

網路計量學

Webometrics

蔡明月

Ming-yueh Tsay

淡江大學資訊與圖書館學系教授

Professor, Department of Information and Library Science

Tamkang University

【摘要 Abstract】

本文擬針對網路計量學的主要內容，包括網路連結、網路內容、網路探勘（Web mining）、網路影響因素（Web impact factor）等加以探討，並進一步以數位圖書館為研究主題進行實證研究，以見證網路計量學之應用並總結網路計量學之問題。

This article describes the major studies of webometrics, such as web link, web performance, web mining, and web impact factor. Applying bibliometrics techniques, an experimental research was conducted to investigate the distributions of domain names and of links between web sites for digital libraries. Some problems with webometrics are also discussed.

關鍵字 Keyword

網路計量學 網路連結 網路內容 網路探勘 網路影響因素 數位圖書館

Webometrics ; Web link ; Web performance ; Web mining ; Web impact factor ; Digital library



壹、前言

資訊計量學 (Informetrics) 的研究恆常會受到資訊科技革命的衝擊，新世紀資訊科技最偉大的產物全球資訊網 (World Wide Web--WWW，以下簡稱 Web) 為資訊計量學注入了新的研究契機。Almind 與 Ingwersen(註 1)建議將這種新研究稱之為網路計量學 (Webometrics)，其為資訊計量學的一分支，是 1997 年興起的一種新的研究領域，主要是應用資訊計量學與科學計量學 (Scientometrics) 的方法去研究「網路空間」的種種現象；亦即利用此二種計量學的方法研究全球資訊網及以網路為基礎的傳播。本文擬針對網路計量學的主要內容，包括網路連結、網路內容、網路探勘 (Web mining)、網路影響因素 (Web impact factor) 等加以探討，並進一步以實證研究見證網路計量學之應用並總結網路計量學之問題。

貳、網路計量學探討內容

網路計量學探討的內容相當多樣化，整體而言主要包括網路連結、網路內容、網路探勘、網路影響因素等，茲分述其要義如下。

一、網路連結

利用全球資源定位器 (Uniform Resource Locator--URL) 加以確認的網頁或網址就等同於利用作者或篇名來辨識的文獻。其內容要項包括了地址、伺服器端的機構及作者。利用關鍵字檢索索引摘要資料庫可以編輯一個書目清單；同樣地，利用搜尋引擎亦可在網路上搜集到適當的資料 (Hits or visits)。文獻之間透過參考與引用的關係互相連結；網站與網站之間則透過網路超文件加以連結。連結的網站對製作的網站而言是一種資訊來源，至於被連結的數目則是評量該網站權威性的一種依

據。網頁的計算與內容的分析即與傳統出版品分析相似，例如：自網頁本身與外面連結，稱之為外連 (Outlinks)；自外而連結本身網頁則稱內連 (Inlinks)，這二種連結一如傳統出版品的參考 (Reference) 與引用 (Citation)。Lawrence (註 2) 於 1997 年研究六個主要的搜尋引擎，包括 Alta Vista、HotBot、Northern Light、Excite、Lycos 及 Infoseek，結果發現索引的網頁達 32 億。Rousseau (註 3) 在 1999 年利用 21 週每週的一天以三個常見的字檢索 Alta Vista 及 Northern Light 二種搜尋引擎並加以比較。結果顯示 Northern Light 隨著網頁的成長，檢索到的資料亦隨之增加。Alta Vista 則變化較大，直到 1999 年 10 月 25 日系統重建之後，檢索到的網頁才顯著增加。

二、網路內容

網路內容評估非常重要，可以利用引用文獻分析法探討網頁的重覆性及高度內連 (Inlinked)，相當於被引用 (Cited) 的網站，從被引用情況判斷其品質指標。此外，亦可利用布萊德福分佈律，界定某學科主題之核心網站 (Core sites)。至於網頁的內容分析更不容忽視，Bar-Ilan (註 4) 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 (註 5) 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。

搜尋引擎品質評估主要在探討在某特定時間點上收錄的網頁數量、相關性及其結構，以便觀察是否有大量的變化、不一致性或不規則的現象。Clarke 與 Willett(註 6)採用回現(Recall)率的準則，比較 Alta Vista、Excite 及 Lycos 三個搜尋引擎，結果發現 Alta Vista 的表現比其他二個搜尋引擎好。

三、網路探勘

資訊探勘(Data mining)技術提供網路計量學研究的大好機會。Web 可視為呈指數成長的分散式資料庫，收錄被索引的網頁超過十億，以及一百億個連結，此外尚包含分散在一百萬個伺服器端的資料。(註 7)網路上數以百萬的參與者，隨時不斷在創造、修改、移除網頁與連結，這種分散式、多變化、充滿動力與異質性等特質，將成為網路探勘的沃土，經由網路文獻的連結，可以從中抓取有用的知識，追蹤某一主題發展演變的過程，此種研究或可稱之為網路知識探勘。其中主要研究內容仍不外乎網頁內容及連結結構分析。連結結構與書目計量學中的引用文獻分析有密切關係，網路上大量、複雜但具結構化的超文件透過超連結結構研究可以推論出網路社群，確認出權威網頁、萃取出主題知識，亦可改進搜尋引擎的演算法，網路探勘技術包括：資訊檢索、統計、機器學習、模式辨識及多維尺度等。(註 8)由此可知網路探勘的方法論與書目計量學有許多相似之處，例如：書目計量學的叢集分析，可以視為網路探勘在書目與引用資料庫的應用。

