

科技部補助專題研究計畫成果報告 期末報告

開放取用資源與商業系統之比較研究：以引文系統與機構典藏 為例(第2年)

計畫類別：個別型計畫
計畫編號：NSC 102-2410-H-004-221-MY2
執行期間：103年08月01日至104年07月31日
執行單位：國立政治大學圖書資訊與檔案學研究所

計畫主持人：蔡明月

報告附件：出席國際會議研究心得報告及發表論文

處理方式：

1. 公開資訊：本計畫可公開查詢
2. 「本研究」是否已有嚴重損及公共利益之發現：否
3. 「本報告」是否建議提供政府單位施政參考：否

中華民國 104 年 10 月 23 日

中文摘要：本研究旨在比較商業與開放取用兩種學術傳播機制，透過實際操作商業引文索引資料庫(以WOS與Scopus為例)與開放取用引文系統(以綜合型：Google Scholar、Microsoft Academic Search(MAS)；學科型：物理學的arXiv.org、經濟學的RePEc再加上OAIster與OpenDOAR二個世界重要機構典藏為例)，針對其資料收錄的完整性與重複性情況與資料庫本身之差異性等進行比較研究。本研究利用上述六個引文索引資料庫與系統，以及二個機構典藏系統進行檢索作業，分析比較各系統之檢索介面，再就檢索結果之呈現與輸出結果評鑑各個資料庫與系統。本研究之研究樣本為諾貝爾物理學與經濟學2001年至2012年得獎者之著作，透過於各資料庫與系統分別檢索諾貝爾經濟學獎得主與物理學獎得主之著作，所獲之著作清單透過書目資料剖析，包括排序、比對、刪除、聚集與統計等，最後分別就各資料庫與系統所得之資料進行個別量化統計，並相互交叉比對，以分析各系統收錄資料之完整性、重複性與獨特性。本研究結果期望能提供圖書館選擇引文索引資料庫與建置機構典藏，或未來引文系統發展之參考。此外，藉由對各系統完整性評估結果，提出學術評鑑的指標與工具之建議。

中文關鍵詞：引文索引資料庫, 開放取用系統, 機構典藏, 資料庫完整性, 資料庫重複性, 資料庫獨特性, 學術傳播

英文摘要：In this study, scholarly communication system of commercial services and open access will be examined through comprehensiveness, overlap and database variation of coverage via field operations of commercial citation index databases (WOS and Scopus) and open access citation system (general: Google Scholar and Microsoft Academic Search, MAS; disciplinary: arXiv.org of physics and RePEc of economics; prestigious institutional repository: OAIster and OpenDOAR). Retrievals will be conducted in the six citation index databases and two institutional repositories stated above to analyze and compare their retrieval interfaces, and evaluations of each system will be made as well according to presentation and output of retrieval results. Noble laureates in physics and economic sciences from 2001 to 2012 are selected as samples in this study. Records of their publications over time will be retrieved and downloaded from each system, and a computer program will be developed to perform the analytical tasks of sorting, comparison, elimination, aggregation and statistics. Bibliographic records retrieved from the four databases and two repositories will undertake quantitative analyses and cross references to determine the comprehensiveness, overlap and uniqueness of their system coverage. The results of the study may provide better references for libraries to acquire citation index databases, to build institutional repositories, or to create citation index systems on their own in the future.

Suggestions on indices and tools for academic assessment will be presented based on the comprehensiveness assessment of each system as well.

英文關鍵詞：Citation Index Database, Citation Enhanced Database, Open Access System, Institutional Repository, Scholarly Communication, Database Coverage, Database Overlap

開放取用資源與商業系統之比較研究：以引文系統與機構典藏為例

Comparative Study on Open Access and Commercial Systems: an Approach on Citation System and Institutional Repository

摘要

本研究旨在比較商業與開放取用兩種學術傳播機制，透過實際操作商業引文索引資料庫(以 WOS 與 Scopus 為例)與開放取用引文系統(以綜合型：Google Scholar、Microsoft Academic Search(MAS)；學科型：物理學的 arXiv.org、經濟學的 RePEc 再加上 OAIster 與 OpenDOAR 二個世界重要機構典藏為例)，針對其資料收錄的完整性與重複性情況與資料庫本身之差異性等進行比較研究。本研究利用上述六個引文索引資料庫與系統，以及二個機構典藏系統進行檢索作業，分析比較各系統之檢索介面，再就檢索結果之呈現與輸出結果評鑑各個資料庫與系統。本研究之研究樣本為諾貝爾物理學與經濟學 2001 年至 2012 年得獎者之著作，透過於各資料庫與系統分別檢索諾貝爾經濟學獎得主與物理學獎得主之著作，所獲之著作清單透過書目資料剖析，包括排序、比對、刪除、聚集與統計等，最後分別就各資料庫與系統所得之資料進行個別量化統計，並相互交叉比對，以分析各系統收錄資料之完整性、重複性與獨特性。本研究結果期望能提供圖書館選擇引文索引資料庫與建置機構典藏，或未來引文系統發展之參考。此外，藉由對各系統完整性評估結果，提出學術評鑑的指標與工具之建議。

英文摘要

In this study, scholarly communication system of commercial services and open access will be examined through comprehensiveness, overlap and database variation of coverage via field operations of commercial citation index databases (WOS and Scopus) and open access citation system (general: Google Scholar and Microsoft Academic Search, MAS; disciplinary: arXiv.org of physics and RePEc of economics; prestigious institutional repository: OAIster and OpenDOAR). Retrievals will be conducted in the six citation index databases and two institutional repositories

stated above to analyze and compare their retrieval interfaces, and evaluations of each system will be made as well according to presentation and output of retrieval results. Noble laureates in physics and economic sciences from 2001 to 2012 are selected as samples in this study. Records of their publications over time will be retrieved and downloaded from each system, and a computer program will be developed to perform the analytical tasks of sorting, comparison, elimination, aggregation and statistics. Bibliographic records retrieved from the four databases and two repositories will undertake quantitative analyses and cross references to determine the comprehensiveness, overlap and uniqueness of their system coverage. The results of the study may provide better references for libraries to acquire citation index databases, to build institutional repositories, or to create citation index systems on their own in the future. Suggestions on indices and tools for academic assessment will be presented based on the comprehensiveness assessment of each system as well.

關鍵字

引文索引資料庫,開放取用系統,機構典藏,資料庫完整性,資料庫重複性,資料庫獨特性,學術傳播

英文關鍵字

Citation Index Database, Citation Enhanced Database, Open Access System, Institutional Repository, Scholarly Communication, Database Coverage, Database Overlap

目次

前言	1
研究目的	3
文獻探討	4
一、資料庫與系統檢索介面與檢索結果	4
二、資料庫與系統獨特性(unique)和重複性(overlap)	7
三、商業引文索引資料庫、開放取用引文系統與機構典藏	11
研究方法	21
研究結果	24
一、商業資料庫及引文檢索功能比較	24
二、商業性引文資料庫與開放取用引文系統完整性分析	26
三、商業性引文資料庫與開放取用引文系統重複性分析	27
四、商業性引文資料庫與開放取用引文系統獨特性分析	30
結論與建議	32
參考文獻	34

表次

表 1 各資料庫簡介比較.....	19
表 2 諾貝爾經濟學獎得主，2001-2012 年	21
表 3 諾貝爾物理學獎得主，2001-2012 年	22
表 4 各資料庫及系統檢索功能比較.....	25
表 5 諾貝爾經濟學獎得主之作品清單於商業性引文資料庫及開放取用引文系統之完整性比較.....	26
表 6 諾貝爾物理學獎得主作品清單於商業性引文資料庫及開放取用引文系統之完整性比較	27
表 7 諾貝爾經濟學獎得主作品清單於商業性引文資料庫及開放取用引文系統之重複性比較	28
表 8 諾貝爾物理學獎得主作品清單於商業性引文資料庫及開放取用引文系統之重複性比較	29
表 9 諾貝爾經濟學獎得主作品清單於商業性引文資料庫及開放取用引文系統之獨特性比較	30
表 10 諾貝爾物理學獎得主作品清單於商業性引文資料庫及開放取用引文系統之獨特性比較	31

前言

期刊是學術傳播的主要媒介，大學圖書館引進大量期刊資源，在學術傳播環境中扮演相當重要的角色。長久以來，商業出版機構卻對學術期刊加以壟斷，研究者在學術期刊刊登研究成果須付費、圖書館購買學術期刊和資料庫亦須付費，甚至研究者要取用其發表文章要再次付費，出版商藉多方收費以獲取高額的利益，阻礙了學術傳播分享學術資源的意義。尤有甚者，學術期刊訂費高漲不下，例如：美國研究圖書館學會(Association of Research Libraries, ARL)統計於 1986-2006 年共 20 年間，圖書館的經費增加 3 倍，但期刊供應力只成長 51%，形成新世紀的學術傳播危機 (黃鴻珠，民 99)。有鑑於此，學術界的因應方案主要是推動開放取用(Open Access, OA)。

網際網路與資訊科技的快速發展，使得資訊的傳播更加方便，也加速了開放取用資源的成長。隨著開放取用呼聲的越見高漲，人們對於自由取用資訊的需求亦愈迫切，許多政府、學者、組織相繼宣布支持公開取用的權利宣言，例如：2002 年布達佩斯會議宣言(Budapest Open Access Initiative, 2002)，首先提出期刊開放(open access journal)全球以合理的價格或免費使用，以及倡導自我典藏(self-archiving)，建議著者將自己的著作典藏於所屬機構的網站或進行學科典藏。其他尚有：Research Councils UK, 2006；Academy of Sciences of South Africa, 2006；US Public Access to Federally Funded Research Act of 2006 等。

教育部於 2003 年 10 月 16 日左右，首度針對全國 154 所大學校院，採用具國際公認水準的 SCI - Science Citation Index(科學引文索引資料庫)、SSCI - Social Sciences Citation Index(社會科學引文索引資料庫)以及 EI - Engineering Index(工程索引資料庫)，統計我國公、私立大學及技職校院教師發表的學術論文篇數，做為學術評鑑之重要參考。調查結果在全國各大學校院引發巨大震撼。自此，學術評鑑成為大學或研究機構整體或自我評比之重大議題。在學術評鑑要項中，研究出版與發表佔相當大的比例，是研究表現的核心基礎，學術評鑑可自發表文獻的數量與品質——即研究產出及其影響力加以評鑑，該項指標愈來愈常被用作評鑑及決策依據之參考。世界大學排名系統中，研究表現的評比是重要項目之一，例如：英國《時報高等教育補篇》(2011) (Times Higher Education Supplement, THES)的「世界大學排行榜」(World University Rankings)共有五項指標：teaching - the learning environment (30%)；research - volume, income and reputation (30%)；citation - research influence (30%)；industry income -

innovation (2.5%); international outlook - staff, students and research (7.5%)，其中即以教師研究投入與聲譽、論文被引用率做為研究表現的評比項目，且各占 30%之高比例。目前大範圍的學術評量，大都採量化之統計數字，尤其是針對期刊論文的發表，其中生產量的多少往往是第一根據。

但事實上，個人著作「被引用」的實際情況才是真正反映研究成果之「impact」的具體指標。是以，近年來國際、國科會、教育部、各大學與研究機構，有關理工農醫等自然科學學科之學術評鑑，無論是個人或團體，大都朝此方向實施。

1965 年 Eugene Garfield 創造了科學引文索引資料庫(Science Citation Index, SCI)，SCI 樹立了引文索引資料庫的典範，提供了完整之被引用數據，包括個人、文章、機構等被引用次數、自我引用、H index、被何人引用、在哪裡被引用及什麼內容被引用，進而可建置學科的知識地圖並預測研究發展趨勢，引文索引資料庫因而受到研究者與評鑑者廣泛的注意。引文索引可以反映出其研究理論是否被證實或批評；研究方法是否被修正或改進；建議事項是否被嘗試；研究結果是否被加以應用(Garfield, 1983)。研究有沒有被引用，在國際上有沒有造成回應，是學術研究非常重要的一項判準。引文索引所提供的引用數據，不是其他資料庫可以提供的。作者被引用的情形則更是直接且透明的評量作者成就的憑藉。

引文索引資料庫除了提供研究者諸多有力的支持外，對於圖書館作業與資訊服務上也有許多助益，引文索引資料庫中包含的各項統計資料，可協助圖書館判斷某一學科領域之核心作者或核心期刊，以作為圖書館館藏發展選擇的依據。圖書館亦可利用商業引文索引資料庫，建立讀者個人檔進行使用者新知通報或研究趨勢追蹤等個人化服務，進而提升圖書館服務品質。開放取用引文系統更可協助圖書館為讀者提供免費且多元的服務。

繼 SCI 成功後，其他商業機構也開始引文索引資料庫的製作，如 Elsevier 建置的 Scopus，亦是引文索引資料庫發展十分成功的例子。然而，這些商業機構卻不斷提高其產品之價格，阻礙圖書館和研究者使用。在此時機應運下，促成了開放取用引文系統的發展。研發者紛紛發展了在網路上可自由傳播供研究者取用的檢索平台。較具代表性者有：1997 年 Lee Giles、Kurt Bollacker 與 Steve Lawrence 在美國 Princeton, New Jersey 的 NEC Research Institute 共同研發的 CiteSeer 引文系統、Tim Brody 於 2005 年在英國 University of Southampton 所建置的物理、數學與電腦科學主題的 CiteBase 系統等，皆是開放

取用平台提供引文索引資源的例子。此外，Google Scholar 亦是一個開放性的學術傳播平台，其提供文章被引用資源受研究者普遍重視。商業引文索引資料庫與開放取用引文系統發展至今，仍有一些限制，例如：Elsevier 的 Scopus 系統及 Google Scholar 全文回溯的年代以及許多系統亦無法提供全文等。

2000 年後，資訊科技的持續發展與學術傳播模式改變，除了加速開放取用期刊與引文系統之發展，也促成機構典藏的興起。機構典藏開始受到國內外圖書館、博物館與檔案館的共同關注，成為一個新應用的計畫。機構典藏的定義是以儲存及保管一個機構其組織內的數位資源及知識資產為主，並作為一機構文獻典藏之長期使用為目的(Branin, 2009)。機構典藏的內容是否與商業引文索引資料庫或其他開放取用引文系統相同，其差異為何，彼此之間是否可以相互取代或輔助補充？是一值得關注的議題。

隨著傳播媒體的改變，學術傳播管道亦不斷演進，如何精確嚴謹的生產、檢索、使用、並快速的傳播與評鑑學術資訊，是一大課題。商業與開放取用之學術傳播模式，目的、功能和方法差異甚大，因此引發本研究加以探討。本研究旨在透過商業性引文資料庫、開放取用引文系統以及機構典藏，深入比較此兩種學術傳播機制的特色及所涵蓋資源的完整性、重複性與獨特性。以期能提供圖書館選擇引文索引資料庫與建置機構典藏，或未來引文系統發展之參考。此外，藉由對各系統完整性評估結果，提出學術評鑑的指標與工具的建議。

