的遮蔽個人敏感欄位並不足以完全去除被識別的可能，而需要進一步的設計。另一方面，在資料去識別化的過程必然涉及原始資料的更動，更動完的資料是否能在保障個人隱私的前提

[3] Arvind Narayanan and Vitaly Shmatikov, "Robust De-anonymization of Large Sparse Datasets," paper presented at *2008 IEEE Symposium on Security and Privacy*, Oakland. pp. 111-125.

下提供有意義的統計分析結果，亦是一大取捨。隨著技術的進展，學者亦開始思考是否有可能在不提供原始資料的情況下，進行跨組織的資料共同分析，例如本地差分隱私[4]，以及聯盟式學習 (federated learning; 或中譯為聯合式學習、聯邦式學習)[5]。

本文首先將從當前各國的法規及國家討論出發，以理解在資料含有個人資料的情況下，能在合乎各國規定的前提下進行跨組織資料分析的技術可行途徑，再對各個可能的技術途徑進行介紹與優缺點分析及比較，以說明在現行的法律架構下，不同的跨組織分析需求可以採取如何的技術途徑來達成，以兼顧資料分析的需求以及個人資料保護規範的要求。表格型資料仍為目前組織間合作進行資料分析之主要資料形式，因此本文之討論將以表格型資料之討論為主。在研究資源的限制下，暫不討論多媒體資料如影像、影片，以及聲音資料。

貳、個人資料的保護及運用限制

各先進國家針對個人資料保護事宜立法，均已經有相當的歷史。以我國為例，最後一次針對個人資料保護法的修法完成於民國 104 年。我國的個人資料保護法第二條第一款定義個人資料為：「指自然人之姓名、出生年月日、國民身分證統一編號、護照號

[4] Úlfar Erlingsson, Vasyl Pihur and Aleksandra Korolova, “RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response,” paper presented at 2014 ACM SIGSAC Conference on Computer and Communications Security, pp. 1054-1067.

[5] Tian Li, Anit Kumar Sahu, Ameet Talwalkar and Virginia Smith, “Federated Learning: Challenges, Methods, and Future Directions,” *IEEE Signal Processing Magazine*, Vol. 37, pp. 50-60.

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」。第五條則規定：「個人資料之蒐集、處理或利用，應尊重當事人之權益，依誠實及信用方法為之，不得逾越特定目的之必要範圍，並應與蒐集之目的具有正當合理之關聯。」關於公務機關以及非公務機關對個人資料的蒐集、處理、應用，則分別於第二章及第三章明定之。影響所及，目前各公務機關以及較具規模的非公務機關，在收集個人資料時均會備妥相關聲明供使用者同意，表明相關資料僅將使用於特定用途，以避免相關爭議。

而政府開放資料或企業間的跨組織資料交換行為，皆屬於個人資料保護法中定義的「特定目的以外之運用」。除了一些法律中規定的特定情況，如法律明定之外，多需要另外徵詢提供資料的個人同意後方能為之。但大數據分析所需要的大規模數據整併及後續分析，如果跟每個機構合作都需要一一取得當事人的同意，則不僅需要耗費大量的成本，且能否取得足夠多使用者的同意，恐怕也有相當的難度。而從國際競爭以及國家安全的角度而言，如果大數據及人工智慧的發展因為相關法規的限制，而使其發展速度遠遠落後於個人資料保護觀念薄弱的國家，恐也不符社會的最大利益。因此，不論是政府或非政府機關，均在保障個人隱私的前提下，持續尋找洽當的資料聚合方式，以兼顧隱私保障與科技發展的需求。過去十年來針對此部份之技術解決方案，學界大多數朝資料去識別化的方向進行研討，也有一定的進展。而近五年來聯盟式學習技術的興起，則提供了跨組織資料共同分析

的另一種可能性。以下將分別就相關技術所能滿足的個人資料保護規範，以及技術本身的潛力及限制進行探討。

參、將個人資料轉化為非個人資料：資料去識別化

一、各國法律對「去識別化資料可視為非個人資料」之認可

如前所述，我國個人資料保護法明定其適用範圍為個人資料，且在定義中已經明定，個人資料為可直接或間接識別個人的資料。換言之，如果原始資料包含個人資訊，但已經透過適當的處理，使之不能直接或間接識別個人，則經過處理後的資料可視為非個人資料，而不需要受到個人資料保護法的約束。另個人資料保護法在第十六條第五款以及第二十條第五款中，亦有「公務機關或學術研究機構基於公共利益為統計或學術研究而有必要，且資料經過提供者處理後或經蒐集者依其揭露方式無從識別特定之當事人。」之規定，呼應去識別化後的資料，即可視同非個人資料。以上可知，資料如果能夠達到去識別化，在我國個人資料保護法的架構下，是可以視為非個人資料而進行揭露或交換的。

歐盟 GDPR 對於個人資料的認定，在前言第二十六條 (Rectical 26) 有較詳細的說明[6]。歐盟的用詞為「擬匿名化」以及「匿名化」，其中「擬匿名化」係指直接識別欄位經過編碼，但在需要時仍可以逆向解開編碼而識別個人的匿名化方法。擬匿名化的資料如果在增添外部資料比對的情況下可以識別出個人，則

