

# 全球研究報告 資料的分類： 方法選擇與產出

Martin Szomszor、Jonathan Adams、  
David A. Pendlebury 及 Gordon Rogers



# 作者簡介

**Martin Szomszor 博士** 為現任 Institute for Scientific Information (ISI) 總監。他具有資訊工程專業背景，專精工程、機器學習及自然語言處理。他與英國高等教育資助委員會 (HEFCE) 合作創建了 REF2014 影響力案例研究資料庫 (REF2014 impact case studies database)，因此被評為 2015 年英國資訊時代 (Information Age) 前 50 名資料領導者。ORCID: <https://orcid.org/0000-0003-0347-3527>.

**Jonathan Adams** 是 Institute for Scientific Information (ISI) 首席科學家，同時也是英國倫敦國王學院 (King's College London) 政策研究所的客座教授。由於在高等教育和政策研究領域的卓越貢獻，他在 2017 年被英國艾希特大學 (University of

Exeter) 授予榮譽理學博士學位。ORCID: <https://orcid.org/0000-0002-0325-4431>

**David Pendlebury** 是 ISI 的研究分析主管。自 1983 年以來，他一直致力於使用 Web of Science 資料來探尋科學研究的結構和動態。他與 ISI 創辦人 Eugene Garfield 共事多年，並與 Henry Small 共同開發了 Web of Science Essential Science Indicators™ (ESI) 資料庫。ORCID: <https://orcid.org/0000-0001-5074-1593>.

**Gordan Rogers** 是 ISI 資深數據科學家。他在書目計量學及數據分析領域有超過 10 年的經驗，協助全球客戶評估機構的研究產出及策略。ORCID: <https://orcid.org/0000-0002-9971-2731>.

## 奠基過去，展望未來

### 關於 Institute for Scientific Information (ISI)

科睿唯安旗下的 Institute for Scientific Information™ (ISI) 自成立伊始，半個多世紀以來一直引領著全球科學資訊的研究。今天，ISI 致力於推動研究誠信，改善科學資訊的檢索、解釋和應用。作為科睿唯安學術研究業務的知識研究機構，ISI 透過活動、會議與出版

物對外進行知識傳遞，同時進行基礎研究，讓知識研究庫可以持續擴展和更臻完善。更多資訊請參考：<https://clarivate.com/webofsciencegroup/zh-hant/solutions/institute-for-scientific-information/>

# 引言

這份全球研究報告說明我們如何確認知識與研究的自然劃分，更具體的說法是我們如何分類出版品以使用於探索、分析、管理及制訂政策。分類法一直以來多半採用「由上而下」的原則，我們除了說明其中的歷史發展，也舉例說明科睿唯安產品之中全球各地分類系統的特性，並介紹可自行運用研究資料的由下而上全新方法 Citation Topics。我們也探討將前述方法應用於國家和機構資料的分析結果，並關注不同分類主題對文獻數及引文影響力計算產生的效應。

這份報告不僅提供說明，也推動對資料管理最佳實務的需求，協助各界以負責任的態度使用研究指標。針對研究資料選擇分類系統，並不是價值中立的決策：Charles Handy 在《*The Age of Unreason*》(Handy, 1989 年, 第 17 頁) 一書中寫道：「文字是吹起社會變革的號角。隨著語言改變，行為也很快會有所變化。」研究中的主題及學科領域標示就是如此。瞭解研究出版品分類方法的特性及限制，對研究管理相當重要，因為這會影響我們對既有及創新研究主題的看法、

分析研究活動與績效評估的方式，甚至影響我們成立機構單位進行研究的方式。有許多方式可以整理此類資訊，其中大部分都以合理清楚的方式建立關聯，但可能是這些分類方式的開發只適用於特定的用途。如果只是偶然且隨意地選擇分類主題，而不是依據充分資訊做出決定，或是將其用於預設之外的用途，則產生的結果就可能因為資訊不足而造成解讀錯誤。

**研究時首重專業，而這種專業即是呈現「科學領域中的某種自然規律」的「隱形學院」……我們依據描述詞或其他詞彙製作論文索引的方法，幾乎確定與這樣的自然規律不符……如果我們能夠成功定義自然規律，就會建立一種巨大的科學論文語意地圖，可即時維護用於分類及監控各種產生的發展成果。」**

- Derek J. de Solla Price (1980)

# 分類方法的相關研究

科學歷史學家 Derek J. de Solla Price 認為並沒有能夠直接展現所有知識的通用範本。知識是一種連續，不過早在很久以前就已經發現，如果以推論方式訂定事物之間的界線，並依此整理資訊，不但十分便利，其必要性也持續升高。柏拉圖在西元前 4 世紀，就將知識區分為算術、幾何、音樂及天文學。

民族生物學家 (Ethnobiologists) 認為人類傾向分類是一種天性。其中似乎有一種普世規則用於辨識人類以外的動物，最早出現的是鳥 / 魚 / 蛇等類別，然後又逐漸加入蟲 / 哺乳動物 (Atran, 1990 年)。

隨著知識基礎逐漸擴展，人類需要針對周遭資訊進行整理及排定優先順序的訊息也更加強烈。16 世紀的自然哲學家有可能知道大部分的研究領域，不過所謂的專業分類出現後，各種分類法持續成長。19 世紀出現正式的關聯式圖書館索引這項創新成果，因為當時文獻數量激增，如果沒有按照規則上架，就無法歸檔及使用。

## 19 世紀：國會圖書館分類法與杜威分類法

杜威十進圖書分類法 (Dewey Decimal Classification, 1876 年) 是其中一種版本，國會圖書館分類法 (Library of Congress Classification, 1897 年) 則是另一種版本。Paul Otlet 及 S. R.

Ranganathan 則分別推廣通用十進分類法 (Universal Decimal Classification, 1904 年) 及冒號式分類法 (Colon Classification, 1933 年)；此外還有許多其他分類法，包括 E. Wyndham Hulme 的分類法在內 (請參閱側欄)。杜威及國會圖書館分類系統建立了階層式資訊結構的概念。國會圖書館的 Q 類代表科學，而其中的 QK 子類則代表植物學。杜威也採用類似系統，不過是以數字取代字母作為類別名稱。其中有些類別已經過時：畢竟知識分類系統是他們當時的產物！這類分類系統受歡迎的原因，在於使用者能夠依據文獻的主要內容，迅速存放及檢索各種既有及新發表的文獻。

## 20 世紀：Science Citation Index

研究知識持續成長。期刊數量激增、各種既有領域演進融合、全新領域興起，而知識的擴展幅度已經大到研究人員無法立即掌握。各界開始需要與資訊有關的資訊。

到了 1950 年代，美國科學資訊研究所 (Institute for Scientific Information, ISI) 創辦人 Eugene Garfield，認為展現頂尖工作成果的能力，取決於研究人員目前對本身領域認知的探索成果。在第二次世界大戰後，需要大量人力的傳統索引作業，遠落後於科學文獻的成長速度。像是化學摘要 (*Chemical*

*Abstracts*) 等既有索引，提供的資訊已經過時好幾年。

Garfield 決定要彙整最新重要期刊的目錄頁，並以 *Current Contents* 為名出版；這是每週出版的特定學科公報，可讓研究人員檢閱其關注的研究領域最近出版的期刊。

不過 Garfield 也發現特定學科分類的缺陷，特別是使用控制詞彙及主題標題的傳統索引方法。他提倡所謂的引用索引法 (Garfield, 1955 年)，並表示這是嘗試「打破主題索引的屏障」(Garfield, 1957 年)。Garfield 在 1964 年製作第一份 Science Citation Index (SCI)™，顯示可透過引文辨識出相關領域的論文；這種導覽方法利用引用參考文獻的形式，將作者在論文中加入的專家知識置於其中 (Garfield, 1964 年)。

Science Citation Index 並未提供分類。每份期刊會歸類於一個以上的領域類別，以協助資訊檢索，讓使用者能夠依據領域分類進行檢索。為了打造涵蓋所有領域的統一科學索引，Garfield 開始尋找各個領域中最具影響力的標題，然後諮詢專家意見並分析各個領域引用數最高的期刊，以延伸語意庫 (Pudovkin & Garfield, 2002 年)。

SCI 在設計上並不是作為百科全書，而是要精挑細選，收錄當時在國際間具影響力的期刊 (最初約 600 份)。

之後 SCI 收錄的期刊已數位化，也就是所謂的 Web of Science™ (1997 年)，分類為學科、領域及子領域 (Web of Science 期刊類別)。Web of Science 目前收錄兩萬種以上期刊，分為 254 種類別。所有學科類別中的每份期刊的每個項目都會完整收錄，並隨知識流動演進持續審查收錄範圍。科睿唯安也使用 Essential Science Indicators (ESI)™ 資料庫，其中擁有同樣以期刊定義的 21 種廣泛主題類別，並外加多學科類別，共 22 種分類主題；其中並未收錄藝術或人文學科領域。以上系統之間的差異除了收錄範圍和精細度，還有 Web of Science 學科是以包容為原則，所以期刊可能出現在一種以上的學科領域，而 ESI 類別則是採用排他原則。Web of Science 也提供「研究領域 (Research area)」，是與 Web of Science 學科一樣採用包容原則，有一種以上的分類。

## 研究管理及評鑑

自 1990 年代起，研究管理在國家和機構層級更為活躍，並在全球各地出現新型系統用於分類研究活動。大學重新規劃系所結構，可說是其中預告產生變化的訊號：例如植物學和動物學系所消失，並興起基因學和生物化學等系所，進而組織成為生物科學學院。商學院大幅擴展，涵蓋各種新型及複雜的系所組合。醫學院則裁撤歷史久遠的專科，納入核子醫學等新型技術。

經濟合作暨發展組織 (The Organisation for Economic Cooperation and Development,



### E. Wyndham Hulme

英國專利局圖書館館長  
(1894 至 1919 年)

E. Wyndham Hulme 提出一套系統，依據內容及出版品屬性分類書籍 (Dousa, 2017 年)。他值得關注的部分除了分類理論之外，也包括建議計算各個類別長期下來的研究數量，以反映各個類別及新興專業領域的知識累積情形。他寫道：「書籍分類就是

書架分類，而書架分類執行到最大限度時，就必須統一書架分類的擴展及定義。如果在書籍分類中加入書籍時間順序，則分類的分類主題就會取得新值：其中可呈現每個期間的書目，相當於人類心智活動的對應成長情形」(Hulme, 1923 年)。這種量化方法在半個世紀前就預期會產生科學計量研究，讓 Hulme 成為具有遠見的先驅者。

OECD) 則以全球觀點開發出 Frascati 分類標準 (1963 年，於 2007 年修訂)。這項標準與大學結構完全不同，因為 OECD 類別是以作用而非內容為重點。研究理由各不相同，Frascati 將其分為以下項目：**基本**：目標是取得新知識，並未引導作為任何特定用途；**應用**：透過研究取得知識，並引導發展特定實用目標；或**實驗開發**：以既有知識為基礎的全面研究。前述三種研究類型的活動，可指派給六個以 OECD 主題為基礎的科學領域 (FoS)：自然科學；工程與技術；醫療與健康科學；農業科學；社會科學；以及人文學科。FoS 分類是多種分類方法之一，可讓科睿唯安 InCites™ 平台的使用