研究目的

本研究基於開放取用資源的興起，旨在比較商業與開放取用兩種學術傳播機制，透過實際操作商業引文索引資料庫(以 WOS 與 Scopus 為例)與開放取用引文系統(以綜合型：Google Scholar、Microsoft Academic Search(MAS)；學科型：物理學的 arXiv.org、經濟學的 RePEc 再加上 OAIster 與 OpenDOAR 二個世界重要機構典藏為例)，分二年進行，針對其資料收錄的完整性與重複性情況與資料庫本身之差異性等加以研究，期望達成下列目的：

- (1) 探討商業引文索引資料庫與開放取用引文系統與機構典藏之優缺點，及各資料庫與系統之建置背景、收錄年代、資源類型、語言、學科等方面之差異性。
- (2) 比較商業引文索引資料庫與開放取用引文系統在收錄資源的完整性、正確性、重複性與獨特性。

(3) 針對 WOS、Scopus、Google Scholar、MAS、RePEc 及 arXiv.org 六個引文系統以及 OAIster 與 OpenDOAR 二個機構典藏系統分析其優劣，並根據分析結果加以提出建議，以期能對國內圖書館之作業與服務，機構典藏或學術評鑑上，提供更完整的背景資料作為參考及選擇之依據。

文獻探討

本研究以諾貝爾經濟學獎得主與物理學獎得主之著作為研究樣本，透過檢索不同商業引文索引資料庫及開放取用引文系統，以其檢索結果評鑑兩者在收錄資源的完整性、正確性、重複性與獨特性之差異。因此以下就資料庫與系統檢索介面與檢索結果、資料庫與系統獨特性和重複性、不同商業引文索引資料庫、開放取用引文系統與機構典藏等相關文獻進行論述。

一、資料庫與系統檢索介面與檢索結果

根據柏克萊大學圖書館網頁 Evaluating Web Pages: Techniques to Apply & Questions to Ask UC Berkeley - Teaching Library Internet Workshops (2012)上之說明，一個好的搜尋引擎應該具備以下條件：(1)網路文件資料庫：資料庫的大小(Size)、資料庫更新度、文本的完整度、被提供的形式、速度和協調性。(2)搜尋引擎的功能：檢索欄位的設計、檢索語言的限制等等。(3)結果展示：排列及展示方式的多元性。林佳怡(民 100)針對 Web of Science(WoS)、Scopus、中文社會科學引文索引(CSSCI)、中國引文資料庫(CCD)、臺灣社會科學引文索引資料庫(TSSCI)、臺灣人文學引文索引資料庫(THCI)以及學術引用文獻資料庫(ACI)等七個引文索引資料庫，經由實際檢索與觀察，分析並討論各系統之來源文獻收錄情形、查詢功能、索引欄位、檢索結果輸出欄位以及引文分析功能與個人化功能等，並輔以資料庫使用者、圖書館館員以及資料庫建置專家之訪談，藉此瞭解引文索引資料庫之現況與不足之處。黃鴻珠(民 99)於《變動中的圖書資訊服務》一書第二章針對開放取用期刊之建立背景、期刊整理與檢索方式、期刊的學術傳播模式與流程加以整理與分析。

eTesting Labs(2000)以 Google、AltaVista、Northern Light 及 MS Search、DirectHit、Fast 六個搜尋引擎為研究對象，利用簡易單字(simple one-word)、簡易多字(simple multi-word)、複雜多字(complex multi-word)、自然語言(natural language)及首頁標示(home page target)五種查詢測試評鑑其優劣。此研究評分是以檢索到的網頁相關程度佔 10 分，網頁是否是活的(live)佔 5 分、第一筆輸出的資料相關程度

佔 2 分、第二筆輸出的資料相關程度佔 1 分為依據。最後根據每日每種查詢方式被使用的次數比重得知，各種查詢結果為：簡易單字與簡易多字查詢各佔 30%、複雜多字與首頁標示查詢各佔 20%，自然語言查詢則佔 10%。

Bar-Ilan (2000)利用六個大型的搜尋引擎，檢索「informetrics」這個詞彙之後，進一步採用內容分析法分類了近 800 個獨立的網頁，結果發現約 40% 的網頁可以粗略地歸類為參考書目，至於其它類別尚有國際科學與資訊計量學會、虛擬圖書館以及網路計量學。Allen(1999)採用專家評鑑法，調查科學網站的可信度。該調查以生物進化與基因改造二個問題為範例，每一個問題由二位專家分析評估最初找到的 500 個網站，直到每一位專家個別審查完大約 60 個與主題相關的網站。每個問題的 60 個網站再以三種方式評分。若網站包含錯誤的資訊，則評為「錯誤」，若網站胡亂解釋科學，則評為「誤導」。至於提供的是無同儕評審的資訊則評為「無參考」。研究結果顯示，在生物進化方面的網站錯誤與誤導率達 87.8%；基因改造的網站錯誤與誤導百分比為 82.8%。至於「無參考」的網站在二個問題的分佈率都在 48% 以上。由此可見網站資訊的正確性有待加強。

吳樹華(2008)以產業界熟悉 Metasearch 的介面設計專家、學者與不同學科背景之研究生為對象，對 Metasearch 檢索介面進行使用評估研究。研究結果發現，受試者普遍認同 Metasearch 一次查全功能的便利性。但一般受試者較重視檢索結果排序與組織，希望系統能自動集中重複的結果，呈現相關排序並給予適當的群集分類，以方便篩選與判斷。故作者認為系統應考量哪些功能的需求頻率較低，不符合使用情境；哪些重要或具有特色的功能，隱藏在使用者不易查覺的地方，不便於快速直覺的操作；哪些工作是使用者希望系統主動完成。若可改善被動、複雜或不符合使用習慣的設計，將有助於提升 Metasearch 之接受度。

Miller 與 Pellen(2009)針對 Central Search、Google Scholar 及 Windows Live Search Academic 進行了搜尋特徵(Search Features)與檢索選項(Retrieval Options)兩個面向的比較。搜尋特徵之比較項目包含：標題、段落、布林邏輯、作者、全文、關鍵字、主題、摘要、國際標準書號(ISBN)、國際標準期刊號(ISSN)、同儕評審、出版社、時間及學科領域等 14 個項目。檢索選項之比較項目則包含：學術期刊、全文、全文過瀘、以相關度排序或群集、按日期排序、按作者排序或群集、按期刊排序、資料匯出、以及語言限制等 10 項。比較結果顯示，兩個面向的比較都以 Central Search 各包含 11 項和 8 頁為表現最好。

馬建玲與田曉陽(2010)，分析 Google Scholar(GS)搜尋引擎的引文檢索功能，其研究發現(1)GS 沒有作者權威檔，故在作者姓名的檢索上無法得到精確的檢索結果；(2)缺少作者機構限制的檢索功能，易檢索到不相關的結果；(3)不支持主題和分類檢索，亦無刊物及人名索引，影響結果的準確性；(4)文獻類型標示不夠清楚，讀者無法確定所使用之文獻類型為何；(5)沒有清楚標明檢索結果為付費資源或是開放取用資源；(6)檢索結果的排序只有依關聯性或時間排序，選擇不多。由此可見 GS 在檢索功能的設計上還有很多可改善的空間。

Bar-Ilan(2010)，以 Leo Egghe 和 Ronald Rousseau 的著作「Introduction to Informetrics」為樣本，針對 Web of Science(WOS)、Scopus 及 Google Scholar(GS)三個資料庫進行該書被引用次數之比較研究。研究結果顯示，WOS 及 Scopus 的引用次數相互涵蓋性達 90%，在此 90%裡，有 30%是 GS 未涵蓋的。但 GS 所涵蓋的 108%卻是 WOS 及 Scopus 都未涵蓋的。由此被引用文獻的引用次數分析研究可看出各資料庫之相同及相異之處。

Johnson(2010)整理了 D. Stone 等人於 2005 年對使用者介面所提出之建議，包含以下幾點：(1)可視性(Visibility)：使用者可快速認識系統與操作方式；(2)示能性(Affordance)：讓使用者知道介面之用途；(3)回饋性(Feedback)：讓使用者知道發生了什麼事、什麼結果；(4)簡單性(Simplicity)：畫面盡可能做到最簡單並把焦點放在使用者所給的指令；(5)結構性(Structure)：理性的內容組織；(6)一致性(Consistency)：介面上相同的功能之設計應有一致的用法；(7)容錯性(Tolerance)：防止出現錯誤並協助修復；(8)取用性(Accessibility)：可供所有使用者使用，並克服阻礙、取用設備或環境的限制。

Adriaanse 與 Rensleigh(2011)，以 2004 至 2008 年南非環境科學學術期刊(The Environmental Sciences Scholarly Journals in South Africa)為樣本，針對 WOS、Scopus 和 Google Scholar(GS)進行了宏觀與微觀之比較研究。宏觀研究以 26 個評價面向比較三個資料庫，其中包含內容、取用、服務、介面、搜尋、搜尋結果、費用、引文及分析工具以及連結功能等。微觀研究則為資料庫收錄資料領域及資料完整度之比較。宏觀研究結果發現 Scopus 優於 WOS 和 GS；微觀研究則發現 WOS 優於 Scopus 及 GS。根據以上兩個面向之研究結果可知，GS 在該研究樣本的檢索無法取代需費用使用的 WOS 和 Scopus，只能做為補充檢索之用途。

二、資料庫與系統獨特性(unique)和重複性(overlap)

重複性和獨特性的研究範圍十分廣泛，各種資料產生來源包含出版社、資料庫及搜尋引擎，及各種資料來源類型如期刊文獻、專利文獻、網路資源等，也都是重複性與獨特性研究的範疇，重複性和獨特性的多寡，也是圖書館選擇索引與摘要資料庫的重要依據。圖書資訊學界探討期刊範圍與索引摘要資料庫重複性的研究始自 Bradford(1937)。重複性是指相同的文章被二個或二個以上資料庫同時收錄之數目的比率。1960 年代中期，Martyn(1967)收集了特定學科的書目資料，檢索摘要期刊(當時還沒有電腦資料庫)中的作者欄位，計算數個摘要期刊中的參考書目，最後進行分析比較索引與摘要期刊間的重複性。其研究結果顯示，摘要期刊的收錄範圍，大約檢索到 70%的相關資料，且超過一半的文獻被收錄在一個以上的索摘期刊中。Martyn 總結說明了索引與摘要資料庫重複性雖具有優點，然而其亦會造成索摘製作者之間的對立與競爭。

Wood(1972)針對「生物學摘要(Biological Abstracts, BA)」、「化學摘要(Chemical Abstracts Service, CA)」與「工程索引(Engineering Index, EI)」三個資料庫進行紙本出版品與機讀文獻之間的關係，以及收錄資料的重複性研究。觀察三個資料庫截至 1970 年 5 月 1 日為止所收錄的 14,592 種期刊，並針對三個資料庫中的標題欄位進行分析。研究結果顯示，有 1%的期刊同時被三個資料庫收錄，27%的期刊被兩個資料庫同時收錄，而有 72%的期刊僅為一個資料庫收錄。此研究所得之資訊能作為三個資料庫未來合作計畫與政策實施協調之參考依據，以提供使用者社群更有用的資訊資源與服務。

Larson(1976)以生命科學資料庫(BioSciences Information Service, BIOSIS)所收錄之期刊文章的紙本書目著錄為對象，比較其形式、格式與資料內容的異同。研究結果發現，生物學摘要(Biological Abstracts, BA)、化學摘要(Chemical Abstracts Service, CA)與工程索引(Engineering Index, EI)三個子資料庫共使用 15 個不同的書目著錄單元，其中有 8 個為三個子資料庫所採用，4 個為其中兩個資料庫所用，剩餘的 3 個僅由其中一個資料庫使用。三個資料庫的紙本書目著錄在內容上的差異並不會妨礙使用者對收錄書目的辨識。

Doszkocs(1977)認為在大規模操作多種書目資料庫檢索系統中，當相同的引用文獻不只出現在一個資料庫時，即為收錄範圍重複。辨識與確認資料重複的程度對系統的設計、管理與使用是有意義的。本研究開發 DUPCHECK 程式，測試 CANCERLINE、MEDLINE 以及 TOXLINE 三個資料庫資料重複

率。藉由統計抽樣方法與技術，研究結果呈現多個資料庫重複程度的最小估計值。再者，並提出各資料庫如何正確比較欄位，以及線上檢索引文獻檔案的書目範本，供後人參考。

Poyer(1984)亦表示當相同的期刊文章被 2 個或 2 個以上的資料庫收錄並製作索摘，即謂之為期刊文獻重複性。該研究收集 70 篇醫學學位論文的參考文獻，針對 Index Medicus、Science Citation Index(SCI)、Biological Abstracts(BA)及 Chemical Abstracts(CA)四個資料庫進搜尋比對，檢視收錄期刊文章的重複性。在 7,969 篇被引用的期刊文獻中，有 92%被 2 個或 2 個以上的資料庫所收錄；其餘文獻中有 591 篇僅被 1 種資料庫收錄，55 篇並未被上述任一資料庫所收錄。作者歸納資料庫間重複性受資料庫更新時間、期刊標題選擇及資源收錄政策影響。

Nicholls(1989)透過搜尋教育文獻資料庫(Educational Resource Information Center, ERIC)、資訊科學摘要(Information Science Abstracts, ISA)、圖書館學與資訊科學資料庫(Library and Information Science Abstracts, LISA)以及圖書館文獻索引(Library Literature, LL)四個圖書資訊學資料庫中所涵蓋，1970 年至 1989 年間鐳射影碟應用於圖書館的相關文獻。此外，說明了四個資料庫包含不同的收錄範圍，例如：Information Science Abstracts 重視研討會論文和專利文獻；LISA 包含大量的外國語文資料及資訊科學期刊，反而較少圖書館學期刊；Library Literature 則包含較多書籍、書評期刊；ERIC 則是收錄最完整的資料庫。

Mychko-Megrin(1991)選擇醫學領域 41 種主題共 7,281 篇論文與專書書目(1,655 為蘇聯出版，5,626 為他國出版)，研究 5 個生命科學資料庫的涵蓋範圍，比較各種書目摘要資料庫 1970-1988 年間的學科收錄情形。書目來自 3 個不同的樣本，但涵括所有蘇聯的醫學主題文獻。本研究提供 5 個資料庫中生物醫學期刊的國家分布情形，並建議有 12 種索引可用以評估生物醫學資料庫的品質。最後，研究結果呼籲減少重複性，始可增加各資料庫中的涵蓋範圍。