四、網路影響因素

網路影響因素與期刊影響因素一樣，其在探討

一個網站或網頁被引用的情況，換言之，主要在計算的網站的連結強度。Ingwersen(註 9)利用搜尋引擎 AltaVista，計算國家層級的網域與個人網站的影響因素，結果發現平均網域自我連結分數為 0.5，若為.com 的商業網域自我連結為 0.59。平均外連網頁的影響因素為 0.39。此外個人網站的網路影響因素較之國家級網域的網路影響因素不可靠。因此，英格爾森的網路影響因素主要在計算與網域連結的網頁數目，該指標較適合評量國家層次的網域，較不適用於機構或個人層級的網域。

除了網路影響因素，將書目資料的叢集(Clustering)法研究，擴展到網頁資訊的分析，也是網路計量學重要的研究重點。此外，將書目計量學中「共被引概念」應用至網站或網頁的互相連結亦是一研究主流，這種關係 Rousseau(註 10)稱之為互連網(Co-citation)。網路超連結可以清清楚楚的觀察內文與內文的連結，這種相當實體化的引用連結，配合互連網分析一組高度被連結(Highly-sited)的網站，可以探討網路空間(Cyberspaces)的智識結構。(註 11)

參、網路計量學實證研究

數位圖書館為近十年來圖書資訊學界熱烈探討的議題。本節擬採用書目計量學的研究方法，就網路上有關該主題且由台灣地區自建的網站與網頁(以下行文以網頁統稱)，進行計量分析，以布萊德福—齊夫定律(Bradford-Zipf's law)、洛卡定律(Lotka's law)及普萊斯根號定律(Price's square root law)，分別界定核心網站與檢驗網站生產力分佈，進而觀察數位圖書館網頁的發展變化。

本研究(註 12)於 2002 年五月五日至七日，利用 google 搜尋引擎，以「電子圖書館 or 數位圖書館 or 數位博物館」，並配合設限網域 site:tw 的



檢索策略，查尋台灣地區建立的有關數位圖書館主題的網頁。最後共得 5850 個網頁，其中 907 個為商業 (.com) 機構所製作，4259 個為教育 (.edu) 網站，政府 (.gov) 單位製作的有 387 個，網路事業 (.net) 單位製作了 23 個；另外 246 個是由組織機構 (.org) 所製作，最後個人 (.idv) 製作的網頁亦有 28 個。

進一步分析第三層網域，由表一可見，5850

個網頁共由 392 個單位所製作，其中 174 個單位只生產一個網頁，佔全部單位數之 44%。超過 100 個網頁的單位有 11 個，所製作的網頁共計 2889 個，約佔全部網頁的 50%。換言之，有關數位圖書館之一半網頁是由這些單位所生產，其中前九個單位均為教育機構，另外二個分別為商業與政府機構。

表一：數位圖書館網頁與製作單位及其累積數

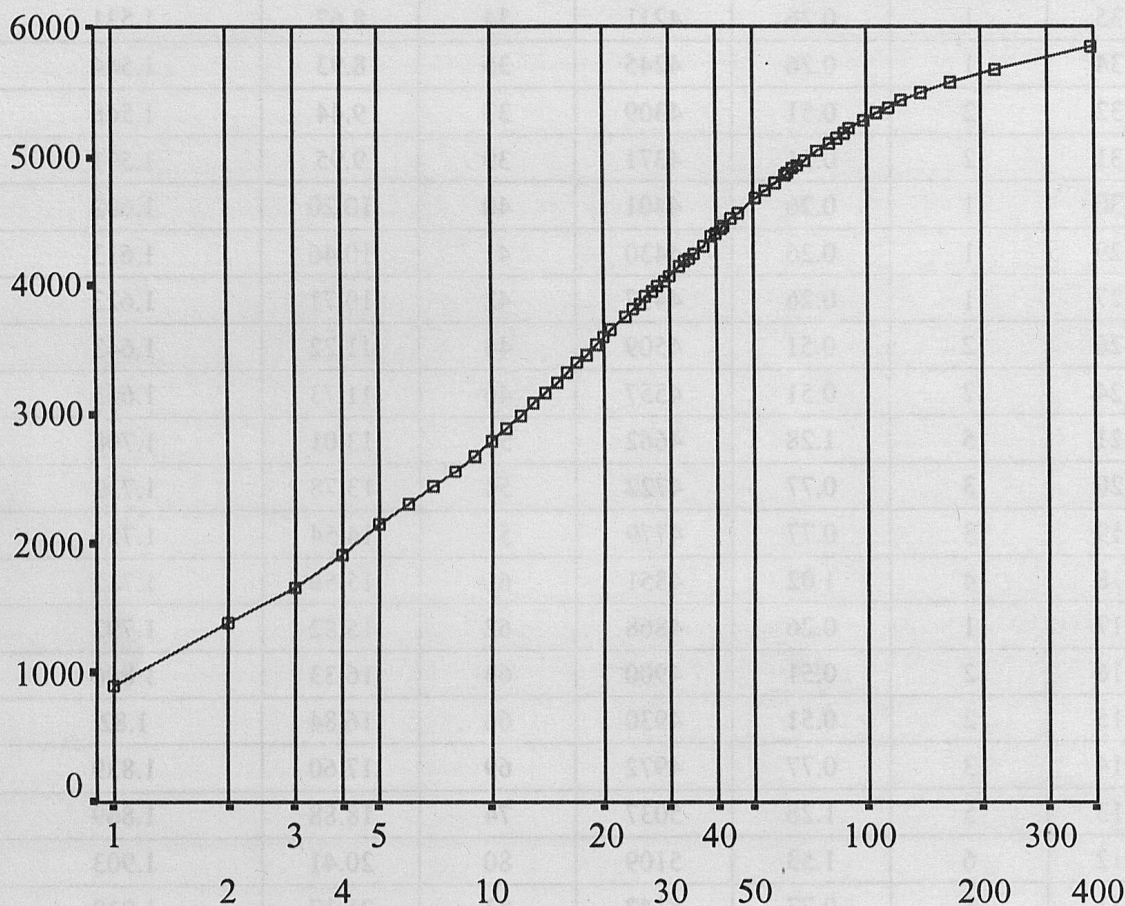
網頁數	單位數	%	累積網頁	累積單位	%	累積單位數之對數值
908	1	0.26	908	1	0.26	0
476	1	0.26	1384	2	0.51	0.301
273	1	0.26	1657	3	0.77	0.477
254	1	0.26	1911	4	1.02	0.602
248	1	0.26	2159	5	1.28	0.699
156	1	0.26	2315	6	1.53	0.778
130	1	0.26	2445	7	1.79	0.845
115	1	0.26	2560	8	2.04	0.903
114	1	0.26	2674	9	2.30	0.954
113	1	0.26	2787	10	2.55	1
102	1	0.26	2889	11	2.81	1.041
98	1	0.26	2987	12	3.06	1.079
95	1	0.26	3082	13	3.32	1.114
85	1	0.26	3167	14	3.57	1.146
80	1	0.26	3247	15	3.83	1.176
78	1	0.26	3325	16	4.08	1.204
73	1	0.26	3398	17	4.34	1.23
70	1	0.26	3468	18	4.59	1.255
61	1	0.26	3529	19	4.85	1.279
60	1	0.26	3589	20	5.10	1.301
56	1	0.26	3645	21	5.36	1.322
54	2	0.51	3753	23	5.87	1.362
52	1	0.26	3805	24	6.12	1.38
49	1	0.26	3854	25	6.38	1.398

