

多元視野開啟創新研究

徐茉莉*



因為我的研究生涯有點異於傳統，常從事一些比較特別的研究與教學，所以我很榮幸可以獲得「科技部傑出研究獎」，我也想透過這份殊榮向跨領域研究、國際合作還有不傳統的看法致敬。

我在以色列長大，讀書，服兵役。1994年我從海法大學的心理學系和統計學系畢業 (summa cum laude)。之後我在以色列科技學院學習統計。1997年拿到碩士學位，2000年獲得博士學位。由於我在工程學院讀書跟研究，因此培養出工程的思維模式。2000-2002我跟先生一起去了美國，我在Carnegie Mellon大學的統計系當客座助理教授。在那裡，我跟行銷學系、流行病學系、機器學習系一起合作研究。我們有一個關於疾病爆發早期發現的計畫，用當時不常見的資料，例如網路搜索的關鍵字或是其他資料等，而十年以後Google的研究人員才做出GoogleFluTrends，利用Google的關鍵字搜索在早期發現流感。

* 國立清華大學服務科學研究所特聘教授

2002-2012 年，我在 Maryland 大學的 Robert H Smith 商業學院，先從助理教授開始，最後當上副教授。在那裡，比起以前的同事跟學生，新的環境不同極了。那時我跟一位統計博士 Wolfgang Jank 一起開始了一個全新的研究，用統計學分析網路拍賣與購物，例如 ebay。我們開發了新的分析方法，因為每一樁 ebay 拍賣的下標時間都不一樣，而且時間分布不平均，這種資料的結構非常有特色，不可以用一般的時間序列預測方法。因此，我們使用特別的函數型數據分析 (functional data analysis) 與資料視覺化 (data visualization) 的方法發現新的可能。

我們在 2005 年開始舉辦一個研討會叫做 Statistical Challenges in e-Commerce Research (SCECR)，主題是關於用統計方法分析網路資料，研討會每年定期舉辦，持續到現在。我跟 Jank 教授也一起出版了許多文章跟專書，同時，我開了一堂關於資料探勘的新課程。因為在那個時候沒有適合商學院學生的課本，所以我跟其他專業人士又合作撰寫一本課本 *Data Mining for Business Intelligence*，到現在這本課本仍被許多學校使用。

2008-2009 年我在不丹休假，一邊研究、一邊在不丹最有名的資訊科技學院當義工。我選擇到不丹，是因為那裡非常偏遠，可以有自己思考的時間，不會被打擾。那年寫了我目前最有影響力的文章 “To Explain or To Predict”，我研究這個題目好幾年，也在各地演講很多次。文章內容是關於資料分析。在資料分析時，因果解釋跟預測是兩種非常不一樣的模式，但是大部分的人不清楚他們的不同。我的研究說明這兩個分析過程中不同的地方，從第一步研究設計跟資料收集，到最後一步結果的使用。比如說，在一樁 ebay 拍賣開始時，如果要預測最後的價格，不可以使用關於下標數量的資料，因為這還沒出現。相反的，如果要解釋什麼會影響最後價格，就可以使用之前下標的數量。由於這項研究對各個領域都非常重要，所以現在這篇文章已經被很多人引用了。

中間我一度回到美國，在 2010-2014 年跟我的家人又再次回到不丹，繼續我的志工服務。我跟我先生在學院裡創辦了一個 R&D Lab，舉辦科技和資料分析相關的工作坊，並開發一些科技應用於教學的計畫。舉例來說，我們收集了各種非線上的學習資源，並免費提供給不丹的民眾使用。同時，我也持續研究。那個時候，大數據和商業分析 (Business Analytics) 在亞洲剛開始變得熱門，所以，我決定把學術生涯從美國搬回亞洲，因此辭掉了美國 Maryland 大學的工作，開始在印度商業學院 (Indian School of Business) 當副教授暨特聘教授。在那裡，我除了教書、做研究，也擔任 SRITNE 研究中心主任，負責學院的產學合作並提供研究資金。

2014年我們來到臺灣，我在清華大學服務科學所當特聘教授。為什麼我要在那裡工作呢？雖然這裡的教授各有自己的研究領域（像信息系統、行銷、商業、機器分析等），但是全部都與「人」息息相關。我也非常關心許多公司使用大數據時，只在乎自己的利益，對於其他事情卻毫不關心。我在其他國家教過很多的商學院學生使用資料探勘跟統計方法幫忙解決公司的問題，但是我感覺越來越不舒服，因為這些商業學院的教授跟學生對於商業領域之外的議題毫不在乎，所以我很高興可以在清大服務科學所從事教書與研究！現在，我的研究跟教學都非常注重「人」。我試著說明大數據分析對公司、顧客、人類、社會、環境等的影響。我更出版一篇文章“Research Dilemmas Using Behavioral Big Data”，探討研究人員面臨的各種問題，例如道德困境等。

我將現在的研究命名為“Crossing Modeling Borders”：使用預測模型作因果解釋以及使用因果解釋模型作預測。這個研究的目的是顛覆因果解釋模型與預測模型過去的使用方法，用因果模型作預測，並將預測模型使用在因果解釋上，以發展出適合大數據的方法論。聯合這兩種科學基礎讓我覺得十分興奮，因為這是一個重要而且頗具挑戰的研究題目。

在我二十年的研究裡，我看過很多種不同的研究，在不同的研究團隊中，我學到了四件事。第一，在傳統的領域中，例如統計學、經濟學跟人類行為科學，研究非常嚴謹，可是有時候卻無法適應快速轉變的技術，像現在的大數據。相反的，較新的領域，例如機器學習跟信息系統等能更快地利用新技術，適應不同的研究方法。我覺得這是商業積極與這些創新領域合作的理由之一。因此建議所有研究人員，傳統與創新領域的研究皆需要認識。

第二，即使你的研究非常優秀，在最好的學術刊物出版跨領域的研究也不是一件簡單的事。我建議大家多演講，還有常跟成功的同事分享你的研究。如果你的文章被拒絕，也不要失望！你應該利用刊物的回饋讓你的文章更好，然後再試圖投稿至別的刊物。同時，多多撰寫文章，因為在出版的過程中你可能有多一點機會，所以文章越多越好。

第三，選研究題目的時候，有沒有興趣非常重要。即使比較有經驗的研究前輩告訴你不值得，你仍應該選一個你真正有興趣的題目，因為一個題目值不值得研究還不能太早下定論。而且做研究要花非常多的時間跟努力，如果你對這個研究不夠有興趣，很難持續下去，所以跟隨你的興趣吧！

最後，現在有一些不同的領域也使用大數據做研究。我跟別的研究人員因為有技術背景（比如統計、機器學習等），所以能用強大的分析方法發現有意思的答案，可是我們只關心技術表現，並不著重人類道德原則的研究。相反地，

有些研究人員具有行為科學背景，知道怎麼進行人類研究，但是不知道怎麼使用大數據方法。所以，我建議每位研究人員都應該跟不同領域的人合作，開啟不同的研究視野，才能達到雙贏的局面。

誌謝

特別感謝我的中文老師孔繁好和蘇九如，盡心盡力地幫助我完成這篇文章。