Sievert 及 Andrew(1991)研究資訊科學摘要(Information Science Abstracts, ISA)的重複款目，以探討此資料庫索引的一致性。研究結果顯示將近一半的索引完全一致，另一半則否。ISA 的索引政策規定每篇文獻給定 1 個主要標目與 1 至 2 個次要標目。標目數量的限制與 ISA 本身過少的主題詞表，可能是造成兩極化分布的原因。從控制詞彙表中選出的敘述詞(descriptors)之索引一致性最高，而從自然語言中選出的識別語(identifier，又稱關鍵字)之索引一致性最低。根據所得資料顯示，當一篇文獻的標目

越多，其索引的一致性就會越低。

Bharat 及 Broder(1998)以 HotBot、AltaVista、Excite 與 Infoseek 四個搜尋引擎為對象，使用統計方式探究這 4 個搜尋引擎於 1997 年年中與 1997 年 11 月收錄網頁的總數，並以百分比方式呈現容量大小與重複性的估計值。結果顯示這四個引擎截至 1997 年 11 月收錄的網頁容量大小分別為 77M、100M、32M 與 17M，總頁數達 1 億 6 千多萬頁。靜態公開網站截至 1997 年 11 月止之網頁超過 2 億多頁。4 個搜尋引擎的重複性相當低，僅 1.4%、220 多萬個網頁同時為 4 個搜尋引擎所收錄。

許多研究曾試圖探討索摘資料庫在不同時間的收錄情形，包括學科範圍、完整性與一致性等。線上與光碟資料庫使蒐集資料與確定收錄情形的過程更加容易。Jacso(1998)使用進階檢索方式於 DIALOG 資料庫的「期刊名稱查詢(Journal Name Finder)」子資料庫，利用 Dialog 複本刪除系統，藉以精簡 42 種圖書資訊學(library and information science, LIS)重要期刊，於 1966 年至 1996 年間在六個資料庫中的收錄情形。使用多個資料庫比較同一段時間內各資料庫收錄期刊的範圍與重複性，是相當困難的工作。其研究結果發現約有 10% 左右的重複性。Jacso 的研究方法可用以建立期刊收錄情形的資料庫，應用在資料庫評鑑、書目計量研究、資料庫採購與授權合約簽訂等。

Hood(1998)調查線上書目資料庫的書目記錄分布情形，以 1965 年至 1993 年間與模糊集合論(Fuzzy Set Theory)相關的主題文獻為調查對象。由 Dialog 系統下之所有資料庫下載包含，「fuzzy」的資料，並選出和「Fuzzy Set Theory」相關的文獻，約由 114 個不同的資料庫中選出超過 30,000 筆資料。每個資料庫的收錄筆數與資料庫間重複的筆數皆會進行計算。該研究開發 6 種不同的計算方法，用以查找並處理資料庫間重複紀錄的問題。若將重複記錄計算在內，最佳的資料庫收錄筆數佔總數的 19%；若去除重複的記錄，則最佳的資料庫收錄筆數佔總數的 37%。結果顯示，資料庫的記錄分布符合「布萊德福—齊夫定律(Bradford-Zipf)」的雙曲線分布。

Read 及 Smith(2000)以 DIALOG 資料庫版本的 Information Science Abstracts, (ISA)以及其他兩個競爭對手：圖書資訊學摘要(Library and Information Science Abstracts, LISA)與圖書資訊學文獻索引(Library Literature and Information Science, 又稱為 Library Literature)為對象，研究三者 1999 至 2000 年間索摘文獻的主題範圍、重複性與資料類型等收錄情形。在 20 個圖書資訊學的主題檢索中，Library Literature 在其中 16 次搜尋回傳最多的筆數，LISA 有 3 次回傳最多筆數，有 1 次兩者持平；ISA 則是所有主題

檢索中回傳筆數最少的資料庫，這和其原本收錄資料筆數較少有很大的關係。就任一主題來看，三者收錄重複文獻的百分比最高僅達 21%，顯示至少須搜尋 1 個以上的資料庫才能獲得較完整的結果。收錄資料的類型主要為期刊文獻與特稿。

Walters 和 Wilder(2003)針對 7 個特殊學科資料庫和 5 個跨學科資料庫，有關「晚年遷徙(later-life migration)」主題的相關文獻，進行內容研究，以評鑑 12 個不同資料庫索引摘要該主題文獻的表現。研究問題包括單一學科和跨學科資料庫，哪一種資料庫涵蓋範圍較完整？學科中的核心文獻是否也會被較多資料庫索引？以及資料庫之間的重複性是否可以反應學科之間的相似性？研究結果顯示，佔該領域文獻出版量 40% 的 5 種期刊中，3 種屬社會老人學(social gerontology)、1 種屬鄉村社會學(rural sociology)、1 種為區域科學(regional science)。出版最多晚年遷徙文獻的學科不見得是索摘收錄最完善的學科。4 個跨領域資料庫較其他單一學科資料庫的索摘收錄更完整。不同資料庫之間的平均重複率約在 45% 左右，12 個資料庫間相對低程度的重複率，顯示欲研究晚年遷徙此一主題的學者需要倚賴更多種書目工具。

Jasco(2005)比較 Web of Science、Scopus 與 Google Scholar 三個引文索引資料庫，他以資訊科學之父 Bush 之名著 *As We May Think* 為例，檢索三個資料庫，結果發現因該文章發表在非學術性質的期刊 *Atlantic Monthly*，因此 WOS 與 Scopus 均未收錄，故較少被引用，然而 Google Scholar 卻有被引用連結。WOS 檢索結果顯示共被引用 712 次，其中 90% 集中在 1975 年至 2005 年，在 1999 年被引 45 次達到高峰，其餘 10% 分布在 1945-1974 年。反觀 Scopus 只檢索到 267 次，因其收錄年代自 1995 起。2008 年，Ahmad 與 Jeelani 針對生物技術主題，使用 20 個查詢結果，進行 Google、AltaVista、HotBot、Scirus、Bioweb 五個搜尋引擎的重複性研究。該實證性之研究顯示 HotBot 與其他搜尋引擎之重複性最高，其次是 Google。此外，該研究發現複合和複雜式的檢索查詢會產生更多的重複結果 (Rather, Lone & Shah, 2008)。

陳薇竹(民 97)選取商業引文索引資料庫(SCIE 與 Scopus)以及開放取用引文系統(Google Scholar 與 CiteSeer)進行實證研究，選取電機學界的諾貝爾獎「杜林獎」，在 1966 年到 2006 年間得獎的 50 位得主為樣本，查詢四個引文索引資料庫與系統所收錄 50 位作者於 1990 年至 2005 年發表的著作，再藉由評比檢索結果筆數及檢索細項、四個引文索引資料庫與系統個別與交叉重複性及獨特性以及重複資料來源，比較四個引用文獻索引資料庫之資料完整性及優劣。其研究結果發現：傳統商業引文索引資料

庫具有較完整的檢索機制；商業引文索引資料庫有較完整的書目著錄格式；Google Scholar 的資料收錄較齊全，然而 SCIE 在學術出版收錄較完整；美國計算機械學會出版品及網路資源，佔電機學科文獻的重要位置。

2011 年，Esmaeil、Mohammad、Kiaie 與 Kiaie 的研究在比較搜索引擎系統及資料庫的重複性。此研究群體包含六個開放取用的搜尋引擎系統，這些搜尋引擎來自 searchenginewatch.com 網站所介紹之最常用的搜索引擎系統。該研究利用數據分析和 Microsoft Excel 計算其頻率分布，再使用百分比和平均數來繪製表格和圖表。研究結果發現，在不同的搜索引擎系統中，Yahoo 與其他搜尋引擎系統之重複率最高，約達 40%。Curry Guide 搜尋檢索物理學門資訊的回現率為 77.1%。此外，該系統與其他系統有約 43.7% 的重複率。因此，以物理學門上的重疊程度可以說 Meta Search Engine 算最好的搜尋引擎 (Esmaeil, Mohammad, Kiaie M. & Ketab, 2011)。

Wang(2012)等人的研究認為目前網際網路發展迅速，醫療健康資訊的取得也逐漸普及。此研究對 Google、Yahoo!、Bing 和 Ask.com 四個搜尋引擎的檢索研究，以乳腺癌為檢索詞彙來進行搜尋。其研究結果為四大搜尋引擎中對乳腺癌的六大標準的檢索結果都有擠身前 30 名。而就有效性而言，Google 為最好的，其次為 Bing 與 Ask.com，最差為 Yahoo!。此外，於搜尋引擎的重複性來說，其兩兩搜尋引擎的資料都約有高於 50% 的重複性。再者，就使用者滿意度來說，四大搜尋引擎皆有其強調的不同類型之內容。對於使用者而言，Bing 的滿意度最好，其次為 Yahoo! 與 Google，最差的是 Ask.com。此研究認為如果使用者需要檢索醫療資訊，應該依據不同的資訊類型去檢索適合的搜尋引擎才能事半功倍 (Wang, Wang, Michael, Yong & Xu, 2012 May)。

魯佑敏(2013)以 1990~2011 年 30 位杜林獎學者 1990~2012 年(10 月)約 23 年的作品書目資料為樣本，以作者名稱至商業性引文資料庫 SCIE、Scopus 及開放取用系統 Google Scholar、Microsoft Academic Search、CiteSeer 檢索，發現 SCIE、Scopus 收錄資源是以學術研究為主，Google Scholar 資料多元，SCIE 和 Scopus 收錄資源重複性非常高。

三、商業引文索引資料庫、開放取用引文系統與機構典藏

商業性引文資料庫仰賴於營利性出版商的建置和維護，而商業性引文資料庫具有資料完整、學科範圍多元且選擇資料政策嚴格等特色，因此其資料具有權威性。以下介紹本研究所採用的 Web of

Science 及 Scopus 兩種商業性引文資料庫的背景及現況。

(一)Web of Science

Web of Science(簡稱 WOS)為美國 Thomson Reuters 公司於 1997 年間建置之網際網路版引文索引資料庫，提供使用者理、工、醫、農、人文、及社會科學等各學科領域之文獻書目、作者摘要、及引用文獻等資料。該系統收錄期刊超過 1 萬 3 千種，每週更新其內容，每年提供超過 110 萬筆書目及 2300 萬筆引用文獻資料。本資料庫特有之文獻間相互引用關係檢索，不但提供研究人員更完整之研究參考資訊，更可依循前人研究軌跡，進一步瞭解同儕及競爭對手之研究(財團法人國家實驗研究院，2010)。

Eugene Garfield 博士認為引文索引資料庫的期刊收錄需以經濟方面考量，製作一個無用的文獻索引與一個有用的文獻索引所花的成本是一樣的。因此，期刊收錄範圍就成為商業考量。索引必須嚴格限制它的收錄範圍，盡可能地只收錄那些可能對使用者有用的資訊。Garfield 博士於 1961 年著手建立了 ISI 編輯發展部，其期刊選擇模式沿用至今。ISI 的期刊選擇模式包括四個面向：期刊出版標準、編輯內容、國際化多樣性和引文分析。以上四面向無法比較孰重孰輕，每一項評價都有其作用與價值，每個考察因數都有其側重面(ISI Web of Knowledge，2013)。

Web of Science 共涵蓋下列三大資料庫：(1)Science Citation Index Expanded(SCIE)，SCIE 收錄科學技術類期刊計 8,250 餘種(較光碟版約多 2,000 餘種)，涵蓋主題約 150 餘種。目前回溯資料到 1900 年，1991 年起之資料含作者摘要。(2)Social Sciences Citation Index(SSCI)，收錄社會科學類期刊計 2,850 餘種，亦從近 3,300 種科技類期刊中挑選相關資料收錄，涵蓋主題約 50 餘種。收錄年限自 1965 年起，1992 年起之資料含作者摘要。(3)Arts and Humanities Citation Index (A&HCI)，收錄藝術人文類期刊計 1,540 餘種，亦從近 7,000 種科學和社會科學類期刊中挑選相關資料收錄，主題涵括藝展評論、戲劇音樂及舞蹈表演、電視廣播等。收錄年限自 1975 年起，1999 年起之資料含作者摘要(ISI Web of Knowledge, 2013)。

(二)Scopus

荷蘭著名的 Elsevier 出版集團(www.elsevier.com)，是一個以科學、技術和醫學資訊產品的出版社。集團每年出版 1,800 多種期刊和 2,200 本新書，以及一系列創新性的電子產品，例如：Science

Direct, MD Consult、摘要型資料庫、線上參考書目及特定學科入口網等。此集團在 2004 年 11 月推出了 Scopus 摘要資料庫。該資料庫具有不同以往的資訊檢索、科學分析及導航功能(馮玉, 2007)。

Scopus 最大的特色之一是收錄來自研究機構的機構典藏與學術網頁等網路資源及專利文獻, 其獨有的網頁引用與專利引用資訊可提供研究者不同於 Web of Science 的學術期刊引用資訊, 讓學術研究內容在期刊引用外的網頁與專利引用資訊也能被發現。而其更與全球最大的線上科技搜尋引擎 Scirus 做整合, 以便藉由蒐集這些網頁的參考文獻來提供網頁引用分析的功能(馮玉, 2007)。

Scopus 的文章收錄原則為依據使用者意見為主, 由 Scopus 系統開發組, 根據市場調查研究結果, 進行期刊選擇。使用者也可直接在網頁上推薦期刊, 再由內容選擇委員會(Content Selection Committee) 進行評選, 內容選擇委員會也會去調查研究者和圖書館所使用的新資料來源, 以尋求最適合研究者之完整的資料。Scopus 收錄期刊的原則包含: (1)由於 Scopus 界定為學術索引摘要資料庫, 故被收錄者必須為學術性刊物; (2)期刊需有同儕審查制度, 以確保文章的品質; (3)文章必需提供英文摘要, 以協助讀者初步了解內容; (4)期刊必須定期出版, 以保持資料庫收錄文獻的穩定性(Scopus, 2013)。

Scopus 收錄來自全球 5,000 個出版公司, 至 2014 年 5 月收錄經同儕審查的期刊數量高達近 20 萬種, 內含 3,800 萬筆, 記錄 2,480 萬筆專利資料及 2 億 5 千萬筆參考文獻, 並每天新增 5,500 筆資料。同時也收錄會議論文集以及專利資料等相關學術資源, 以確保其廣泛的跨領域特性。此外, Scopus 不僅提供文獻之引文追蹤, 並利用其清晰又簡便之介面, 直接整合搜尋網頁資源引用及專利引用。Scopus 還可直接連結到全文文章、圖書館的資源、及其他書目管理軟體等應用, 所以比其他任何文獻搜尋工具, 速度更快, 使用更簡便, 涵蓋範圍也更廣泛。其收錄範圍及學科領域有物理與工程、健康科學、社會科學、心理學與經濟學、生命科學為多數(Elsevier, 2014)。