48	1	0.26	3902	26	6.63	1.415
45	1	0.26	3947	27	6.89	1.431
42	1	0.26	3989	28	7.14	1.447
39	1	0.26	4028	29	7.40	1.462
38	1	0.26	4066	30	7.65	1.477
37	2	0.51	4140	32	8.16	1.505
36	1	0.26	4176	33	8.42	1.519
35	1	0.26	4211	34	8.67	1.531
34	1	0.26	4245	35	8.93	1.544
32	2	0.51	4309	37	9.44	1.568
31	2	0.51	4371	39	9.95	1.591
30	1	0.26	4401	40	10.20	1.602
29	1	0.26	4430	41	10.46	1.613
27	1	0.26	4457	42	10.71	1.623
26	2	0.51	4509	44	11.22	1.643
24	2	0.51	4557	46	11.73	1.663
21	5	1.28	4662	51	13.01	1.708
20	3	0.77	4722	54	13.78	1.732
19	3	0.77	4779	57	14.54	1.756
18	4	1.02	4851	61	15.56	1.785
17	1	0.26	4868	62	15.82	1.792
16	2	0.51	4900	64	16.33	1.806
15	2	0.51	4930	66	16.84	1.82
14	3	0.77	4972	69	17.60	1.839
13	5	1.28	5037	74	18.88	1.869
12	6	1.53	5109	80	20.41	1.903
11	3	0.77	5142	83	21.17	1.919
10	4	1.02	5182	87	22.19	1.94
9	3	0.77	5209	90	22.96	1.954
8	8	2.04	5273	98	25.00	1.991
7	9	2.30	5336	107	27.30	2.029
6	8	2.04	5384	115	29.34	2.061
5	9	2.30	5429	124	31.63	2.093
4	15	3.83	5489	139	35.46	2.143
3	29	7.40	5576	168	42.86	2.225
2	50	12.76	5676	218	55.61	2.338
1	174	44.39	5850	392	100.00	2.593



圖一為根據表一所繪製的數位圖書館網頁之布萊德福—齊夫分佈圖，其橫軸為單位排名之自然對數，縱軸為累積的網頁數。由圖一可見，其分佈曲線與典型的布萊德福—齊夫分佈圖頗為相近。首先，排名前十名的網站形成一非線性的曲線，繼之

為一線性的區域，當排名超過 55 左右之後，網頁累積之曲線斜率開始變小，形成一尾端下滑的偏垂區段。位於曲線部分前十名的網站則形成所謂的核



圖一：數位圖書館網頁之布萊德福—齊夫分佈圖

進一步利用洛卡定律與普萊斯根號定律檢驗網頁生產力分佈，結果發現，如前所述，只生產一個網頁的單位佔全部單位的 44%。此數值與洛卡原始定律中發表一篇文獻的作者人數約佔全部作者數之 60%，有頗大的差異；此外，利用最小方差的方法求得洛卡定律中的直線斜率 n 值為

-0.7，與洛卡原始定律之 n 值(-2)差異亦相當大。足見本研究網頁製作單位的生產力分佈與洛卡定律不符。

普萊斯根號定律的要義為：在同一主題中發表全部文獻數量一半的科學菁英為全部科學研究人員的平方根。根據此定律，本研究全部網頁製作機

構為 392 個，開根號得 19.8 個。換言之，最多產的 20 個單位所製作的網頁應達全部網頁之半數，亦即 2925 個網頁；實際上，如表一所示，前 20 個單位製作的網頁總量為 3589 個，佔所有網頁的 61%。因此，本研究結果與普萊斯根號定律亦不符