者用於研究績效分析。

OECD FoS 看來像是傳統學術工具，但其實並非如此。農業領域表示這實際上是一種經濟分類法。OECD 有興趣衡量的項目，是投入特定產業部門研發工作或其中產生的支出、人員、活動及結果。例如學術界將微生物學視為獨立類別，但 OECD 則想瞭解這究竟是微生物生理學 (自然科學)、微生物疾病 (醫療及健康) 或微生物土壤生態學 (農業)。此外像是化學科學部門及化學產業之間，也同樣存在類似的比對問題：產業會使用工程、數學及其他項目。這樣以作用區別的方式完全合理，但必須能讓使用者適當瞭解。

# 國家層級的分類法系統

除了 OECD、Web of Science 及 ESI 以外，還有許多其他成熟的分類法。以下說明的範例都是 InCites 資料庫中所收錄的分類法，因此我們提出這些分類法如何對應至 Web of Science。

這類系統的開發目標各不相同，中心原則通常是要讓政府（研究資助者）及研究人員（提供研究結果者）感到合理，以便建構對雙方都公平公正的評鑑流程。每種系統都是由大量的政策工作、利害關係人審核、先導專案及經驗所造就的產物。沒有一種系統是隨意產生，所有系統都獲得使用者共同支持，並妥善運作提供結果，不過並不是所有系統的運作內容都相同。不同分類方式可能產生不同結果（稍後說明），但其中並沒有對與錯：因為產生各項結果的目標各不相同。

## 英國 - 研究評鑑

英國首次的研究選擇評鑑 (Research Selectivity Exercise) 是在 1986 年進行。大學提交本身的研究活動組合，由 150 個學科小組以同儕審查方式進行評鑑；1992 年的研究評鑑作業 (Research Assessment Exercise, RAE) 則縮減為 72 個評鑑單位 (UOA)，2008 年 RAE 為 67 個 UOA，2014 年的研究卓越架構 (Research Excellence Framework, REF) 則為四個主要小組的 36 個 UOA。RAE/REF UOA 結構是仿效大學系所。這項結構是依據評鑑需求所建立；大學需要向小組提交可供比較的內容；評鑑等級將成為資助公式中的加權因數，

並告知機構。30 年間類別數逐漸從 150 個減少至 36 個，顯示過度精細的系統可能幫助不大：在分析 2014 年 (36 個 UOA) 所提交的產出結果時，類別可能在內容及期刊使用方面出現重複，不過小組在精細的 UOA 結構中，可能會變成高度自我參考。

UOA 對應至 Web of Science 資料庫的方法，是參照學術機構提交特定期刊文章進行評鑑的頻率。這樣可為各個 UOA 建立包容的核心涵蓋範圍，以及規模雖小但盡量減少重疊學科之間的模糊地帶。

## 澳洲與紐西蘭 - 評鑑與政策

澳洲與紐西蘭標準研究分類法 (ANZSRC) 最初是由澳洲統計局擬定 (1993 年)，並於最近更新，橫跨經濟與學術分界。ANZSRC 的使用者為政策組織，在研究分析方面相當熱門。其中包含三種面向：活動類型 (TOA)，源自 OECD；17 項社經目標 (SEO)；以及 22 項研究領域 (FOR)。SEO 及 FOR 系統為六位數階層，結合杜威類型概念及現代的主題群組。因此 SEO 分類 86 (製造) 具有群組 8607 (農業化學) 及實體目標 860702 (化學肥料)，同時 FOR 分類 09 (工程) 則有群組 0901 (航太工程) 及領域 090101 (空氣動力學)。(我們在註釋中將分類縮寫為 L1，群組則縮寫為 L2)。澳洲研究委員會 (Australian Research Council, ARC) 採用 ANZSRC，進行澳洲卓越研究 (Excellence in

Research for Australia, ERA) 的週期評鑑流程。其中會參考專家指派的期刊清單，以便向各個 FOR 提交資料。每個 FOR 都有指定及專屬的內容。

FOR 期刊清單會直接對應 Web of Science 收錄的期刊。這類清單僅供澳洲使用，並未取得區域外研究，可能會在直接分析 Web of Science 資料時，造成比對國家或機構減少的相符結果。

## 巴西：CAPES – 評鑑研究人員

巴西高教人才培育協調 (CAPES) 分類主題，是由 CAPES 基金會所建立，並與教育部連結，以支援評鑑巴西高教人才並強化技能。這是階層式分類法，結構分為三層：9 個廣泛領域；49 個評鑑領域；以及 121 個更精細的子領域。較高階層領域的研究績效資料，包含下屬層級領域的研究績效資料。例如健康科學 (代碼 6) 研究將包含所有九個下屬評鑑領域 (6.1 至 6.9) 的研究資料。體育 (6.8) 將包含所有三個下屬子領域 (6.8.1 至 6.8.3) 的研究資料。

為了對應 Web of Science，CAPES 分類主題內部的類別元素，會與 Web of Science 期刊類別的範圍註釋進行比對。每個 Web of Science 學科會指派給 121 個 CAPES 子領域主題代碼的其中一個 (或偶而一個以上)，而每個 CAPES 類別則會囊括多個 Web of Science 期刊類別。

## 巴西：FAPESP - 研究投資

聖保羅研究基金會 (FAPESP) 分類主題，是用於評鑑巴西聖保羅州 (São Paulo) 的科學與技術發展成果；聖保羅州以投資研究為優先工作。這項分類為階層式，共有兩層結構：9 個高層類別及 72 個更詳細的類別。高層領域的學科分類，包含下屬領域的學科分類。例如健康科學 (4) 將包含所有七個下屬主題領域 (4.1 至 4.7)。

FAPESP 分類包含所有 Web of Science 學科。大部分期刊類別僅對應至一個 FAPESP 類別，而多個類別可對應至相同的 FAPESP 類別。

## 義大利 - 大學管理及審查

義大利國家大學和研究評估中心 (ANVUR) 分類主題，是以義大利大學研究及教學的正式學術領域和學科清單為依據，約分為 17 個廣泛類別。ANVUR 採用的方法，是將績效評鑑連結主要任務：科學研究、教學 (僅適用於大學) 及社經影響。其中包括協調機構獨立評鑑單位 (Nuclei di Valutazione)；一般分析、統計報告和基準；分享最佳實務；以及實地訪視。義大利議會在 2009 年針對所有公共機構推行新的結果導向績效評鑑。

其中指派政府透明及誠信機構 (CIVIT) 作為獨立評鑑委員會，負責監控及評鑑策略規劃、績效與責任。ANVUR 在 2013 年接手 CIVIT 工作，負責義大利教育大學研究部 (MIUR) 轄下的公立大學及研究機構。

ANVUR 與科睿唯安進行合作專案，以便將 ANVUR 分類對應至 Web of Science 學科，並在 2013 年建立書目計量分析基礎，由 ANVUR 負責執行。這項研究擬定了國際研究排名指標，其中 ANVUR 是以 ANVUR 評鑑研究品質 (ANVUR Evaluation of Research Quality) 架構，對大學研究品質進行評鑑。

## 日本 - 補助金分析

KAKEN 類別定義是以科學研究補助金 (KAKENHI) 資料庫為基礎，目標是擬定研究人員主導計畫，範圍涵蓋人文學科、社會科學及自然科學的基本至應用研究。補助金針對富有創意及先驅的研究專案提供研究資金補助；這類專案預期將成為社會發展的基礎。研究專案是以同儕審查流程進行評選，由專業領域與申請人相近的多位研究人員負責篩選。KAKENHI 的投稿大約分為 300 個類別，依據階層方式組織為四個層級。

2007 年版在第 3 層擁有 66 個類別，在第 2 層擁有 10 個類別。前述兩個層級的類別會透過研究對應至 Web

of Science 學科；前述研究會彙整論文，並透過日本國立情報學研究所 (NII) 營運的資料庫連結 KAKENHI 補助金。

## 中國大陸 - 大學教學與研究

中國國務院學位委員會 (SCADC) 的分類依據，是 SCADC 及中華人民共和國教育部在 2018 年發佈的學位授與及學術訓練指南。因此其中具備強烈的學術課程結構導向，不過目前也用於研究評鑑。SCADC 分類為階層式，共有兩層。其中有 13 個廣泛層級類別，以兩位數代碼表示，另有 96 個更精細的類別，依據各自所屬的廣泛類別進行編號；例如生物學 (0710) 的上層就是自然科學 (07)。

科睿唯安與 SCADC 合作，開發 Web of Science 至 SCADC 分類的期刊對應。部分類別不願意轉移至書目計量分析，原因可能是比較少出現在出版品中，或是因為主題領域重疊。有 15 個下屬 SCADC 類別可能不會轉移至國際對應類別，另外有一個廣泛類別 (軍事科學) 並未獲得國際文獻認可。

# 分類的方法論

科學計量研究人員對分類法特別有興趣，因為分類是計算績效衡量的基本要素 (Sjögårde 2019 年；Waltman 2016 年)。雖然會收集引文作為研究影響力的證據，但是為了公平比較各個出版領域及年份之間的論文，原始引用數必須正規化。每個領域都有其引用模式與特性，而比較久遠的論文會有更多時間累積引用數。所以為了針對每份論文計算正規化或相對化的引用數，就需聚焦於以論文出版年份，從類似引用模式的論文中取得參考文獻。參考文獻的平均引用數，將作為每份特定論文引用數的基準 (Moed, 2005 年；Waltman & van Eck, 2019 年)。

驗證參考文獻系統並沒有所謂的基本真理或金科玉律。各種不同方法都有各自的優缺點、便利性和不便之處。目前有三種主要方法：**期刊定義領域**：在期刊發表的論文將歸屬於該期刊的相關領域；**監督資訊檢索**：會透過關鍵字、作者姓名、期刊及其他屬性，以及引用連結等項目彙總類似論文；以及論文層級的**演算法建構分類** (Zitt 等人，2019 年)。

**期刊定義領域** (例如 Web of Science 學科) 會結合引用分析及明智判斷，以定義合理接近領域和子領域的類別。這種方式易於瞭解、方便取得且便於使用，資料分析複寫也相當簡單。科學計量學傳統上偏好使用這類系統，因為可在各項研究間進行比較 (Glänzel & Schubert, 2003 年；Glänzel 等人，2009 年；Leydesdorff & Rafols, 2009 年)。

在多種期刊定義分類系統中，Web of Science 學科分類一直以來可能仍是最常使用的系統。不過像是《科學》(Science) 及《自然》(Nature) 等多學科期刊的內容，並無法指派至一個或數個類別。解決方法是將這類期刊的每篇論文，依據其參考文獻及引用至特定領域的頻率來指派類別 (Glänzel 等人，1999 年)。ESI 就是以這種方式處理多學科期刊論文，以便指派至 ESI 學科領域。對其他背景的分析師而言，這並不是容易的任務，一般只有科學計量人員能夠建構系統進行這類重新指派的分類工作。

