

InfoSpring 浩然數位圖書館建置計畫

柯皓仁

國立交通大學計算機中心副教授／系統支援組組長

新竹市大學路 1001 號

TEL:(03)5712121 EXT. 52807

EMAIL: claven@cc.nctu.edu.tw

張瑞川

國立交通大學計算機中心主任暨圖書館館長

新竹市大學路 1001 號

TEL:(03)5712121 EXT. 52801

EMAIL: rc@cc.nctu.edu.tw

摘要

完整而迅捷地取得研發資料是產官學界持續保持領先的重要關鍵。由於網際網路（Internet）與全球資訊網（World Wide Web, WWW）的盛行，使得世界知名的研發資料提供廠商陸續在 WWW 上建立 24 小時全年無休的全文或書目參考資料查詢系統。然而這些系統多以歐美地區為基地，因為受限於網際網路的國際通訊頻寬，使得國內甚至於亞太地區的產官學界無法充分利用這些服務。

本篇論文的主旨在於介紹由國科會科資中心、國立交通大學與財團法人交大思源基金會所合力籌設的「InfoSpring 浩然數位圖書館」。InfoSpring 浩然數位圖書館的成立目的是在國立交通大學建立各研發資料提供廠商的亞太地區資料檢索伺服器，透過交通大學快速的網路，使國內以及亞太地區的產官學界能夠迅捷地獲得所需的研發資料。

一、前言

所謂的數位圖書館[1]，乃是運用電腦技術將書籍、聲音影像（Audio-Video）及圖片等資料數位化，以電腦的儲存設備（如硬碟、磁帶等）來儲存這些經數位化後的資料，並配合功能強大的資料庫檢索系統，透過通訊網路，提供讀者作資料之搜尋、擷取與處理。近年來，由於電腦硬體技術的成長與大量生產，使得硬體的效能大幅提高、價格普遍降低，再加上通訊網路（尤其是 Internet）的逐漸普及、網路的傳輸速度也較數年前為高，在這些因素的影響下使得「數位圖書館」得以實現。

1994 年，美國的 DARPA、NASA 以及 NSF 等單位共同投資二千四百萬美金，預計以四年的時間協助史丹佛、卡內基美濃等六所大學 [2,3,4] 及多家電腦軟硬體廠商進行數位圖書館先導計畫，研究探討數位圖書館的相關技術，如：數位化資料的製作、資料的保存管理與檢索、智慧財產權與權限管制等。

這幾年來，國內各大學及研究圖書館的經費逐漸減少，而各種期刊論文的訂費卻相對上揚，在這種狀況下，如何充分運用有限的資源來提供讀者更多更新更好的期刊資料，便成為各大圖書館必須探討的課題。雖然圖書館間的館際合作是一種解決的方法，但是館際合作往往需要三天到一個星期以上，因此在時間上緩不濟急。

對於產業界而言，各項研發及專利資料是產業界創新研發、持續在市場上保持優勢的重要參考資訊。囿於經費限制，國內產業界大多沒有或僅有少數與研發及專利的相關資料，所以，他們對於此類資料的需求更是十分殷切。

自 90 年代以來，由於 Internet 以及 WWW 的日漸普及，使得國外著名的研發資料提供廠商，如 Elsevier[5]、Cambridge Scientific Abstracts (CSA) [6] 等都紛紛開始建立 WWW 上全文或書目參考資料的檢索、閱覽系統，讀者可以直接透過 WWW 線上閱讀並列印期刊資料。與目前紙本期刊或透過圖書館館際合作相比較，運用 WWW 檢索閱讀各項 R&D 與專利資料相對地具有價格合理、便捷迅速、24 小時全年無休的優點。但是由於絕大多數的資料提供廠商並未在亞太地區設置資料檢索伺服器，使得國內的產官學界無法忍受緩慢的 Internet 跨國傳輸速度。對於資料提供廠商而言，為了亞太地區讀者而在亞太地區架設並維護一個資料檢索伺服器，往往需要百萬美金以上的人力、物力、機器設備，而這也是各資料提供廠商不敢投資的主要原因。

在另一方面，國內大學圖書館如要自行建立大規模的數位圖書館，則會面臨下列問題：

A. 電腦軟硬體及網路設備投資龐大

目前全文電子期刊式資料庫大都是以影像或 PDF 檔儲存，通常一頁就需要 80K 左右的儲存空間。以 Elsevier Science 的電子期刊為例，每年的資料量達 200G 左右，加上索引檔和其它相關軟體的建置，需要購置高性能的伺服器和大量資料儲存系統 (Mass Storage System)，故需投資龐大經費購買電腦軟硬體設備。以目前國內大學和研究機構經費緊縮的狀況，各大學圖書館幾乎不可能從事相關的計畫。

