

大數據分析下的資訊計量學： 以世界大學排名系統為例

文·圖／黃慕萱

資訊計量學的研究由來已久，最早有文獻記錄出現的是1917年Cole與Eales透過統計分析解剖學的出版品，1969年Pritchard以書目計量學（Bibliometrics）一詞定名，並正式定義為「將數學與統計的量化方法應用在圖書與其他形式的傳播媒體上」的一門學問。隨著此種量化分析方法被應用在不同資訊載體上，陸續出現資訊計量學、科學計量學、專利計量學等相關的詞彙。資訊計量學是以代表基礎研究的科研論文，或者以產業技術的專利為研究對象，研究資訊記錄的出版、傳播與使用。近20年來資訊科技的進步，促使各種書目記錄、專利記錄等資料庫收錄範圍越來越廣，引文索引資料庫（Web of Science、Scopus等）功能越來越強大，使得資訊計量學的研究與大數據（Big Data）分析概念不謀而合，都是透過客觀大量的數據，消弭少數人的偏見，發掘當中隱而未顯的意義及趨勢。

在學術環境中，學術評鑑的運用甚廣，小至期刊論文審查、期刊評鑑，以至於教職的聘任、升等、系所評鑑、學校經費補助、學生選校、校際合作對象的選擇等，往往與學術評鑑相關。一般來說，學術評鑑可分為質化與量化兩種方式，常見的質化評鑑為同儕審查（Peer Review），透過相關領域的專家，在制式或非制式的表單上，針對各種指標提出評分及看法，這在期刊論文審查、教職聘任、升等、系所評鑑等較為普遍應用，但此種評鑑方式耗費的時間、金錢較多，速度較慢，而且不論在東西方，均有裙帶關係或性別歧視等問題發生。

量化評鑑的方式包含問卷調查、資訊計量學等方法，在期刊評鑑、世界大學排名等較為普遍使用。好處是能夠以較經濟的方式，取得較大量數據，避免少數人偏見。但問卷調查容易成為聲譽調查，在填寫問卷時，容易傾向勾選知名度較高的選項，而不是因為認為有相當品質。資訊計量學方式的缺點則在於以量採計每一個單位時，通常以均等方式計算，但每一筆科研論文或專利的品質並不一，因此透過多元指標的方式試圖弭平這些問題。

2007年開始的NTU Ranking（世界大學科研論文質量評比計畫），為國內首度自行研發、製作的全球大學排名系統，其影響力不僅在於新興及發展中國家，更深受全球科研管理者與學界矚目。此排名系統單純以資訊計量方法，從科研論文著手，不僅考量科研論文的發表數量，更著重於產出品質，有助於打破一般認為論文數等同研究力的迷思。並以學術生產力、學術影響力與學術卓越性三大面向為評估構面，而指標設計同時兼顧短期與長期，對於新興及發展中國家欲轉型為研究型大學的學術機構，設計一套具體可行的指標。

表1：大學排名中使用的資訊計量指標及權重

大學排名系統	資訊計量指標及權重
NTU Ranking	SCI、SSCI、ESI (100%)
上海交大世界學術排名(ARWU)	N&S- Nature, Science論文數(20%) PUB- SCI、SSCI論文數(20%)
泰晤士報大學排名 THE Rankings	WOS 5年常規化論文平均被引次數(30%) 平均教員/研究員發表論文數(6%) 國際合著文章比(2.5%)
QS Rankings	SCOPUS 5年教師論文平均被引次數(20%)
US News 最佳世界大學排名	學術文章數(12.5%) 引用影響力(10%) Citation(10%) 高被引文章數(12.5%) 高被引文章比(10%) 國際合著文章比(10%)