[6] 歐盟，<一般資料保護規則>《Complete guide to GDPR compliance》<https://gdpr.eu/recital-26-not-applicable-to-anonymous-data/?cn-reloaded=1>(2021 年 6 月 8 日查詢)。

仍需要視同個人資料加以保護，此部份規範於前言第二十六條第二款。而匿名化資料係指已經無從識別個人，且已經無法逆向還原的資料，則非 GDPR 的保護範圍，此部份規範為前言第二十六條第五及第六款。此部份之定義則與我國個人資料保護法對個人資料的定義雷同。

而與我國較為鄰近的日本，也在 2016 年修正其「個人情報保護法」後，於 2017 年正式開始施行去識別化資料的相關規定，稱為「於非識別加工資訊」[7]。至 2018 年末，有 380 家企業依法進行資料的去識別化，將資料販售給外部研究單位進行分析，或與其他單位進行合作分析，各事業單位對於可以在降低洩漏客戶隱私的情況下進行資料的進一步運用，均表滿意。法規規定相關事業必須設立窗口接受有疑慮者之申訴，惟截至 2018 年底，尚無透過去識別化進行資料再利用之業者收到申訴。從我國、歐盟及日本的個人資料保護法規觀之，資料如果經過處理後已經不能再識別出特定的個人，即可免於適用個人資料保護法規，應是目前各先進國家的共識。

二、間接識別資料認定問題

就目前各國法規的定義，直接識別資料部分較無疑義。諸如姓名、身分證字號、手機號碼、電子郵件、個人照片、甚至網路上的暱稱等資訊，無論是公務機關或是非公務機關均已具有相當的敏感度，在處理資料時多以謹慎的態度保密之。間接識別資料的部份，我國的個人資訊保護法施行細則第三條中規定係指保有

[7] 范姜真媺 、周逸濱，《日本個人資料保護相關法制之匿名（非識別）加工研究委託研究計畫結案報告》（台北：國家發展委員會，2019 年）。

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的情況。以下我們以經典的 Netflix 案例說明，在去除或遮蔽直接識別資訊的情況下，有意侵犯個人隱私者，透過與其他公開或非公開的資料進行交叉比對，仍有許多手法可以辨識出資料中獨特的個人。

Netflix 於 2007 年發起一項名為 Netflix Prize 的資料科學競賽，並準備了大量的使用者租片資訊供參賽者分析。競賽所提供的資料僅有四個欄位，分別是使用者編號、電影編號、使用者給予電影的評分、以及評分給予的日期，不含任何直接識別資訊，而一般人也無從從使用者編號中識別出參賽者的身分。參賽者的任務為利用 Netflix 所提供的資料，預測特定訂戶對某個影片的評分為何。若其預測準確率能高於原本 Netflix 自有演算法的準確率，就有機會獲獎。由於釋出的資料欄位有限，且使用者以及電影均僅提供編號，Netflix 認定此資料不含有隱私資訊，因此用於公開競賽之用。德州大學研究生 Arvind Narayanan 及學者 Vitaly Shmatiko 針對此資料進行研究[8]，證明即使在資料有限的情況下，透過與其他線上電影社群資料之評分資料進行比對，仍可透過比對評分的紀錄在一定程度上辨識出用戶及其所觀看的影片為何。由於在電影社群中有的用戶會揭露的自己的真實身份，這樣的辨識工作在大量進行累積後，足以建立電影編號與實際代表的電影的對照表，使得特定人觀看特定電影的情況曝光。Netflix 在 2007-2009 進行了三次第一代競賽，但由於美國有規範錄影帶出

[8] Arvind Narayanan and Vitaly Shmatikov, “Robust De-anonymization of Large Sparse Datasets,” paper presented at 2008 IEEE Symposium on Security and Privacy, Oakland. pp. 111-125.

租紀錄保密的相關法律，Netflix 也遭到使用者的提告，並以和解收場。在 2010 年 Netflix 試圖推出第二代的競賽活動，但同樣因為有無法解決的隱私疑慮，最後遭到取消。

由以上的案例我們可以得知，由於技術的進展以及各機構相關資料管理狀況的動態變化，在釋出或分享資料當時認為資料並無被間接識別之虞，並不保證在未來也可以確保資料沒有被間接識別的可能。但若因任何已被處理過的資料未來均有被再識別的風險，而否定所有去識別化資料的應用，將大幅減少大數據技術應用之空間。關於此部分，目前各國的文獻均認為去識別化不可能絕對毫無風險，僅能檢視其是否已排除「合理可能」識別特定個人之程度。因此，資料保有者須評估所釋出之去識別化資料有無合理之可能性[9]。如果資料與「其他資訊」結合後足以識別個人，即應視為具備間接識別資訊而需要納入個人資料保護的範圍，因此如何定義「其他資訊」為其中關鍵。根據英國、歐盟、及澳洲採行之「有心侵入者測試」（the motivated intruder test），其假設窺探隱私者缺乏先前知識，在進行隱私窺探時僅能查詢各項線上或線下的公開資料，透過與去識別化資料的串連來識別個人。而這樣的認定方式，應該也是目前較為適切定義間接識別資料的的方式。意即，個人資料保護法規在認定間接去識別化資料時，主要是判斷其與公開可取得之資料串連後是否可辨識出個人，而不考慮與一些在法規或倫理上已經要求專業人員需要妥善保存的資料的串連例如稅務人員可接觸之稅務資料，或醫療