本研究所選取之開放取用引文系統總共有 Google Scholar、Microsoft Academic Search、RePEc、arXiv.org、OAIster、OpenDOAR 共六個系統。其中, 可分為開放取用期刊論文系統及機構典藏系統二大類, Google Scholar、Microsoft Academic Search 為綜合性學科搜尋引擎系統, RePEc 與 arXiv.org 為專門性學科開放取用期刊論文系統, 兩者可歸為開放取用期刊論文系統; 機構典藏系統則有 OAIster、OpenDOAR, 其中 OpenDOAR 為機構典藏集合系統。以下分別介紹各開放取用引文系統:

(一)Google Scholar

Google 創辦人 Larry Page 和 Sergey Brin 在 1995 年結識於史丹佛大學。1996 年，他們聯手打造了原名為 BackRub 的搜尋引擎系統，利用連入連結來判斷個別網頁的重要性。1998 年正式改名為 Google。Larry Page 曾說過：「完美的搜尋引擎系統必須確實理解使用者的意思，提供恰如所需的資訊。」Google 在他發表這番言論後逐漸成長茁壯，服務範圍不再侷限於搜尋，但是追求完美的精神始終不變。從搜尋到 Chrome 到 Gmail，Google 提供的技術一應俱全，目的是讓研究者輕鬆找到所需資訊，以及順利完成各種工作(Google, 2013)。

Google Scholar 於 2004 年開始營運，它是一個免費的引用文獻系統，利用 Google 強大的搜尋引擎進行學術資源的檢索，雖然 Google Scholar 是免費使用的工具，但檢索到的資源並非全都是免費資源，屬於付費系統內的文獻內容只能在購買者的區域網路內才能使用，通常購買付費系統由各大學圖書館或研究機構所負責，只要所屬大學或研究單位的圖書館有購買須付費的學術系統，機構內的同仁皆可利用 Google Scholar 檢索利用付費系統的內容。Google Scholar 可搜尋到的學術資源有學術文章、同儕評鑑文章、預印本、圖書、學位論文、技術報告等，資料來源為學術資料庫、各大學機構典藏系統、預印本資料庫以及各學門的相關學會，並包含 Online Computer Library Center(簡稱 OCLC)的 WorldCat 資料庫中學術或非學術的書目資料，此外，研究者也可主動申請讓自己學術文獻網頁被 Google Scholar 檢索到。

在經營的模式上，Google Scholar 利用搜尋器在各合作者的資料庫和網路中漫遊，其搜尋器被放置在多種不同資料來源，包含(1)許多大型且知名的出版社和大學機構的資料庫中，例如：美國電子電機工程師學會(Institute of Electrical and Electronics Engineers, IEEE)、美國計算機協會(Association for Computing Machinery, ACM)、Macmillan 與 Wiley 出版社、University of Chicago 等；(2)雲端伺服器，例如：HighWire Press、Meta Press、Ingenta；政府及其他學術組織和政府基金會，例如：美國物理學會(the American Physical Society, 簡稱 APS)、國家衛生研究院(National Institute of Health, NIH)、美國海洋暨大氣總署(National Oceanic and Atmospheric Administration, NOAA)；(3)預印/重印服務者，例如：arXiv.org、Astrophysics Data System、Research Papers in Economics 及 CiteBase。此外 Google Scholar 也歡迎個人使用者將其作品及個人的資料系統加入 Google Scholar 搜尋器的清單中，讓使用者在 Google Scholar 能夠搜尋更多更豐富的資料(Jacsoó, 2005)。

Google Scholar 做為一個開放取用的學術資源搜尋系統，在操作上仍有以下缺點，在使用時必須注

意：(1)有些全文資料的取得來源描述不一致，使用者檢索到的文獻會包含許多不同的版本，在取用時須檢查來源是否正確；(2)所收錄的學術資源主題以科技及醫學領域為主，社會科學或人文學的文獻較為缺乏；(3)年代久遠的期刊文獻不容易找到，其中包含 JSTOR 的期刊檔案，雖然可檢索的資料非常多，數量也很龐大，但對部分重要文獻的檢索結果並不完整；(4)檢索功能較為簡易，使用者無法使用更進階詳細的檢索策略；(5)引用數據並非根據文獻的參考書目，從檢索結果可發現許多近期的文獻反而被早期的文獻引用，提供錯誤的引用資料並以此數據進行檢索結果的排序將對使用者造成誤導，可能會讓使用者對文獻的影響力有錯誤的認知，在選擇文獻時也可能因此錯過更重要的相關文獻(吳孟瑾，2013)。

(二)Microsoft Academic Search

Microsoft Academic Search 是微軟亞洲研究院 2009 年開發的開放取用學術搜索引擎系統。為研究員、學生、圖書館員和其他使用者查找學術論文、國際會議、學術期刊、作者和學術等文獻，多達 3800 萬的學術文獻和 1900 萬個學術作者。微軟學術搜索系統為增加內容及檢索功能效率，其內容比原先預期的範圍更小。與傳統的搜索引擎系統比較，微軟學術搜索系統參考 H 指數排名，因此系統能夠查找到該領域內 H 指數較高的科學家或期刊文獻等，以獲得學術圈青睞。內容上以權威性文獻目標，使檢索結果有更高的精確率，並以此提高使用者滿意度。該系統更能查找某領域的經典、當前正受歡迎論文及正在崛起的學術界作者。Microsoft Academic Search 的資料範圍發展始於電腦科學和資訊科學，至今已成為綜合性的搜尋引擎，包含農業科學、人文藝術、生物科學、化學、電腦科學、環境科學、地理科學、社會科學、醫學以及藥學等(Microsoft Academic Search, 2013)。

(三)RePEc

Research Papers in Economics(簡稱 RePEc)是由全球的志工共同建立的開放取用期刊論文系統。成立於 1997 年，其目標在於促進經濟學及相關學科研究成果的廣泛傳播與交流，進而帶動經濟學領域相關研究。該系統收錄超過 1.4 億筆資料，其資料的主要內容有研究手稿、期刊文章、圖書以及各類型電子資源。機構組織領域涵括經濟學門，科學研究院及政府組織。其中，研究手稿可免費取得全文；期刊也可以檢索和瀏覽的方式查找，部分能夠免費取得全文，但商業期刊仍需要訂購(RePEc, 2014)。

RePEc 系統支援的主要資料來源有下列機構及其系統：(1)Elsevier；(2)Wiley Blackwell；(3)AgEcon Search (US)；(4)Springer；(5)Federal Reserve System (Fed) in Print (USA)；(6)Taylor & Francis Journals；(7)National Bureau of Economic Research (NBER, USA)；(8)Cambridge University Press；(9)Programme National Persée (France)(10)Oxford University Press；(11)Munich Personal RePEc Archive (MPRA, Germany)；(12)NEREUS: Economists Online Archive；(13)WOPEBI (Canada)；(14)University of Chicago Press；(15)ECONSTOR (ZBW, Germany)等，共 53 個相關機構。因此，RePEc 的規模可謂為全球最大的經濟學門期刊論文系統，其資料包含近 44 個國家、經濟類的 100 多種期刊、研究報告、軟體等。其中，研究報告 138,000 篇、期刊文章 144,000 篇、軟體 1,100 種、書籍 750 種，以及提供 5,500 位作者與出版商與 8,100 個研究院所的相關資訊(RePEc, 2013)。

RePEc 系統共包含 15 個子系統，其中 Munich Personal RePEc Archive 及 RePEc Author Service 系統為作者上傳及管理資料的 RePEc 子系統，Socionet、SPZ、CollEc、CitEc、LogEc、PEc Plagiarism Committee、EDIRC、NEP、EconAcademics.org、RePEc Biblio、RePEc Genealogy、EconPapers 皆 RePEc 後端維護功能，包含作者身平建置、索引維護、統計分析及檢索結果排序機制等。使用者所使用的介面為 IDEAS，IDEAS 包含檢索及瀏覽功能，瀏覽功能包含作者、論文、期刊、軟體、圖書、圖書章節瀏覽等，皆有製作學科指引及作者索引，特別注意的是，IDEAS 有特別建置諾貝爾經濟學獎得主所著作的圖書、期刊及各類文獻(RePEc, 2014)。

(四) arXiv.org

arXiv 電子預印本典藏庫原名 xxx.lanl.gov，是由美國 Los Alamos National Laboratory (簡稱 LANL) 的物理學家 Paul Ginsparg 於 1991 年 8 月所發起建置，目的在提供一個公平、一致的全球電子預印本檢索機制。目前 arXiv 由 Cornell University Library 維護。其所蒐集之電子預印本原是以高能物理理論領域為主，後來亦加入物理學、數學、電腦科學及定量生物學等領域的研究文獻。收錄文獻形式則包括研討會論文、同儕評閱期刊論文、期刊論文預印本等等(arXiv, 2013)。

Ginsparg 曾提到，1991 年創設 arXiv 時尚未存在任何物理領域的電子期刊，所以其起始目的並非要取代學術期刊，而是希望透過 arXiv 提供一個公平的、一致的全球預印本檢索機制。此外根據 arXiv 的網站簡介，該電子預印本典藏庫是一個高度自動化的研究文獻電子典藏與傳佈系統。其宗旨在促進

科學研究成果的交流與共用，幫助研究者追尋本學科最新的研究進展，對於避免研究工作的重複有很大的助益，對科學研究者具有高度的參考價值(Ginsparg, 2003)。

arXiv 預印本典藏庫目前涵蓋多種學門領域，主領域包括物理(physics)、數學(mathematics)、電腦科學(computer science)、非線性科學(nonlinear sciences)以及量化生物學(quantitative biology)。之所以說「目前」，乃因其涵蓋領域並非自始不變，隨著學術研究發展轉變或是子典藏庫的併合，涵蓋領域乃有所變動。在其網站的最新消息資料中，也記錄了涵蓋領域及各子領域的變遷歷史(arXiv, 2013)。

而在檢索與瀏覽方面，arXiv.org 使用者可根據時間選擇瀏覽最新發表文獻(new, most recent mailing)、最近一週發表文獻(recent, past week)、或是新近發表文獻(current, month's)，也可以瀏覽最近修正取代的文獻清單(replacements)。可查詢欄位包括作者(Authors)、題名(Title)、摘要(Abstract)、完整記錄(Full record)、附註(Comments)、期刊附記(Journal-ref)、主題分類(Subject-class)以及電子預印本編號(Report-no)。並可以領域、子領域、年代等與上述查詢項進行組合查詢或列表。

(五)OAIster

OAIster 以全球機構典藏系統為範圍，於 2002 年建置，提供 OAI 資料庫的檢索服務。OAIster 系統為美國以密西根大學圖書館(University of Michigan Library)所建置之機構典藏資源聯合目錄系統，主要提供數位資源搜尋服務，其典藏雖非自身產出之數位內容，但其目標是在建立開放取用的數位資源，讓散布在網路世界中的學術性數位資源可被方便地被使用。

OAICat 開放源碼軟體項目是一個 Java 網頁應用程式，提供了一個符合 OAI-PMH V2.0 資料庫的框架。這個框架可以支援任意格式數據儲存，可實施通過若干 Java 接口。為擴大及增加本系統的使用率及曝光度，2009 年 10 月，密西根大學將 OAIster 資料移至 OCLC(Online Computer Library Center)，使全球使用者可透過 OCLC 的 WorldCat.org 來檢索 OAIster 的資料。其貢獻包括：(1)可以設置自己的檢索進度和自行定義資料集；(2)圖書館也可以將 WorldCat 資料收錄至各館館藏之中；(3)通過 WorldCat 的官網，提高各館機構典藏的能見度。

OAIster 收錄各類主題及各種類型資料，包括：圖書、學位論文、期刊文章、聲音檔、圖片、影像、統計資料等。OAIster 包含世界上 1,100 多家圖書館及機構提供的數位資源，是全球最大的開放取

用機構典藏系統，資料多達 2,300 多萬筆並有部分電子全文。該系統涵蓋開放取用檔案資源 (OAster,2014)。

(六)OpenDOAR

OpenDOAR 旨在利用全球性開放取用系統進行資料的蒐集、組織及傳遞，提高開放取用資源的取用和使用效益，推動開放取用的發展。該機構典藏及合系統係由開放社會研究所(Open Society Institute)、英國聯合資訊系統委員會(Joint Information Systems Committee，簡稱 JISC)、英國大學學術圖書館聯盟(Consortium of University Research Library，簡稱 CURL)、歐洲學術出版與學術資源聯盟(Scholarly Publishing and Academic Resources Coalition Europe，簡稱 SPARCEurope)的資助，由英國諾丁漢大學(University of Nottingham,UK)和瑞典蘭德大學(University of Lund,Sewden)於 2005 年共同創辦，由英國諾丁漢大學維護 (OpenDOAR, 2014)。

此外，OpenDOAR 和開放取用期刊指南(Directory of Open Access Journals 簡稱 DOAJ)並列為當前開放取用學術資源檢索的主要平台，其資料類型包含期刊論文、會議論文、學位論文、技術報告、專利資料、學習資源、多媒體資源、資料集、工作手稿及資料、預印本等(何琳，2009)。

OpenDOAR 的功能在於匯集資料庫，其含括 1118 個資料庫，其中有 895 個資料庫為機構典藏資料庫。其收錄的資料在美國、德國及英國的機構典藏資料庫為主，占總數的 408 個機構典藏。OpenDOAR 收錄的學科領域可分為 29 類，計有：心理學、管理學、圖書資訊學、法律及政治、教育、商業與經濟學、語言學、歷史及考古學、機械工程與材料、電子工程、電腦科學、土木工程、建築學、數學與統計學、化學與化學工程等。由於每個學科中又涉及多個學科，所以一般學科分別也會分別收錄在不同的領域之下(何琳，2009)。截至 2013 年 10 月，該機構典藏系統已有超過 2,481 個機構參與系統(OpenDOAR, 2014)。

上述各商業性引文資料庫及開放取用引文系統之成立宗旨、收錄範圍、學科主題、建置年代及語言整理如表 1。商業性引文資料庫之成立宗旨是為企業、專業人士與學術社群所建置，開放取用引文系統則是為全球使用者建置。除 Google Scholar 無明確收錄範圍，其餘商業資料庫及開放取用系統皆有敘明其收錄範圍。在學科主題上，除了 RePEc 與 arXiv.org 收錄織文獻內容具特定主題之特性外，其餘皆是綜合學科之商業性引文資料庫或開放取用引文系統。在年代上，arXiv 建置於 1991 年最早，