現有的世界大學排名系統中，或多或少採用資訊計量指標，表1所示，上海交通大學世界大學學術排名（ARWU），資訊計量指標占40%，分別為近5年Nature及Science發表論文數量，以及過去一年SCI與SSCI論文數量，各占20%。英國泰晤士報大學排名（THE Ranking）的資訊計量指標共占38.5%，分別為30%的近5年論文平均被引用次數、6%的教師平均論文數、2.5%的5年國際合著論文比例。英國的QS排名（QS Ranking）的資訊計量指標僅有20%的近5年教師論文平均被引用次數。美國US News最佳世界大學排名的資訊計量指標65%，包含近5年機構論文數、論文被引用次數、常態化論文平均被引用次數、高被引論文數、高被引論文數比例、國際合著論文數比例，分占10%到12.5%。可以發現，資訊計量使用學術論文客觀分析資訊，是各學術排名系統共同採納的。

表2：2014年NTU Ranking的指標及權重

構面	指標	權重	
學術生產力	近11年論文數(2003-2013)	15	25
	當年論文數(2013)	10	
學術影響力	近11年論文被引次數(2003-2013)	15	35
	近2年論文被引次數(2012-2013)	10	
	近11年論文平均被引次數(2003-2013)	10	
學術卓越性	近2年h指數(2012-2013)	10	40
	高被引文章數(2003-2013)	15	
	高影響期刊論文數(2013)	15	

NTU Ranking的指標全為資訊計量指標，單純使用學術論文客觀分析大學的研究表現，為純粹評估學術研究的大學排名系統。表2顯示其分別為屬於學術生產力構面的近11年論文數（15%）與當年論文數（10%）兩項指標。屬於學術影響力構面的近11年論文被引次數

(15%)、近2年論文被引次數(10%)、近11年論文平均被引次數(10%) 3項指標。屬於學術卓越性構面的近2年h指數(10%)、高被引文章數(15%)、高影響期刊論文數(15%) 3項指標。另外考量學校教師人數的多寡容易造成校際規模的不一，學校規模大者在科研論文評比中通常較具優勢，依各校全職教師數進行正規化處理，提供較為平衡的參考排名。

各種大學排名系統皆有其設計的理念，雖然各有所長，但難免有偏頗，因此在解讀使用時，需注意其指標的設計及評估面向，瞭解此排名系統要呈現的內容為何。執行排名的過程中，亦經常從中挖掘出各種新的問題探討，如大科學(big science)的發展，已經出現一篇文章有3000多位作者的現象，因此在排名的研究之外，陸續探討過創新指標的設計研擬、合著的議題、作者序的問題等資訊計量學較為創新的議題。

大學的使命之一是追求知識的創新與出版一流的研究論文，以促進學術傳播。全球化時代中，知識創新與學術傳播為一體兩面，歷史上學術傳播開放的國家通常具有最頂尖創新的知識，封閉自守的學說、語言、文化則往往沒落衰退。追求本土化的過程中，全球化開放的趨勢無法避免，學術研究仍需持續的創新開放，才能與時俱進，在全球化潮流中競爭。[圖]

(本期專題策畫／臺灣文學研究所洪淑苓所長)

參考文獻：

- [1] Cole, F. J., & Eales, N. B. (1917). The history of comparative anatomy. Part I: A statistical analysis of the literature. *Science progress*, 11, 578-596.
- [2] NTU Ranking. (2014). Performance ranking of scientific papers for world universities 2014 –Related Resources. Retrieved from <http://nturanking.lis.ntu.edu.tw/Others/Link-enus.aspx>
- [3] Pritchard, A. (1969). Statistical bibliography or bibliometrics. *Journal of documentation*, 25, 348.



黃慕萱小檔案

國立臺灣大學圖書資訊學系特聘教授，曾任文學院副院長、圖書資訊學系系主任。近年來主要從事資訊計量學的研究，將資訊計量學與評鑑概念導入專利資訊分析，除 NTU Ranking，研究主題包含人文社會學者評鑑、學術合著議題、評估指標研擬及修正、專利分析、自我引文分析、合著計次議題、國際合著議題、集中度關係、甚至以數據支援國家決策，這些研究論著持續發表於國內外期刊。曾獲得 2013 年美國資訊科學學會 (American Society for Information Science) 的 Best JASIST Paper Award，臺灣大學特聘教授、國科會傑出研究獎、臺灣大學傅斯年獎等獎勵肯定。