

展望下一個十年的校務研究

■ 文／傅遠智・國立臺北科技大學人文與社會科學學院副院長、臺灣校務研究專業協會副秘書長
張育慈・國立臺北科技大學技術及職業教育研究所碩士生

我國高等教育將校務研究納入校務運作的一環迄今已近十年，無論是校務評鑑、招生專業化，以及教育部推動的各項競爭型計畫，都能看到校務研究的身影。歷經十年的發展，各校也已逐步將校務研究融入作為校務決策的重要組成部分。展望下一個十年，臺灣校務研究仍然有許多未解的挑戰，不僅需要校務研究領域同仁的努力，更需要所有高等教育行政主管凝聚共識，才能在數據時代，讓校務研究為校務決策者提供更專業的服務，本文即著眼於此提出推動方向的建議。

如何取得優質的數據這一個議題仍舊困擾著

各校校務研究單位，教育部大學校院校務資料庫已經維運了近二十年，這一個數據資料庫是以各校依據表單定義自行填報的方式所組成，對於教育部以及外部利害關係人了解各校的辦學情形是一項重要的管道。該資料庫對於我國高等教育數據建設有一定的貢獻，也確實部分解決了高等教育數據收集與儲存的問題。但是，真正能夠讓學校用來發現本身的問題並進行自我改善，目前校務資料庫的資料是遠遠不足的。

首先，這種加總性數據填報的方式，令各校在應對每年例行的填報作業上，仍是由校內各職能單位各自填報加總性數據方式進行，內部橫向數據未能串接，校內仍存在著數據孤島的問題。其次，學校內部受到各種申請作業期程的關係，例如學生辦理休退學等，數據本身其實

是動態更新。校內僅有加總性數據的結果，往往也造成數據正確性難以核對的問題。其

三，大學校院校務資料庫經過數年的累積，數據表單本身要填報的各項欄位眾多，但

到底有多少是能直接反映大學本身辦學成效也是值得商榷。

最後，要促進學校自我改善，除了內部的資料外，外部的資料，包括學生生源組



成、學生畢業流向、甚至大學研究對於產業的影響，這些能反映各大學服務對象的組成以及經由高等教育後對於社會產生的影響，才是校務治理真正需要知道的反饋訊息。

針對上述的問題，過去十年來各校在全國性大學校長會議或是教務會議中都曾多次呼籲教育部協助各校取得畢業生流向資料以供學校了解並改善教學之用。事實上，教育部也曾多次試圖建立一個能夠串接個人數據資料的教育數據處理中心。雖然至今並未實現，但過程中經過許多教育先進的摸索與努力，教育體系需要一個數據處理中心來提升教育政策與校務治理的品質的想法獲得越來越多人的支持。

從大學校務治理層面來看，大學端最迫切需要取得的數據無非是三種，第一是學生生源數據，第二是學生畢業就業數據，第三則是大學研究表現數據。首先，在學生生源數據方面，由於招生專業化的關係，各校近年執行很多生源分析的工作，也都取得不俗的成果，但囿於學校取得的數據往往是學生

入學後在校填報的數據，學生在考試階段，例如甄選入學或聯合登記分發時選填志願的過程，甚至更往前為學生在高中階段專業科目的成績等，都難以為各校校務研究單位取得。面對少子化的威脅，大學的校務治理層往往需要知道校內各個科系是屬於全國性招生，或者是地區性招生，例如熱門的校系學生往往來自全國各地，而非熱門的校系往往生源會比較集中在部分的地區。在確認本身校系招生的地位後，便需要評估哪些校系在市場中是競爭相同生源的校系，而這一個部分便能直接從分析學生在選填志願的過程中獲得。即便目前各校沒有這方面的數據，也有數據分析技術比較好的校務研究單位藉由交叉查榜的方式去發現各個校系的競爭者，但總得來說，能做到這個部分學校並不普遍。

若是從科技大學端來看，由於與技術型高中群科對應比較緊密，如果校系本身是屬於地區性招生，其實很容易就可以了解正在和多少相同屬性的大學校系競爭多大規模的高中生源。例如位於南部地區且屬於地區性招生的科技大學，所招收的學生就很有可能也是位於南部地區的技術型高中學生。我國整體學齡人口雖然逐年減少，但地區間人口消長的情形卻有極大的差異，例如新竹縣市周圍便出現既有的高中無法容納學齡人口的情況，即便是同一個縣市，例如桃園，也在國小階段出現舊城區學校生源不足，但新的都市規劃區學校容量不足的情況。這些分析的結果能有助於各大學正確評估本身的招生規模，從另一個角度想，面對學齡人口正在增加的地區，也能及早因應。當前的校務研究往往僅能去看自己內部的情形，



但事實上，全國性產業聚落板塊以及都市重劃帶動的人口移動，都正在改變著大學校園外部的環境，但包括教育主管機關以及各大學校務研究單位都還未能充分利用生源數據，來省視全國教育資源配置的合理性。