合。

除了網頁的分佈定律之外，本研究亦針對製作 30 個網頁以上的高生產機構進一步加以分析。表二顯示第二層網域名稱及其第三層網域的分佈狀況。

表二：網頁數超過 30 之第二層網域及其第三層網域名稱與網頁數

第二層 網域名稱	第三層網域名稱 /單位名稱		網頁數
Edu	sinica	中央研究院	908
Edu	ntu	台灣大學	476
Edu	nctu	交通大學	273
Edu	ncl	國家圖書館	254
Edu	ks	高雄縣教育網路	248
Edu	tp	台北市教育網路中心	156
Edu	tku	淡江大學	130
Edu	tpc	台北縣政府教育局	115
Edu	nthu	清華大學	114
Com	webcation	光復威柏網	113
Gov	stic	科資中心	102
Com	pczone	PCZONE 討論區	98
Edu	ntnu	台灣師範大學	95
Edu	ptl	國立台中圖書館	85
Edu	yzu	元智大學	80
Com	kimo	Yahoo!奇摩-個人網頁	78
Com	sina	新浪網	73
Edu	nia	台北藝術大學	70
Edu	fju	輔仁大學	61
Gov	nsc	國科會	60
Edu	nsysu	中山大學	56
Edu	pu	靜宜大學	54
Com	udnnews	聯合新聞網	54
Edu	ncku	成功大學	52
Edu	kh	高雄市教育網路	49
Edu	mcu	銘傳大學	48
Com	pchome	電腦家庭	45



Edu	nmns	科學博物館	42
Com	softking	軟體王	39
Edu	chc	彰化縣教育網路中心	38
Edu	nchu	中興大學	37
Edu	ncu	中央大學	37
Gov	taipei	台北市政府	36
Edu	nccu	政治大學	35
Org	gaya	香光資訊網	34
Edu	mit	明志技術學院	32
Com	brightideas	頑石創意有限公司	32
Edu	ncnu	暨南大學	31
Com	taconet	章魚網	31
Edu	tnc	台南縣教育網路中心	30

深一層進入各類型機構的第三層網域分析，則可歸納出下列情況：

一、.com 網域

商業性網域搜尋出第三層網域共有 163 個單位、907 個網頁，網頁數超過 50 的有光復威柏網（webcation）113 個、PCZONE 討論區（pczone）98 個、Yahoo!奇摩-個人網頁（kimo）78 個、新浪（sina）73 個、及聯合新聞網（udnnews）54 個。

光復威柏網規劃一數位圖書館的單元，分為台灣、文學、歷史、地理、社會科學、自然科學、應用科技、醫學、休閒育樂等類，將與其相關的報導資訊綜合整理於此，目前只蒐集文字及圖片資料，將來將繼續增加動畫、影片及聲音等資料，以方便使用者找尋資料。

PCZONE 討論區是一個專為電腦相關技術設置的討論專區，其中所討論的多與電腦軟體、硬體、網路、數位技術等有關的問題回覆、訊息交流、功能討論等，例如：推薦電子圖書館 eLib 2.01 軟體。其與圖書館界所討論的電子圖書館、數位圖書館、數位博物館的文獻或計劃有所差別。

包含在 Yahoo!奇摩-個人網頁中的網頁較前二者（光復威柏網、PCZONE 討論區）複雜，因為本身即是一個搜尋引擎網站，故所含蓋的資料較多樣性，加上本身亦提供個人網頁的架設，所以也包含不少個人網站的資料，故此類的網頁大多數的內容多是相關網頁的分類目錄、相關報導、相關連結、會議訊息公告，或個人站內的相關連結與訊息報導等。此外，由於本身網站即是搜尋引擎，因此有不少連結的網頁內容是已被移除的、或是網頁改版而目前並無存在相關的資料。新浪網搜尋引擎與 Yahoo!奇摩-個人網頁的網站性質相似，故含括在內的網頁內容，同樣具有多樣及多變的性質，且搜尋出的網頁數量與 Yahoo!奇摩-個人網頁相去不遠。

搜尋數量排名第五高位的是聯合新聞網，因為該網站是以新聞報導為主，故所找出的網頁，多為電子圖書館、數位圖書館、數位博物館之新聞或新知報導。此外，在網頁中會列出報導的時間，以顯示其時效性，此為其與一般網頁不同的地方。

二、.gov 網域



政府網域搜尋出第三層網域有 51 個單位、387 個網頁，網頁數最高的是國家科學委員會科學技術資料中心（STIC）有 102 個相關網頁，其次是國科會，佔 60 個網頁。

第三層網域 STIC 包括三個單位。行政院國家科學委員會科學技術資料中心（<http://www.stic.gov.tw/>）有 52 個網頁，其中列表的計畫及書目資料計 33 個網頁、全國學術電子資訊資源共享聯盟資料六個網頁、科學技術資料中心簡訊 10 個網頁、新書通報或資料庫查詢三個網頁。科資中心有 50 個網頁，其中國內資料庫（<http://nr.stic.gov.tw/>）34 個網頁、全國館際合作系統（<http://ill.stic.gov.tw/>、<http://ulej.stic.gov.tw/>）12 個網頁、圖書館資料庫檢索（<http://cdnet.stic.gov.tw/>）二個網頁、文獻傳遞服務資源入口網站（<http://ddsportal.stic.gov.tw/>）二個網頁。