B. 缺乏資訊相關專業人員

建構電子期刊和全文資料庫，需要對電腦系統和資訊相關技術有相當程度的了解和實務經驗。目前各圖書館均缺乏此類技術人員。

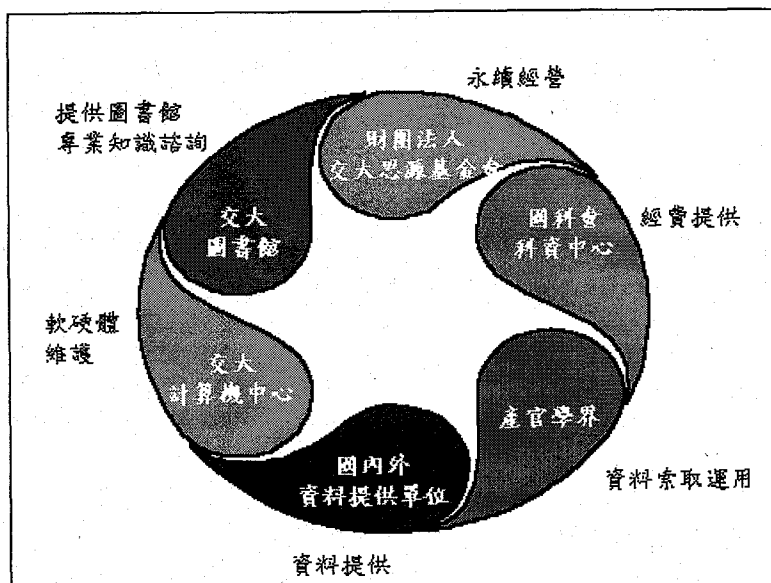
C. 電子出版品授權費用昂貴

電子出版品雖已日漸普及，但提供校園各網路使用的授權費用仍非常昂貴，以 ISI 公司的 SCI (Science Citation Index) 資料庫為例，回溯 5 年的租金達 12 萬美元。在圖書經費日漸緊縮的狀況下，如何籌備經費購買足夠電子出版品是第一 大問題。

基於上述原因，各大學及研究圖書館均希望能由國科會 科資中心協調各相關單位，成立共用數位圖書資訊中心建置電子期刊、全文資料庫 和參考資料庫，透過網際網路供全國各圖書館在獲得適當授權下使用。國立交通大學 圖書館和計算機中心將利用過去兩年來發展數位圖書館的技術經驗和設備配合國科會 科資中心補助款在三年內共同設立「InfoSpring 浩然數位圖書館」，並負責營運管理，以提供國內各圖書館一流的 數位圖書資訊服務。

本篇論文將闡述 InfoSpring 浩然數位圖書館的成立宗旨、經營策略與執行現況。論文組織如下：第二節中說明 InfoSpring 浩然數位圖書館的經營策略；第三節中介紹參與本計畫的資料提供廠商；第四節中敘述本計畫的硬體與網路架構；第五節則描述本計畫的執行現況；最後是結論。

二、InfoSpring 浩然數位圖書館的經營策略



上圖所示為「InfoSpring 浩然數位圖書館」計畫中各合作單位所扮演的角色。其中國科會科資中心為軟硬體經費的主要來源，產官學界則為本數位圖書館成立之後的使用者。接下來我們便逐一敘述交通大學圖書館與計算機中心、國內外資料提供單位與財團法人交大思源基金會在本計畫中所扮演的角色：

A. 交通大學圖書館/計算機中心

- 在硬體方面：負責檔案伺服器及對外網路等資訊基礎建設之建置與維護。本計畫中所用到的硬體與網路資訊基礎建設將在第四節硬體與網路架構中詳細敘述。
- 在軟體方面：負責各資料提供廠商的資料庫管理系統與 WWW 查詢檢索系統之管理維護。
- 在內容方面：負責資料的更新。

B. 資料提供廠商

- 負責資料庫管理系統與 WWW 查詢檢索系統程式之撰寫與修改。
- 負責其擁有之各項期刊論文、專利資料內容之提供與定期更新。
- 與其代理商共同負責爭取台灣及亞太地區之用戶。