[9] 林裕嘉，〈公務機關利用去識別化資料之風險評估及法律責任（下）〉，《司法週刊》，1853 期。

人員可接觸之醫療資料。

此外，在窺探者嘗試再識別資料中的個人時，可能採取不同的技術手法。不同的技術手法在推論個人的身分時，其信心程度也不同。在某些情況下窺探者可以完全確定再識別的個人為何，在其他情況下則只能做出不同機率把握的猜測。理論上，即使窺探者只是做出任意的猜測，仍有猜中之可能。如果有此情況發生，是否能夠被視為違法個人資料保護法規？目前英國的資訊特任官辦公室 (UK's Information Commissioner's Office, ICO) 的資料匿名化的實務指南 (Anonymisation: managing data protection risk code of practice) 中[10]，較明確的認定資料的再識別必須有一定的確定性。該指南的第二十六頁之「識別以及有所本的猜測」(Identification and the educated guess) 乙節可試翻譯如下：「資料保護法規所重視者為足以識別個人的資訊，亦即此資訊要能在相當水準的確定性下區辨個人。所謂識別並不僅是基於可得資訊進行猜測，因為猜測可能是不準確的。這樣的猜測有可能會會帶來隱私風險，但如果資料已經經過洽當的去識別化而使得個人資料並未為猜測者所知，就並不屬於資料保護的議題。即使猜測者基於匿名化資料識別出某些個人，也並不代表個人資料遭到洩漏。」

可知，如果僅是有所本的猜測 (educated guess)，雖然亦可能侵犯到特定個人的隱私，但卻非因為個人資料的外洩而造成，因此不應在個人資料保護的範圍內處理之。此一說法對後續判斷各類去識別化機制在個人資料保護上的效用有相當的意涵，將於技

[10] UK's Information Commissioner's Office(ICO), 2012. "Anonymisation: managing data protection risk code of practice", <https://ico.org.uk/media/for-organisations/documents/1061/anonymisation-code.pdf>.

術探討時進行。

三、資料去識別化相關技術

在資料去識別化技術中，目前為資料科學學界探討較多、發展較成熟者，為 k 匿名 (k-anonymity) 以及非互動式差分隱私 (non-interactive differential privacy)，以下將分別對技術進行說明並探討其與現行個人資料保護法規之間的關聯性。

（一）k 匿名 (k-anonymity)

k 匿名一詞首見於 Samarati and Sweeney 於 1998 年的研究[11]，其研究動機源自於美國法律對投票人名單須公開的規定。窺探者可利用這份公開的名單中的資料欄位，與其他已遮蔽識別欄位的資料表對照，就未遮蔽的欄位進行比對，進而推知已遮蔽識別欄位的資料是屬於哪個特定個人所有[12]。

k 匿名主要運用於表格型態之資料，以罹病資訊為例，若有一表格含有身高、體重、性別，以及特定病症的罹病資訊，有意竊取敏感性罹病資訊者若因種種因素持有特定病人身分證字號以及身高、體重、性別資訊，則可以利用這些資訊的交集，比對出特定病人的罹病與否。舉例而言，假設患病檢查資料表中只有一筆病患身高為為 150cm、體重為 64kg 且性別為女性，而窺探隱

[11] Pierangela Samarati and Latanya Sweeney, "Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression," *Harvard Data Privacy Lab.* < https://epic.org/privacy/reidentification/Samarati_Sweeney_paper.pdf>.

[12] Latanya Sweeney, "K-Anonymity: A Model For Protecting Privacy," *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, Vol 10, No.5, pp. 557-570.

私者又可以確知某位具有這些屬性的女性必定在其中的話，窺探者就可以據此推論該女性之患病狀況。要避免這種情況發生，k 匿名機制會檢視各種欄位組合的獨特性，以確認一份資料受到此類攻擊的機會為何。所謂「k」值，即為在相同的欄位組合下，可以找到多少相同的資料。當 k 越大時，窺探隱私者就越難進行準確的猜測。但另一方面，即使 k 的數值夠大，但若 k 位受檢者的患病狀況均相同，窺探者仍可確知該名女性已經罹病。欲解決此問題後續有學者對 k- 匿名進行調整[13]，增加檢查在 k 名病患中患病的狀態是否有差異，有差異才能確保其隱私性。

k 值的大小係由資料處理者挑選確認。進行 k 匿名檢測時，若發現敏感屬性交集資料數量有低於 k，或數量高於 k 但敏感欄位狀態均一致時，通常有兩種處理方式，一種稱為一般化 (generalization)，一種稱為遮蔽 (suppression)：

1. 一般化係指將某些欄位進行「概化」，例如將數字改成數字範圍，或是將類別型欄位以更上位的概念來涵蓋，例如將音樂家與畫家合併為藝術家。
2. 遮蔽係指將某些欄位加以遮蔽，過去文獻上均以填入「*」來處理之。使用這樣的處理方式基本上已是完全放棄該欄位所蘊含的資訊。

上述的兩種處理方式在概念上並不難理解，但實作上卻有其困難之處。當我們辨別出一些不符 k 匿名要求的資料時，可以有

[13] Machanavajjhala, Ashwin, Daniel Kifer, Johannes Gehrke and Muthuramakrishnan Venkitasubramaniam, “L-diversity: privacy beyond k-anonymity,” paper presented at 22nd International Conference on Data Engineering (ICDE'06), Atlanta, GA, USA. pp. 24-24.