第三層網域 nsc 包含三個單位，網頁數最多的是 NSC 行政院國家科學委員會（<http://www.nsc.gov.tw/>）有 55 個網頁，其次為電信國家型科技計畫（<http://www.ntpo.nsc.gov.tw/>）佔四個網頁，最後是行政院國家科學委員會身心障礙諮詢網（<http://cares.nsc.gov.tw/>）佔一個網頁。NSC 行政院國家科學委員會 55 個網頁中，18 個為國科會數位博物館網站內容；12 個是國科會出版品之中華民國科學技術年鑑，其中報導與電子圖書館、數位圖書館、數位博物館相關之文章；五個網頁表列電子圖書館、數位圖書館、數位博物館相關計畫；四個網頁是人文及社會科學發展處的站內數位博物館相關連結。此外，國科會新聞資料、90 年國家技術科學發展計畫摘要、以及國際合作處（<http://www.nsc.gov.tw/int>）亦各有四個網頁。至於只有一個相關網頁的名稱者計有：NSC 主站內的數位博物館專案計畫站內連結、90 年度碩士論文獲獎名單，其題目為「以 XML 為基礎之數位博物館展示資料模型與內容管理」、工程科學類的計

畫任務「電子圖書館文獻與博物館藏品數位化」、迎新千禧系列報導等。

行政院國家科學委員會身心障礙諮詢網中條列國內外身心障礙研究計畫，其中有一題目為「適合視障者使用之電腦界面技術與系統設計(I)－子計畫五：有聲電子圖書館的自動化技術與工具(I)」。電信國家型計畫有四個網頁，內容涉及電信相關產業中數位圖書館的報導，例如：「為因應國際電信發展趨勢修正規劃方向中，期達成提昇網路服務品質（QoS；Quality of Service）以提供各種多媒體寬頻普及應用」的目標裡列有數位圖書館一項。

三、.org 網域

組織機構網域共搜尋出第三層網域 183 個單位、246 個網頁，其中佔前四名的網頁分別是香光資訊網（gaya）、工研院圖書館（itri）、文化環境網（ceformosa）、及財團法人資訊工業策進會（iii）。

香光資訊網佔了 34 個網頁，其中有 25 個網頁為出現數位圖書館、數位博物館、電子圖書館等關鍵字的《佛教圖書館館訊》之相關網頁文獻，例如：電子資源的長久保存；另外九個網頁是香光尼眾佛學院圖書館之資料，內容包括：佛教期刊論文使用指引，列有「佛教數位圖書館暨博物館」資料庫連結，例如：佛教數位圖書館暨博物館期刊原文、佛教數位圖書館暨博物館書目檢索等的超連結。

工研院圖書館 27 個網頁，其中四個網頁為相關超連結，例如：麻省理工學院數位圖書館、交大數位圖書館、工研院數位圖書館、數位圖書館的核能安全資訊網站；五個網頁多為媒體光碟百科全書的相關報導，如：Online Library 的數位圖書館服務、工研院資訊服務中心圖書視聽館館訊，報導如何擁有數位圖書館及美國主要線上百科全書數位圖書館期刊索引比較表等；另外四個網頁為圖書館導覽站內之數位圖書館連結。



文化環境網有 20 個網頁，皆為中時/聯合報系的新聞報導資料，例如：國立自然科學博物館即將完成建構「阿里山山脈與鄒族文化數位博物館」、「虛擬博物館/數位博物館的經營型態」、「國科會昨起兩天國際研討會推介十二個數位博物館」、「眷村博物館實體工程規劃設計及第一期眷村數位博物館建置計畫」、「國科會數位博物館專案計畫」等的相關報導。

第三層網域為 iii 者有 19 個網頁，單位有五個，其中 14 個網頁是資策會資訊資料服務中心 (<http://www.cisc.iii.org.tw>) 的數位圖書館服務資料庫市場報告(一個網頁)、相關文獻(四個網頁)、相關連結(二個網頁)、相關報導(四個網頁)、站內連結(三個網頁)。二個網頁是財團法人資訊工業策進會 (<http://www.iii.org.tw>) 的個人數位圖書館與多媒體數位圖書館介紹。另一個網頁則是資訊工業策進會教育訓練處 (<http://www.iiiedu.org.tw/>) 電子圖書館軟體廠商(優碩資訊科技股份有限公司)介紹。最後二個網頁，分別是科技法律中心 (<http://stlc.iii.org.tw/>) 以及天空裡的教室 (<http://www.istec.iii.org.tw>) 電子圖書館相關報導。

四、.idv 網域

個人網域共搜尋出 28 個網頁，分屬 18 個人。最高網頁數為許弘網站(hsuhung)、雲岫(icloud)、與徵蔚小築(wei1105)。許弘網站作者是東海大學附屬高級中學國文教師兼教學組長，其中的三個網頁資料，都是列出相關的超連結，例如：在網站中網路資源下的圖書館分頁裡列出相關圖書館、博物館、書店的網路連結，其中就有國科會數位博物館 (<http://www.nsc.gov.tw/y2k/dml/>) 一項；在網路資源下的文學分頁中亦列有中國文學電子圖書館計畫 (<http://cell.cs.uh.edu/>) 的網路連結；至於其在創作集錦下作文分頁中則有超連結至中國文學電子圖書館計畫 (<http://cell.cs.uh.edu/>)。雲岫網

