

專題演講

資訊科技與全球化數位圖書館發展 對書目資料庫合作發展的意義

IT and Global Digital Library Development
What this means to the cooperative development
of bibliographic database

陳 劉 欽 智
美國西蒙斯學院圖書資訊研究所

很高興能有機會在這裡，而且更樂於看到兩岸的朋友們齊聚一堂。因此，我要特別感謝莊館長堅持邀請我到臺灣來。老實說，我也越來越熟悉這座宏偉的圖書館、這個國際會議廳、和講臺了。

就在十天以前，我自己的 NIT'99 會議（1999 年新興資訊科技第十一屆會議）於 8 月 18 至 20 日假國家圖書館召開，共有 300 多位圖書資訊界人士出席，約有 70 餘位來自於 20 幾個國家，本屆會議主題包括新資訊科技所有的相關論題，但更重視數位圖書館的議題。華文書目資料庫並非我個人之專長，希望藉此機會講一些未來的數位圖書館或新資訊技術的發展情形。NIT'99 會議的成果出版了一本厚達 525 頁的書，書名為資訊圖書館與全球化數位圖書館發展（IT and Global Digital Library Development），您們也許會有興趣參考其中一些資訊。

壹、技術發展快速更迭

技術在短短的時間內改變是多麼的快速呀！我很清楚的記得在 1995 年嶺南基金會請我在中山大學講學，正逢兩岸數據庫研討會在廣州舉行，四年前的當時我談到 WWW 及 Internet，大家都很好奇，但今日已經是非

國家圖書館
漢學研究中心 編印

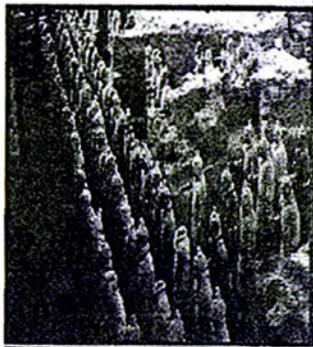
中華民國八十九年三月



常的著焉了。今天所提到的技術在短時間內亦會成爲事實，容我提出來給大家作爲一個參考。

我在 NIT'99 會議上發表的文章談到現在的 Internet、下一代的 Internet，在座的某些（並非全部的）華文書目資料庫專家學者也許不甚瞭解。

話說十五年前，早在 1984 年我做秦始皇多媒體的計畫（圖一），也對華文研究造成一個長遠的影響，當時秦始皇計畫是美國人文基金會支持最大的項目之一，是技術與人文的結合，結合東西方文化在一起，沒有時空的距離，故宮博物院副院長應有記憶 1985 年在國家圖書館展示，用滑鼠一點文字、影像、聲音都出來了，這在 15 年前是很新的技術觀念，一般人除非親眼所見，否則很難相信。



圖一 Project Emperor-1：互動式影像及後來的中國第一位始皇帝多媒體光碟計畫

資料庫只是一個指標，不是內容的本身，如何與資訊內容聯合起來，是一個很重要的方向。在過去十年中，我發表了好幾篇關於這方面的文章，其中一篇發表在 IBM 的 Multimedia Today 期刊上，由秦始皇的項目講到數位知識源、知識庫的發展趨勢，表示以後大家需要的是知識，並不只是目錄（圖二）。目錄是需要的，但更需要知識。

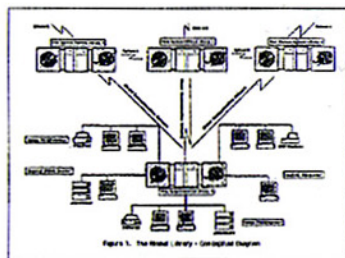
我也曾經在這個講臺上展示秦始皇影碟，其中包含了十八萬張圖片，一小時的影片，中英文雙軌的配音，在當時是很了不起，但以後技術發展很輝煌，由大光碟發展成爲小的 CD-ROM，數位多媒體技術可以讓我們將六小時的資料全部納入一片 CD-ROM 中。數據庫的建設與內容的充實才是真正的重點，書目資料量雖然大，需要有結構的組織起來，有一致的標示，尤其是可以立刻檢索到內容，才真正切合讀者的需要。書目或任何形式資料庫的建設是最基層的工作，是很重要，但是



圖二 Multimedia Today, 2 (2): 66-71 (March 1994)