更重要的是，期刊定義領域並未完整呈現研究領域，因為其中所屬的論文，通常是在指派至其他領域的期刊中發表。事實上，高被引論文通常出現在高影響力的多學科期刊之中。此外，雖然學科中期刊獲得妥善定義，但是期刊的重點及引用密度仍可能不同。在這種情況下，如果將論文引用數與學科基準比較，可能會對論文嚴重不利 (或有利) (van Eck 等人，2013 年)。近期有一項研究是以中國科學引文資料庫 (Chinese Science Citation Database™) 的資料為依據，其中表示在指派論文至期刊定義學科時，正確率約為 50% (Shu 等人，2019 年)。最後，由於期刊定義領域 (例如 Web of Science 學科) 一般是設計用於協助資訊檢索，因此如果用於不同用途，指派就可能產生所謂「索引者效應」(indexer effects) 的現象 (Rafols & Leydesdorff, 2009 年)。總而言之，期刊定義領域確實強大，但並不能完善表示領域及引

用特性 (Leydesdorff & Bornman, 2016 年；Milojevic, 2020 年；Wang & Waltman, 2016 年)。

**監督資訊檢索**是第二種方法，需要技術主題的專業知識，較為麻煩且鮮少使用 (Haunschild 等人，2018 年；Lewison, 1996 年)。這種方法更容易找到高度類似的內容，並可能由同質性的集合提供完善的參考文獻資料集用於引用正規化。但其中的代價是：只有具備檢索和主題專業的少數人能夠建立這類自訂集合；而且只能使用一次，不太可能重複使用。

**演算法分類**使用論文而非期刊群組之間的關係定義結構及領域，通常會使用階層方式。領域描繪最常以引用網路分析為依據，不過也可使用詞庫特徵及其他屬性，並提出混合系統 (Boyack & Klavans, 2020 年；Janssens 等人，2009 年；Yu 等人，2017 年；Zitt & Bassecouard, 2006 年)。這種由下而上的方法，是仰賴提出第一份引文索引時描述的「聯結想法」概念 (Garfield, 1955 年)。

Henry Small 曾經擔任 ISI 首席科學家多年，在 1970 年代率先提出共被引這種方法，以便定義科學的專業結構。共被引可依據頻繁成對引用連結相關出版物 (Small, 1973 年)。Small 與費城卓克索大學 (Drexel University) 的 Belver Griffith, 在 1974 年使用共被引集群建立研究專業地圖，並採用多維標度法，依據計算所得的相似度訂定集群縱座標。

集群演算法及視覺化軟體目前成為重要的研究活動，在所謂的演算法建構分類系統中使用論文間引用關係，包括書目耦合、共被引和直接引用以及其他技巧(Ahlgren & Colliander, 2009 年; Ahlgren 等人 2020 年; Sjögarde & Ahlgren, 2018、2020 年; Traag 等人 2019 年; Waltman 等人 2020 年; Waltman & van Eck, 2012 年)。有特定特性的數據會比其他特性受到歡迎，例如集群數量及大小的範圍 (Perianes-Rodriguez & Ruiz-Castillo, 2017 年; Ruiz-Castillo & Waltman, 2015 年; ubelj 等人 2016 年)。不過很難斷言這類方法都一定有利或受到偏好，因為在較上位的分類階層中，我們可以顯示論文層級與期刊定義的分類和績效衡量之間大致相同。其中明顯缺點就是演算法的黑箱性質：不同研究群組無法重新建立分類主題；而不同演算法使用相同資料及對應的產出變化，則讓人質疑這樣的結果是否為演算法所操控 (Gläser 等人, 2017 年)。

## 類別精細度的效應

建構任何分類主題及影響分析時，分類法的精細度是非常重要的。Zitt 等人 (2005 年) 提出學科正規化引文影響力 (CNCI)，有可能會因為任何正規化所產生的層級 (以「縮放」描述) 而改變。其他學者也曾提出類似問題，例如 Hirst (1978 年) 有關「學科影響因子」(Discipline Impact Factors) 的部分；Schubert & Braun (1993、1996 年) 評論書目計量指標在各個領域之間的比較結果；而 Glänzel & Moed (2002 年) 也曾對不同分類後的層級所產生的效應表示意見。

Adams 等人 (2008 年) 測試了「Zitt 縮放」對研究績效指標的效應，其中分析向 2001 年英國 RAE 提交評估的論文，在不同正規化層級情況下的相對影響力 (表 1)。在三個評鑑單位 (UOA13 心理學、UOA14 生物科學及 UOA19 物理學) 獲得最高評等前三名 (4、5 及 5\*) 的大學系所，

其資料顯示專家審核和引文影響力之間，於部分層級中具有正相關。如果引用數在期刊層級進行正規化，則任何等級之間的影响力指標差距非常小，不過如果正規化與 Web of Science 學科分類或整個 UOA 有關，評等較高的單位就會在統計上大幅提升相對影響力，支持 Zitt 等人 (2005 年) 的分析結果。

評鑑選擇採用較精細的分類進行評估的風險，即是分類會出現自我參照的現象。排名較低單位所提交的資料，無疑是源自平均影響力較低的期刊，而頂尖單位提交的資料，則源自平均影響力較高的期刊。相對於期刊，論文與其發表媒介之間的影响力是相似的。只有著眼於更高的分類層級，例如 Web of Science 學科，我們才會發現排名高的單位的論文絕對被引次數明顯更高。

**表 1：**英國大學研究人員在 1996 至 2000 年期間所發表論文及綜述文獻的學科正規化引文影響力 (CNCI) 平均值，對象為 2001 年研究評鑑作業 (RAE 2001) 獲得評等 4、5 或 5\* 的單位。其中顯示了三個評鑑單位 (UOA) 的資料，包括每項評等的單位數量，以及出版品的 CNCI，而引用數則以三種不同精細度進行正規化：出版的期刊；Web of Science 期刊類別；以及整個 UOA 的資料集。(Adams 等人, 2008 年)

	UOA13 Psychology				UOA14 Biological sciences				UOA19 Physics			
	Average CNCI				Average CNCI				Average CNCI			
Grade at RAE 2001	Number of units	Journal	Web of Science	UOA	Number of units	Journal	Web of Science	UOA	Number of units	Journal	Web of Science	UOA
Grade 4	17	1.22	1.40	0.80	17	1.29	2.35	1.89	15	1.28	1.84	1.98
Grade 5	17	1.18	1.80	1.05	30	1.11	2.33	2.33	23	1.47	2.51	2.96
Grade 5*	12	1.32	2.38	1.63	11	1.18	2.53	2.93	5	1.82	3.32	3.75

# 全新方法 Citation Topics：由下而上的分類法

所謂的引用型分類法，是一種由下而上的分類系統，其中個別主題會逐漸連結成為更大的分類，並依據基礎資料特徵享有共同特性。

科睿唯安目前在 InCites 推出引用型分類法；這是與荷蘭萊頓大學 (Leiden University) 科學與技術研究中心 (Centre for Science and Technology Studies, CWTS) 的頂尖科學計量學術團隊合作開發。這種方法的目的，是運用論文層級演算法分類的優點：

- 更準確表示微觀資料集或專業主題；
- 提升內容同質性；以及
- 改善引用正規化。

這項創新方法與過往的期刊分類法完全不同。資料導向方法會依據論文中所有書目資訊產生分類，而不是使用人類概念。這樣的類別可以微調，例如增加或減少集群數量 (精細或粗略分類)，而類別如何出現的基礎機制則是源自資料模型。

科睿唯安為了評估這種方法的可行性及可能用途，與 CWTS 合作建構演算法，依據 Web of Science 的引用網絡建立類別結構。演算法支援階層式分類法。層級定義為微觀 (建立的最精細集群)、中觀 (將類似微觀集群聚集在一起的中階彙總層級) 和宏觀 (聚集中觀集群的最上位彙總)。

雙方共同探索其中的各項挑戰，針對最終資料的產生及更新方式做出明智決策。

**時間尺度** - 針對 1980 年之後的內容進行分析，以便與 InCites 中的資料相互呼應，並反映縱向書目計量分析中最常用的時間範圍。

**文獻類型** - 雖然大部分書目計量分析執行時僅使用論文及綜述文獻，但所有文獻類型均包含在分析中。請注意，只有和其他文獻之間具有引用連結的文獻 (亦即引用或被引) 可納入分類之中，因此可能並未包含期刊、書籍章節和會議論文集。

**集群大小** - 演算法可微調以產生任何數量的集群，不過重點是要考量各個集群的數量，及其對引用數正規化的效應。InCites 可製作學科正規化引文影響力 (CNCI) 及百分位數指標。每個類別 / 年份組合都必須訂定適當的最小數量，否則正規化基準就會不穩定。分析顯示在這樣的限制下，可建立 2,000 個以下的微觀集群，如果數量更多就會反正會過於發散。

**相似度衡量** - 可用於透過引用網路判定相似度的衡量方式為：

- **書目耦合** (依據共同引用參考文獻判定相似度)；
- **共被引** (依據成對文獻引用判定相似度)；以及
- **直接引用** (任何連結)。

共被引在分類時會排除過多文獻。書目耦合與直接引用之間的取捨問題與學科動態有關。如果透過書目耦合衡量相似度，文獻位置為固定 (引用參考文獻清單永遠不變)。不過如果是包含引用文獻的直接引用解決方案，由於引用文獻會因加入更多文獻及引用而隨時間增加，因此指派的文獻集群有可能演進發展。這是有利現象，讓既有文獻可在新主題出現時重新指派。

**更新** - 分類靜態資料可說是簡單明瞭，不過真實世界中的解決方案必須定期更新，所以我們必須考量由下而上分類法應如何隨時間演進發展。基於實務理由，建議使用以下兩種的更新形式：

**1. 每月** - 吸取新資料後，文獻就會依據本身的引用參考文獻指派至現有微觀集群。

**2. 每年** - 集群演算法會每年重新執行。此時新的微觀集群可能出現，或是有部分文獻可能在不同集群間游移 (游移可能會影響近期文獻，因為這類文獻在出版後幾年具有最高的被引數)。演算法已微調，可在每年更新期間適度建立新的中觀或宏觀層級集群。

可用性是推動此項解決方案發展的動機。其中產生的分類主題應相對固定，這樣使用者在每次年度更新後，就不必完全重新學習其中局勢。同時分類主題應提升演進發展的靈活度，以便建立新興主題。CWTS 為此納入緩衝標準，確保每年更新期間變更集群的論文不會超過 5%。

**標示集群名稱** - 演算法結果就是將文獻指派至微觀集群；建立階層將微觀集群聚集為中觀集群；然後將中觀集群聚集為宏觀集群。集群會進行標示，以便讓初步使用者辨識內容。其中會由領域專家提供資訊，加上集群的摘要資訊，透過專家編輯的方式，為宏觀及中觀層級集群建立標示，例如常用的作者關鍵字及 Web of Science 學科。微觀集群