其次是畢業生就業資料，在過去十年，教育部能與勞動部串接學生的就業資料並提供各校以校系為單位的學生平均就業薪資。但即便是在例行的校務評鑑，各校還是習慣於使用自行透過校友電訪調查而來的畢業生就業資料。這種自行調查而來的畢業生就業資料，其正確性以及完整性存在很明顯的問題。各校如果能取得學生的投保紀錄，究竟能為大學校務治理帶來哪些有助於回饋辦學的訊息呢？

首先，一個人的勞保投保紀錄會包含以下幾個基本欄位：投保人、投保時間、投保狀態、投保金額、公司統一編號以及公司所在縣市等。公司統一編號還可以進一步串接經濟部公司行號註冊資料，以獲得公司有哪些營業登記項目、資本額以及公司規模等。所以對串接後的勞保投保紀錄進行數據清理後，可以幫助校務研究人員了解校友畢業後在各個時間點的投保薪資、就業天數，工作地點以及所投入的行業，但我們無法從數據中得到的就是該名校友在該公司實際從事的職位。即便是如此，我們已經獲得許多重要的訊息，例如，某一個科系校友畢業後就業狀態是否穩定，薪資水準是否高於平均水準，薪資水準是否受到地區性的影響，畢業校友進入到某一個行業的人數等。

這些細部性的資料不僅有助於大學掌握畢業校友的就業情形，藉由了解學生就業的目的地以及行業的特性，在課程以及人才培育上做比較及時的調整，從教育主管機關的角度，在人力資源培育以及媒合上也能進行整體的評估。舉例來說，

隨著臺灣在半導體產業進入另一個飛躍性發展的階段，許多縣市紛紛設立科學園區來因應半導體產業的擴張，但新設立的科學園區往往設在土地尚未充分開發的都市邊緣，而周圍即使有大學，大學科系本身也是在1990年代左右設立，與當前的周邊產業結構已出現脫節。當前教育部對於新設系所的審查，仍舊是由各校在總量管制下提交個案審查的方式進行，另外就是配合國家政策對於特定領域的系所進行擴張或縮減。但無論是個案審查或是基於國家政策，都鮮少有當地產業人力需求的事實性數據作為布局的基礎。

最後是大學研究表現數據，大學研究表現數據最核心也最為人所熟悉的就是大學教師的論文發表數據，專業術語稱為書目資料。事實上，書目資料並不僅能呈現某位教師在某一年發表多少論文這種簡單的訊息，更包括但不限於反映了一位教師的研究方向以及科學網絡。以一所大學來說，學校教師所有發表的書目資料反映了這所大學的研究方向，以及在這個領域是否具備技術優勢，這對於一個國家的技術創新生態系統來說是一個非常寶貴的訊息，例如國內的高科技產業想要尋求技術突破的合作夥伴，這就是一個非常重要的指引。近年來受惠於數據科學的發展，機器閱讀的技術使得我們不僅僅知道學術論文之間的相互引用的關係，我們也能知道學術論文被專利引用的路徑，接著再串接每一個專利證號所登載的技術領域分類碼以及專利所有權人等資訊後，甚至對於每一個專利所有權人在專利登記資料中，我們還能進一步確認專利所有權人所屬的國籍。在這些資料的整合下，我們已經能知道大學的科學成果是如何促進技術的發展，甚至是哪些國家的公司及技術領域直接受惠於我國大學的研究成果。

國內大學對於研究表現數據的應用普遍還停留在因應大學排名的競爭，作為校內研究獎勵的基礎。即便是教育主管機關，對於大學研究表現數據也鮮少去了解這些大量的資料所蘊含的訊息。目前校務資料庫裡與研究相關所呈現是加總後的論文發表量，獲得政府機關補助的科研經費以及各校獲得的專利數量，然而，這些訊息僅僅是書目資料中所呈現最表層且極有可能被刻意扭曲的訊息。在對於整個科研經費資助、研究發表成果、研究成果對專利形成的影響、專利持有公司，以及公司所在地等資訊充分的掌握下，不僅有助於檢討科研經費分配的有效性，在學校層級也能盤點本身的研究優勢，在國內外尋找可以共同投入技術開發的公司，達到釋放大學研究能量促進產業升級，並增加大學內部經濟收益的目的。

上述這三種數據都有助於教育主管機關以及各大學進行校務策略性的規劃，並使得校務決策的方向更貼近社會與產業發展的脈絡。然而，我國當前校務研究面臨的是如何取得數據以及

是否具備分析這些數據的能力。首先，在如何取得數據上，個人資料保護法的存在是為了讓個人資料在符合公共利益的前提下保障個人資料的使用，並不是阻止個人資料被使用。我們這幾年面臨的矛盾在於我們被期望透過數據分析提供良好的決策品質，但我們往往也被限制使用好的個人層級數據來進行數據分析，就好像期待一個數據分析人員在沒有取得正確數據的條件中獲得一個可靠的發現，無疑是緣木求魚。若是數據治理是一個國家必須走的方向，透過科學的數據才能有效的協調資源的配置，那在個人資料保護法的框架下解決各校取得優質的數據該當是教育主管機關的首要任務，而不能僅僅是一句有違個人資料保護法的疑慮擱置了議案的討論。