Microsoft Academic Search 建置於 2009 年最晚。各商業性引文資料庫及開放取用引文系統皆為全球性，故在語言上皆為多種語言，並以英文為主。

表 1 各資料庫簡介比較

資料庫名稱	宗旨	收錄範圍	學科主題	建置年代	語言
Web of Science	為企業和專業人士提供情報資訊	收錄期刊超過 13,000 種，每週更新其內容，每年提供超過 110 萬筆書目及 2300 萬筆引用文獻資料	理、工、醫、農、人文、及社會科學等	1961(1997-網路版)	多種
Scopus	促進學術傳播，創造研究社群	收錄來自全球 5,000 個出版公司，至 2012 年 5 月收錄經同儕審查的期刊數量高達 19,500 種，內含三千萬筆記錄及余兩億筆參考文獻	物理與工程、健康科學、社會科學、心理學與經濟學、生命科學等	2004	多種
Google Scholar	協助使用者從全球的學術研究中找到最相關的研究	無明確數量	屬於綜合性學科資料庫	2004	多種
Microsoft Academic	開發的線上免費學術搜索引擎系	多達 3800 萬的學術文獻和 1900 萬	綜合性學科，包含農業科學、人文藝	2009	多種

Search	統	個學術作者	術、生物科學、化學、電腦科學等		
RePEc	促進經濟學及相關學科研究成果的廣泛傳播與交流，進而帶動經濟學領域相關研究	收錄超過 28 萬筆資料，主要文獻類型為研究手稿、期刊文章、圖書以及各類型電子資源	經濟學門、科學研究院及政府組織資源	1997	多種
arXiv.org	提供一個公平、一致的全球電子預印本檢索機制	收錄超過 100 萬筆資料，主要文獻類行為研討會論文、同儕評閱期刊論文、期刊論文預印本	物理學、數學、電腦科學及定量生物學等領域	1991	多種
OAIster	建立可公開取用的數位資源，讓散布在網路世界中的學術性數位資源可被方便地取用	包含世界上 1,100 多家圖書館及機構提供的數位資源，記錄數量多達 2,300 多萬並有電子全文。	收錄全球機構典藏	2002	多種
OpenDOAR	提高開放取用資源的取用和使用效益，推動開放取用的發展	含括 1118 個資料庫，以英國、美國及德國資料庫為主	為綜合性資料庫，資料包括：圖書、學位論文、期刊文章、聲音檔、圖片、影像、統計資料等	2005	多種

研究方法

本研究之研究對象為 WOS、Scopus、Google Scholar、Microsoft Academic Search、arXiv.org、RePEc 六個引文索引系統以及 OAIster 與 OpenDOAR 二個機構典藏系統，透過檢索作業，分析比較各系統之檢索介面，再就檢索結果之呈現、輸出等項目評鑑各個資料庫與系統。本研究之研究樣本選定為諾貝爾物理學(All Nobel Prizes in Physics, 2012)與經濟學(All Nobel Prizes in Economic Sciences, 2012) 2001 年至 2012 年之得獎者(見表 2 與表 3)，針對此期間之歷年得獎者，分別於各系統進行檢索其歷年來之著作，並加以下載其書目資料，再進一步執行書目資料剖析，包括排序、比對、刪除、聚集與統計等。最後將自各資料庫與系統中所搜尋出之書目資料進行個別之量化統計，並相互交叉比對，以分析各系統收錄資料之完整性、重複性與獨特性，進而得到研究結論並提出建議。

由於本研究是利用作者查詢檢索 WOS、Scopus、Google Scholar、Microsoft Academic Search、arXiv.org 及 RePEc 六個引文索引資料庫與系統以及 OAIster 與 OpenDOAR 二個機構典藏系統，以諾貝爾物理學與經濟學(2001 年-2012 年)得獎者歷年來的著作為樣本。由於作者同姓名或同縮寫的情況十分普遍，且網路資源全自動書目著錄常在著錄過程發生錯誤，書目著錄和實際文獻內容會有不相符的情形，因此，在資料檢索後的篩選與查證十分重要。再者，不同資料庫有其個別的資料篩選方式，WOS 和 Scopus 具有作者權威檔，可以在檢索後點選作者觀看作者的資料以作為篩選標的；但 Google Scholar、MAS、arXiv.org、RePEc 及 OAIster 與 OpenDOAR 都未設置作者權威檔，必須先了解各作者的生平相關資料，分析出作者完整姓名及慣用筆名、生平、專長及研究領域、求學學校、任職機構、e-mail、文獻來源出處等作為篩選條件，並以網路上收集到之作者目錄作為參考根據，再進一步參考資料庫中提供的文章全文，藉此判斷所檢索出的資料是否確實為該作者之著作。

表 2 諾貝爾經濟學獎得主，2001-2012 年

年份	得獎原因	諾貝爾經濟學獎得主
2012 年	對「穩定分配理論和市場設計實踐」研究有所貢獻	Alvin E. Roth
		Lloyd S. Shapley
2011 年	研究「經濟政策與總體經濟間的因果關係」	Thomas J. Sargent
		Christopher A. Sims

2010 年	提出「搜尋理論」	Peter A. Diamond
		Dale T. Mortensen
		Christopher A. Pissarides
2009 年	對「經濟治理」有效研究	Elinor Ostrom
		Oliver E. Williamson
2008 年	其「創新國際貿易與世界性都市化理論」、「貿易政策研究」研究有所貢獻	Paul Krugman
2007 年	提出「機制設計理論」	Leonid Hurwicz
		Eric S. Maskin
		Roger B. Myerson
2006 年	提出「附加 Phelps 曲線」	Edmund S. Phelps
2005 年	對「不合作的賽局理論」有所貢獻	Robert J. Aumann
		Thomas C. Schelling
2004 年	提出「動態經濟研究」	Finn E. Kydland
		Edward C. Prescott
2003 年	提出「自我回歸條件異質變異數模型」和「共整合」理論	Robert F. Engle III
		Clive W.J. Granger
2002 年	將經濟學結合心理學領域，將經濟學也帶入科學領域	Daniel Kahneman
		Vernon L. Smith
2001 年	提出「資訊不對稱市場分析」	George A. Akerlof
		A. Michael Spence
		Joseph E. Stiglitz

資料來源：All Nobel Prizes in Economic Sciences. Oct. 24, 2013. Retrieved form:

http://www.nobelprize.org/nobel_prizes/economics/laureates/.

表 3 諾貝爾物理學獎得主，2001-2012 年

年份	得獎原因	諾貝爾物理學獎得主
2012 年	表彰在「量子光學領域的卓越貢獻」	Alvin E. Roth
		Lloyd S. Shapley
2011 年	表彰「透過觀測遙距新星而發現宇宙加速膨脹」的貢獻	Saul Perlmutter
		Brian P. Schmidt
		Adam G. Riess
2010 年	表彰「有關二維材料石墨烯的開創性實驗」	Andre Geim
		Konstantin Novoselov
2009 年	對「光在光纖中傳輸方面所取得的開創性成就」與「發明了一種成像半導體電路，CCD(電荷耦合器件)傳感器」	Charles Kuen Kao
		Willard S. Boyle
		George E. Smith
2008 年	在「亞原子物理領域裡發現了自發性對稱破缺的機制」與「發現了對稱破缺的起因，並據此預言自然界中至少存在三族夸克」	Yoichiro Nambu
		Makoto Kobayashi
		Toshihide Maskawa
2007 年	發現「巨磁電阻效應」	Albert Fert
		Peter Grünberg
2006 年	發現了宇宙微波背景輻射的基本形式和它在不同方向所具有的微小差別	John C. Mather
		George F. Smoot
2005 年	對「光相干量子理論的貢獻」與「雷射精密光譜學發展的貢獻」	Roy J. Glauber
		John L. Hall
		Theodor W. Hänsch
2004 年	發現了「高能下的夸克處於漸近自由狀態」	David J. Gross
		H. David Politzer
		Frank Wilczek
2003 年	發展了「超導理論」，並解釋了「超流動性」中的一種	Alexei A. Abrikosov
		Vitaly L. Ginzburg
		Anthony J. Leggett

2002 年	對天體物理學特別是在宇宙中微子探測方面做出了開創性的貢獻；以及對天體物理學的開創性貢獻，此貢獻導致了宇宙 X 光源的發現	Raymond Davis Jr.
		Masatoshi Koshiba
		Riccardo Giacconi
2001 年	在「鹼性原子的稀薄氣體中獲得了玻色-愛因斯坦凝聚和對這額凝聚體特性的早期基礎研究」	Eric A. Cornell
		Wolfgang Ketterle
		Carl E. Wieman

資料來源：All Nobel Prizes in Physics. Oct. 24, 2013. Retrieved from:

http://www.nobelprize.org/nobel_prizes/physics/laureates/.

研究結果

針對七個商業性引文資料庫及開放取用引文系統之檢索功能探討，以及透過交叉分析其完整性、獨特性與重複性，評鑑各商業性引文資料庫或開放取用系統與其他商業性引文資料庫及開放取用系統之優劣，茲敘述分析結果如下：

一、商業資料庫及引文系統檢索功能比較

各資料庫及引文系統檢索功能整理如表 4，可發現每個引文資料庫或引文系統皆有一般檢索功能。其中，檢索功能最多者為商業性資料庫 Web of Science 及 Scopus，各有四種檢索功能。一般檢索為所有商業性資料庫及引文系統的基本檢索功能，進階檢索除了 RePEc 和 OpenDOAR 外，其他引文資料庫及引文系統皆有此功能，但 RePEc 和 OpenDOAR 分別有編號檢索、機構典藏檢索以及清單瀏覽。推測機構典藏系統沒有進階檢索，可能為其資料利用其他檢索方式成效較佳。其中 Web of Science、Scopus 及 RePEc 皆另外提供作者檢索，以辨別作者差異。Web of Science 則以引用文獻檢索做為其特色功能，是其他資料庫或系統較難取代的功能之一。

就標題檢索而言，所有的資料庫及引文索引系統在鍵入題名後皆能找到所需的資料。就一般檢索欄位而言，Google Scholar、RePEc、OAIster 及 OpenDOAR 皆為簡易檢索，不提供檢索欄位限制。其他商業性引文資料庫及開放取用系統則提供較多欄位限制檢索功能。但就操作實務而言，Google Scholar、Microsoft Academic Search 提供較強的容錯機制，即使輸入欄位有問號(?)、

減號(-)或其他特殊符號，也會檢索其他相近文獻並找到其資料。然而，RePEc、OAIster 及 OpenDOAR 在檢索上，即使有特殊符號，若完全相符便會自動檢索到書目。唯獨 Web of Science、Scopus 和 arXiv 若有特殊符號便可能造成檢索失敗，檢索上較為不便。

作者檢索於資料庫或引文索引系統檢索時，往往會因技術性問題造成無法正確檢索，儘管商業性引文資料庫 Web of Science、Scopus 及開放取用期刊論文系統 RePEc、Microsoft Academic Search 有建置作者索引典。但 Web of Science、Scopus 及 RePEc 所建置的作者索引詞有所不同，以 Joseph E. Stiglitz 為例，即有「Joseph E. Stiglitz」、「Joseph E. S.」或「Joseph E.」等不同著錄方式，各資料庫雖然建置索引典可供使用，但不同的姓名著錄標準確實會影響使用者進行作者檢索與判斷。此外，單一商業性引文資料庫及開放取用系統內部的索引詞也未必統一，檢索到的資料更未必齊全。基本上，作者檢索功能除 RePEc 較為完整外，其他商業性引文資料庫及開放取用系統上，作者檢索功能可能作為輔助之檢索功能較適宜。

表 4 各資料庫及系統檢索功能比較

資料庫	一般 檢索	進階 檢索	作者 檢索	引用 文獻 檢索	服務 機構 檢索	編號 檢索	機構 典藏 檢索	清單 瀏覽
Web of Science	●	●	●	●				
Scopus	●	●	●		●			
Google Scholar	●	●	●					
Microsoft Academic Search	●	●	●					
RePEc	●		●			●		
arXiv.org	●	●	●			●		
OAIster	●	●	●					
OpenDOAR	●						●	●

二、商業性引文資料庫與開放取用引文系統完整性分析

以諾貝爾經濟學獎與物理學獎得主之作品清單，分析商業性引文資料庫及開放取用引文系統之完整性、重複性與獨特性，2001-2012 年諾貝爾經濟學獎得主作品清單共有期刊書目資料 1888 筆。諾貝爾物理學獎得主作品清單共有期刊書目資料 6094 筆。為顯示不同學科主題之文獻於不同商業性引文資料庫及開放取用引文系統中，其書目完整性、獨特性與重複性之差異，諾貝爾經濟學獎得主與物理學獎得主作品清單之檢索結果，分別敘述如下。

諾貝爾經濟學獎得主之作品清單於各商業性引文資料庫及開放取用引文系統期刊書目資料筆數及完整性如表 5，總共有 1888 筆；以 RePEc 完整性最高，有 1883 筆書目資料，完整度為 99.7%；其次為 Google Scholar，有 1791 筆，完整度為 94.9%；而 OpenDOAR 之完整性 28.8% 最低，只有 544 筆書目。整體而言，若以諾貝爾經濟學獎得主之書目資料來看，RePEc 收錄最為完整，其次為 Google Scholar。RePEc 收錄完整的原因可能為子系統 IDEAS 依照諾貝爾得獎者作品清單進行收錄，若排除 RePEc，可發現 Google Scholar 及 Microsoft Academic Search 之完整度均達約 94%，顯見搜尋引擎系統的資料完整性比其他資料庫及系統還要豐富，但商業性引文資料庫的資料完整性又優於機構典藏系統。

表 5 諾貝爾經濟學獎得主之作品清單於商業性引文資料庫及開放取用引文系統之完整性比較

	Web of Science	Scopus	Google Scholar	Microsoft Academic Search	RePEc	OAIster	OpenDOAR
期刊書目資料筆數	1034	1117	1791	1767	1883	852	544
期刊書目資料完整度	54.77%	59.16%	94.86%	93.59%	99.74%	45.13%	28.81%

諾貝爾物理學獎得主之作品清單於各商業性引文資料庫及開放取用引文系統期刊書目資料筆數及完整性如表 6，總共有 6094 筆；以 Google Scholar 完整性最高，有 5897 筆書目資料，完整度為 96.76%；其次為 Microsoft Academic Search，有 5541 筆，完整度為 90.93%；而 OAIster 之完整性 41.47% 最低，只有 2527 筆書目。由於 Google Scholar 利用搜尋器索引網路世界文獻，其文獻來源包括各大學術機構典藏與開放取用引文索引系統等，因此其書目完整性最高。若以系統或資