會自動獲得分配標示，依據為與集群中文獻有關的最重要作者關鍵字。

## Citation Topics

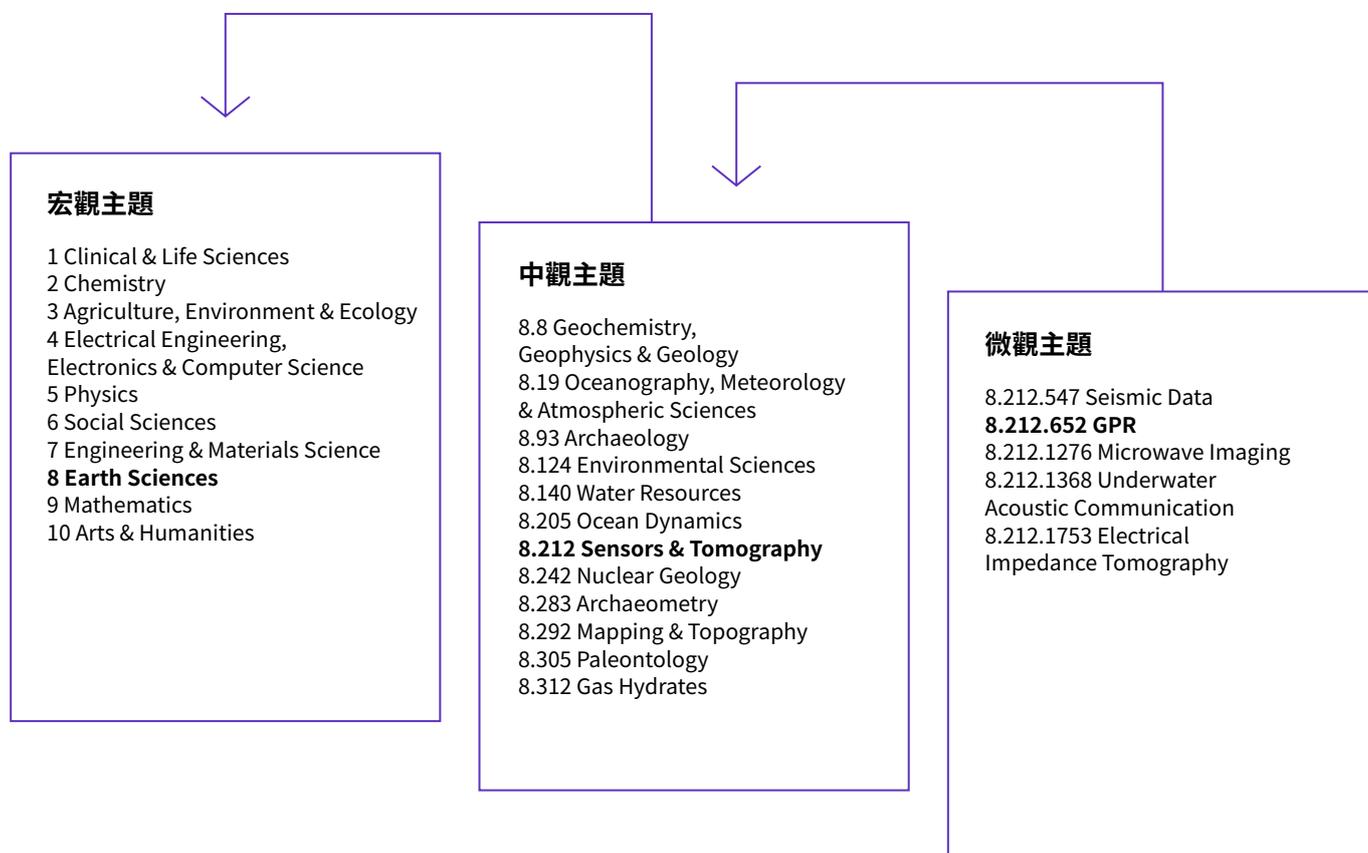
新型的分類主題「Citation Topics」，是以 CWTS 方法為基礎，並於 2020 年 12 月加入 InCites 之中。使用「Citation Topics」這個名稱，是因為分類系統是以引用網路建立，而產生的集群性質各不相同。其中部分集群與學科標示相符（例如眼科學），其他集群則著重於特定疾病、物質或分析技巧。「Topic」（主題）這個詞彙似乎比類別或等級更適合，具有正式結構的意涵。

目前實作（2020 年 12 月）是由 10

個宏觀主題、326 個中觀主題及 2,444 個微觀主題所組成。其中分析的文獻達到 6,000 萬份以上，並向 5,000 萬份以上文獻指派 Citation Topics。對實質研究文獻類型（論文及綜述文獻）而言，1980 年之後的完整資料可能會有高比例文獻指派至主題（92% 論文及 96% 綜述文獻），而前述比例在最近五年則進一步提升（95% 論文及 99% 綜述文獻）。

我們在圖 1 列出 10 個廣泛的宏觀主題，並以「地球科學」為例，說明這項主題如何由 12 個中觀主題彙總而成；其中有一個中觀主題（「感測器與斷層掃描」）是由五個微觀主題彙總而成，而這五個微觀主題的重點則是 GPR（透地雷達）。

圖 1：以地球科學為例說明 Citation Topics 的三層階層



## 比較 Citation Topics 與其他分類法

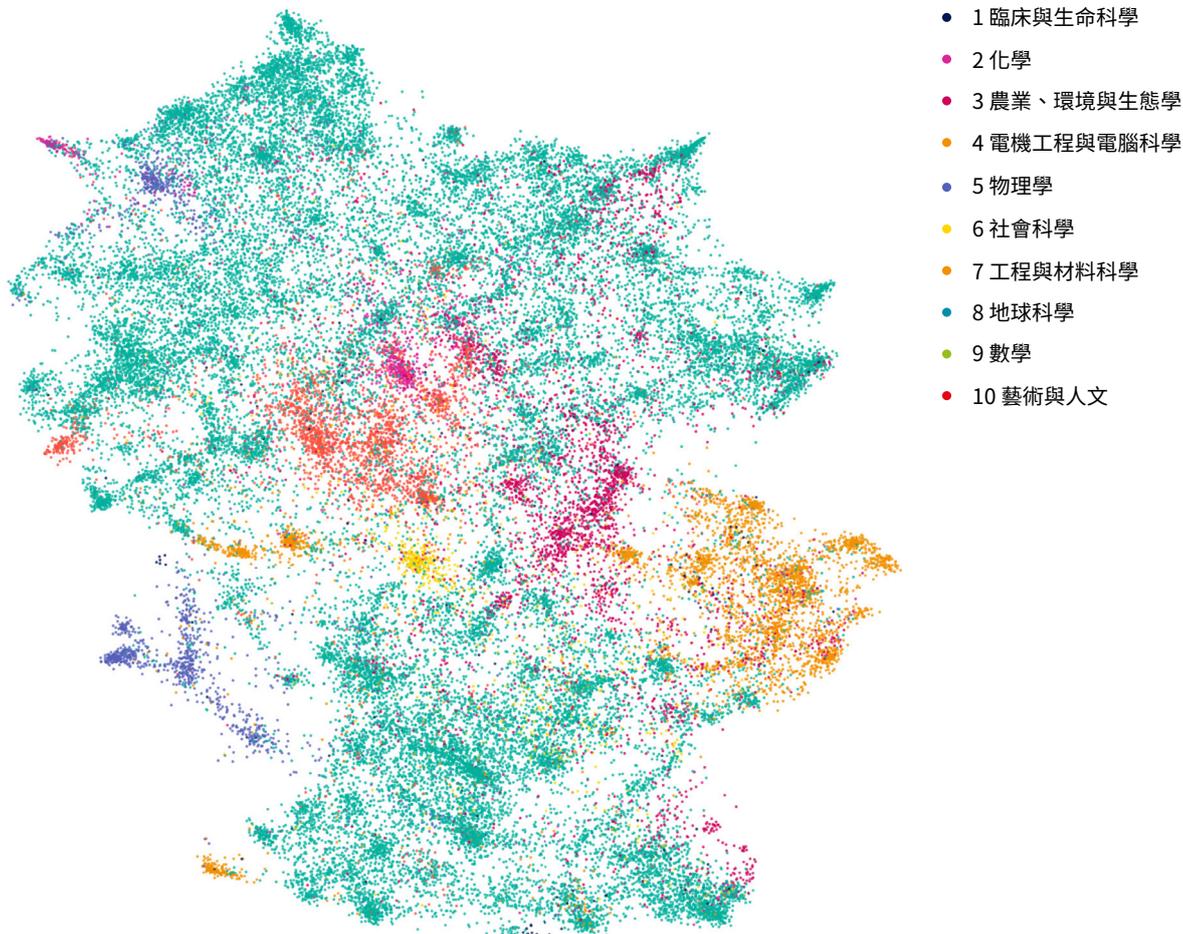
為了比較 Citation Topics (論文層級分類) 與既有期刊分類的類別結構, 我們建立了一系列研究地圖, 以視覺化其中的相似度及差異 (圖 2.A-C)。這些地圖是以 ESI 類別「地球科學」為重點, 僅包含 2015 年發表的 48,000 份論文及綜述文獻。文獻在圖中位於共同的地圖局勢中, 亦即每張圖使用相同配置, 而每份文獻的位置, 則是以其參考文獻與所有其他文獻的相似度為依

據 (書目耦合)。我們利用所謂的均勻流形逼近 (Uniform Manifold Approximation, UMAP) (McInnes & Healy, 2018 年), 將特徵空間投影至兩個維度, 以便輕鬆視覺化文獻空間。然後每張圖會應用不同色彩的分類主題, 顯示相同集合的「ESI 地球科學」文獻在該項分類系統中的分類方式。三項範例系統為: 宏觀 Citation Topics; 中觀 Citation Topics; 以及 Web of Science 學科。

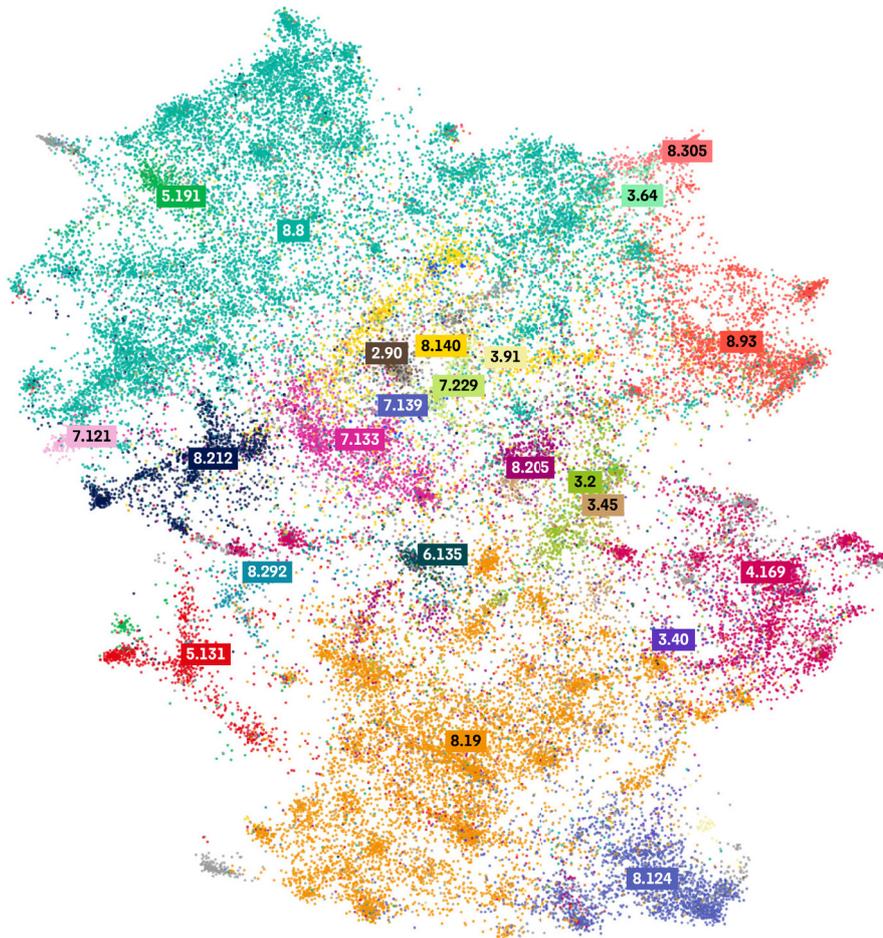
地圖顯示特定分類主題針對相同內容建立群組或集群的方法存在不同之處。雖然由下而上 (引用集群) 與