站中，在「九年一貫教學資源網」主題下有三個相關網頁的連結。

五、.net 網域

網路事業網域共搜尋出 23 個相關主題的網頁，第三層網域裡有七個單位。財團法人台灣網路資訊中心(TWNIC)佔 13 個網頁，其中 12 個網頁內容是討論群的相關文章，討論 TWNIC 所授權的網域名稱(Domain name--DN)受理註冊機構問題，並列出公開銷售之網域名稱，如：數位天地.tw，而數位博物館.tw 亦被列在這 12 個討論公開銷售網域的討論群中。雖是 12 個網頁，其實皆為同一主題的討論；另外一個網頁則是報告“Academia-Sinica Digital Library”的簡報資料。

六、.edu 網域

教育網域搜尋出第三層網域有 92 個單位、4259 個網頁，其中網頁數超過 200 的分別為中央研究院(sinica)、台灣大學(ntu)、交通大學(nctu)、國家圖書館(ncl)和高雄縣教育網路(ks)。

中央研究院有 908 個網頁，除了本身(www.sinica.edu.tw)佔了 416 個網頁外，還包括九個單位的網頁，分別是 FreeBSD(freebsd, 288 個網頁)、數位博物館辦公室(dmpo, 138 個網頁)、資訊科學研究所(iis, 26 個網頁)、歷史語言研究所(ihp, 16 個網頁)、台灣魚類資料庫(fishdb, 11 個網頁)、國家典藏數位化計畫電子辦公室(ndao, 數位博物館辦公室是其前身, 八個網頁)、近代史研究所檔案館(archives, 二個網頁)、語言學研究所(ling, 二個網頁)和中小企業管理答客問(euler2, 一個網頁)。中央研究院本身的 416 個網頁中，幾乎都是「臺灣老照片數位博物館網站」的相關網頁，包括網頁首頁、老照片介紹…等。FreeBSD 的網頁資料則包括即時新聞、討論群介紹等，「電子圖書館」一詞，大都出現在討論群介紹



中，因此，網頁資料與電子圖書館並無直接關連。數位博物館辦公室和國家典藏數位化計畫電子辦公室實際上屬於同一單位，只是因為歷史沿革，單位名稱改變罷了，兩者的網頁中主要包括了許多國科會數位博物館專案計畫網址的超連結、國科會數位博物館先導計畫，例如：搜文解字系統首頁、「希望無窮—加拿大國家數位典藏成功的故事」小冊子之電子版、國科會數位博物館計畫的相關網頁內容介紹，以及以數位博物館為主題的相關文章，例如「虛擬原住民藝廊：薩克其萬印地安文化中心藝術典藏數位化的故事」的翻譯文章，另外還包括中研院週報消息報導、研究員個人經歷介紹…等零星網頁。資訊科學研究所網站的 26 個網頁主要為臺灣社會人文電子影音數位博物館網站、相關文章、消息報告、徵才公告和研究員個人經歷介紹。歷史語言研究所的網頁主要是數位圖書館相關資源的超連結，此外尚包括新書目錄與檢索結果等零星網頁。近代史研究所檔案館是兩篇相關文章。語言學研究所則是兩個相關資源超連結的列表。

台灣大學的 476 個網頁包括 30 個單位，其中網頁數超過 50 的分別是圖書館(lib, 107 個網頁)、計算機及資訊網路中心(cc, 96 個網頁)、圖書資訊學系(lis, 77 個網頁)和電機工程學系(ee, 50 個網頁)。圖書館的網頁主要包括「數位典藏國家型科技計畫介紹」、相關文章、「中小學教師數位博物館訓練班」的課程講義、數位博物館專案介紹與相關資源的超連結。計算機及資訊網路中心則以相關文章、討論群的討論內容、相關資源的超連結和消息報導的網頁為主。圖書資訊學系包括相關資源的超連結、「國科會數位博物館專案計畫」的介紹、期刊目次中的文章名稱、系所老師個人經歷與研究內容，以及以數位博物館為篇名的參考書目。電機工程學系則以討論群討論的內容、個人經歷介紹及相關資源的超連結為主。

交通大學有 273 個網頁，除了本身(www.

nctu.edu.tw)佔了九個網頁外，還包括 18 個單位，前四名的單位分別為計算機與網路中心(cc, 101 個網頁)、圖書館(lib, 87 個網頁)、浩然數位圖書館(e-lib, 21 個網頁)和資訊科學學系(cis, 18 個網頁)。計算機與網路中心的網頁內容，包括個人經歷與研究、相關訓練、浩然圖書館的說明與消息報導。圖書館則以「台灣民間藝術家數位博物館首頁」、「蘇森墉音樂館網」、相關資源的超連結、相關數位博物館計畫介紹與消息公告為主。浩然數位圖書館的網頁包括「楊英風數位美術館首頁」、「楊英風數位美術館」相關網頁介紹以及討論群的討論內容。資訊科學學系主要包括數位圖書館相關計畫與相關文章的說明與介紹。

國家圖書館的 254 個網頁中，除了本身(www.ncl.edu.tw)的 127 個網頁外，還包括 12 個單位。其中中國圖書館學會(lac, 62 個網頁)、全國博碩士論文摘要檢索系統(datas, 26 個網頁)和資訊圖書館(infolib, 12 個網頁)為網頁數最多的三個單位。國家圖書館本身的 127 個網頁，主要為資訊素養叢書目錄與介紹、會議紀要與紀實、相關文章的呈現、消息報導與公告等。中國圖書館學會則以數位博物館之出版與研究資訊以及相關資源的超連結為主。全國博碩士論文摘要檢索系統包括相關資源的超連結、相關參考書目的列表、電子報內容。資訊圖書館全部是相關資源的超連結。