料庫特性觀察其完整性，在商業性引文資料庫中 Scopus 的完整性高於 Web of Science，但搜尋引擎的完整性表現又優於商業性引文資料庫；在開放取用引文系統中，具有學科主題特性的 arXiv 完整性最高，但其收錄的完整程度仍無法和商業性引文資料庫或搜尋引擎相較。

表 6 諾貝爾物理學獎得主作品清單於商業性引文資料庫及開放取用引文系統之完整性比較

	Web of Science	Scopus	Google Scholar	Microsoft Academic Search	arXiv.org	OAster	OpenDOAR
期刊書目資料筆數	4962	5423	5897	5541	2988	2527	2725
期刊書目資料完整度	81.43%	88.98%	96.76%	90.93%	49.03%	41.47%	44.71%

經濟學文獻與物理學文獻於 Web of Science 和 Scopus 中，兩者的完整性有較大差異，諾貝爾經濟學獎得主之著作於 Web of Science 和 Scopus 中，其完整性較低，僅有 54.77% 和 59.16%，但諾貝爾物理學獎得主之著作於這兩大商業性引文資料庫中，其完整性高達 81.43% 和 88.98%，推測其原因可能與資料庫出版商與學者投稿期刊收錄於資料庫情形兩者有關：儘管 Web of Science 和 Scopus 皆為大型綜合學科資料庫，但 WOS 資料庫的 SCIE 資料庫所收錄之科學技術類期刊已回溯至 1900 年，其回溯範圍大於 SSCI(收錄年限自 1965 年)；而 Scopus 的出版商為 Elsevier 集團，該集團是以科學技術和醫學資訊產品為主的出版社，推測此為諾貝爾物理學獎得主之著作於這兩個商業性引文資料庫中完整性較高之原因。

三、商業性引文資料庫與開放取用引文系統重複性分析

諾貝爾經濟學獎得主作品清單於商業性引文資料庫及開放取用引文系統之書目資料重複性上(見表 7)，Google Scholar 對 Microsoft Academic Search、Web of Science 對 RePEc、Scopus 對 RePEc 及 Google Scholar 對 RePEc 的重複性最高，為 100%。推測原因為 Google Scholar 及 Microsoft Academic Search 皆為搜尋引擎，互相抓取的資料範圍大致相同。而 Web of Science、Scopus 及 Google Scholar 對 RePEc 的重複性達 100%，其原因為 RePEc 的期刊書目資料完整度達 99.74%，故重複性高達 100%。

表 7 諾貝爾經濟學獎得主作品清單於商業性引文資料庫及開放取用引文系統之重複性比較

(橫欄之資料庫與與直欄之資料庫比對書目資料)

	Web of Science	Scopus	Google Scholar	Microsoft Academic Search	RePEc	OAIster	OpenDOAR
Web of Science	x	69.05%	98.07%	96.23%	100%	50.39%	31.43%
Scopus	63.82%	x	97.14%	92.57%	100%	45.84%	31.60%
Google Scholar	56.62%	60.58%	x	100.00%	100%	45.95%	29.42%
Microsoft Academic Search	56.31%	58.52%	95.87%	X	99.94%	46.07%	29.54%
RePEc	54.91%	59.32%	95.11%	93.79%	x	45.14%	28.78%
OAIster	61.15%	60.09%	96.60%	95.54%	99.77%	X	38.85%
OpenDOAR	59.74%	64.89%	96.88%	95.96%	99.63%	60.85%	x

諾貝爾物理學獎得主作品清單於商業性引文資料庫及開放取用引文系統之書目資料重複性上(見表 8)，arXiv、OAIster 與 OpenDOAR 分別比對 Google Scholar 的重複性皆為 100% 最高，但諸如 Web of Science、Scopus 和 Microsoft Academic Search 比對 Google Scholar 的重複性也都高達 97%~98%，arXiv 和 OAIster 比對 OpenDOAR 的重複比率更高達 98~99%，而 Web of Science 和 Scopus 對其他各開放取用引文系統也都有 90% 以上的重複比率。

檢視 OpenDOAR 的書目資料來源以「Europe PubMed Central」、「CERN Document Server」和「arXiv」為主，而 OAIster 的書目資料來源也大都來自學術機構典藏，因此可能造成彼此之間的書目重複甚至依存程度較高，使諾貝爾物理學獎得主之作品清單於 arXiv 和 OAIster 比對 OpenDOAR 的重複比率較高。而觀察其他資料庫與系統比對 Google Scholar 書目重複性高的原因，除了因為 Google Scholar 書目完整性最高，也因為 Google Scholar 的資料來源豐富，尤其於檢

索過程中發現，其中 Google Scholar 主要的書目資料來源為「The SAO/NASA Astrophysics Data System(ADS)」書目資料庫，由於該書目資料庫囊括 Astronomy and Astrophysics、Physics 和 arXiv 等書目資料庫，推測其為 arXiv、OAIster、OpenDOAR 對 Google Scholar 重複性高達 100% 的原因。

表 8 諾貝爾物理學獎得主作品清單於商業性引文資料庫及開放取用引文系統之重複性比較
(橫欄之資料庫與與直欄之資料庫比對書目資料)

	Web of Science	Scopus	Google Scholar	Microsoft Academic Search	arXiv.org	OAIster	OpenDOAR
Web of Science	x	89.44%	98.41%	92.74%	57.54%	48.59%	50.73%
Scopus	81.84%	x	98.54%	91.26%	48.3%	41.99%	58.74%
Google Scholar	82.81%	90.63%	x	91.29%	45.09%	39.06%	55.58%
Microsoft Academic Search	83.05%	89.74%	97.61%	x	43.2%	39.14%	53.94%
arXiv.org	95.54%	98.51%	100%	89.6%	x	81.19%	99.5%
OAIster	95.43%	98.86%	100%	93.71%	93.71%	x	98.86%
OpenDOAR	92.37%	97.19%	100%	90.76%	80.72%	69.48%	X

相較於經濟學文獻，物理學文獻在各資料庫與系統中，尤其是 Web of Science 對 Scopus、Scopus 對 Web of Science，以及各開放取用引文系統對上述兩個商業性引文資料庫的重複性比率皆遠高於經濟學文獻，究其原因在於諾貝爾經濟學獎得主之著作於 Web of Science 和 Scopus 中，其完整性較低，僅有 54.77% 和 59.16%，但諾貝爾物理學獎得主之著作於兩資料庫的完整性卻有 81.43% 和 88.98%，由於完整性較高，因此其他開放取用引文系統比對 Web of Science 和 Scopus 的重複程度較高。

四、商業性引文資料庫與開放取用引文系統獨特性分析

由表 9 可知諾貝爾經濟學獎得主作品清單於商業性引文資料庫及開放取用引文系統之書目資料於獨特性上，各商業性引文資料庫及開放取用引文系統對 OpenDOAR 的獨特性最高，至少為 61.15%。以各資料庫及系統對 RePEc 最低，為 0%。推測其原因為 OpenDOAR 所擁有資料量最少，故對應後獨特性最高，而 RePEc 資料量最齊全，故獨特性為 0%。

表 9 諾貝爾經濟學獎得主作品清單於商業性引文資料庫及開放取用引文系統之獨特性比較
(橫欄之資料庫與與直欄之資料庫比對書目資料)

	Web of Science	Scopus	Google Scholar	Microsoft Academic Search	RePEc	OAIster	OpenDOAR
Web of Science	x	30.95%	1.93%	3.77%	0%	49.61%	68.57%
Scopus	36.08%	x	2.86%	7.43%	0%	54.16%	68.40%
Google Scholar	43.38%	39.42%	x	5.42%	0%	54.05%	70.58%
Microsoft Academic Search	43.19%	41.48%	4.13%	x	0.06%	53.93%	70.46%
RePEc	45.09%	40.68%	4.89%	6.21%	x	54.86%	71.22%
OAIster	38.85%	39.91%	3.40%	4.46%	0.23%	x	61.15%
OpenDOAR	40.26%	35.11%	3.13%	4.04%	0.37%	39.15%	x

諾貝爾物理學獎得主作品清單於商業性引文資料庫及開放取用引文系統之書目資料獨特性上(見表 10)，各商業性引文資料庫及開放取用引文系統對 OAIster 的獨特性最高有 60.94%，最少為 30.52%。而商業性資料庫(Web of Science、Scopus)或搜尋引擎(Google Scholar、Microsoft Academic Search)分別比對開放取用系統(RePEc、OAIster、OpenDOAR)，其書目獨特性比率也都偏高，介於 40%~60%間。而商業性資料庫和搜尋引擎其資料筆數、完整性和重複性皆高，因此各資料庫與系統比對商業性資料庫和搜尋引擎，其獨特性皆低。

表 10 諾貝爾物理學獎得主作品清單於商業性引文資料庫及開放取用引文系統之獨特性比較

(橫欄之資料庫與與直欄之資料庫比對書目資料)

	Web of Science	Scopus	Google Scholar	Microsoft Academic Search	arXiv.org	OAIster	OpenDOAR
Web of Science	x	10.56%	1.59%	7.26%	42.46%	51.41%	49.27%
Scopus	18.16%	X	1.46%	8.74%	51.7%	58.01%	41.26%
Google Scholar	17.19%	9.38%	x	8.71%	54.91%	60.94%	44.42%
Microsoft Academic Search	16.95%	10.26%	2.39%	X	56.8%	60.86%	46.06%
arXiv.org	4.46%	1.49%	0%	10.4%	X	18.81%	0.5%
OAIster	4.57%	1.14%	0%	6.29%	6.29%	X	1.14%
OpenDOAR	7.63%	2.81%	0%	9.24%	19.28%	30.52%	X

結論與建議

學術研究是國家發展的重要基礎，在學術發展的過程中，從研究問題到完成研究，此過程包含資訊的檢索、蒐集以及引用等。如何獲取具有品質與完整的資料，是學術人員面臨的問題。引文索引資料庫幫助學術研究人員搜尋資料、解決問題，並讓學術研究人員了解學術發展的環境與趨勢，因此引文索引資料庫的發展對學術研究人員至關重要。本研究針對八個商業性引文資料庫及開放取用引文系統的檢索功能、檢索程序評估，分析各資料庫及系統的資料完整性，並進一步交叉分析兩兩資料庫的重複性及獨特性，最後綜合整理分析結果，提出下列結論：

一、商業性引文資料庫檢索功能較為多元

Web of Science 及 Scopus 皆提供三種以上的檢索功能，供使用者利用不同檢索模式進行檢索查詢，且商業性引文資料庫相較於開放取用引文系統，為增強其商業競爭能力，皆開發出具特色的工具或功能，如 Web of Science 有 Citation Analysis 提供使用者分析引用文獻及被引用文獻情況，Scopus 則建立完善的作者索引典。

二、開放取用引文系統書目格式較為簡略

開放取用引文系統在書目格式上較為簡單，尤其以 Google Scholar、Microsoft Academic Search、OpenDOAR 及 OAIster 的書目格式甚為簡單，大多有提供題名及部分摘要，但有部分書目甚至無明確出版年代、資料來源甚至作者都無法考察，需要藉由內文或進一步連結方能確定正確的書目資料。如此的書目資料呈現較不適合要求詳細參考資源的學術研究環境。

三、Google Scholar 及 Microsoft Academic Search 容錯功能較強

Google Scholar 以及 Microsoft Academic Search 因為有多元的海量學術資料，其系統能夠至海量資料內比對並提供檢索回饋，因此回饋的資料較為正確。商業性引文資料庫 Web of Science 及 Scopus，由於已建立詳盡的書目資料並且強調其精確性，但資料庫無海量資料所提供的檢索回饋，容錯機制較差，若標題本身有特殊符號便難以檢索到書目資料。

四、以經濟學文獻書目資料而言，RePEc 及 Google Scholar 的資料最為多元且完整；而物理學文

獻書目資料則是搜尋引擎(Google Scholar 和 Microsoft Academic Search)所涵蓋的書目資料較完整

RePEc 針對諾貝爾經濟學獎得主的作品建立一套完整書目，此書目除了圖書、期刊外，更包含其他作品，例如：報紙文章、研究計畫等。Google Scholar 因包含各種子系統，且搜尋器索引網路世界文獻書目資料，因此可檢索到較多元豐富的資料。

從上述研究結論可知，與商業性引文資料庫相比，開放取用系統在檢索上略顯不足，其原因可能為開放取用系統之建置目標及經營方式與商業性引文資料庫根本之差異，商業性引文資料庫有商業營利之需求，故在檢索功能上必須開發其他系統無可取代之功能，同時也挹注相對較高的成本經營、維護與研發資料庫功能。

以下針對商業性引文資料庫與開放取用引文系統兩者提出加強及改善的方向，提供圖書館資料庫選購之建議：

一、商業性引文資料庫(Web of Science、Scopus)

雖然商業性引文資料庫在檢索功能上相較多元且完整，但商業性引文資料庫仍著重於期刊文獻之彙整，在圖書書目資料的完整性仍有待加強。建議商業性引文資料庫可以擴大圖書書目資料以及網路資源等其他資料類別的收錄範圍。

資料著錄格式是便於使用者檢索與快速辨識資料的依據，更會影響匯出書目資料與參考文獻的完整性和正確性。相較於利用網頁爬梳、撈取資料的開放取用引文系統，商業性引文資料庫在資料著錄的完整程度與正確性勝過開放取用引文系統，因此商業性引文資料庫應更注重收錄資源的書目品質。對於回溯資料的著錄格式應力求統一，並加強回溯書目注錄與建檔，以補救部分回溯資料書目不完整與不正確的情形。

三、開放取用引文系統(Google Scholar、Microsoft Academic Search、RePEc、OpenDOAR、OAIster)

開放取用引文系統應該加強檢索功能，提供特有之檢索機制或加值內容，方能與商業性引文資料庫競爭抗衡，如 Web of Science 的引文分析功能及 Scopus 的作者索引典都是值得學習之處。

開放取用引文系統也應落實書目品質，利用各種篩選機制將不適用或不完整的書目資料刪除，以維護系統內的資料品質。且開放取用引文系統應克服無效連結或錯誤連結之情形，以避免檢索結果文不對題，或檢索結果之書目資料無法連結至來源網站之情形。

