由「刷臉」進校園事件，談生物特徵的個資保護

大學講師 ㄧ 魯明德

摘要

報載近期某市議員接獲民眾陳情，指出某女中宿舍購置人臉辨識系統，有侵犯隱私權的疑慮。此新聞讓人聯想到，是否可妥善運用資訊科技來深化門禁管理，卻又不會侵害到隱私權的兩全其美做法。

生物特徵和個人資料

生物特徵是某個人特有的生理（Physical）或行為（Behavioral）特徵。生理特徵包含指紋、掌紋、掌型、虹膜、面容、聲紋及 DNA 等，行為特徵則有走路姿勢、心跳及簽名筆跡等。由於生物特徵通常具有獨特、不易改變的特性，因此被廣泛用於個人辨識系統，如門禁管理、上下班打卡等.....。

根據報紙報導，新北市某國中小在 108 年就推行刷臉入校，以臉部辨識的方式辨識學生進入校園，學生們直呼「好潮」、「上學更新鮮」；然中部某女中在 109 年 9 月購置人臉辨識系統後，卻引發家長對隱私權的擔憂。

在資訊界工作的小潘看到這些新聞後，思考著個人隱私跟資訊科技有沒有可能取得平衡？於是在每月一次的師生會上，就立刻提出他的疑問。司馬特老師聽完了這個大哉問之後，喝了口咖啡，緩緩回應小潘：這個問題可分成二個層面來看，一個是生物特徵、一個是個人資料。

在《個人生物特徵識別資料蒐集管理及運用辦法》第 2 條中，定義生物特徵識別資料是「指具個人專屬性而足以辨識個別身分之指紋及臉部特徵資料」。而在《個人資料保護法》第 2 條中，定義個人資料是「指自然人之姓名、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」。

由這個定義可以看出來生物特徵主要指的是指紋及臉部特徵，而個人資料則是除了正面表列的項目外，還包含其他得以直接或間接方式識別該個人之資料，廣義來看，臉部特徵未在條文中列出，但仍然屬於可以識別出個人的資料。

去識別化（De-identification）：個人隱私跟資訊科技之平衡

小潘聽到這裡，馬上想到一個問題：有沒有方法不要識別出個人，又能有效的做門禁管理？司馬特老師想了想，喝口咖啡接著說，當我們存放在資料庫裡的資料不是人臉照片或指紋，再加上把資料去識別化（De-identification）後，就可以做到門禁管理又不會侵害隱私了。

小潘聽得一頭霧水，資料庫裡不放人臉照片，要怎麼比對呢？司馬特老師拿出 2 張圖來說明，每個人的臉部或指紋上的特徵都不一樣，我們可以取得的是這些特徵點的座標及其特徵值（eigenvalue），這是一個多維的資料，再把每個人的這些位置座標與特徵值以演算法做成各自的特徵向量（eigenvector），存放到資料庫，作為日後門禁管理比對的基準。

舉例來說，在進行門禁管理時，若有一天有個小強要進門，人臉辨識系統便會根據小強的這些特徵，分別讀取它的特徵值，做成特徵向量值後，再與資料庫裡的特徵向量做比對；若在資料庫中找到有相同的特徵向量，就表示小強是合法的使用者，可以開門放行；如果資料庫裡沒有相同的特徵向量，則表示小強不是合法的使用者，不會開門讓他進來。又因為門禁的資料庫裡只有特徵向量，並沒有小強的名字，即使看到特徵向量，也沒有辦法辨識出哪一個是小強，因此，就沒有洩漏行蹤的隱私權問題，也沒有個人資料外洩的問題

小潘接著問：那二個特徵向量要如何做比對？司馬特老師喝了口咖啡，繼續說下去，不論是存在於資料庫中的特徵向量，還是要做比對的特徵向量，都存在於多維度的特徵空間（eigenspace），只要把它們做餘弦（cosine）運算，如果 $\cos\theta$ 運算的值為 1，則二個向量就可以視為是相同的向量。

小潘聽完司馬特老師的解說點頭如搗蒜，但是，反應快速的小潘立刻又想到另外一個問題，如果資料庫中只有特徵向量，有一天萬一機房發生問題，要找出誰曾經進去過，豈不是就找不到人了？

司馬特老師非常高興小潘能夠舉一反三，喝口咖啡接著說下去：這就是公司管理的問題囉，在建置員工的臉部或指紋特徵時，一定會有員工的姓名做對照，不然怎麼知道這個特徵向量是誰的，但是，放到門禁管理系統的特徵向量則是經過去識別化的，也就是只有把特徵向量放過去，這樣一旦有一天有異常資料要比對時，自然可以回到內部找出該特徵向量是屬於誰的。

小潘聽完老師一席話恍然大悟，原來資訊科技不只是電腦軟硬體，我們小時候念了半天不知其所以然的向量、三角函數，也有這種用途啊！華燈初上，這次的師生會就在焦糖瑪琪朵的香味中進入尾聲，小潘帶著滿意的答案吹著口哨離開。