多種方法進行一般化及遮蔽，都可以使整個資料集符合 k 匿名的標準，但哪一種能夠保留最佳的統計效果則難以確定。雖然在實施上仍有許多問題有克服，但 k 匿名的理論基礎較容易為社會大眾所了解，也經常被用來對一般人說明資料如何進行匿名化，可說是除了資料遮蔽外較容易對社會大眾說明的去識別化方法。在資訊及通信國家標準技術委員會的「個人資料去識別化驗證標準規範說明[14]」中，即提及 k 匿名技術。值得注意的是，在進行 k 匿名處理時需要先區別直接識別欄位、間接識別欄位以及敏感欄位後，才能進行各項資料去識別化程序，而此部份各項欄位的定義在相當程度上須依賴資料處理者的主觀判斷。隨著時空的不同，人們對於敏感資料的認知有所改變，此時資料處理者可能會需要調整相關欄位的認定，重新以 k 匿名化程序進行資料處理，盡可能兼顧資料可用性及資料隱私。

(二) 非互動式差分隱私 (non-interactive differential privacy)

過去往往認為，如果不以與原始資料類似的形式釋出資料，而僅僅釋出如平均值、特定數值組合總數等統計值，就不會洩漏個人資料。但在一些特定的情況下，手上持有相當背景資料的窺探者可以從統計數值中，判斷特定人是否包含在該資料集中，進而推得敏感資訊。例如有一個資料庫可供查詢特定地區的收入平均數，使用者只要輸入某市、某區，系統即會回應該區的收入平均數。此時有一名窺探者，手上也擁有一份一週前的資料，亦知

[14] 資訊及通信國家標準技術委員會，〈個人資料去識別化驗證標準規範說明〉，《經濟部標準檢驗局》<https://www.bsmi.gov.tw/wSite/public/Data/f1447142556627.pdf>（2022 年 1 月 16 查詢）。

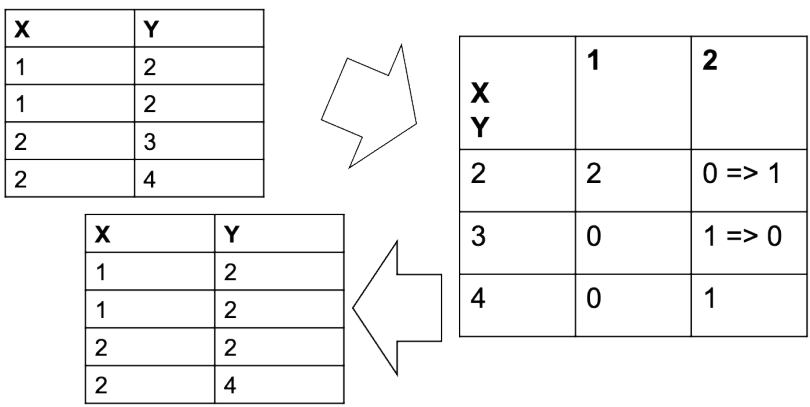
道社區中有一名人員將遷入，此時窺探者若想知道該人的收入，則可向該系統查詢最新的平均值，再與手上資料所計算的平均值進行運算，即可計算出該人員的收入。

誠然，由於窺探者顯有特殊的資料來源管道，顯示原本該保護該資料不外流者有未盡責任的情形，才造成該人員收入資料外洩的風險。因此上述可供查詢的資料庫的設置，依據英國 ICO 的提示，未必有違反個人資料保護法規的問題存在。但這樣的情形也顯示，即使只是統計資訊，在特定情境下，仍可能具有間接識別的可能，而造成個人權益受損。為了避免上述的問題發生，學者 Dwork 提出差分隱私 (differential privacy) 的觀念[15]，指出上述的系統在回覆統計數值時，首先必須在回覆值上加上經過計算的雜訊數值，目標為使新住戶加入前與加入後資料庫系統的回應數值相似，以使窺探者難以推估新住戶的收入。顯而易見地，這樣的設計會使回覆的統計數值偏誤，而影響到統計結果的可用性。但因 Dwork 的提議具有嚴謹的數學機率基礎，對隱私洩漏的定義又極其嚴苛而幾乎足以涵蓋所有的隱私外洩情況，因此目前廣獲接受。

上述所述的資料庫查詢情境，稱為互動式差分隱私。為了因應開放資料以及資料交換等需要提供資料表方可滿足的使用情境，資料管理者可運用差分隱私的原理來產出合成資料。其原理為設計一連串的資料庫查詢，以符合差分隱私要求的資料查詢對

[15] Cynthia Dwork, “Differential Privacy” in Bugliesi M., Preneel B., Sassone V., Wegener I. eds., Automata, Languages and Programming. International Colloquium on Automata, Languages, and Programming 2006. Lecture Notes in Computer Science, Vol 4052.