C. 財團法人交大思源基金會

- 負責 InfoSpring 浩然數位圖書館之永續經營。

事實上，「InfoSpring 浩然數位圖書館計畫」對各合作參與單位都有實質上的益處，以下是我們的分析：

A. 國科會科資中心

- 推動國家級數位圖書資訊中心之建置，達成國內各大圖書館資源共享的目標。
- 具體實現「亞太媒體中心」。

B. 交通大學

- 能夠以最快的速度取得各資料提供廠商之最新資訊，有助於交通大學學術研究質與量之提昇。

- InfoSpring 浩然數位圖書館將是亞太地區第一個擁有各家期刊論文與圖書資料 之中心點 (Hub)，有助於提高交通大學在此方面的聲譽。
 - 由浩然數位圖書館的經營經驗中探索與數位圖書館相關的研究課題，如中文檢索、Data Mining[7]等，進而發展相關的技術。
- C. 資料提供廠商
- 在節省人力、電腦設備以及經費的情況下，建立亞太地區之資料檢索伺服器，增強其對亞太地區用戶之服務。
 - 運用交通大學在電腦與網路方面的優勢，吸引更多用戶。
- D. 財團法人交大思源基金會
- 實現成為「亞太科技與管理產學資訊中心」的目標。
- E. 產官學界
- 無須建置各項軟硬體設備與聘僱系統維護人力便可取得最新研發與專利資料。
 - 能夠以合理的價格取得高品質之研發與專利資料。
 - 透過交通大學快速的網路連線（詳第四節），無須忍受 Internet 跨國連線塞車之苦，可以在最短的時間內取得所需之資料。

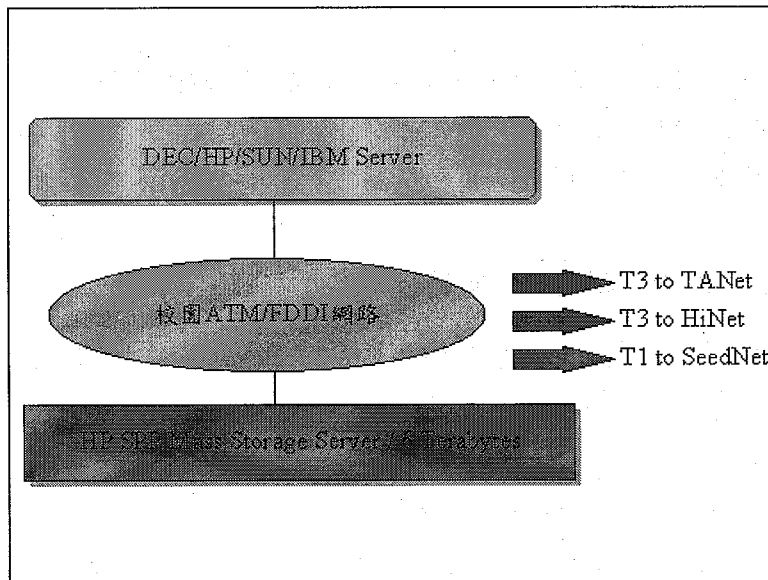
三、資料提供廠商

為推動本計畫，各大學及研究圖書館 館長已組成一委員會，負責選擇適合的全文及參考資料庫提供全國各圖書館使用，並結合群體的力量爭取優惠的資料庫使用費率。在此委員會的積極聯繫下，目前正 與多家全文及參考資料庫提供廠商進行洽談：

- A. 線上電子期刊 (Electronic Journals Online)：
- Elsevier Science, Springer-Verlag, Academic Press, Swet
- B. 參考文獻書目資料庫 (Reference Databases System)：
- CSA, OVID, EBSCO, Silverplatter, UMI, IAC, ISI
- C. 科學與專利資料 (Scientific and Patent Information)：
- Derwent

四、硬體與網路架構

我們將整個硬體與網路架構分為資料庫檢索與 WWW 伺服器 (Database Query & WWW Server)、檔案伺服器 (File Server) 與連外網路 (Networking) 三部份來描述。下圖所示為整體的硬體與網路架構示意圖。



A. 資料庫檢索與 WWW 伺服器

本計畫為每家資料提供廠商量身定作，採用與各廠商在歐美地區的資料庫檢索與 WWW 伺服器同級或更高檔的機器來提供資料庫檢索與 WWW 服務。下表為針對 EES、CSA、OVID 與 ISI 所採用的資料庫檢索與 WWW 伺服器。

資料提供商	硬體規格
EES	Sun UltraII CPU*2, 256MB RAM
OVID	Sun UltraII CPU*2, 256MB RAM
CSA	DEC Alpha Server 4000, 128MB RAM
ISI	Sun UltraII CPU*2, 512MB RAM

B. 檔案伺服器

要儲存所有資料提供廠商的數位圖書資料，經估計每年約需 1200GB 的儲存容量，若全部以磁碟機來儲存，儘管可以得到較佳的效能（Performance），然而所需經費過於龐大。所以在本計畫中我們採用折衷的方式——以 HP Convex 的 HSM（Hierarchical Storage Management）方案來取得效能與價格間的平衡點。HP Convex 的 HSM 系統稱為 Unitree，由兩部份所構成，第一部份是硬碟機，作為資料暫存區（Cache）；第二部份則為一 Tape System，作為所有資料的永久儲存區。當讀者查詢資料時，系統會先檢視硬碟的資料暫存區中是否有要索取的資料，如果有，系統便直接由暫存區中將資料傳給讀者；否則系統會由 Tape System 中將資料調出存至資料暫存區中並傳回給讀者。當資料暫存區快滿載的時候，系統會合併考量檔案的使用頻率與大小等因素，自動將部份資料移回 Tape System。