由上而下 (期刊類別) 兩者之間存在方法論差異, 不過就整體情勢而言, 這些分類群組顯然是相符。這可能代表專家判斷及引用連結中固有的引導原則, 是由所謂的「自然規律」作為基礎。如果要解釋其中差異, 可以解讀某個學科或領域與其他同源領域的關聯方式。Citation Topics 的明顯效益, 就是以更精細的分類層級 (微觀主題) 讓新群組有機會出現, 舊有的期刊分類就無法做到這一點。

**圖 2A : 宏觀 Citation Topics。**第一張圖片顯示 ESI 地球科學領域的 2015 年文獻, 如何依據宏觀主題排列及著色。其中大部分 ESI 內容都是由宏觀主題 - 地球科學 (綠色) 的引用連結形成集群。圖中右側有大部分文獻組成分配至工程 (橘色), 而左側則有另一個集群指派至物理 (紫色)。

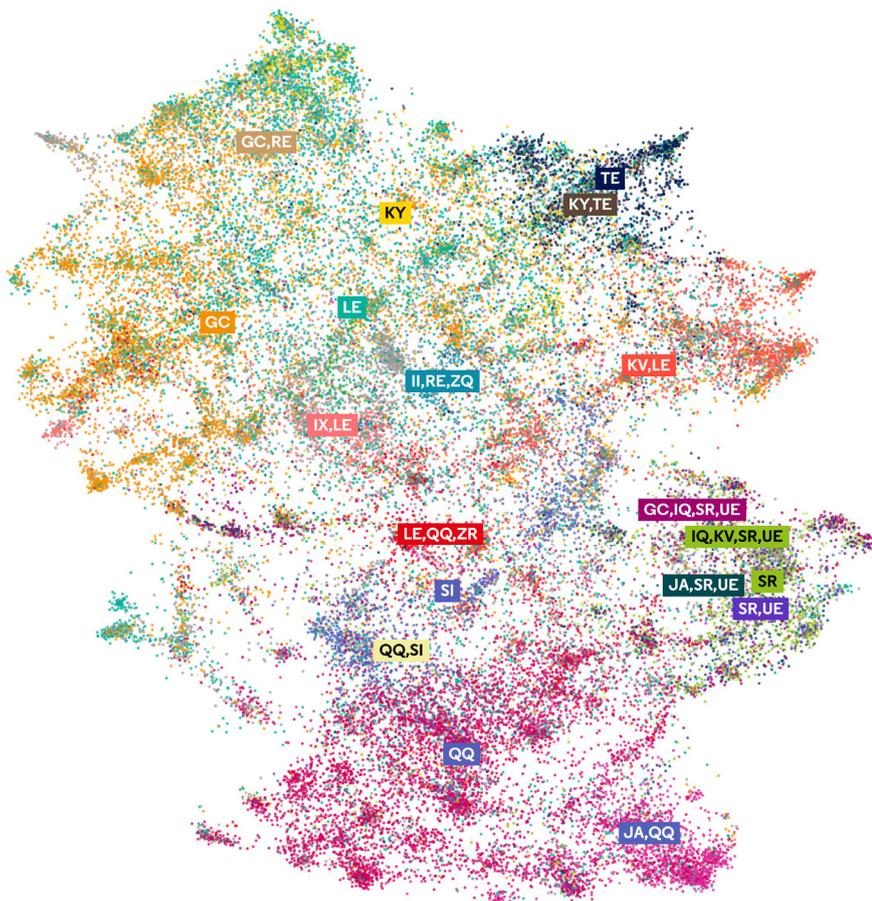


**圖 2B：中觀 Citation Topics。** 第二張地圖顯示中觀主題，其中遵循相同以引用為基礎的論文層級分類法，不過層級更為精細。其中很明顯是由「地球化學、地球物理學及地質學」(8.8，頂端，綠色)和「海洋學、氣象學及大氣科學」(8.19，底部，橘色)等主要群組區分內容。



- 2.90 水處理
- 3.2 海洋生物學
- 3.40 林業
- 3.45 土壤科學
- 3.64 系統發育與基因體學
- 3.91 汙染與植物修復
- 4.169 遙測
- 5.131 氣象與大氣科學
- 5.191 太空科學
- 6.153 氣氛變遷
- 7.121 混凝土科學
- 7.133 岩土工程
- 7.139 能源與燃料
- 7.229 礦物與金屬加工
- 8.124 環境科學
- 8.140 水資源
- 8.19 海洋學、氣象學和大氣科學
- 8.205 海洋動力學
- 8.212 感測器與斷層掃描
- 8.292 製圖和地形
- 8.305 古生物學
- 8.8 地球化學、地球物理和地質學
- 8.93 考古學

**圖 2C : Web of Science 學科。**最後一張圖片呈現相同的論文集但以 Web of Science 學科分類顯示。這是在期刊層級製圖，許多期刊都指派至多個類別。因此色彩可能會對應至類別組合，而相似色彩會讓地圖比較不容易解析，特別是在出現多項工程類別的右側。地圖頂端的主要群組為地球科學、多學科 (綠色, LE)、地質化學及地質物理學 (橘色, GC) 和地質學 (黃色, KY)，與中觀主題 8.8 的地質化學、地質物理學和地質學廣泛重疊。同樣地，底部的的主要群組組成包括氣象學和大氣科學 (紫色, QQ)，以及 QQ 搭配環境科學 (粉紅色, QQ+JA)。觀察力敏銳的讀者會注意到地圖下方群組有一部分並沒有包括氣象學內容 (紫色及粉紅色)，而是與對應的中觀集群 8.19 海洋學、氣象學及大氣科學形成群組。此項內容特別與海洋學 (SI) 有關，突顯指派標示之中反映出真實但微小的精細度差異。



- [LE] 地球科學, 多學科
- [GC] 地球化學與地球物理學
- [QQ] 氣象學與大氣科學
- [KV, LE] 地理、物理地球科學、多學科
- [SI] 海洋學
- [JA, QQ] 環境科學、氣象學和大氣科學
- [KY] 地質學
- [TE] 古生物學
- [SR] 遙測
- [LE, QQ, ZR] 地球科學, 多學科、氣象學和大氣科學、水資源學
- [GC, IQ, SR, UE] 地球化學與地球物理學
- [GC, RE] 地球化學與地球物理學、礦物學
- [GC, RE] 能源與燃料、工程、石油
- [IX, LE] 工程、地質、地球科學, 多學科
- [JA, SR, UE] 環境科學、遙測、影像科學與攝影技術
- [IQ, KV, SR, UE] 電機工程
- [KY, TE] 地質學、古生物學
- [II, RE, ZQ] 工程、化學、礦物學、採礦與礦物加工
- [SR, UE] 遙測、影像科學與攝影技術
- [QQ, SI] 氣象與大氣科學, 海洋學

# 分類法的影響

InCites 是科睿唯安的書目計量分析產品，提供使用者多種由上而下的資料分類法，目前也納入我們與 CWTS 合作開發的由下而上 Citation Topics 分類法。選擇使用的資料分類法，會對使用的資料範圍、建立資料群組的方式及精細度造成多大影響？會如何變更基準及樣本內容？

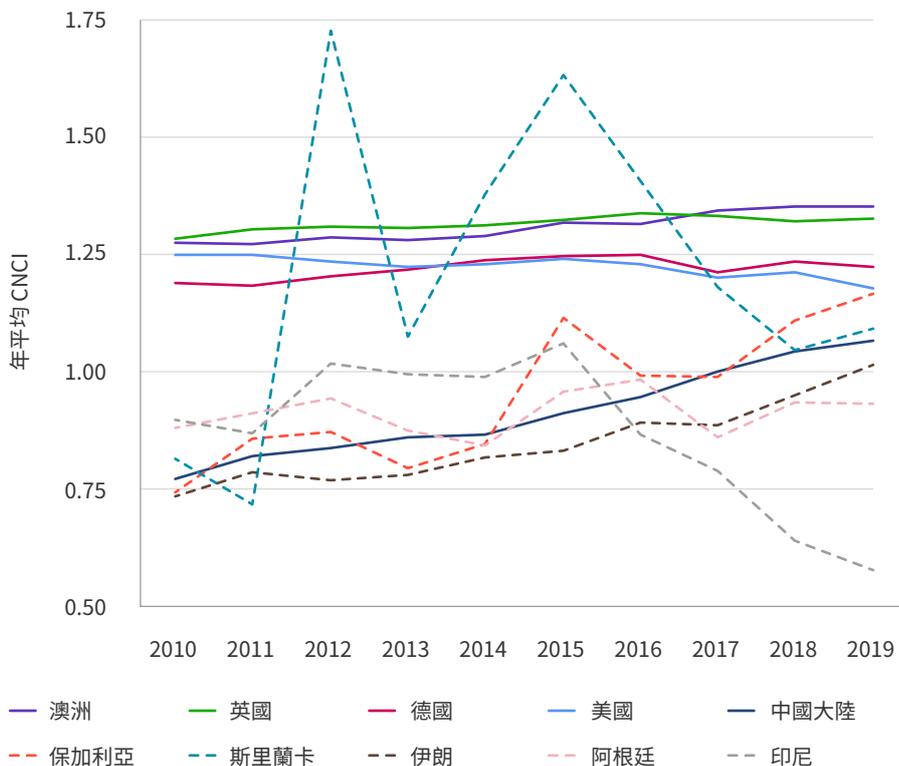
為了探索前述問題，我們比較了以 Web of Science 254 種期刊分類 (涵蓋所有主題領域) 為基礎的分析結果，以及使用其他分類法的分析結果。摘要指標 (表 2) 說明 10 個國家的出版量 (論文及綜述文獻數量) 和學科引文影響力 (CNCI)。其中有五個國家擁有大規模且資金充裕的研究經濟體 (美國、中國、英國、德國及澳洲)，另外五個國家雖然持續進步，但目前在資金方面相對薄弱，研究產出規模也較小。