貳、全球化數位圖書館

這次的會議讓我回想起 1993 年 4 月當時中央圖書館慶祝 60 週年的國際研討會上，我談到「全球數位圖書館未來的前景」，我用一個非常簡單的「全球化圖書館」(GDL) 概念模式（圖三），當時我提到有一天全球的通訊網路能夠將全世界的圖書館串聯在一起的時候，我們將會有一座全球化的數位圖書館。六年後的今天，我們已經有了全球化的網路 Internet，而且我們也開始認識了全球化數位圖書館。若每個國家寶貴的資料，用電腦儲存在一起，或者，以網路將全球的資訊連起來；例如，今天在此談華文，如果在這裡的主機上有一個機制連結全球其他華文主機，就會有一個全球華文知識庫雛形系統，滑鼠一點就去，那麼我們可以自全球各地得到所需的資訊。這也是我目前全球化數位圖書館雛形系統企圖展現我們可以連結整個世界資訊資源在一起，只要它們是以數位化的形式。



圖三 全球圖書館概念模式
(The Conceptual Model of the Global Library)
Presented at NCL in 1993.

雖然技術已經發展至如此進步，資訊內涵仍遠付闕如。目前的網頁技術使用滑鼠點去網頁上盡是指引性資訊 (directory information)，並沒有知識的本身。這是因爲建立真正的知識庫是非常浩大的工程 (labor-intensive)，而且許多人搜尋資訊不再滿足於印刷式的資訊，他們要的是知識庫的多媒體資訊。

我身爲柯林頓總統高科技委員會委員 (Presidential Information Technology Advisory Committee, or abbreviated as PITAC)，在今年 2 月 15 日委員會向柯林頓總統提出我們的報告「資訊科技研究：投資我們的未來」(Information Technology Research: Investing in Our Future)。我們深論資訊科技是如何改變了我們的社會，而且臚列資訊科技造成改變的十項主要的方面：

- 我們通訊的方式 (The way we communicate)，
- 我們處理資訊的方式 (The way we deal with information)，

- 我們學習的方式 (The way we learn) ,
- 醫療的方式 (The practice of health care) ,
- 商業的本質 (The nature of commerce) ,
- 工作的本質 (The nature of work) ,
- 我們設計和建造的方式 (How we design and build things) ,
- 我們進行研究的方式 (How we conduct research) ,
- 我們瞭解環境的方式 (Our understanding of the environment) , 以及政府 (Government) 。

以這種情形來看，圖書資訊顯然也會大幅地改變。現在，高速和高頻網路 (the high-speed and high-bandwidth network) 以及新世代網際網路 (the Next Generation Internet) 比目前的網際網路快 100 或 1000 倍使得發展及利用數位圖書館成為可能。但是，還是有許多困難之處，譬如：

- 定義未明；
- 角色不明；
- 技術問題-基礎設施 (infrastructure)、交互運作性 (interoperability)、比例性 (scalability)、永續性 (sustainability)，例如，是否有足夠的經費可以持續不斷的繼續保存下去等等。同樣是中文，在臺北、香港、美國、大陸都用不一樣的中文，以 NIT'99 的論文為例，同樣是中文 WORD 處理，就是不能交互運作 (interoperable)，其他的困難就更不用說了；

- Organization - metadata, indexing, 在組織方面，數據庫也會講到數據的根源、它的基礎、怎樣檢索、meta 的制定；

- 資訊存取 (Access)-怎樣索引及檢索 (indexing and retrieval) ；
- 數位典藏管理與組織 (Digital collection management and organization)
- 數位化那些資料？
- 使用者介面和人機互動 (User interface and human-computer interaction)-機器上用那些介面、機器與人之間互動 (interaction) 的情形是如何？以汽車為例，裝置的位置是固定的，任何開車的人都知道鑰匙插入的位置、排檔的位置等，啟動車子就可以開上路，但是網頁並沒有規範，進入每一個網頁查資料要靠想像力，何處該放什麼資訊，目前大家位置不統一，呈現一片混亂。知識源如何應付這樣的情況是很大的問題；