高雄縣教育網路，本身(www.ks.edu.tw)網頁數為三個，此外尚包括 15 個單位，其中瑞祥高中佔了 226 個網頁，如此高比例的網頁數乃起因於瑞祥高中的網頁每週都有推薦網站，其中尤以「淡水河溯源數位博物館」的超連結最多。

總之，以教育類而言，電子圖書館、數位圖書館或數位博物館的主題出現在網頁內容中約有下列數種情況：

1. 電子圖書館、數位圖書館或數位博物館的首頁，或是其中的網頁。例如：輔仁大學的「織



- 品服飾數位博物館」<http://www.tcdm.fju.edu.tw/index.html>；中央圖書館台灣分館的「點字、有聲電子圖書館」<http://elib.ncltb.edu.tw/>。
2. 電子圖書館、數位圖書館或數位博物館的計畫說明。例如：台灣大學的「國科會數位博物館專案計畫--Ross」<http://ross.lis.ntu.edu.tw/>。
3. 以電子圖書館、數位圖書館或數位博物館為主題的圖書、文章、紀要或會議紀實。例如：國家圖書館的中文編目文獻課程中的「編目相關書籍」<http://datas.ncl.edu.tw/catweb/3-1-01.htm>；中原大學張靜愚紀念圖書館館刊的「電子圖書館服務與相關法律問題」http://www.lib.cycu.edu.tw/lib_pub/news125.html；國家圖書館館訊的「日本數位圖書館發展現況參訪紀要」<http://www.brightideas.com.tw/ncl/b/86/03.htm>。
4. 消息報導或利用指導，其內容計有：(1)對某一電子圖書館、數位圖書館或數位博物館的介紹，例如：浩然圖書資訊中心的「國內博物館」<http://www.lib.nctu.edu.tw/2001/conference/inside/inSide1.html>，類似路徑示意圖(pathfinder)。(2)簡要陳述某一電子圖書館、數位圖書館或數位博物館的建立或相關主題研討會，例如：國家圖書館的「國內圖書館消息」<http://lac.ncl.edu.tw/lac-info/107/dom107.htm>。(3)報導最新或相關消息，例如：國家圖書館的「電子報」<http://datas.ncl.edu.tw/catweb/epaper03.htm>。(4)教導讀者如何使用電子圖書館、數位圖書館或數位博物館，例如：淡江大學的 IDEAL 線上電子圖書館的指引<http://www.lib.tku.edu.tw/study/guide/guide26.htm>。
5. 以電子圖書館、數位圖書館或數位博物館為關鍵字進行資訊檢索，例如：中華民國期刊論文索引光碟系統 WWW 版的「線上指引」<http://libmaster.lib.ndhu.edu.tw/ncl3web/guide.htm>。
6. 路徑示意圖，亦即列出 URL 做為連結途徑，此種情況相當多見，例如：淡江大學的「國內 WWW 博物館」<http://www.lib.tku.edu.tw/museum/netmuseum/domestic.htm>；資訊圖書館的「電子期刊」<http://infolib.ncl.edu.tw/info/in3-3.asp>；國家圖書館中文編目文獻課程中的「編目相關論文」<http://datas.ncl.edu.tw/catweb/3-1-03.htm>。
7. 出現於題名中，其類型計有：期刊目次，例如 ISBA 全國新書資訊網的「新書資訊月刊目次」<http://isrc.ncl.edu.tw/isbn/frmIndex.asp?Year=89&Month=12>；文章的參考書目，例如：國家圖書館「中文古籍 Metadata」<http://datas.ncl.edu.tw/catweb/2-1-17.htm>；學位論文名稱，例如：成功大學的「論文及其資料下載」<http://www.ncku.edu.tw/TANET2000/download/>；學校考古題，例如：中興大學的「圖書資訊學研究所 90 年考古題」<http://www.nchu.edu.tw/~gilis/test-90.htm>。
8. 課程規劃或系所介紹中提及，例如：清華大學資訊系統與應用研究所的「課程設計」<http://www.isa.nthu.edu.tw/chinese/coursedesign.php>；淡江大學資訊與圖書館學系的「簡介沿革」<http://www.emls.tku.edu.tw/Intro/Intro.htm>。
9. 出現於個人學經歷介紹，或所進行與完成的計畫名稱中，例如：清華大學歷史研究所的「黃一農」<http://www.hist.nthu.edu.tw/~hcs/teachers/huang.html>；清華大學電機資訊學院產學合作聯盟的「張俊盛」。<http://eecsia.cs.nthu.edu.tw/eecsia/academic/projectdetail.php?profid=18>。



肆、結語

誠如前述，網路環境雖然賦予資訊計量學研究新的生命，然而經由第三節的實證研究以及文獻上的探討可以觀察出許多網路計量學的新問題。問題往往是新研究的起點，茲總結於後，供未來發展思考。