斯里蘭卡的平均 CNCI 真的與美國相當嗎？畢竟美國的出版量是斯里蘭卡的 500 倍以上！如果伊朗擁有最高的論文被引用率，但平均 CNCI 卻是第二低，又代表什麼意義？換句話說，這類單項指標是如何解釋及計算？CNCI 趨勢資料 (圖 3) 顯示在這 10 個國家之中，至少有 2 個國家的平均指數不可靠：斯里蘭卡的平均指數不僅偏高，也不穩定。這是否會讓人對更穩定的數值表示懷疑？我們絕對需要進一步瞭解各個經濟體在指標中使用的出版量資訊。單一指標是否能代表數百萬份的出版品及數千萬項引用？我們顯然需要更多資訊以便適當解讀表 2 及圖 3，並探索選擇特定分類系統會如何影響其中結果。

**表 2：**10 個地區在最近 10 年期間 (2010 至 2019 年) 研究產出的摘要指標 (Web of Science 收錄的論文及綜述文獻數量) 和績效 (學科正規化引文影響力平均值，CNCI 世界平均值 = 1.0)。地區依據 CNCI 進行排名。

	論文量	CNCI	引用總數	被引用百分比 (%)
英國	1,981,903	1.41	26,932,154	65.6
澳洲	888,127	1.41	12,626,406	72.4
美國	6,838,175	1.31	90,031,964	63.9
斯里蘭卡	13,068	1.31	170,284	63.6
德國	1,615,968	1.30	23,029,125	71.1
保加利亞	38,366	1.01	360,385	60.2
中國大陸	3,743,888	0.99	39,306,476	71.5
阿根廷	121,077	0.96	1,321,844	71.4
伊朗	362,748	0.91	3,428,680	77.9
印尼	85,885	0.81	342,576	39.1

**圖 3：**10 個地區 (5 個成熟的研究經濟體，5 個成長中的研究經濟體) 在最近 10 年期間 (2010 至 2019 年) 的學科正規化引文影響力 (CNCI) 年度趨勢。



## 分類及數量

如果選擇使用 InCites 提供的六種分類法，全部的 Web of Science 出版品資料集，通常會比實際指派至各

個國家的論文數量更多 (表 3)。其中有部分資料集，特別是 ANZSRC FOR 的期刊清單，會讓印尼等國家的可用資料減少最多一半。即使是美國及英國，出版品集也減少 20%

(L1 部門類別) 或 35% (更具體來說是 L2 群組類別)。相較之下，英國 REF 的分類主題，以及巴西 CAPES 及 FAPESP 使用的分類主題，基本上則運用完整的來源資料。

**表 3：**表 3 所列 10 個地區透過 Web of Science 期刊學科分類指派的論文數量，與 InCites 其他 6 種分類法指派論文數量之間的比例。涵蓋的文獻範圍若有不同，將影響任何後續引文影響力正規化計算的分子及分母引用數。

	ESI	FOR L1	FOR L2	REF2014	CAPES49	FAPESP
美國	0.85	0.80	0.64	1.00	1.00	1.00
中國大陸	0.78	0.71	0.54	1.00	1.00	1.00
英國	0.81	0.80	0.65	1.00	1.00	1.00
德國	0.84	0.77	0.61	1.00	1.00	1.00
澳洲	0.86	0.84	0.67	1.00	1.00	1.00
伊朗	0.90	0.80	0.64	1.00	1.00	1.00
阿根廷	0.90	0.82	0.66	1.00	1.00	1.00
印尼	0.34	0.35	0.26	0.98	1.00	1.00
保加利亞	0.75	0.60	0.48	0.98	1.00	1.00
斯里蘭卡	0.76	0.72	0.59	0.99	1.00	1.00

## 分類及影響

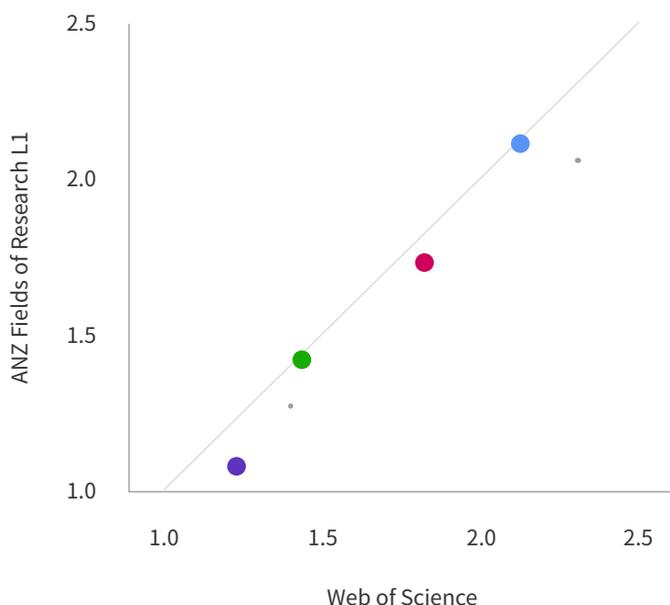
分類主題也會影響 CNCI。我們要再次保證，依據不同分類系統正規化的引用數所取得的 CNCI 值之間具有高度關聯性，不過關聯性並非完美。Y 軸截距及斜率可能都會出現差異；前者可能讓所有值向上或向下，後者則可能會以較低或較高的平均影響力，對機構造成不同影響。如

果要讓評鑑的所有當事方之間都維持公平，數據分類就必須與評鑑目標相符。圖 4 顯示 86 間英國大學的 CNCI 平均值 (2015 至 2019 年)；CNCI 值取自多項不同分類系統的所有學科類別。由 Web of Science 期刊分類移往 FOR 第 1 級的效應，就是會壓抑大部分機構的 CNCI，不過這大多低於世界 CNCI 平均值，而且分佈上端部分幾乎可忽略不計。其中有一些明顯的離群值，所以完全

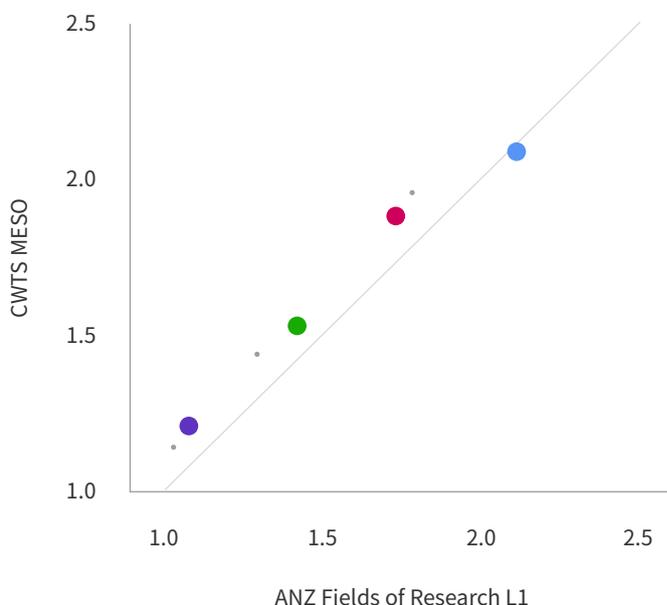
不算是均勻效應。Web of Science 學科與 CWTS Citation Topics 集群所建類別的 CNCI 值之間，具有更密切的關聯性。比較使用 CWTS 中觀 (Meso) 類別及 FOR1 類別的 CNCI 後，顯示 FOR 分類法會壓抑 CNCI 值。如果轉移至更精細的層級，使用 CWTS 微觀 (Micro) 及 ANZ FOR 第 2 級類別，會產生類似效應，而分佈低 CNCI 部分受到的壓抑效應更大。(圖 4)

**圖 4.A-D**：在使用不同分類法在分類機構資料及建立整體基準的情況下，86 間英國大學 (2015 至 2019 年) 在 Web of Science 索引文獻的學科正規化引文影響力 (CNCI) 平均值之間的關聯性。所有關聯性均為高度相關，不過特定機構的迴歸變異有所不同。為了對此進行追蹤，其中以固定色點突顯具有獨特研究歷史和組合的四間大學。(Web of Science 學科將期刊對應至 254 個領域；ANZSRC 研究領域 (FOR) L1 = 期刊對應至 24 個廣泛組別，L2 = 212 個 L1 之中的特定群組；CWTS MESO 及 MICRO 代表萊頓大學 CWTS 開發的中觀及微觀 Citation Topics 類別)。

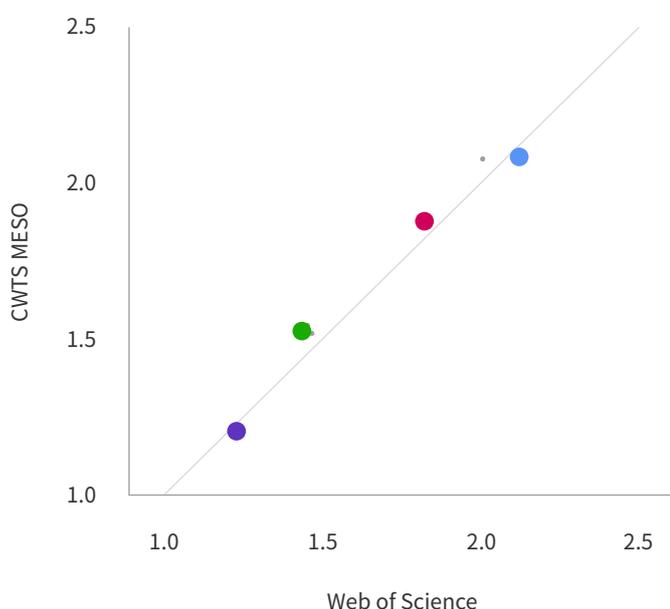
**圖 4A**：FOR1 對比 Web of Science，關聯性 = 0.97



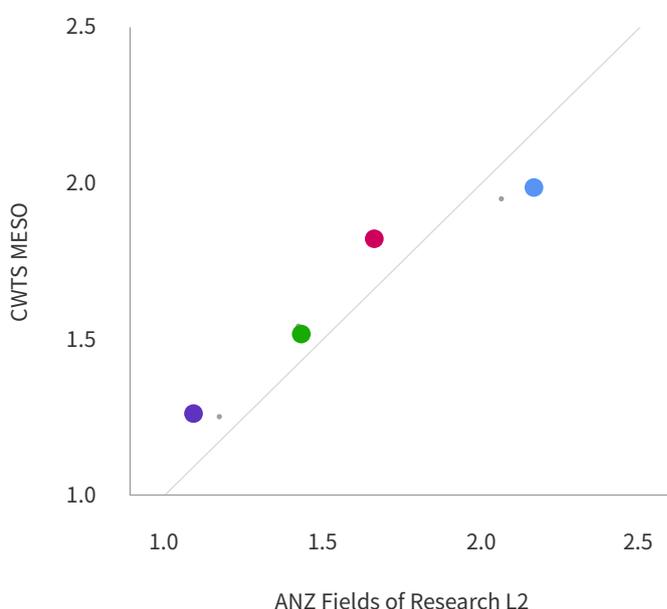
**圖 4B**：CWTS 中觀對比 FOR1，關聯性 = 0.95



**圖 4C**：CWTS 中觀對比 Web of Science，關聯性 = 0.99



**圖 4D**：CWTS 微觀對比 FOR2，關聯性 = 0.93



根據追蹤的四所大學相對位置變化，顯示存在相當大的殘餘變異，因為四所大學在不同的分類體系下轉移所呈現的結果都不一致。有 6 所大學在使用 Web of Science 期刊分類進行正規化時，平均 CNCI 為 1.7，不過如果使用 FOR L1 分類，這 6 所大學的 CNCI 值可能介於 1.45 至 1.85 (圖 4.A)。在四所追蹤的大學中，績效最佳的大學由 CWTS-MESO 轉移至 FOR1 (圖 4.B) 及 FOR2 (圖 4.D) 時出現成長，不過其他三所大

