

# 建置大學數據資料庫 北醫大奠定全國校務研究基礎

■ 文／施純明·臺北醫學大學主任秘書  
陳錦華·臺北醫學大學校務研究辦公室生物統計中心主任

臺灣校務研究（Institutional Research, IR）始於104學年度教育部補助大學提升校務專業管理能力計畫，在詭譎多變的時代中，期望以數據為本之決策制定，協助學校發現問題，解決問題，確保高教品質保證。實施以來，各校前進的速度不一，但可以確定的是，每一所大學都全力經營校務研究領域，並互相學習。

近兩年來，臺北醫學大學也隨各個參訪團至美國及日本取經，參加各項分享會及研討會，以及邀請外國IR辦公室成員到校分享經驗，加上教育部與臺灣校務研究協會的不吝指導，至今雖還無法趕上美國普渡大學（Purdue University）的腳步，但觀諸包含北醫大在內的各大學校務研究發展，短短兩年已有豐碩成果，則是有目共睹的事實。

## 引領AI風潮 透過多樣化數據創造價值

隨著科技產業的進步，近十年來在資料量急速成長、資料型態的多樣性、儲存資料成本下降、雲端概念成熟、各項軟體之技術提升下，已帶動大數據風潮及AI人工智慧（Artificial Intelligence）的崛起。

臺北醫學大學十分重視大數據及AI的發展，北醫大及附設醫院、萬芳醫院、雙和醫院及臺北癌症中心等附屬醫院引領風潮，率先引進人工智能

癌症治療輔助系統（IBM Watson for Oncology），作為癌症治療診斷的輔助工具，為醫療帶來另一波新的革命。但在這些AI的背後，其實蘊含著多樣化的數據，唯有一再反覆訓練與檢視，才能創造AI真正的價值。

因此，要發展校務研究，必須利用適當的數據予以分析，臺北醫學大學遂將建置資料庫作為校務研究的基石，以創造校務研究之價值。本文前段以資料庫建置經驗為例分享經驗，後段針對未來校務研究之發展，以個人拙見提出看法。

## 可供教學研究使用的校務研究資料庫

臺北醫學大學之校務研究資料庫涵蓋「教學研究資料庫」及「臨床資料庫」，前者為校務相關事務之資料庫，後者則為三家附屬醫院之臨床資料庫。資料庫之建置以支持學校決策、提升教學及研究能量、增進社會福祉為目的。

在此前提下，資料庫可供教學研究使用，北醫大於104學年度已制定「校務研究資料庫管理辦法」及「校務研究資料庫申請使用要點」，規範資料庫之管理、申請、使用規則，保護資料安全性，以去連結之方式提供使用，並設置「臺北醫學大學資料分析中心」，申請者使用資料庫時，必須於分析中心使用，分析結果只能攜出經統計後之匯

整數據，無法攜出具單筆之資料格式，須經審查後才能將資料攜出。這些嚴格的規定，是為保護資料安全及個人隱私，也是資料庫使用之基礎。

## 資料匯入

「教學研究資料庫」之建置涉及校內各處室人員，工程浩大，參與人員除了校務研究辦公室同仁外，還有賴各處室之種子人員、資訊處、生物統計中心等各司其職，在各級主管支持及各單位同仁協助下，才能完成此任務。資料入庫前有六大步驟，以確保資料是可分析及具正確性，並定期定時予以更新，以利後續分析使用。

- **步驟一：**各單位先規劃資料檔案，了解資料欄位的定義，及何謂資料字典。
- **步驟二：**資料檔案從各系統分別匯出資料（此部分資料可能來自平台匯出，也可能包含手邊維護資料），供各單位種子人員進行資料清理（資料清理若涉及複雜之資料，須利用程式協助者，由生物統計中心予以協助）。
- **步驟三：**根據原始資料，將不必要的欄位進行修正或是新增欄位，同時對應資料字典的撰寫原則，根據原則及資料型態進行資料字典的撰寫。
- **步驟四：**各單位種子人員完成前三項步驟之後將資料交給資訊系統組，進行資料匯入資料庫前之驗證工作（資訊處協助），包含格式、欄位之正確性。
- **步驟五：**待驗證工作完成，將資料匯入資料庫後，由學校管理發展中心負責管理（此步驟可於視覺化平台呈現資料，以進行資料正確性驗證）。
- **步驟六：**以上五個步驟皆完成後，資料進入