目前 Unitree 系統是架設於 HP Convex SPP-1600 伺服器上（共有 8 顆 CPU）。資料暫存區總計有 100GB 的硬碟。在 Tape System 方面，共有 3 個磁帶機，256 捲磁帶，每捲未壓縮容量為 10GB（壓縮後容量可達 30GB），合計約可容納 6000GB 左右的資料。我們並計畫在明年度添購 IBM 的 ADSM（亦為一 HSM 系統），用以滿足日增的儲存容量需求。

上述之資料檢索與 WWW 伺服器是透過 NFS 連接檔案伺服器。

C. 連外網路

目前我們以 FDDI 與 FastEthernet 連接前述之檔案伺服器以及資料庫檢索與 WWW 伺服器(明年度將以 ATM 網路連接)。對外網路方面,目前交通大學使用 T3 線路連通 TAnet,思源基金會則有 T3 線路連通 HiNet、T1 線路連通 SeedNet。

五、InfoSpring 浩然數位圖書館設置時程

1. 交通大學已於 85 年 5 月與 Elsevier Science 簽約, 成為亞洲第一座加入 Elsevier Electronic Subscription (EES)的圖書館。目前我們正在安裝 EES 所出版的 1200 種電子期刊, 預計於 7 月中開放測試, 9 月份正式對國內與亞太地區所有圖書館展開服務。
 - URL : <http://ees.infospring.nctu.edu.tw>
2. 於 86 年 1 月與 Cambridge Scientific Abstracts (CSA) 簽約, 開始建置 CSA 所提供的 IDS 書目資料庫服務, IDS 系統在 86 年 4 月已經完成建置, 並對外服務。IDS 系統共提供 80 多種的科技性期刊資料庫, 收錄範圍包含生物科技、醫學、工程科學、環境科學等。
 - URL : <http://www.csa.com.tw>
3. 於 86 年 4 月由交通大學、思源基金會、Elsevier Science 與 Cambridge Scientific Abstracts 共同舉辦 EES 與 IDS 之啓用典禮與數位圖書館研討會, 會中介紹多家資料提供廠商的系統, 共有全國近二百位圖書資訊專業人員與會。
4. 於 86 年 4 月建置 OVID 所提供的 Inspec、Compendex、Medline、Wilson、Eric、ABI/Inform 等全文或書目資料庫服務, 範圍包括理工、教育、人文、商管、醫學等領域。已經於 86 年 6 月完成系統建置。
 - URL : <http://ovid.infospring.nctu.edu.tw>
5. 於 86 年 5 月組成委員會, 委員由各大學及研究圖書館館長組成, 負責選擇適合的全文及參考資料庫提供全國各圖書館使用, 並負責結合群體的力量爭取議定優惠的資料庫使用費率。
6. 於 86 年 5 月底完成 InfoSpring 浩然數位圖書館總網頁建置, 介紹浩然數位圖書館中所有的資料庫。
 - URL : <http://www.infospring.nctu.edu.tw>
7. 預定於 86 年 11 月底設立 080 專線與諮詢櫃檯 (Front Desk), 提供國內產官學界諮詢服務, 與種子教師培訓課程。

第六節、結語

以合理的價格即時取得最新最好的研發與專利資訊是產官學界從事各項尖端研究與產品開發時最迫切需要的。國科會科資中心、交通大學、財團法人交大思源基金會為滿足國內各產官學界的研發需求, 乃與多家著名的期刊論文與專利資料提供廠商合作, 建置 InfoSpring 浩然數位圖書館。

InfoSpring 浩然數位圖書館的雙贏策略將使國科會科資中心、交通大學、財團法人交大思源基金會、各資料提供廠商以及國內產官學界 在最節省人力、物力與經費的前提下各取所需。對近年來正努力推動「亞太營運中心」之中華民國而言，更是「媒體中心」的具體實現。

參考文獻與相關 URL

- [1] Special Issue on Digital Library, *Computer*, 29(5), May 1996.
- [2] Stanford University Digital Libraries Project, <http://www-diglib.stanford.edu>
- [3] Carnegie Mellon Informedia Project, <http://www.informedia.cs.cmu.edu>
- [4] University of Michigan Digital Library Project, <http://www.si.umich.edu/UMDL>
- [5] Elsevier Science, <http://www.elsevier.com>
- [6] Cambridge Scientific Abstracts, <http://www.csa.com>
- [7] Special Issue on Data Mining, *Communications of the ACM*, 39(11), Nov 1996.