網路是一種由各種人生產並檢索的非結構化且高度聚集的複雜資訊，與結構化的資料庫，並經由指令語言檢索的資訊系統迥然而異。換言之，網路上資料數據的收集完全取決於各種搜尋引擎的檢索特性。遺憾的是多數搜尋引擎的檢索功能太過簡單，造成數據收集的困難，以至於無法進行進一步的網路計量學的分析。更有甚者，許多與主題相關的網頁，並非真正建立一個具體的網站，而是類似小型入口網站，僅提供相關網頁的連結。如此一來，網頁的計量將無法辨識真正的生產力分佈。此外，同義異名的單位名稱無法加以控制，例如：行政院國家科學委員會科學技術資料中心、科學技術資料中心、科資中心等三個名稱均是同一單位，必須加上人工整理與判讀的工作。再且，在沒有像科學引用文獻索引(Science Citation Index--SCI)資料庫指令式的檢索功能以顯示檢索結果筆數之前，引用次數的計算是不可能的。這種情況發生在現在大多數的搜尋引擎，搜尋引擎大都不會將全部的網頁加以索引，其收錄範圍都不相同且變動大缺乏穩定性。龐大的網路分析工作，完全仰賴電腦線上執行，實有不足。如此一來，網路計量學之研究樣本的決定與處理將變得重要且困難。

網路是各種資訊攜帶者所造成高度複雜的集

合物，網路上產生連結的原因無法掌握，它不若文獻引用的動機明確。換言之，連結的原因與傳統的引用行為全然不同。傳統的參考永遠是指向過去，但是網路的動態化與及時傳播的特性經常顯現網頁是同時彼此互連的，這是傳統引用世界不可能存在的現象。此外，Web 包含任何的連網，任何人都可自建網頁，由於缺乏同儕審查(Peer review)制度，會產生資訊品質或知識價值的問題。

網路重塑學術傳播的方式，產生了新的學術出版模式，正進行中的研究、初期的草稿，甚至於評審制的論文，均可立即共享，任何人可以做任何的連結，可以以任何的方式存取資料、資訊，甚至知識，網路造就了一個資訊自由的真境界。隨之而來的另一面問題是：網路的不確定性變成一種信任的問題。

除了網路計量學之外，其他名稱尚有 Bossy (註 13)所提出的 *netometrics*，主要在探討以網際網路(Internet)為媒介的科技互動現象；另一新名詞則是期刊《Cybermetrics》的刊名，該期刊是由西班牙馬德里的科學資訊及文獻中心(Centro de Informacion y Documentacion Cientifica-CINDOC)刊行。《Cybermetrics》為電子期刊，雖刊載以科學計量學、資訊計量學及書目計量學的研究為主，但仍然特別強調該三種計量學與網際網路的相互關係。總之，網路計量學雖然充滿了許多不確定性，但可以肯定的是，其未來的發展將為資訊計量學開啓新的研究契機。

(收稿日期：2002 年 7 月 18 日)

註 釋

註 1：T. C. Almind, and P. Ingwersen, "Informetrics Analyses on the World Wide Web : Methodological Approaches to Webometrics !", *Journal of Documentation* 53 : 4 (1997), pp.404-426.

註 2：S. Lawrence, and C. L. Giles, "Searching the World Wide Web", *Science* 280(1998), pp.98-100.



- 註 3 : R. Rousseau, "Daily Time Series of Common Single World Searches in Alta Vista and Northern Light", Cybermetrics 2/3 paper2, <<http://www.cindoc.csic.es/cybermetrics/articles/v2ip2.html>> (visited 08.11.2000)
- 註 4 : J. Bar-Ilan, "The Web as Information Source on Informetrics? A Content Analysis", Journal of the American Society for Information Science 51 : 5(2000), pp.432-443.
- 註 5 : Allen, E. S., J. M., Burke, M. E., Welch, L. H. Rieseberg, "How Reliable is Science Information on the Web?", Science 402(1999), p.722.
- 註 6 : S. J. Clarke, and P. Willett, "Estimate the Recall Performance of Web Search Engines", Aslib Proceedings 149(1997), pp.184-189.
- 註 7 : Frawley, W. J., Piatetsky-Shapiro, G., and C. J. Mathews, "Knowledge Discovery in Databases : An Overview," in G. Piatestsky-Shapiro and W. J. Frawley ed., Knowledge Discovery in Databases, ed. by G. Piatestsky-Shapiro and W. J. Frawley (Menlo Park, CA : AAAI Press, 1991).
- 註 8 : 同註 7。
- 註 9 : P. Ingwersen, "The Calculation of Web Impact Factors", Journal of Documentation 54 : 2(1998), pp.236-243.
- 註 10 : R. Rousseau, "Sitations: An Exploratory Study", Cybermetrics :International Journal of Scientometrics, Informetrics and Bibliometrics 1 : 1(1997). 或見<<http://www.Cindoc.csic.es/cybermetrics/articles/v1ilpl.html>>
- 註 11 : R. R. Larson, "Bibliometrics of the World – Wide-Web: An Exploratory Analysis of the Intellectual Structure of Cyberspaces," In Steve Hardin eds., Proceedings of the American Society for Information Science(ASIS)/59th Annual Meeting, edited by Steue Hardin 1996, October pp.21-24, Baltimore, MD, Medford, NJ: Information Today Inc. for ASIS, pp.71-78,1996.或見<<http://sherlock.berkeley.edu/asis96/asis96.html>>
- 註 12 : 本研究網頁計量數據由鄭琚媛、陳憶玲、賴淑靜協助收集。
- 註 13 : M. J. Bossy, "The Last of the Litter: 'Netometrics'", Solaris Information Communication 2(1995), pp.245-250,或見<<http://www.info.unicaen.fr/bnum/jelec/Solaris/do2/2bossy.html>>