▲臺北醫學大學資料分析中心。（臺北醫學大學提供）

資料倉儲，未來以此資料進行議題分析及視覺化查詢平台建置重要指標，並可開放申請研究使用。

## 資料清理

在資料匯入的過程中，很多同仁反映，資料既已存在資料庫中，直接利用這些資料進行分析即可。這個建議雖好，但在未了解資料品質、資料型態、資料內容前，仍需要有所堅持。

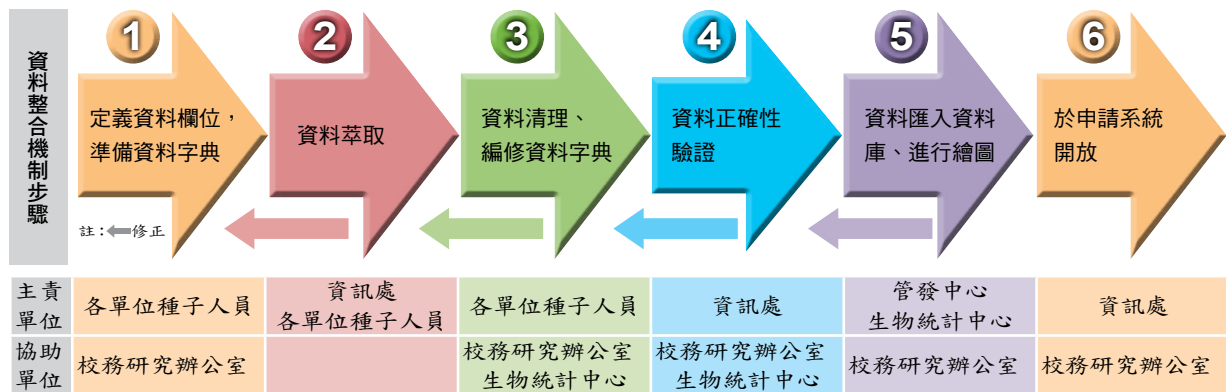
例一：資料庫中關於學生訊息，包括父母的電話或家中地址，並不是需要分析的欄位，因此應予以刪除。地址部分則會保留地區別。

例二：由於系統及政策變化，處室維護資料時，常有「備註欄位」，此類型的欄位可利用程式以規則性原則予以清理，通常可清理達90%，特殊部分再請同仁補齊。

例三：在分類時，有多種名稱代表同一類。例如畢業高中欄位，北一女中、北一女、臺北市立第一女子高級中學、一女中等名稱必須統一，這也是資料清理的重要部分。

例四：系統資料有時會出現重覆的筆數，須將重覆資料識別出來並予以刪除，以免重覆計算。

以上例子說明資料分析前之資料品質的重要性，若不在意品質，難免會出現「垃圾進、垃圾



\*資料檔串接方式

學生面：學號、e-mail帳號、身分證號（依教育部要求）等欄位

教師面：員工編號、e-mail帳號等欄位

圖一 資料整合之機制

出」（garbage in, garbage out）。為避免因提供錯誤資訊而導致錯誤決策，北醫大已在近兩年針對資料進行深入探索，接下來將進入「資料收集2.0版」，以減輕同仁負擔。

### 資料分析與應用

至2017年7月，共有15單位參與資料匯入，包含83個資料表，資料皆進入資料倉儲。利用此資料庫，北醫大每年將針對重要議題進行分析，以學生學習成效為主，並產生報告及建議；另外也建置視覺化查詢平台系統，包含學校重要指標、議題模組資料，以使用者為出發，設計連動式圖形呈現，作為資料查詢使用。

#### ● 視覺化查詢平台系統

此查詢平台系統之建置目的有二，一是驗證資料正確性，確保資料品質；另一則是建置查詢平台，供教學及行政面使用，以達學校資訊公開化。

資料收集後，為驗證資料的正確性，將以視覺化平台系統予以驗證。可將資料匯入，利用雲端軟體，繪出各類數據，並以圖型呈現，觀察資料之可能出錯之處，並和資料所屬處室討論，儘可

能發現數據之問題，來回修正數據，以確保資料品質。

另外，此平台之功能可開放校內學生、教職員、長官依不同權限使用，目前除了重要指標如註冊率、在校人數等之外，也針對校務研究議題所使用之資料，建立主題式之查詢系統，在校務研究應用上更為方便，以達資訊開放流通、重要資料可公開使用的目的。