學仍然下滑。

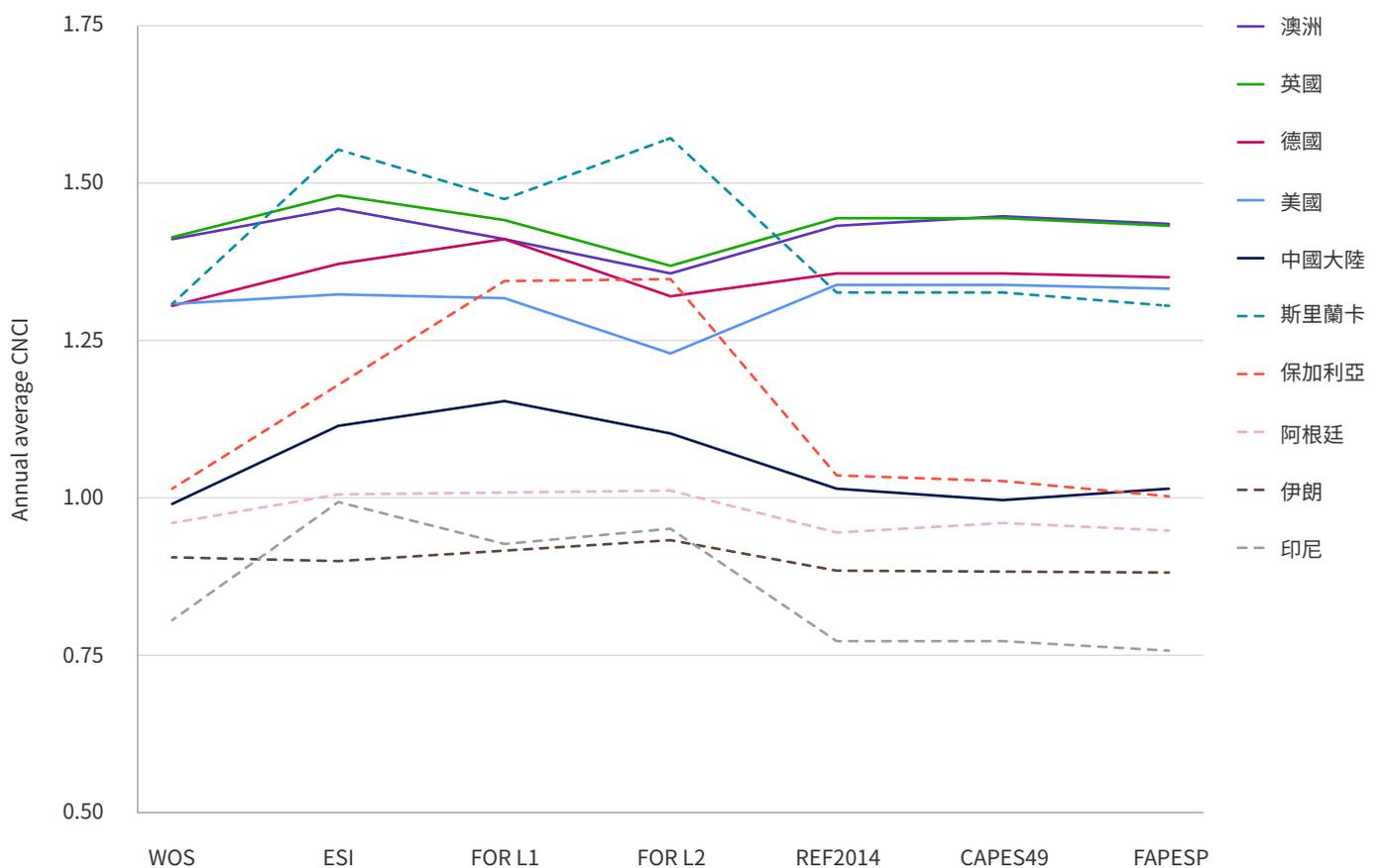
不論前述轉移原因是主題混合 (因為各項分類法所指派期刊至類別的情形不同造成整體基準改變) 或其他因素，都會實質影響相對的機構結果。

我們現在可以在國家績效指標比較不同分類主題的結果。表 2 (依據 Web of Science 期刊分類) 顯示斯里蘭卡的 CNCI 與美國及德國相近。圖 5 顯示使用 ESI 或兩種 ANZ FOR 分類主題所產生的結果，斯里蘭卡

將成為世界第一。如果使用以上分類主題，印尼的 CNCI 也會上升，不過大部分國家 CNCI 受到的影響非常小，但美國、英國、澳洲及德國在 FOR 第 2 級仍然受到略微壓抑。

以上這些解答都沒有錯。我們在其中學到的教訓，就是在選擇分類主題進行資料選擇和彙總時，將會影響分析和解讀，不過沒有一種分類主題是隨意或在沒有規劃的情況下實作，而分析及之前的發展都呈現合理且基於事實的結果。

**圖 5：**在資料依據 7 種不同分類主題正規化的情況下，針對 10 個地區所計算的學科正規化引文影響力 (CNCI) 平均值。各種分類主題間用於計算 CNCI 的出版品數量各不相同，如表 3 所示。圖中直線不代表各項分類主題之間的任何連結，僅用於視覺輔助。



# 合理指標的意涵

負責任的書目計量資料使用者，需要明確瞭解自己的資料是否與本身提出的評鑑問題有關；使用者需要事前瞭解自己如何使用資料，以及所選擇應用的方法；而選擇方法時必須仔細瞭解用於建立資料群組和基準的分類主題。

如果評鑑人員完全仰賴表 2 摘要資訊，判斷相對或絕對的研究實力（甚至涵蓋整個國家），顯然就是一種輕率的行為。如果評鑑人員審查相同國家的機構列表，或是尋求資金贊助的研究團體，這項問題就更加重要，但這種情形經常發生。

高度精細的類別系統，會將研究論文分為小型的自我參考區塊，大幅提升研究表面的相對引用績效，但其實相同研究在熟悉的主題彙總方法中，引用績效相當差。普遍而言，選擇學科 / 主題類別以彙總出版品及正規化引用時，會出現雙重效應。首先，如果國家的國內研究基礎發展程度較低，且引用的國內研究產出程度較差，在使用排他性類別系統（例如 ESI）的情況下，出版數會較少。第二，由於這種類別會挑選重點期刊，因此會忽略國家活動被引最差的部分，造成平均 CNCI 上升。因此，雖然斯里蘭卡、保加利亞和印尼在 ESI 分析中的出版品數

量比 Web of Science 分析減少許多，仍然產生更高的平均 CNCI。

我們完全同意 Henk Moed 教授的觀點 (Moed, 2020a、2020b)；這位教授認為需要評鑑架構，在其中將評鑑的背景及目的做為最重要的考量因素。引用本身就具有價值觀，是透過社會及研究加權建構而成。任何彙總的引用數、後續透過正規化和分級進行的資料管理，以及選擇之後應用的分析方法，都必須提供進一步的主觀修訂，以便由原始資訊轉為格式化指標。

使用者在規劃研究評鑑時，應留意以下摘要重點：

- **目的：**資料分類符合使用者需求：有些使用者注重學術績效，其他使用者則可能注重經濟或社會效益；有些使用者會以研究領域為準，其他使用者則以標準課程為準。
  - › 使用者需要考量本身目標是否符合分類法的設計原則。
- **類別：**有許多分類法用於指派期刊及 / 或個別出版品至學科類別，沒有一種分類法是唯一正確的選擇。
  - › 使用者在選擇資料來源時，應考量研究人員的產出組合。

- **精細度：**在依據整體基準對引用數進行正規化時，會選擇廣泛或狹隘的重點，以便進行比較或彙總各個年份及學科之間的資料。

› 使用者需要留意精細度，並選擇適當的分類層級。

- **涵蓋範圍：**評鑑單位研究中的所有主題，並沒有獲得所有分類主題公平涵蓋，而學科及地區都同樣需要獲得公平對待。

› 使用者應判定分類法是否針對所有利害關係人公平擷取資料。

本報告已經證實數據分類在研究政策、管理及評鑑方面，是具有參考價值。如果評鑑人員明確掌握目標、需要解決的問題、前述問題的書目計量相關性、可用資料的性質，以及書目計量分析在整體評鑑架構中的地位，就應該繼續處理各項問題，並判定自己是否完全瞭解前述項目意涵，以及在本身目的與資料脈絡下所產生的結果。