參考文獻

- CONCERT(2010)。Web of Science 中文使用手冊，檢索日期：2013 年 10 月 18 日，檢自：
<http://ppt.cc/Px1Z>。
- Elsevier Taiwan(2013)。Scopus 操作說明書，檢索日期：2013 年 10 月 18 日，檢自：
<http://taiwan.elsevier.com/htmlmailings/Scopus-QRG-2010.pdf>。
- Elsevier(2014)。SCOPUS 資料庫簡介，檢索日期：2014 年 06 月 04 日，檢自：<http://ppt.cc/a~zv>。
- Google(2013)。Google 簡介，檢索日期：2013 年 10 月 18 日，檢自：
<http://www.google.com.tw/about/company/>。
- IBM SPSS Statistics (2013)。IBM SPSS 產品說明，檢索日期：2013 年 12 月 28 日，檢自：
<http://www-01.ibm.com/software/tw/analytics/spss/products/statistics/>。
- ISI Web of Knowledge(2013)。ISI Web of Knowledge 公司網站，檢索日期：2013 年 12 月 28 日，檢自：<http://www.isinet.com/selection/>。
- Microsoft (2013)。Excel 分析新功能，檢索日期：2013 年 12 月 14 日，檢自：
<http://office.microsoft.com/zh-tw/excel/>。
- 毛慶禎譯(2008)。Open Access-開放取用機會及挑戰手冊。Library & Book，German 頁 5。
- 毛慶禎譯(2012)。布達佩斯公開取用倡議 Budapest Open Access Initiative (BOAI)，在圖書館學與資訊科學大辭典，檢索日期：2013 年 10 月 18 日，檢自：
<http://terms.naer.edu.tw/detail/1678959/>。
- 何琳(2009)。OpenDOAR 和機構知識庫發展現況。圖書館工作與研究，156，30-33。
- 吳孟瑾(2013)。引用文獻索引資料庫-Google Scholar，在認識大學排名系列報導。檢索日期：2013 年 10 月 18 日，檢自：<http://www.lib.nctu.edu.tw/html/categoryid-34/id-165/>
- 吳明德(1991)。館藏發展。台北市：漢美。
- 吳惠林(2013)。諾貝爾經濟學家的故事(修訂二版)。台北市：五南文化事業。
- 李治安、林懿萱(2007)。從傳統到開放的學術期刊出版：開放取用出版相關問題初探。圖書館學與資訊科學，33(1)，頁 39-52。
- 李治安、林懿萱(2007)。從傳統到開放的學術期刊出版：開放取用出版相關問題初探。圖書館學與資訊科學，33(1)，頁 39-52。
- 邱炯友(2006)。學術傳播與其刊出版。遠流出版，台北，頁 222~223。
- 翁竣霖(2014)。商業性引文資料庫與開放取用引文系統之比較研究—以諾貝爾經濟學獎得主為例。政治大學圖書資訊與檔案學研究所，未出版碩士論文，台北市。
- 財團法人國家實驗研究院(2010)，Web of Science 簡介，檢索日期：2013 年 10 月 18 日，檢自：
<http://concert.stpi.narl.org.tw/fdb/Web of Science/index.html>。
- 財團法人國家實驗研究院科技政策研究與資訊中心主編(2009)。開放取用的機會與展望。財團法人國家實驗研究院科技政策研究與資訊中心，台北，頁 3。

- 陳亞寧(民 93)。開放式資訊取用之現況發展分析。圖書與資訊學刊，51，89-108。
- 陳薇竹(2007)。引用文獻索引資料庫之比較研究。政治大學圖書資訊與檔案學研究所，未出版碩士論文，台北市。
- 馮玉(2007)。Scopus 文摘數據庫及其檢索利用。情報科學，25(3)，401-403, 413。
- 劉琮(2005)。美國 25 名諾貝爾獎得主致國會公開信支持學術論文的開放訪問。圖書館雜誌，24(6)，頁 63-64。
- 魯佑敏(2013)。商業性引文資料庫與開放取用系統之引文索引比較研究。政治大學圖書資訊與檔案學研究所，未出版碩士論文，台北市。
- ArXive (2013). arXiv introduction. Oct. 18, 2013. Retrieved form: <http://en.wikipedia.org/wiki/ArXiv>.
- Berkeley library (2013). Search engine criteria. Oct. 23, 2013. Retrieved form: <http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/SrchEngCriteria.pdf>.
- Borgman, C. L. (1990). Editor's introduction, *Scholarly Communication and Bibliometrics*. Newbury Park, Sage, p13-16.
- Bourne, C. P. (1969). Characteristics of coverage by the bibliography of agriculture of the literature relating to agricultural research and development, In *Clearinghouse for Federal Scientific and Technical Information*(Va.22151). Springfield.
- Budapest Open Access Initiative(2013). Budapest open Access initiative. Oct. 2, 2013. Retrieved form: <http://www.budapestopenaccessinitiative.org/>.
- Budd, J. M. (2012). The communities of the academic library. In *The changing academic library*(2nd ed), Unite States: Association of College and Research Libraries(ACRL).
- Calhoun, K. (2008).Online catalogs: what users and librarians want? An OCLC report. Oct. 23, 2013. Retrieved form: <http://ppt.cc/cq4F>.
- Chu, H. (1996). Search engines for the world wide web: A comparative study and evaluation methodology. In *ASIS 1996 Annual Conference proceedings*.
- Doszkocs, T. (1977). Estimating citation overlap in a multi-file retrieval system, In *Information Management in the 1980's: Proceedings of 40th ASIS Annual Meeting* (Chicago, vol.14), White Plains, N. Y. :Knowledge Industry Publications for American Society for Information Science, x.
- Esmail & Mohammad S. & Kiaie, R. M. & Ketab, F. (2011). A comparison between search engines and Meta-Search engines in retrieving information related to physics and the extent of their overlap. *Library and Information Studies*, 22. 3, p130-140.
- Getz, M. (2009). Open access scholarship and publishing, In *Encyclopedia of Library and Information Sciences*(Third Edition). Oct. 18, 2013. Retrieved form: <http://www.tandfonline.com/doi/pdf/10.1081/E-ELIS3-120044492#.UmDK0IBkObY>.
- Google (2013). Google scholar. Oct. 20, 2013. Retrieved form: <http://scholar.google.com.tw/>.
- Hirsch, J. E., (2005). *An index to quantify an individual's scientific research output* , PNAS 102(46):16569-16572.
- Hood, W. W. & Wilson, S.(1998). The distribution of bibliographic records in databases using different counting methods for duplicate records. *Seientometrics*, 46(3). 473-486. Dec. 14, 2013. Retrieved form: http://download.springer.com/static/pdf/257/art%253A10.1007%252FBF02459605.pdf?auth66=1387171099_525062b391909a564bcad9b80df0dea3&ext=.pdf.

- IDEAS (2013). IDEAS-Search. Oct. 20, 2013. Retrieved form: <http://ideas.repec.org/search.html>.
- Jacsoó, P. (2005). Google scholar: The pros and cons. *Online information review*, 29(2): 208-214.
- Julian, R. L. (1976). Comparison of printed bibliographic descriptions distributed by BIOSIS, CAS, EI. *Journal of the American Society for Information Science*, 27, 46-52.
- LaBorie, T. (1985). Library and information science abstracting and index service: coverage, overlap, and context. *Library and Information Science Research*, 7(1), 183-195.
- Lynch, C. A. (2003). *Institutional repositories: Essential infrastructure for scholarship in the digital age*. ARL: A bimonthly report. 226, pp.1-10. Oct. 2, 2013. Retrieved form: http://scholarship.utm.edu/21/1/Lynch_IRs.pdf.
- Martyn, J. (1967). Tests on abstracts journals: Coverage overlap and indexing. *Journal of Documentation*, 23, 45-70.
- Microsoft academic search. (2013). Microsoft Academic Search Help. Dec. 14, 2013. Retrieved form: <http://academic.research.microsoft.com/About/Help.htm>
- Nicholls, P. T. (1989). Bibliometrics of the laserdiscs applications literature. *Laserdisk Professional*, 2, 106-109.
- Nobelprize(2014). All Prizes in Economic Sciences. May. 28, 2014. Retrieved form: http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/.
- Nobelprize(2014). All Nobel Prizes in Physics. May. 28, 2014. Retrieved form: http://www.nobelprize.org/nobel_prizes/physics/laureates/.
- Okerson, A. (1989). Of making many books there is no end: Report on serial prices for the association of research Libraries, In *Report of the ARL Serials Prices Project*.
- OAIster(2014). The OAIster® database . Feb. 19, 2014. Retrieved form: <http://www.oclc.org/oaister.en.html>.
- OpenDOAR(2014). The directory of open access repositories . Feb. 19, 2014. Retrieved form: <http://www.opendoar.org/>.
- Paisley, W. (1989). Bibliometrics, scholarly communication, and communication research. *Communication Research*, 16(5), 701-717.
- Poyer, P. K. (1984). Journal article overlap among index medicus, Science Citation Index, Biological Abstracts, and Chemical Abstracts. *Bulletin of the Medical Library Association*, 72(4), 353-357.
- Rather, R. A., Lone, F. A. & Shah, G. J. (2008). Overlap in web search results: A study of five search engines, library philosophy and practice.
- Read, E. & Smith, C.(2000). Searching for library and information science literature: a comparison of coverage in three database. *Library Computing*, 19, 118-126.
- RePEc (2013). RePEc services. Dec. 14, 2013. Retrieved form: <http://repec.org/>.
- RePEc (2014). RePEc services. Jun. 04, 2014. Retrieved form: <http://repec.org/>.
- Scopus (2013). Scopus in detail: What does it cover? Oct. 23, 2013. Retrieved form: <http://info.scopus.com/detail/what/>.
- Walters, W. H. & Wilder, E. I. (2003). Bibliographic index coverage of a multidisciplinary field. *Journal of the American society for information science and technology*, 54(14), 1305-1312.
- Wang, L. & Wang, J. & Michael, L. & Yong, L. & Wang, Y. & Xu, D. (2012 May). Using internet search engines to obtain medical information: A comparative study. *Journal of Medical Internet*

Research, 14(3).

Wood , J. L. & Flanagan, C. & Kennedy, H. E. (1972). Overlap in the List of Journals monitored by BIOSIS, CAS,Ei. *Journal of the American Society for Information Science*, 23, 36-38.

Yerkey, N. & Glogowski, M.(1990). Scatter of library and information science topics among bibliographic data bases. *Journal of American Society for Information Science*, 41(4), 245-253.

出席第十五屆國際科學及資訊計量學研討會報告

第十五屆國際科學及資訊計量學研討會（15th International Society of Scientometrics and Informetrics Conference）於2015年6月29日至7月4日於土耳其伊斯坦堡(Istanbul, Turkey)舉行。第一天議程為報到與博士生論壇及系統實作研習(workshops)，6月30日至7月2日是正式會議，7月3日至7月4日則是參訪活動。該研討會是由國際科學及資訊計量學學會（International Society for Scientometrics and Informetrics - ISSI）主辦，每二年舉辦一次。此次主辦單位為Boğaziçi University 電腦工程學系(Computer Engineering Department)，協辦單位為Hacettepe University 資訊管理系學系(Department of Information Management)與The Turkish Academic Network and Information Center (ULAKBIM)，贊助廠商計有：Thomson Reuters、Elsevier、EBSCO、Emerald、Springer 等公司。以下就ISSI國際研討會的緣起與發展，本次與會經過及感想與建議加以敘述。

一、ISSI 國際研討會之緣起與發展

書目計量學（bibliometrics）（或稱文獻計量學）是圖書資訊學重要研究課題之一，其起源甚早。1948年印度圖書館學大師阮加納桑（Ranganathan）即提出圖書館計量學（librametrics）作為研究圖書館資料的使用形態與讀者活動的量化分析。同年，Alan Pritchard亦提出書目計量學一詞，其研究內容主要涉及出版文獻的計量與分類。隨著時間的推移，受其他學科知識的影響而產生名稱與內涵的遞變。有關其演變可自ISSI歷屆會議名稱窺其端倪。

比利時書目計量學家Egghe與Rousseau是提倡舉辦ISSI國際研討會的創始者。他們首先於1987年舉行了第一屆會議，並訂其名為書目計量學及資訊檢索的理論研究（Bibliometrics and Theoretical Aspects of Information Retrieval）。當時的研討會名稱只見書目計量學一詞，未見有科學計量學及資訊計量學。直到1993年第四屆研討會時，科學計量學與資訊計量學才出現。該次研討會名稱為第四屆國際書目計量學、資訊計量學及科學計量學研討會（4th International Conference on Bibliometrics, Informetrics and Scientometrics），此次研討會仍將書目計量學一詞置於最前，其次是資訊計量學，再次才是科學計量學。直到1995年ISSI成立，第五屆國際研討會時，才正式訂名為國際科學及資訊計量學研討會（International Conference of the International Society for Scientometrics and Informetrics）。至此，書目計量學一詞消失不見。其實上述各種名詞仍然經常被交替使用，其內涵互有重疊，例如：書目計量學主要探討各種索引摘要資料庫之出版文獻的計量與分類，其基本概念與定

律是其他計量學應用的基礎。科學計量學代表了科學活動與學術傳播的測量以及科學知識的社會意識。資訊計量學則在測量資訊的使用、流動及其產生的影響。Webmetrics 或 cybermetrics 則是以網際網路 (internet) 及全球資訊網 (WWW) 的資訊使用、流動及影響為研究的對象。最近幾年受到網路社群分析技術的影響，進而興起 Altmetrics (Alternative Metrics) 新的研究議題，逐漸發展成一新領域，亦即替代性計量學。