#### ● 議題分析主要目的為確保學生學習成效，善盡大學社會責任

資料庫建置完善將利於執行議題分析，北醫大分析之議題目前著重於招生策略變化、提升學生學習成效、弱勢生之協助與關懷。以弱勢生之分析為例，將各類弱勢生身分別予以區分，並探討學習成效、社團參與等變數和身分別之相關性，以及獎助學金措施之效益。

臺北醫學大學善盡落實社會責任，在參與社會關懷活動、社區營造、偏鄉服務與國際醫療面上不遺餘力。在獎學金挹注方面，每名學生於學雜費減免外所獲得之金額，自101學年度到104學年度為止提升40.8%，從51,770元增加至72,907元，補助比例由101學年度的81%，成長至88%，大幅

改善學生之經濟狀況，使學生得以安心就學，此即是由本研究結果得到的驗證。

上述各項作業讓筆者深感校務研究工作之繁瑣，須有賴各類專業人員協助才得以完成。在此特別感謝北醫大參與校務研究之所有同仁，有他們的參與，才有今日累積的成果。

### 校與校之資料交換機制

校務研究之目的主要在協助校務治理，如何將這些分析及應用成果轉換為校務績效報告，以落實校務研究之目的，必須先了解績效報告呈報之重點。臺灣可借鏡美國的現況，除了著重於學校治理基本面外，亦需呈現學校對社會公共責任落實之程度，涉及層面十分廣泛。

因此，若要充實績效報告內容，單就校內建置之資料庫是不夠的，因校內資料庫內容只能呈現學校目前樣貌，在議題解釋或社會責任之落實上，仍須輔以其他數據，例如：問卷調查資料、高中學習歷程、大學端學習歷程、勞動部資料等，才能完備績效報告之內涵。

尤其特別的是在社會責任落實部分，並非只看績效指標達成率，而是對社會所產生之實質貢獻與助益，以增進國家進步為目的。例如：美國大學之績效報告中，社會責任部分呈現增進當地就業率之百分比，以就業率提升將造就社會之安定並促進社會發展，顯示出這項指標之重要性；由於此部分之議題呈現十分多元化，各類資料庫間之串連就顯得十分重要。目前在臺灣，為充實績效報告書之材料，讓校務研究發揮實質效用，可

由短期、長期予以思考。


#### ● 短期之發展

1. 建置跨校資料整合平台，以完整學生之學習歷程，增加議題分析解釋之深度。在美國已有「全國學生資料庫」（National Student Clearinghouse, NSC）的非營利組織，制定資料交換之規範，以協助落實校與校之資料交換機制，可進一步了解學生在大學間之學習及流動。

2. 思考共同評估、衡量工具之制定與發展，以減輕各校IR辦公室人員之工作負擔，例如：以評量指標（Rubric）方式作為評估工具，評量學生學習成績；調查問卷之開發，除了省去問卷開發之時間，調查結束後，亦能互相比較各校之差異。

#### ● 長期之發展

政府部分，在資料使用上應持開放態度，此態度間接影響未來校務研究及高教深耕計畫之發展，若能進行各部會資料間之串連，將更深入地了解大學治理之相關問題，對於國家政策面也有實質之助益。在大數據及AI風潮下，數據開放與倫理規範間如何平衡等議題，應在促使社會進步下全面重新思考，若仍以舊思維處理新事務，將無法有效解決問題，更何況政府要以發展AI為主題，資料如何釋出使用、管理機制，將是無法逃避之問題，政府相關單位應予以正視，制定資料使用辦法，才是合理途徑。

未來臺灣發展校務研究，勢將乘著大數據及AI的風潮持續發展，站在這路口，您已經準備好了嗎？想必各大學皆已擬定未來的地圖，奮力向前邁進！

#### ◎ 參考資料

李欣宜(2015年4月)。一次搞懂大數據(上)。數位時代，251。取自<https://www.bnext.com.tw/article/35807/bn-2015-03-31-151014-36>

傅遠智(2017)。借鏡美國經驗對下一階段我國發展校務研究的啟發。評鑑雙月刊，69，31-35。

臺北醫學大學(2017)。105學年度教育部補助「大學提升校務專業管理能力計畫」期中成果報告。臺北：臺北醫學大學。