原始資料取得答案後，再以這一連串的答案重新組成資料表，以供使用。理想上，以這樣的方式產出的資料表，具備差分隱私的隱私保證，在統計上應該也會跟資料庫中的原始資料相差不遠。以 DPTable[16] 為例，係以「次數查詢」為主軸對資料庫進行重複的查詢後取得各種屬性組合在資料庫中出現的次數，再利用這些次數資訊重構資料庫，其過程可簡化後圖示如圖一。



圖一：以資料庫中的計數查詢結果產出合成資料

(資料來源：李思壯、黃彥男，2019)

位於左上角的表格即原始資料集。右側黃色的列聯表的內容，是各種 (X,Y) 組合的數量。如果 X 或 Y 有其他可能的數值，如 (1,3)，只是剛好在原始資料集中未出現的話，在黃色列聯表中的 X 軸也需要有 3 的這個可能數值，並在列聯表的對應欄位上填上計數為 0。由於黃色列聯表中的每個格子都是一個對原始資料

[16] Rui Chen, Qian Xiao, Yu Zhang and Jianliang Xu, "Differentially Private High-Dimensional Data Publication via Sampling-Based Inference," paper presented at 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. pp. 129-138.

表的次數查詢 (counting query)，因此黃色列聯表其實就是原始資料表的一種呈現形式，而這些結果正好類比於互動式差分隱私的資料庫查詢。因此，我們需要在黃色列聯表的各種 (X,Y) 的次數查詢結果上依照差分隱私原則加入雜訊，使得列聯表內的次數發生改變，再依據黃色列聯表的新次數產生一張新的合成資料表，就可以得到一個以差分隱私為基礎所產生的合成資料表。

由上述的描述可知，非互動式差分隱私機制與 k 匿名機制十分不同。由於資料的產生是基於加入過雜訊的統計數值，即使資料表中產生與真實個人相似的資料，窺探者也無法完全確定該資料屬於個人。由於差分隱私所產出的資料表係由機率過程產出，窺探者僅可以從表中猜測特定個人的特定屬性數值，但永遠無法確認，也因此依此原則所產出的資料集，根據英國 ICO 之標準，似可相容於現行個人資料保護規範之需求。但值得特別討論的是，差分隱私具備一個參數 epsilon，可用於調整加入雜訊的大小。加入之雜訊小時，所產出的資料偏移量較小，反之則資料偏移量較大。前者將使窺探者的猜中機率提高，但資料的統計精確性較佳。後者將使窺探者的猜中機率降低，但資料的統計精確性亦隨之下降。若資料處理者為了資料精確度而選擇一很大的 epsilon 值，導致差分隱私所輸出的統計數值有很高的機率貼近真實數值，使其具有極強的間接識別能力，是否仍應如 ICO 目前建議，因其機率機制的特性而鬆綁其此類資料之分享？有待進一步討論。

肆、另闢蹊徑：聯盟式學習

資料去識別化在過去很長一段時間內，被認為是可能兼顧個人資料保護以及跨組織資料分享需求的唯一方法，但近期的聯盟式學習技術則開啟了另一個可能性，可以在不違反個人資料保護法規的情況下，讓多個組織得以共同訓練人工智慧模型。以下討論聯盟式學習的技術機制，以及其與個人資料保護法規之契合性。

一、聯盟式學習技術概要

聯盟式學習係指在進行機器學習之模型訓練時，不需要先將訓練資料聚集，而只需要先由各方約定好欲採用之人工智慧模型架構，由各方各自在自有節點上進行模型訓練，將訓練完成之參數彙報至一共同的中央節點進行彙總運算後，再將綜合各方節點結果的模型參數佈達給各個節點，此模型的效果將優於各方自行訓練的模型。而此過程也可以再次啟動，讓各方能夠基於彼此的資料進一步提昇模型效果。聯盟式學習的提出始於 Google 在開發輸入法時，希望能夠運用散佈世界各處數以千萬計之使用者的輸入，來讓其他使用者能夠透過輸入文字之預測，來讓文字輸入工作更快速的完成[17]。 由於聯盟式學習最先被使用於多設備共同機器學習的情境，研究者後將跨組織資料分析的情境命名為 cross-silo，多設備協同訓練模型的情景稱為 cross-device，並探討

[17] Timothy Yang, Galen Andrew, Hubert Eichner, Haicheng Sun, Wei Li, Nicholas Kong, Daniel Ramage and Françoise Beaufays, "Applied Federated Learning: Improving Google Keyboard Query Suggestions," ArXiv, abs/1812.02903.