二、與會經過

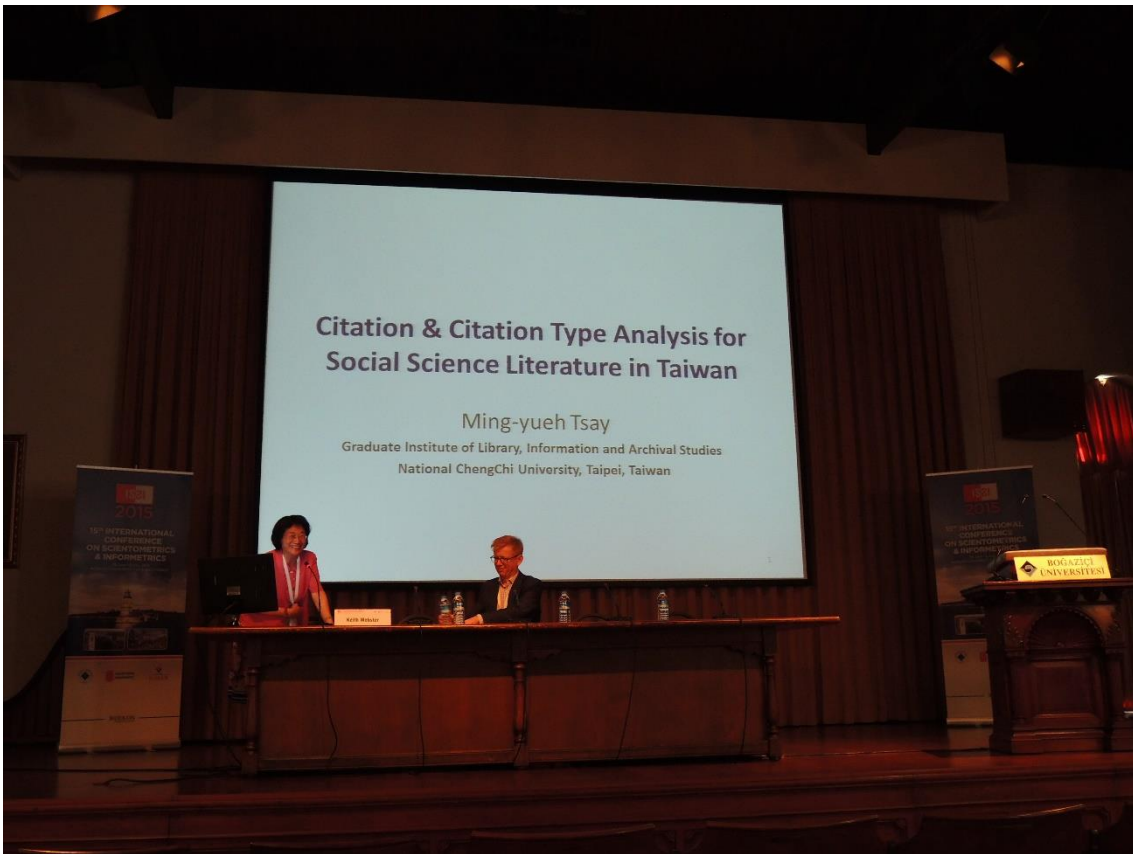
第十五屆 ISSI 國際研討會共有來自全世界 46 個國家，302 位代表參與盛會。出席國家計有：美國、英國、加拿大、澳洲、紐西蘭、印度等英語系國家。亞洲國家有中國大陸、日本、韓國、台灣、馬來西亞等。歐洲國家包括：西班牙、丹麥、法國、挪威、瑞士、瑞典、德國、匈牙利、荷蘭、義大利、比利時、捷克、芬蘭、波蘭、奧地利等，其他國家尚包括：墨西哥、肯亞、南非、南斯拉夫、愛沙尼亞等。由此可見其規模之大，全球各地幾乎都有代表參加，突顯資訊計量學研究在全世界的普遍性。此次會議投稿文章共 354 篇，最後通過同儕審查得以在會議上發表者只有 207 篇，其中 123 篇為口頭報告(oral presentation)，71 篇為海報展示(poster)。口頭報告又分為全文型(full paper)與進展中的研究(research in progress paper)二類。其中發表最多作者人數的國家依次為中國大陸、土耳其。每篇投稿論文經過至少二位評閱者審查，足見其嚴謹性。此次研討會議程共計有 4 個系統實作研習(tutorials)、5 個工作坊(workshops)、一天博士生論壇(doctoral forum)及大會研討會。大會研討會共計有三場特邀演講(keynote sessions)、32 個同步並列場次(parallel sessions)、4 個特別場次(special sessions)、4 個海報展示場次(poster sessions)、2 個贊助者發言場次(sponsored sessions)。此外尚有各種研究工具、軟體、與產品的展示。由此觀之，足見此次國際研討會規模之龐大多元。

該次研討會包羅了前述各種計量學的範疇，涵蓋主題自引文分析至開放取用，詳細內容計有：一、引用模式；二、相同的數據不同的結果；三、開放科學(opening science)；四、全球性共同合作研究(global collaboration)；五、科學的社會系統；六、專利分析；七、學科領域計量常規化(field normalization)；八、研究產出與影響力；九、人文學指標；十、數據的正確性；十二、科學計量工具；十三、推文(Tweet)在科學領域的應用；十四、社群媒體計量學(social media metrics)；十六、機構計量；十七、研究評鑑等。

除了上述主題之外，該次研討會另外提供了三個大會專題特別演講，其中特別值得一提的是由荷蘭阿姆斯特丹大學傳播研究所榮譽教授 Loet

Leydesdorff 教授主講，其講題為：「知識的生產、財富的創造與規範建立的三螺旋關係：創新生態系統的進化模式(The Triple Helix of Knowledge Production, Wealth Generation, and Normative Control: A Neo-evolutionary Model of Innovation Ecosystem)」，三螺旋關係是指大學—產業—政府三方在創新中密切合作、相互作用，同時每一方都保持自己獨立身份的一種創新模式，三螺旋模型概念由 Etzkowitz 於 1997 年首次提出，並用以解釋大學、商業和政府三者間在知識經濟時代的新關係。Leydesdorff 針對此概念進行發展並提供該模型的理論系統模式，描述知識商品化的不同階段，不同創新機構（公共、私人 and 學術）之間的多重互反關係。另二個大會贊助機構的專題演講，分別是：(一)由 Thomson Reuter 公司報告，其講題為：「整合型研究分析與資訊管理工具：如何驅使決策跨越研究的生命週期(Integrated Research Analytics and Information Management Tools to Drive Decisions across the Research Lifecycle)」。其主要內容在介紹 Thomson Reuters 公司開發的各種系統與資料庫，協助研究人員與學術機構建立各種知識管理與學術評鑑的機制。(二)由 EBSCO 公司演講，講題為：「從你最近一篇文章我可以知道你在作什麼：利用 Plumx 即時評量(I Know What You did with Your Last Article: Real Time Measurement with Plums)」，主要介紹 EBSCO 推出的 Altmetrics 資源工具 Plumx，Plumx 提供強大的學術評鑑新模式，追蹤所有形式的研究成果，補強傳統以引用(Citation)為學術評鑑指標的不足，採用 Cave 主張之 Altmetrics 指標分類體系，以梅花的五片花瓣(PLUM Print)各自代表使用率(Usage)、擷取(Capture)、提及與談論(Mention)、社群媒體(Social Media)與引用(Citation)等五種 Altmetrics 資料指標，花瓣的大小視覺化呈現不同資料類型的影響力，Plumx 目前提供的 Altmetrics 影響力分析不僅限於單篇文章，另可應用至機構典藏系統，針對研究者、系所、學校、機構或單一期刊進行 Altmetrics 分析。

本人發表論文題目為：「台灣社會科學文獻之引用類型分析(Citation Type Analysis for Social Science Literature in Taiwan)」，為全文型論文(full paper)，需於大會會場口頭報告(oral presentation)20 分鐘，此次會議全文型論文計有 164 篇論文投稿，僅接受其中 82 篇，其接受率為 50%。本人發表之論文被安排在 7 月 2 日下午 3:30--5:00 於大會的主會場報告，該場次的主題為「Genre Differences」，場次主持人為美國卡納基美隆大學(Carnegie Mellon University)圖書館館長 Keith Webster【圖 1】，發表論文內容主要聚焦於具各國國情與文化特色的資訊計量分析，尤其偏重社會科學與人文學主題。本人文章發表將台灣的社會科學家資訊使用與引用行為與世界各國與會者分享，引起多位研究同好的熱烈討論並交換研究心得，不啻為最具體的國際交流表現。



圖一 本人於 ISSI 2015 年會進行論文發表，右方為場次主持人美國卡納基美隆大學(Carnegie Mellon University)圖書館館長 Keith Webster

值得一提的本人得於會議中場休息時間與多位國際知名資訊計量學大師，包括 Ronald Rousseau、Loet Leydesdorff、Wolfgang Glanzel 等交談請益，真是獲益匪淺。

此外，此次會議 The Derek de Solla Price 傑出貢獻(memorial medal) 獎的得獎者為英國伍爾弗漢普頓大學(University of Wolverhampton)教授 Mike Thelwall 獲得。再且，另一令人關注的訊息是，2017 年的國際科學與資訊計量學研討會將由中國大陸武漢大學信息管理學院主辦，因此，該單位有多位代表出席本次會議，以便學習觀摩，深入了解，預作準備。

三、感想與建議

本次會議主題仍不出歷年研討會主題之範圍，雖然涉及的研究對象多有差異。唯較為熱門與深具時代意義的論題是網路社群媒體計量(social media metrics)，茲就此議題論述個人感想如下：

網際網路與資訊通信科技的先進與蓬勃，致使網路資訊尋求與使用行為日趨頻繁，社群媒體網站的多元化發展，造成學術傳播模式的轉變，傳統以引用次數為主的學術評鑑方式，已無法全然反應數位時代網路環境之下真正的學術影響力，結合傳統與創新的學術評鑑方式因而誕生，在開放取用、數位出版與大數據的衝擊下開啟網路社群媒體計量的興起。往後在網路社群媒體計量研究方面，可針對理論面進行新計量指標及傳統計量學指標之間的關係研究，實證面則研究計量數據來源之各式網站與社群媒體的學術傳播模式與資料內容特性，應用面則著重於開發搜集和分析網路社群網站、開放取用資料的應用分析工具。

出席國際學術研討會發表論文是研究同儕之間面對面討論、交換心得研究、刺激研究靈感最直接有效的方法。學術研討會是孕育新研究工作的泉源，更是啟發創新思考及促進共同研究的催化劑。藉由出席本次 ISSI 2015 的國際學術研討會，不但得知各相關領域的研究現況與未來發展趨勢，更結交許多國際知名學者，建立彼此之間學術交流的管道，促進國際間合作研究，達到學術傳播的目的，是值得從事學術研究者積極從事的一種有意義的活動。

蔡明月

政治大學圖書資訊與檔案學研究所

民國 104 年 7 月 15 日

科技部補助計畫衍生研發成果推廣資料表

日期:2015/10/23

科技部補助計畫	計畫名稱: 開放取用資源與商業系統之比較研究: 以引文系統與機構典藏為例
	計畫主持人: 蔡明月
	計畫編號: 102-2410-H-004-221-MY2 學門領域: 資訊科學
無研發成果推廣資料	

102年度專題研究計畫研究成果彙整表

計畫主持人：蔡明月		計畫編號：102-2410-H-004-221-MY2					
計畫名稱：開放取用資源與商業系統之比較研究：以引文系統與機構典藏為例							
成果項目		量化			單位	備註（質化說明： 如數個計畫共同成果、成果列為該期刊之封面故事...等）	
		實際已達成數（被接受或已發表）	預期總達成數（含實際已達成數）	本計畫實際貢獻百分比			
國內	論文著作	期刊論文	0	2	100%	篇	
		研究報告/技術報告	0	1	100%		
		研討會論文	0	0	100%		
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（本國籍）	碩士生	0	2	100%	人次	
		博士生	0	1	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
國外	論文著作	期刊論文	0	2	100%	篇	
		研究報告/技術報告	0	0	100%		
		研討會論文	0	1	100%		
		專書	0	0	100%	章/本	
	專利	申請中件數	0	0	100%	件	
		已獲得件數	0	0	100%		
	技術移轉	件數	0	0	100%	件	
		權利金	0	0	100%	千元	
	參與計畫人力（外國籍）	碩士生	0	2	100%	人次	
		博士生	0	1	100%		
		博士後研究員	0	0	100%		
		專任助理	0	0	100%		
其他成果 （無法以量化表達之 成果如辦理學術活動 、獲得獎項、重要國 際合作、研究成果國 際影響力及其他協助 產業技術發展之具體 效益事項等，請以文 字敘述填列。）		無					

	成果項目	量化	名稱或內容性質簡述
科教處計畫加填項目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與(閱聽)人數	0	

科技部補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

達成目標

未達成目標（請說明，以100字為限）

實驗失敗

因故實驗中斷

其他原因

說明：

2. 研究成果在學術期刊發表或申請專利等情形：

論文： 已發表 未發表之文稿 撰寫中 無

專利： 已獲得 申請中 無

技轉： 已技轉 洽談中 無

其他：（以100字為限）

擬投稿二篇國內TSSCI期刊，二篇國外SSCI期刊

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以500字為限）

本研究針對八個商業性引文資料庫及開放取用引文系統的檢索功能、檢索程序評估，分析各資料庫及系統的資料完整性，並進一步交叉分析兩兩資料庫的重複性及獨特性，最後綜合整理分析結果並提出結論如下：

Web of Science及Scopus皆提供三種以上的檢索功能，供使用者利用不同檢索模式進行檢索查詢，且商業性引文資料庫相較於開放取用引文系統，為增強其商業競爭能力，皆開發出具特色的工具或功能，如Web of Science有Citation Analysis提供使用者分析引用文獻及被引用文獻情況，Scopus則建立完善的作者索引典。

開放取用引文系統在書目格式上較為簡單，尤其以Google Scholar、Microsoft Academic Search、OpenDOAR及OAIster的書目格式甚為簡單，大多有提供題名及部分摘要，但有部分書目甚至無明確出版年代、資料來源甚至作者都無法考察，需要藉由內文或進一步連結方能確定正確的書目資料。如此的書目資料呈現較不適合要求詳細參考資源的學術研究環境。

Google Scholar以及Microsoft Academic Search因為有多元的海量學術資料，其系統能夠至海量資料內比對並提供檢索回饋，因此回饋的資料較為正確。商業性引文資料庫Web of Science及Scopus，由於已建立詳盡的書目資料並且強調其精確性，但資料庫無海量資料所提供的檢索回饋，容錯機制較差，若標題本身有特殊符號便難以檢索到書目資料。

以經濟學文獻書目資料而言，RePEc及Google Scholar的資料最為多元且

完整；而物理學文獻書目資料則是搜尋引擎(Google Scholar和Microsoft Academic Search)所涵蓋的書目資料較完整

從上述研究結論可知，與商業性引文資料庫相比，開放取用系統在檢索上略顯不足，其原因可能為商業性引文資料庫有商業營利之需求，故在檢索功能上必須開發其他系統無可取代之功能。

以下針對商業性引文資料庫與開放取用引文系統兩者提出加強及改善的方向，提供圖書館資料庫選購之建議：

雖然商業性引文資料庫在檢索功能上相較多元且完整，但商業性引文資料庫仍著重於期刊文獻之彙整，在圖書書目資料的完整性仍有待加強。商業性引文資料庫在資料著錄的完整程度與正確性勝過開放取用引文系統。然而，對於回溯資料的著錄格式應力求統一，並加強回溯書目注錄與建檔。

開放取用引文系統應該加強檢索功能，提供特有之檢索機制或加值內容，方能與商業性引文資料庫競爭抗衡，如Web of Science的引文分析功能及Scopus的作者索引典都是值得學習之處。開放取用引文系統也應落實書目品質，利用各種篩選機制將不適用或不完整的書目資料刪除，以維護系統內的資料品質。且開放取用引文系統應克服無效連結或錯誤連結之情形，以避免檢索結果文不對題，或檢索結果之書目資料無法連結至來源網站之情形。