

學術名詞（academic term），或稱作術語（terminology），乃指具有領域學術特點，並可構成該領域概念體系的名詞。對於瞭解領域知識，有著舉足輕重的地位。國家教育研究院雙語詞彙、學術名詞暨辭書資訊網現已收錄 150 多類，超過 150 萬則的學術名詞資訊。然而，目前學術名詞主要以領域相關書籍之索引詞彙作為來源，由於書籍出版需時，新興術語難免有遺珠之憾。所幸，隨著網路時代的來臨，知識的分享媒介也逐漸地從紙本轉換為電子媒體，例如維基百科、數位化的學術論文。此類大量且持續迅速增加的網路文本資料，有助於新興術語之發掘。

考量學術名詞資訊網建置目的即是為了學術名詞的編譯、統一、推廣及發展，因此，本研究選擇國內博碩士論文作為參考來源之一，除可瞭解國內目前高等教育使用學術名詞的情形外，更因論文常需同時提供中英文關鍵詞，若能優先收錄論文關鍵詞，並經審譯提供適當的翻譯，則可讓使用者不致於因缺乏正確翻譯參考來源而隨意翻譯，導致品質參差的名詞翻譯四處流傳。而本研究另一詞彙參考來源則是著名且具有豐富資訊的英文維基百科（Wikipedia）。

本研究採用演算法實作方式，共分為三個階段，以從網路資源中擷取特定資訊並進行相關統計，作為學術名詞新詞收錄參考之用：

階段一、取得相關網路術語資源，包括國家教育研究院「電子計算機名詞」類約 96,000 則名詞、英文維基百科約 11,250,000 筆詞目資料、「電算機學門」之博碩士論文約 136,000 個外文關鍵詞，並擷取所需之術語相關資料，例如詞彙，學科領域，年代等等，並建置成網路術語資料庫。

階段二、採用 Dice 係數作為領域詞彙關聯值，建立博碩士論文關鍵詞之關聯性資訊。

階段三、產生新詞推薦清單並輔以各式參數及統計資訊，如詞彙出現頻率、年代，詞彙關聯性等供參考。

本研究以博碩士論文關鍵詞為基礎，交叉比對本院學術名詞資料庫及維基百科收錄現況。經檢視前 300 名的高頻關鍵詞收錄情形，學術名詞資料庫的收錄比例大約 50%，而維基百科則約 90%。比例落差除因兩資料庫收錄總量懸殊外，本研究所選擇之「電算機學門」與名詞資料庫之「電子計算機名詞」類，兩者收錄之領域範圍雖接近但非完全相同亦是造成名詞資料庫收錄比例較低之原因之一。而觀察關鍵詞使用時期分析可發現，雖同屬高頻關鍵詞，但其使用頻率仍在不同時期，如近十年及五年，有明顯增減，此種詞彙使用消長的情形，有助於收詞優先序之判斷。整體而言，透過本研究所提出的詞頻、使用時期的分析方式，可以歸納出最常用以及值得留意的學術名詞清單，並進而排序作為階段性之新詞收錄參考。

另一方面，本研究透過詞彙相關性分析，可提供基礎詞彙關連資訊。經觀察兩資料庫對於特定詞彙的高度相關詞列表後發現確有明顯的差異。因此，本研究建議若欲收錄較傾向概念、技術詞彙，

可用博碩士論文資料庫的排序為主要參考，若想要較全面性的檢視各種相關詞彙的可能性，則建議用維基百科資料庫為主。

最後，本研究提出三點建議供未來研究發展參考：

一、近期可先運用本研究之既有成果於相關領域之學術名詞審譯過程，中長期則可運用本研究方式至其他學科領域。

二、參考資料庫方面，除了博碩士論文及維基百科仍應持續更新外，也應加入國內外研討會以及期刊論文資料，冀能快速且有效地提供各領域之學術名詞建議收錄列表。

三、可考量對既有的學術名詞資料庫進行加值的分析。例如，藉由觀察國內外各種研究領域的趨勢，並提供熱門領域新詞及相關詞彙網絡資訊，抑或是整合簡稱、別名、另名等資訊，都將有助於研究人員或專家學者的教學及研究。

資料來源

吳鑑城、白明弘、彭佳宣（2017）。學術名詞新詞編譯與科技應用之研究（**I**）子計畫一：學術名詞新詞網路蒐集技術之研究—以資訊領域為例。國家教育研究院整合型研究計畫案成果報告（**NAER-106-12-F-1-01-01-1-01**）。新北市：國家教育研究院。