事實上，姑且不論中央教育主管機關，各縣市政府教育主管機關乃至於各級學校，由於長年不重視也不投入資源，對於轄下學生學籍資料及段考成績資料等基本的資料都欠缺維護。這種不重視數據基礎建設的結果，會使得各項教育規劃流於空談，更無法以科學的角度去檢視後續的發展。而形塑出務實不務虛的高等教育治理文化，恰恰好是未來十年身處於高等教育激烈轉型過程中的各個學校都必須面對的問題。

回到校務研究來看，各校直接取得學生個人資料確實也存在一定的風險，這個風險來自於各校校務研究在資料保存以及數據分析的能力上都存在相當顯著的差異。借鑑健保資料庫加值中心的經驗，成立一個教育數據加值中心是一個可行的方案。各校依照意願提供學籍資料後，由教育主



管機關與其他數據源串接後，將原始資料統一存放在這個教育數據加值中心。各校派員至指定的教育數據加值中心處理學校本身的數據資料，並將分析結果攜出。若沒有意願提供學生學籍資料的學校，便按照原有的數據取得方式進行校內的校務研究。

擁有好的數據也要有足夠的能力對於龐大的數據進行分析才能發揮數據真正的價值。當前並非各校校務研究人員都能處理這種堆疊式的行政資料，校內與這種堆疊式的行政資料結構比較相近的是學生的修課紀錄，亦即一個學生在一個學期會有多筆選課資料。但是包括學生在甄選入學階段的多個志願選填，學生進入就業市場後不同時間的投保紀錄，大學教師在不同時間段研究發表紀錄對於不同時間專利的影響等，這些都需要校務研究人員經過相當程度的數據清理訓練，才能從流水帳式的數據結構中萃取出可以被理解的重要訊息。當前社會科學領域的統計課程從未涉及這一個部分的訓練，學生在學習時往往也是處理已經被清理後可以投入分析模型的數據結構，事實上，在真實世界裡並不存在這種已經被清理好的數據結構。即便是許多以大數據為名的數據分析課程，也鮮少有機會讓學生有機會去處理，除了取得這方面資料本身即具有難度外，更重要的是在處理這方面的數據資料還需要對該特定領域有足夠的專業背景才能在數據清理過程中有清楚的方向。

舉勞保投保資料為例，在勞動經濟學領域裡用來衡量一個人在勞動力市場的穩定性，可以以在一定時間的就業天數來測定，若要知道一個人在勞動力市場的價值，可以以一個人在一定時間內的最高月薪來測定等。若要知道大學研究對於技術創新的影響，就必須了解每一個專利均有國際專利分類碼（International Patent Classification,

IPC），而國際專利分類碼有其分類的架構與體系，可以代表該專利是屬於哪一個技術領域。若是處理學生在甄選入學時的數據，就必須清楚知道篩選倍率、面試、志願序以及錄取決定等各個階段的意涵，以及決定的順序。簡言之，若沒有數據清理技術的訓練，以及對於數據本身代表的意涵有深入的認識，即便擁有與學校本身有關的各項數據，校務研究人員也難以做出有意義的分析。

為了使得校務研究能幫助高等教育更好的回應社會的需要，無論是在招生規模與系所設置，人力需求與供給，以及研究成果的社會影響力等，都需要更好的數據以及更專精的數據分析技術。我國校務研究社群可以致力於推動兩個任務，首先，促成教育數據處理中心的設立，各校應與教育主管機關合作，促使與各校有關的學校外部數據能被順利取得並有效利用，共同制定並遵守數據使用規範。其次，校務研究社群可以採集中住宿訓練的方式，邀請有實際處理各項數據經驗的專家，以一週工作坊方式使用共同的程式語言，教導校務研究人員能夠編寫程式語言清理學校自己的數據，並產出一些基本的統計圖表。此外，必須跳脫既有的校務研究議題，邀請專業領域專家，例如勞動經濟學、書目計量、專利智財，甚至是人口區域發展等，介紹該領域專業知識以及新知提供校務研究人員認識與大學相關數據本身的意涵。

校務研究本身就是一項跨領域的學科，而大學本身在社會中扮演一個重要的轉化器，包括促進社會階級流動，將知識轉化成技術，提供周圍產業所需的技術人才等，這些議題不僅與校務發展切身相關，更會決定整個高等教育系統能為國內社會與經濟發展帶來何種程度的貢獻。