- 經濟上、社會上、法律上如何應付？著作權有使用上的限制、隱私權、安全性等；

- 跨多種語言 (multi-lingual) 的處理也是問題；
- 在國內、區內、國際上如何合作？

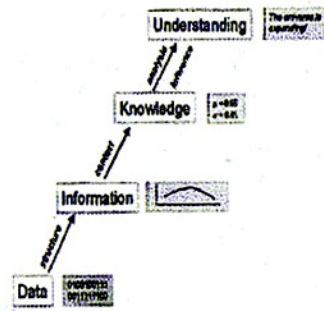
以上列舉的僅僅是一些明顯的問題，但值得我們認真的思考。儘管如此，目前技術的進展一定會推動全球化數位圖書館的發展。

參、新世代網際網路

所謂新世代網際網路 (Next Generation Internet) 指的是建立比現在最快的高速網路還要快 100 到 1000 倍的網路環境。正在發展中的一種稱之為寬頻網路 (vBNS)，臺灣目前已經擁有這項技術了，它以 622MB 的頻寬連

接美國，彼此間可以多媒體形式等應用以超高的速度互相交流。這種網路技術設施最適合發展分散式知識環境，其認知的進化將分階段進行，如圖四所示。

圖四說明數據資料，當它有了結構以後變成資訊，資訊與環境配合變成知識，知識分享以後可促進全球的了解。現在所有的相關研究方向都向這個方向進展。最重要的是，當我們專注於資料庫的建立與發展時，切記這種進化的方式，我們必須了解這種工作是重要的奠基工作，但一切只是剛剛開始而已。



圖四 分散式知識環境下認知的進化
(Evolution of Understanding in a Distributed Knowledge Environment)
(Courtesy of Steve Griffin, NSF)

肆、技術未來發展方向

接著我用幾分鐘的時間，介紹美國資訊科技最新的進展情形。因為時間很有限，請容我舉一個簡單例子說明。領導全世界計算機發展的麻省理工學院計算機中心 (The MIT Laboratory for Computer Science)，今年 4 月慶祝 35 週年慶，該中心主任 Michael L. Dertouzos 發表演說提到以後新技術將「可以做更多事但花更少的力氣」(Doing more by doing less)。該校獲得美國國防部三千八百萬美元 (38 million USD) 的經費資助

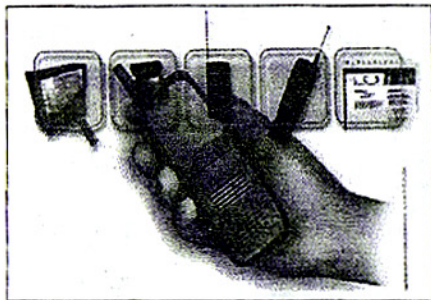
“Oxygen System” 研究計劃，把技術帶入生活，促進人類生活的生產力 (productivity)。

早期的自動化系統是一個 CPU 對很多用戶的分時系統 (Time-sharing)，例如在 Santa Monica 的 Dialog 國際百科檢索公司，擁有一座龐大的主機電腦提供全世界資訊檢索人員使用。後來個人電腦 (PC) 問世以後是一臺電腦對一個使用者。但是，在後個人電腦時代 (the Post-PC era) 裡的現象是，很多臺電腦服務一個使用者，所以需要一個智慧型代理人 (intelligent agent) 透過網路把外面的知識搜尋整合回報。

Scientific American 八月號專題介紹 “Oxygen System” 計畫 (圖五)，Handy 21 (圖六) 技術為該計畫的核心，類似現在的大哥大，但可以同時擁有電話、收音機、電視、呼叫器、無線連網等五個功能。

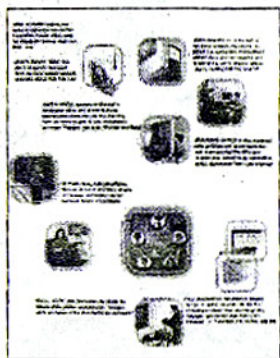


圖五 Scientific American 1999 年八月號封面



圖六 MIT LAC'S Handy 21 圖示

例如，如果有人對秦始皇有興趣，可用 Oxygen 立刻要求某人傳遞需要的資訊到手上顯現。圖七說明某人在巴黎用 Oxygen 呼叫波士頓和紐約的同事索取資訊，訊息立刻傳送到波士頓和另外一位身在紐約街頭慢跑的同事，所需要的資訊經過確認，且都蒐集齊全後，決策會被傳送回到巴黎需求者的手上 (詳情可參閱這樣的技術麻省理工學院已經展示證明可行。)