兩者的異同。在 cross-device 的情境下，各組織的頻寬及運算資源充裕，許多在 coss-silo 情境需要考慮的限制可以放鬆，而可以專注在資料隱私及可用性問題的解決。隨著聯盟式學習的逐步發展，聯盟式學習逐步發展出三個子類別：分別為水平式聯盟學習、垂直式聯盟學習，以及轉移式聯盟學習[18]。

1. 水平式聯盟學習 (horizontal federated learning)：係指共同進行機器學習模型訓練的各方所持有之資料欄位均為一致，例如均有居住地、年齡、性別等屬性。
2. 垂直式聯盟學習 (vertical federated learning)：係指共同進行機器學習模型訓練之各方所持有之資料欄位不一致，但資料來源至少有部分重疊之部分。例如某一方擁有特定顧客的姓名、年齡資訊，另一方則擁有特定顧客的姓名、購買資訊。在此情境下各方所持有之資料其來源必須有所重疊，以個人資料為例，必須至少有部分來自相同的個人。

[18] Peter Kairouz, H. B. McMahan, Brendan Avent, Aurélien Bellet, Mehdi Bennis, Arjun Nitin Bhagoji, Keith Bonawitz, Zachary B. Charles, Graham Cormode, Rachel Cummings, Rafael G. L. D'Oliveira, Salim Y. El Rouayheb, David Evans, Josh Gardner, Zachary Garrett, Adrià Gascón, Badih Ghazi, Phillip B. Gibbons, Marco Gruteser, Zaïd Harchaoui, Chaoyang He, Lie He, Zhouyuan Huo, Ben Hutchinson, Justin Hsu, Martin Jaggi, Tara Javidi, Gauri Joshi, Mikhail Khodak, Jakub Konecný, Aleksandra Korolova, Farinaz Koushanfar, Oluwasanmi Koyejo, Tancrède Lepoint, Yang Liu, Prateek Mittal, Mehryar Mohri, Richard Nock, Ayfer Özgür, R. Pagh, Mariana Raykova, Hang Qi, Daniel Ramage, Ramesh Raskar, Dawn Xiaodong Song, Weikang Song, Sebastian U. Stich, Ziteng Sun, Ananda Theertha Suresh, Florian Tramèr, Praneeth Vepakomma, Jianyu Wang, Li Xiong, Zheng Xu, Qiang Yang, Felix X. Yu, Han Yu and Sen Zhao, "Advances and Open Problems in Federated Learning," ArXiv, abs/1912.04977.

3. 轉移式聯盟學習 (federated transfer learning)：係指將某種機器學習模型訓練的成果運用至其他機器學習模型訓練的工作。例如運用某個青少年群組歷時多年的討論記錄訓練詞彙向量模型，再運用於另一個青少年群組的近期討論資料分群分析。

簡而言之，聯盟式學習讓組織之間可以合作進行機器學習模型訓練，卻不需要交換原始資料，在技術上具有節省傳輸頻寬、分散運算負擔、互相運用機器學習成果等優勢。而由於其不必事先匯聚資料的特性，也被視為解決資料隱私問題的可行途徑之一。

二、聯盟式學習與個人資料保護法規

由前段的技術探討可以得知，進行典型的聯盟式學習時，各個參與聯盟式學習的節點之間，並不需要聚集原始資料，而是共享一個人工智慧模型的訓練架構，先依各自所持有的資料對人工智慧模型進行參數的更新，只將參數送往中央伺服器進行彙整。可知在聯盟式學習的架構下，並不存在透露直接識別資料的問題。而這些參數是否屬於間接識別資訊則是聯盟式學習能否滿足法規需求的要點。與我們先前在非互動式差分隱私的討論雷同，模型參數與一般的資料統計量，均為對資料進行數學運算所獲得之結果，而非一般資料去識別化情境中，提供經過處理的資料欄位的情境。這些統合計算後的數據仍有可能洩漏隱私資訊，但洩漏隱私資料的情境與將來自不同來源的表格式資料近交叉比對十分不同。

以差分隱私定義的攻擊情境為例，窺探隱私者如果想測試某一筆特定的資料是否存在於某一個參與聯盟式學習的節點中，必

須先擁有較該節點多一筆或少一筆之完整原始資料，並確知該聯盟式學習機制的全貌，最後還須有效取得其送往中央伺服器的完整資料進行比對，才能推論某一筆資料存在或不存在於該資料集中。根據前文中所提到的「有心侵入者測試」，除非上述各項資訊均屬一般人能夠便利取得資訊，否則上述資料節點傳遞予中央彙整伺服器之資訊，應不會被認定為屬於間接識別資訊。亦即，聯盟式學習的資料傳遞過程，應不會被認定為涉及傳輸直接或間接識別資訊，而觸犯個人資料保護規範。

另一方面，雖然依「有心侵入者測試」，聯盟式學習的技術程序應不致於被認定違反個人資料保護規範，但如果窺探隱私者確實持有上述的資料，則實際上是可能對針對特定個人的資料進行猜測的。從聯盟式學習的架構來看，能夠取得模型的參數資料，除了有能力攔截參數資料的窺探者，另一個角色就是負責中央聚合分析之中央伺服器。如果要在此兩者之一擁有單一節點完整原始資料的情況下，阻止其推論某單一個人是否存在於資料集中，則須進一步在參數傳遞的過程中增加資料保護機制，例如追加差分隱私雜訊[19]或其他機制，以徹底杜絕被窺探的可能。而相關的措施必須以運算量的增加以及模型效果的降低為代價。

伍、結論及未來研究方向

隨著機器學習及深度學習等相關技術的廣獲認可，不論在實

[19] Kang Wei, Jun Li, Ming Ding, Chuan Ma, Howard Hua Yang, Farokhi Farhad, Shi Jin, Tony Q. S. Quek and H. Vincent Poor, "Federated Learning With Differential Privacy: Algorithms and Performance Analysis," *IEEE Transactions on Information Forensics and Security*, Vol.15, pp. 3454-3469.

