



莊庭瑞  
中央研究院資訊科學研究所副研究員，  
合聘於資訊科技創新研究中心  
以及人文社會科學研究中心。

# 個人資料一釋出難收回

永不遺失的資料讓人心安，卻也帶來另一種困擾。

撰文／莊庭瑞

現代人最怕掉手機。通訊錄、記事本、照片影音記錄等個人資訊以及綁定的行動支付工具，都在行動電話裡。一旦察覺手機不在身邊，心就慌了；再想到手機裡的資訊可能被陌生人窺見，感覺更是不安，只想趕緊找回手機。許多資料雖然可以同步備份在雲端，但要下載到新手機裡畢竟麻煩。而且，存在雲端的個人資料，真的安全嗎？存在雲端的資料不會遺失，這可是資訊服務業者盡力確保的事。但如果這些資料也永遠不會消失呢？

數位資料的特質就是可以快速且低成本地複製並傳遞。某些敏感資料被不自主揭露後，即使事主有能力刪除敏感資料，但有心人還是可以在網路上找到。同樣地，儘管某些個人資料先以失去脈絡的方式處理，才對外流出、供他人使用，但倘若一般大眾知道這個實情，相信也會成為一大精神負擔。

各國的個人資料保護法規都強調，必須在當事人知情同意之下，才能進行個人資料的蒐集與處理，而且排除某些情況，資料僅能用於特定目的。不過，眾人為了使用線上地圖導航，或是方便交友圈的相濡以沫與即時聯繫，卻同意交付行蹤與通訊內容給 Google、臉書或 LINE 這些巨型的資訊服務業者。一方面，若不交付這些個人動態資料就不能使用這些服務；另一方面，因為使用者都同意服務業者的使用條款，業者也就可以持續保有並使用眾人源源不絕產生的個人動態資料。

眾人動態行為產生的巨量資料，衍生了許多商機，例如即時的道路擁塞資訊（公路上眾人的手機都在哪裡）；我們與親友的對談內容，提點了業者可銷售的商品（談到暑假，套裝旅遊廣告即刻上場）。如果這些個人動態資料只供業者內部使用，似乎還不令人太擔心。

不過，有必要永久保存這些個人一時的行蹤與對話資訊嗎？而眾人活動歷程累積的大量資料，經「匿名化」（anonymized）處理後，是否就可放心釋出、供其他目的的使用？

匿名化通常指去除或是模糊資料集裡可識別個人身分的資訊，例如電子信箱、姓名、生日，或是僅保留出生年，沒有日期資訊。然而，匿名化後的資料集通常無法提供完善性的保證。比對其他來源的多項資料集，或是經由個人行為，往往可以在號稱已匿名化的資料集裡，「再次識別」（re-identify）出特定個人的資料。

眾人的活動歷程在匯整並匿名化之後，就算達到無法再次識別特定個人的理想境界，但因為描繪了群眾活動的整體樣貌，即使沒有點出資料集裡的誰是誰，仍舊引發隱私外洩的議論。今年1月29日，運動社群公司 Strava 把 2000 多萬名使用者數年來的運動軌跡經匿名化處理後，得到含有三兆個位置座標的運動路線資料集，並繪製成詳盡的地圖，線上提供給眾人使用。

這地圖揭露了美國一些秘密軍事基地的地點，顯然有軍事駐點人員把跑步軌跡上傳到 Strava 資料平台，暴露了基地的位置與範圍。我們或許可以推測，這些上傳者於註冊使用 Strava 服務時，同意（或是不知道自己已經同意）Strava 把上傳的軌跡資料經匿名化後另做他用。運動軌跡資料一旦釋出，使用者就難以收回；有些上傳者應該十分懊惱，一起跑步但沒使用 Strava 服務的同伴，發覺基地位置遭自己人暴露出去，心中應該也五味雜陳。

在與網路形影不離的生活中，眾人在各類資訊平台上分享自身動態，但你可否留意，你的個人動態資料正如何被使用？[SA](#)

就算眾人活動歷程達到  
無法再次識別個人的  
理想境界，但仍舊引發  
隱私外洩的議論。