圖七 說明 LAC'S Handy 21 使用的方法

此外，麻省理工學院在語音技術方面也有長足的進步；例如，明天要回美國，今天就可以知道波士頓明天的天氣狀況，可以顯示在大哥大螢幕上，並用聲音告知美臺兩地天氣如何。

伍、廿一世紀資訊科技 (IT*2)

讓我再介紹另外一項最新的資訊科技發展，今年初，因為美國柯林頓高科技委員會 (PITAC) 給總統的建議中有一項是廿一世紀資訊科技發展 (a new IT*2 Initiatives)，柯林頓政府新增一筆三億六千六百萬美元 (US \$366 million) 支持研究資訊科技重點工作。下列為可能的進步方向：

- 能說、能聽、能瞭解人類語言的電腦，不但容易操作，而且能夠正確又即時的翻譯各種語言；
- 「智慧型代理人」為我們的資訊需求穿梭來往於網際網路之上，在浩瀚的資料中檢索和摘錄資訊；
- 各式各樣科學和技術上的探索能夠借助於超級電腦模擬，並讓全國的研究人員存取利用；
- 網路可以不斷地連結數以百萬計的電腦，還有百億以上的裝置；
- 比現在超級電腦更快千百倍的電腦問世，或者是另外完全不同技術 (採用生物的或量子計算方式) 發展出來的電腦；
- 新的精密軟體更具穩定性，更容易維護，並且更可以放心的去運轉電話系統、高壓電力輸電線路網、金融市場、以及我們的資訊基礎設施上的其他核心項目。

陸、因應數位時代做法

雖然技術發達到這樣的地步；但是，資訊內容卻仍然很匱乏。在我們圖書館界來說，建置資訊內容，譬如說作廿五史或四庫全書，是花費非常大量時間和人力的 (labor intensive) 工作，但於今之際，如果不大量建置內容的話，就好比有一條高速公路上跑的是牛車或者是小汽車之類的交通工具。

所以，書目資料庫的建立當然是必需要做的，因為它是奠基的工作。

沒有它作為基礎，我們不可能作進一步的發展。但是，使用者需要的是資訊的內容。如果光光只有目錄是不能滿足使用者的需求。再者，現在大家各自建立數位圖書館，就好像手上都有個寶貝，卻沒有合作的想法，也不能分享。全球的技術發展如此的快速，卻沒有什麼內容可用。今年4月號 Database 期刊 M. Ojala 的文章 “The end of online as we know it,” (圖七) 提到，數據庫的發展分成幾個階段：由例如 IBM Punch Card 的批次處理時代，經過 Dialog 系統等的互動時代，到動態紀元 (Dynamic Era) 充斥著各種的資源、通道、存取、和虛擬社區等等；到最終的直覺時代 (the Intuitive generation)，在那個時代裡，查詢工程師將能夠瞭解我們需要什麼。

因此，我相信在短短卅分鐘裡，我很快地向各位介紹了從 1985 年到未來的幾年當中技術方面的發展。我也試圖將這些關連到數位化圖書館方面的發展，以及資料庫的發展。對我們華文研究的專家們而言，有了搜索引擎 (search engine)，找到目錄是不夠的，還要能夠找到資訊內容本身，此點值得我們努力去達成。那麼，華文專家的挑戰是什麼？首先，應該經常舉辦互動式研討 (interactive discussion) 的活動；此外，必須建立國際上的合作，單打獨鬥的時代已經過去，應該大家分享彼此的專長經驗 (share expertise)，在數位環境下，分享數位資源，打破各自建立數位圖書館知識庫的藩籬，共同打造國際化的華文知識庫 (working toward international Chinese knowledge-base)，讓研究人員想要什麼立刻就可以擷取到手。我們應該重視這一點。反之，國外的華裔人士只是知道目錄，知道在北京、臺北、或是香港卻又拿不到內容，反而更痛苦。既然，外界的技術發展速度遠遠超越過我們追趕的地步，沒有內容的話也是枉然。

這項挑戰無與倫比。讓我們以興奮的心情和實際的行動一起迎向未來。

(作者謹向國家圖書館蔣嘉寧編輯為本篇文章翻譯成中文致謝)