務以及學術上，大數據分析都已經佔有舉足輕重的地位。望文生義，大數據分析的基本要件為巨量數據的彙整及聚合。在民主社會中，個人資料的保護重要性亦不遑多讓，但若囿於個人資料保護的需求而遲遲無法發揮大數據分析的效益，恐怕將使國家競爭力難以與擁有巨型企業甚至是獨裁體制的國家抗衡，而落後日多。如何提供解決方案兼顧大數據分析以及個人資料保護之需求，實為重中之重。

去識別化資料為過去十年間因應開放資料以及資料交換需求而生的技術，目前其中 k 匿名方法的思考邏輯廣被規範制定者所採用，在許多實施標準中運用其概念。非互動式差分隱私技術所產出的合成資料，由於其機率生成的特性，與現有的個人資料保護規範亦應相容。在開放資料的應用場合，此二種去識別化技術均能滿足個人資料保護法規的要求，避免在資料開放或交換時，洩漏使用者的直接或間接識別資訊。而資訊科學及相關領域學者也持續在此部份進行努力，避免資料因隱私處理而對統計有效性影響過大。聯盟式學習的技術發展之初並非因為隱私保護的需求，而是為了廣泛整合使用者終端設備在日常生活中所產出的使用資訊進行模型訓練所創造。由於聯盟式學習不需要事先聚合來自各節點的資料，而僅需要由中央伺服器收集各節點之模型參數進行處理，再將彙整計算完成之參數佈達到各個節點，其中固無直接識別資料的傳遞。而模型參數的傳遞按目前一般之認定，亦未達間接識別個人的標準，應可確認聯盟式學習機制目前足以符合一般個人資料保護法規之要求。

去識別化技術與聯盟式學習各有其適用的情境。例如在政府開放資料或企業開放資料以徵求社會大眾參與分析工作的場合，亦或是組織間希望進行共同分析的方法聯盟式學習機制尚未開發完成時，去識別化技術仍是唯一的技術選擇。若是組織之間希望彼此合作增加數據量，以共同訓練機器學習模型時，則可以使用聯盟式學習，以完全避開設計資料交換機制時可能面臨的法規問題。其中亦有若干情境是兩種技術均適用者，此時因為去識別化技術與聯盟式學習均可符合個人隱私規範，則可以兩種機制同步進行跨組織資料分析後，比較兩者之統計有效性、預測準確率等後，取其中較為滿意之成果應用之，應可獲得在當前的技術及法規環境下，相對令人滿意之成果。

本文研析現有各國之個人資料保護規範，並探討了資料去識別化技術以及聯盟式學習技術是否合乎當前各國之規範。惟目前已知有下述議題尚待探討，可在後續研究接續之。首先，以差分隱私技術所產出的合成資料，隨參數設定之不同，可以在一定程度內控制合成資料與原始資料之偏移程度，影響到窺探者的猜中機率。但在英國 ICO 的認定中，目前認可若能達到窺探者無法確知資料正確性的目標，即可視為非個人資料而自由運用。此部份機制是否有其他國家有不同之規定？又需不需要補強，以強化對個人資料之保護，值得未來接續研究。再者，以差分隱私為基礎的機制在本文中尚有亂數回應機制 [20] 因篇幅限制未及討論，因其

[20] Úlfar Erlingsson, Aleksandra Korolova and Vasyl Pihur, “RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response,” paper presented at 2014 ACM SIGSAC Conference on Computer and Communications Security, pp. 1054-1067.

亦屬於不需要事先匯聚原始資料即可進行聯合分析之機制，在未來研究中將與聯盟式學習進行比較討論。最後，因研究資源的限制，本文的研究範圍以跨組織資料分析時以運用最為廣泛的表格型資料為限，惟差分隱私以及聯邦式學習均可以運用於多媒體資料之處理，將在未來研究資源許可時進行進一步研究探討。

誌謝

本研究成果由經濟部科技專案補助計畫（計畫名稱：主動式資安情資與智能偵防技術計畫，計畫編號：110-EC-17-A-21-1702）、科技部計畫 MOST109-2221-E-001-019-MY3，以及中央研究院計劃 AS-KPQ-109-DSTCP 經費補助，特此致謝。（投稿：2021 年 5 月 3；第一次修訂：2021 年 6 月 15 日；第二次修訂：2021 年 12 月 20 日；接受：2021 年 12 月 26 日）

參考資料

一、中文部分

（一）期刊論文

李思壯、黃彥男，2019。〈數位時代之數位隱私保護〉，《國土及公共治理季刊》，第 7 卷第 4 期，頁 30-39 。

林裕嘉，2017。〈公務機關利用去識別化資料之風險評估及法律責任（上）〉，《司法週刊》，1852 期。

林裕嘉，2017。〈公務機關利用去識別化資料之風險評估及法律責任（下）〉，《司法週刊》，1853 期。

（二）研究計畫

范姜真媺、周逸濱 ，2019。《日本個人資料保護相關法制之匿名（非識別）加工研究委託研究計畫結案報告》。台北：國家發展委員會。

二、英文部分

（一）專書、專書篇章

Dwork, Cynthia, 2006. “Differential Privacy” in Bugliesi M., Preneel B., Sassone V., Wegener I. eds., *Automata, Languages and Programming. International Colloquium on Automata, Languages, and Programming 2006.* Lecture Notes in Computer Science, Vol. 4052. Venice, Italy.