# 參考文獻

- Adams, J., Gurney, K., & Jackson, L. (2008). Calibrating the zoom — a test of Zitt's hypothesis. *Scientometrics*, 75(1), 81-95.
- Ahlgren, P., & Colliander, C. (2009). Document-document similarity approaches and science mapping: Experimental comparison of five approaches. *Journal of Informetrics*, 3(1), 49-63.
- Ahlgren, P., Chen, Y., Colliander, C., & van Eck, N. J. (2020). Enhancing direct citations: A comparison of relatedness measures for community detection in a large set of PubMed publications. *Quantitative Science Studies*, 1(2), 714-729.
- Atran, S. (1990). *Cognitive foundations of natural history: Towards an anthropology of science*. Cambridge, UK: Cambridge University Press. ISBN 978-0521438711
- Boyack, K. W., & Klavans, R. (2020). A comparison of large-scale science models based on textual, direct citation and hybrid relatedness. *Quantitative Science Studies*, 1(4), forthcoming.
- Colliander, C., & Ahlgren, P. (2019). Comparison of publication level approaches to ex post citation normalization. *Scientometrics*, 120(1), 283-300.
- Dousa, T. M. (2017). E. Wyndham Hulme's classification of the attributes of books: On an early model of a core bibliographical entity. *Knowledge Organization*, 44(8), 592-604.
- Garfield, E. (1955). Citation indexes for science — a new dimension in documentation through association of ideas. *Science*, 122(3159), 108-111. <http://garfield.library.upenn.edu/papers/science1955.pdf>
- Garfield, E. (1957). Breaking the subject index barrier — a citation index for chemical patents. *Journal of the Patent Office Society*, 39(8), 583-595. <http://www.garfield.library.upenn.edu/essays/v6p472y1983.pdf>
- Garfield, E. (1964). Science Citation Index: a new dimension in indexing. *Science*, 144(3619), 649-654. <http://www.garfield.library.upenn.edu/essays/v7p525y1984.pdf>
- Glänzel, W., Schubert, A., & Czerwon, H. J. (1999). An item-by-item subject classification of papers published in multidisciplinary and general journals using reference analysis. *Scientometrics*, 44(3), 427-439.
- Glänzel, W., & Moed, H. F. (2002). Journal impact measures in bibliometric research. *Scientometrics*, 53(2), 171-193.
- Glänzel, W., & Schubert, A. (2003). A new classification scheme of science fields and subfields designed for scientometric evaluation purposes. *Scientometrics*, 56(3), 357-367.
- Glänzel, W., Thijs, B., Schubert, A., & Debackere, K. (2009). Subfield-specific normalized relative indicators and a new generation of relational charts: Methodological foundations illustrated on the assessment of institutional research performance. *Scientometrics*, 78(1), 165-188.
- Gläser, J., Glänzel, W., & Scharnhorst, A. (2017). Same data — different results? Towards a comparative approach to the identification of thematic structures in science. *Scientometrics*, 111(2), 981-998.
- Handy, C. (1989). *The Age of Unreason*. Boston, MA: Harvard Business School Press.
- Haunschild, R., Schier, H., Marx, W., & Bornmann, L. (2018). Algorithmically generated subject categories based on citation relations: An empirical micro study using papers on overall water splitting. *Journal of Informetrics*, 12(2), 436-447.
- Hirst, G. (1978). Discipline impact factors: A method for determining core journal sets. *Journal of the American Society for Information Science*, 29(4), 171-172.
- Hulme, E.W. (1923). *Statistical Bibliography in Relation to the Growth of Modern Civilization: Two Lectures Delivered in the University of Cambridge in May, 1922*. London: Butler & Tanner, Grafton & Co.
- Janssens, F., Zhang, L., De Moor, B., & Glänzel, W. (2009). Hybrid clustering for validation and improvement of subject - classification schemes. *Information Processing & Management*, 45(6), 683 - 702.
- Kessler, M. M. (1963). Bibliographic coupling between scientific papers. *American Documentation*, 14(1), 10-25.
- Klavans, R., & Boyack, K. W. (2011). Using global mapping to create more accurate document level maps of research fields. *Journal of the American Society for Information Science and Technology*, 62(1), 118.
- Klavans, R., & Boyack, K. W. (2017). Which type of citation analysis generates the most accurate taxonomy of scientific and technical knowledge? *Journal of the Association of Information Science and Technology*, 68(4), 984-998.
- Lewis, G. (1996). The definition of biomedical research subfields with title keywords and application to the analysis of research outputs. *Research Evaluation*, 6(1): 25-36.
- Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology*, 60(2), 348-362.
- Leydesdorff, L., & Bornmann, L. (2016). The operationalization of "fields" as WoS subject categories in evaluative bibliometrics: The cases of "library and information science" and "science & technology studies," *Journal of the Association of Information Science and Technology*, 67(3), 707-714.
- Leydesdorff, L., Comins, J. A., Sorensen, A. A., Bornmann, L., & Hellsten, I. (2016). Cited references and Medical Subject Headings (MeSH) as two different knowledge representations: Clustering and mappings at the paper level. *Scientometrics*, 109(3), 2077-2091.
- McInnes, L., & Healy, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction, *ArXiv e-prints* 1802.03426.
- Milojević, S. (2020). Practical method to reclassify Web of Science articles into unique subject categories and broad disciplines. *Quantitative Science Studies*, 1(1), 183-206.
- Moed, H. F. (2005). *Citation analysis in research evaluation*. Dordrecht: Springer, Chp. 14.3, 187-189.
- Moed, H. F. (2020a). Appropriate use of metrics in research assessment of autonomous academic institutions. *Scholarly Assessment Reports*, 2(1), 1.
- Moed, H. F. (2020b). How evaluative informetrics relates to scientific, sociohistorical, political, ethical, and personal values. *Scholarly Assessment Reports*, 2(1), 9.
- Neuhaus, C., & Daniel, H.-D. (2009). A new reference standard for citation analysis in chemistry and related fields based on the sections of Chemical Abstracts. *Scientometrics*, 78(2), 219-229.

- Perianes-Rodríguez, A., & Ruiz-Castillo, J. (2017). A comparison of the Web of Science and publication-level classification systems of science. *Journal of Informetrics*, 11(1), 32-45.
- Price, D. J. de S. (1965). Networks of scientific papers. *Science*, 149(3683), 510-515.
- Price, D. J. de S. (1980). Foreword. In Garfield, E., *Essays of an Information Scientist*. Volume 3. 1977-1978. Philadelphia, PA: ISI Press, v-ix. <http://www.garfield.library.upenn.edu/essays/v3forewordy1977-78.pdf>
- Pudovkin, A. I., & Garfield, E. (2002). Algorithmic procedure for finding semantically related journals. *Journal of the American Society for Information Science and Technology*, 53(13), 1113-1119. <http://garfield.library.upenn.edu/papers/pudovkinsemanticallyrelatedjournals2002.html>
- Rafols, I., & Leydesdorff, L. (2009). Content-based and algorithmic classifications of journals: Perspectives on the dynamics of scientific communication and indexer effects. *Journal of the American Society for Information Science and Technology*, 60(9), 1823-1835.
- Ruiz-Castillo, J., & Waltman, L. (2015). Field-normalized citation impact indicators using algorithmically constructed classification systems of science. *Journal of Informetrics*, 9(1), 102-117.
- Schubert, A., & Braun, T. (1993). Reference standards for citation-based assessments. *Scientometrics*, 26(1), 21-35.
- Schubert, A., & Braun, T. (1996). Cross-field normalization of scientometric indicators. *Scientometrics*, 36(3), 311-324.
- Shu, F., Julien, C.A., Zhang, L., Qiu, J., Zhang, J., & Larivière, V. (2019). Comparing journal and paper level classifications of science. *Journal of Informetrics*, 13(1), 202225.
- Shu, F., Ma, J., & Larivière, V. (2020). Classifications of science and their effects on bibliometric evaluations. *Scientometrics*, 125(3), 2727-2744.
- Sjögårde, P. (2019). The challenge of categorizing research. *Leiden Madtrics*. <https://leidenmadtrics.nl/articles/the-challenge-of-categorizing-research> Sjögårde, P., & Ahlgren, P. (2018). Granularity of algorithmically constructed publication-level classifications of research publications: Identification of topics. *Journal of Informetrics*, 12(1), 133-152.
- Sjögårde, P., & Ahlgren, P. (2020). Granularity of algorithmically constructed publication-level classifications of research publications: Identification of specialties. *Quantitative Science Studies*, 1(1): 207-237.
- Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24(4), 265-269.
- Small, H., & Griffith, B. C. (1974). The structure of scientific literatures I: Identifying and graphing specialties. *Science Studies*, 4(1), 1740.
- Small, H. (2003). Paradigms, citations, and maps of science: A personal history. *Journal of the American Society for Information Science and Technology*, 54(5), 394-399.
- ubelj, L., van Eck, N. J., & Waltman, L. (2016). Clustering scientific publications based on citation relations: A systematic comparison of different methods. *PLoS ONE*, 11(4), e0154404.
- Traag, V. A., Waltman, L., & van Eck, N. J. (2019). From Louvain to Leiden: Guaranteeing well-connected communities. *Scientific Reports*, 9, article number 5233.
- Van Eck, N. J., Waltman, L., Van Raan, A. F., Klautz, R. J., & Peul, W. C. (2013). Citation analysis may severely underestimate the impact of clinical research as compared to basic research. *PLoS ONE*, 8(4), e62395.
- Waltman, L. (2016). A review of the literature of citation impact indicators. *Journal of Informetrics*, 10(2), 365-391.
- Waltman, L., & van Eck, N. J. (2012). A new methodology for constructing a publication level classification system of science. *Journal of the American Society for Information Science and Technology*, 63(12), 2378-2392.
- Waltman, L., & van Eck, N. J. (2019). Field normalization of scientometric indicators. In Wolfgang Glänzel, Henk F. Moed, Ulrich Schmoch, and Mike Thelwall, Eds., *Springer Handbook of Science and Technology Indicators*. Cham, Switzerland: Springer Nature, 2019, 281-300. ISBN-13: 978-3030025106
- Waltman, L., Boyack, K. W., Colavizza, G., & van Eck, N. J. (2020). A principled methodology for comparing relatedness measures for clustering publications. *Quantitative Science Studies*, 1(2), 691-713.
- Wang, Q., & Waltman, L. (2016). Large-scale analysis of the accuracy of the journal classification systems of the Web of Science and Scopus. *Journal of Informetrics*, 10(2), 347-364.
- Yu, D., Wang, W., Zhang, S., Zhang, W., & Liu, R. (2017). Hybrid self-optimized clustering model based on citation links and textual features to detect research topics. *PLoS ONE*, 12(10), e0187164.
- Zhang, L., Janssens, F., Liang, L., & Glänzel, W. (2010). Journal cross-citation analysis for validation and improvement of journal-based subject classification in bibliometric research. *Scientometrics*, 82(3), 687-706.
- Zitt, M., Ramanana-Rahary, S., & Bassecoulard, E. (2005). Relativity of citation performance and excellence measures: From cross-field to cross-scale effects of field-normalisation. *Scientometrics*, 63(2), 373-401.
- Zitt, M., & Bassecoulard, E. (2006). Delineating complex scientific fields by an hybrid lexical-citation method: An application to nanosciences. *Information Processing and Management*, 42(6), 1513-1531.
- Zitt, M., Lelu, A., Cadot, M., & Cabanac, G. (2019). Bibliometric delineation of scientific fields. In Wolfgang Glänzel, Henk F. Moed, Ulrich Schmoch, and Mike Thelwall, Eds., *Springer Handbook of Science and Technology Indicators*. Cham, Switzerland: Springer Nature, 2019, 25-68. ISBN-13: 978-3030025106

## 關於 ISI 全球研究報告系列

ISI《全球研究報告》(Global Research Reports) 利用我們獨特的行業見解，提供分析、想法和評論，以啟發更多的深入討論。

每一份報告展現了科研資料擁有的巨大潛力，可以為科研評估和科研政策中的管理問題提供資訊，並加速全球科研體系的發展。

已發布的報告：

Profiles, not metrics

研究倫理：瞭解我們的共同責任，營造永續發展的學術生態系統

識別 Web of Science 中的研究前沿：從指標到意義

書目計量資料庫的價值

多作者署名與研究分析

聯合國永續發展目標的研究導覽

## 關於科睿唯安

科睿唯安 (Clarivate™) 是全球領先的專業資訊服務提供者。今天，科睿唯安銳意進取，為使用者提供值得信賴的資訊與卓越的洞見，幫助客戶解決複雜難題，洞察先機，加速創新步伐。我們的專業知識和解決方案覆蓋創新生命週期的每一個關鍵環節，從學術研究和科學發現，到智慧財產權的管理保護，直至實現創新成果的商業化，涵蓋科學研究、生命科學與製

藥、智慧財產權各個領域。更多資訊請參考 [clarivate.com](https://clarivate.com)。

科睿唯安旗下的 **Web of Science™** 集結全球科學研究資訊，以全球最大、出版社中立的引文索引及研究情報平台為基礎，協助學術機構、企業、期刊出版社、政府單位加速科學研究步伐。

科睿唯安 台灣辦公室

台北市信義區松智路 1 號 11 樓

[clarivate.com/zh-hant](https://clarivate.com/zh-hant)