（二）期刊論文

Kitchin, Rob, 2014. “Big Data, new epistemologies and paradigm shifts,” *Big Data & Society*, Vol 1, No. 1, pp. 1-12.

Li, Tian, Anit Kumar Sahu, Ameet Talwalkar and Virginia Smith, 2020. “Federated Learning: Challenges, Methods, and Future Directions,” *IEEE Signal Processing Magazine*, Vol. 37, pp. 50-60.

McAfee, Andrew and Erik Brynjolfsson, 2012. “Big data: the management revolution,” *Harvard Business Review*, Vol. 90, No. 10, pp. 60-68.

Sweeney, Latanya, 2002. “k-anonymity: a model for protecting privacy,” *International Journal on Uncertainty, Fuzziness and Knowledge-based Systems*, Vol. 10, No.5, pp. 557-570.

Wei, Kang, Jun Li, Ming Ding, Chuan Ma, Howard H. Yang, Farokhi Farhad, Shi Jin, Tony Q. S. Quek and H. Vincent Poor, 2020. “Federated Learning With Differential Privacy: Algorithms and Performance Analysis,” *IEEE Transactions on Information Forensics and Security*, Vol.15, pp. 3454-3469.

（三）研討會論文

Chen, Rui, Qian Xiao, Yu Zhang and Jianliang Xu, 2015. “Differentially Private High-Dimensional Data Publication via Sampling-Based Inference,” paper presented at *21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. pp. 129-138.

Ashwin Machanavajjhala, Daniel Kifer, Johannes Gehrke and Muthuramakrishnan Venkitasubramaniam, 2006. “L-diversity: privacy beyond k-anonymity,” paper presented at *22nd International Conference on Data Engineering (ICDE'06)*, Atlanta, GA, USA. pp. 24-24.

Erlingsson, Úlfar, Vasyl Pihur and Aleksandra Korolova, 2014. “RAPPOR: Randomized Aggregatable Privacy-Preserving Ordinal Response,” paper presented at *2014 ACM SIGSAC Conference on Computer and Communications Security*, pp. 1054-1067.

Narayanan Arvind, and Vitaly Shmatikov, 2008. “Robust De-anonymization of Large Sparse Datasets,” paper presented at *2008 IEEE Symposium on Security and Privacy*, Oakland. pp. 111-125.

（四）網際網路

Pierangela Samaratiy and Latanya Sweeney, 1998. “Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression,” *Harvard Data Privacy Lab*, <https://epic.org/privacy/reidentification/Samarati_Sweeney_paper.pdf>.

UK’s Information Commissioner’s Office(ICO), 2012. “Anonymisation: Managing Data Protection Risk Code of Practice,” <https://ico.org.uk/media/for-organisations/documents/1061/anonymisation-code.pdf>.

Kairouz, Peter, H. B. McMahan, Brendan Avent, Aurélien Bellet, Mehdi Bennis, Arjun Nitin Bhagoji, Keith Bonawitz, Zach-

ary B. Charles, Graham Cormode, Rachel Cummings, Rafael G. L. D'Oliveira, Salim Y. El Rouayheb, David Evans, Josh Gardner, Zachary Garrett, Adrià Gascón, Badih Ghazi, Phillip B. Gibbons, Marco Gruteser, Zaïd Harchaoui, Chaoyang He, Lie He, Zhouyuan Huo, Ben Hutchinson, Justin Hsu, Martin Jaggi, Tara Javidi, Gauri Joshi, Mikhail Khodak, Jakub Konecný, Aleksandra Korolova, Farinaz Koushanfar, Oluwasanmi Koyejo, Tancrède Lepoint, Yang Liu, Prateek Mittal, Mehryar Mohri, Richard Nock, Ayfer Özgür, R. Pagh, Mariana Raykova, Hang Qi, Daniel Ramage, Ramesh Raskar, Dawn Xiaodong Song, Weikang Song, Sebastian U. Stich, Ziteng Sun, Ananda Theertha Suresh, Florian Tramèr, Praneeth Vepakomma, Jianyu Wang, Li Xiong, Zheng Xu, Qiang Yang, Felix X. Yu, Han Yu and Sen Zhao, 2019. "Advances and Open Problems in Federated Learning," *ArXiv*, abs/1912.04977.

Timothy Yang, Galen Andrew, Hubert Eichner, Haicheng Sun, Wei Li, Nicholas Kong, Daniel Ramage and Françoise Beaufays, 2018. "Applied federated learning: improving google keyboard query suggestions," *ArXiv*, abs/1812.02903.