

· 電子 TIBCO 爲何會如此受到國際圖書大廠和民營子公司的青睞。心中充滿
對圖書館員的敬意，對於 automated library system 中資訊會被認為何等開

數據資料庫

王梅玲

Numeric Databases

專欄內容：談論圖書館資訊系統的發展，討論其技術、應用及問題。

作者簡介：王梅玲，女，碩士，北京大學圖書館工作，研究方向為圖書館管理。

Mei-ling Wang Chen

Head

Acquisitions Department

National Taiwan University Library

Taipei, R.O.C.

Abstract

In 1979, the International Communication Bureau of R.O.C. connected some U.S. information service centers through the international telecommunication network. Since then, there are Dialog, ORBIT & BRS introduced into this country. However, the users are interested in the bibliographic databases and seldomly know the non-bibliographic databases or the numeric databases.

This article mainly describes the numeric database about its definition & characteristics, comparison with bibliographic databases, its producers, Service systems & users, data element, a brief introduction by the subject, its problem and future, library role and the present use status in the R.O.C.

前 言

古人言：工欲善其事，必先利其器。對於學術界與工商業界之研究人員言，近二、三十年所發展之線上資訊檢索系統，無異是其研究發展工作之主要工具。民國 68 年 12 月起，我國國際電信局經由國際電信網路連接美國資訊

服務中心。旋即有69年6月師範大學圖書館連接SDC公司之ORBIT系統，同年10月國科會科資中心連接Lockheed公司之Dialog系統，自此開啟我國使用線上資訊檢索系統之風氣。

從ORBIT、Dialog系統，國內又陸續引進BRS資訊系統。國人在檢索資訊時，常以其在極短時間內，針對需求之主題，迅速提供完整之文獻資料，覺得十分驚訝與滿意。殊不知線上資訊系統除了提供書目性資料庫，尚有一別開洞天之寶藏——非書目性資料庫，可提供使用者所提問題之直接答案。而其中數據資料庫更占有極大比例，並以其提供之資訊範圍廣泛，內容專精，對研究人員尤有助益，故不避個人淺見，草此簡文，簡單介紹數據資料庫之大要，用供參考。全文依下列重點敘述：線上資料庫之緣起，數據資料庫定義與書目性資料庫之比較，數據資料庫之製作者，服務系統及使用者，數據資料庫之資料內容，各學科數據資料庫之簡介，數據資料庫問題與趨勢，圖書館扮演之角色及國內使用現況。

一、線上資料庫之緣起

(一) 線上資料庫之發展

自十九世紀以來，科技文獻急遽增加，學報由100種，到1960年代末期已增為10,000多種，研究人員以人工方式查尋文獻，十分不便。第二次世界大戰之後科技進步，首將科學電子資料處理技術應用於科學文獻之書目控制，儲存資料在電腦磁帶、磁碟上，於是引起線上資料庫之革命性發展。1951年即有人提出線上查尋計畫，但由於尚缺乏技術，直至1960年，由於電腦分時技術、遠距離終端機設備與通訊能力已臻成熟，才有第一套線上查尋系統公開展示，即System Development Corporation公司推出之交談式系統(Drotosynthes)，可查尋Golden Book Encyclopedia內容。

一般對「線上資料庫」(Online Databases)所給予定義為：「以電腦機讀型式收集資料，並提供使用者可透過遠距離終端機與微電腦，以交談方式獲取資料。」(註一)1960年代中期後，許多科技進步更推動線上資料庫之迅速發展。這些技術包括：1.以電腦機讀型式或電子出版方式收集出版社或學術機構產生之書目、數據或全文資料；2.多功能分時電腦使查尋者透過終端機檢索更為便利；3.交談式電腦程式軟體；4.儲存設備技術進步，並且

成本降低；5.電腦終端機與微電腦技術；6.通訊網路之廣泛應用等。因之，導致 1971 - 72 年間許多線上資料庫系統紛紛出籠，如 Dialog 、 ORBIT 、 ERIC 、 NTIS 、 MEDLARS 等服務系統廣受歡迎，一時風靡世界。

線上資料庫以電腦技術協助研究人員，得以迅速且完整地獲取所需資料，因此受到廣泛使用與重視。其發展與成長速度十分驚人，從表一 Directory of Online Databases 提供線上資料庫統計資料即可顯示（註二），1979 / 80 年收錄資料庫 400 種，而 1988 年增為 3,699 種，不到十年即成長九倍。資料庫製作者從 1979 / 80 年的 221 家，到 1988 年增為 1,685 家，成長將近八倍。線上資料庫系統從原來 1979 / 80 年的 59 家，到 1988 年增為 555 家，成長將近十倍。這樣驚人的成長統計數字，足證線上資料庫之魅力十足，活力充沛，並受到普遍的肯定與重視。對於重視科技研究發展之我國而言，是極其值得深入探討與引進使用的。

表一 Directory of Online Databases 線上資料庫統計

Directory	出版期數	資料庫數	資料庫製作者數	資料庫系統數
1979 / 80		400	221	59
1980 / 81		600	340	93
1981 / 82		965	512	170
1982 / 83		1350	718	213
1983 / 84		1878	927	272
1984 / 85		2453	1189	362
1986		2901	1379	454
1987		3369	1568	528
1988		3699	1685	555

(二) 資料庫分類

依資料性質，線上資料庫可分為書目性資料庫 (Bibliographic Databases) 與非書目性資料庫 (Non-bibliographic Databases) 兩種。一般使用者較熟悉常使用書目性資料庫；非書目性資料庫其實早已存在，而偏重在金融、經濟與科技範疇。其數據資料庫自 1975 年更是大量興起，然以其性質多為專門學科，很難預估全部總數或窺得全貌。

若依照 Cuadra & Associate 公司之 *Directory of Online Databases* : 一書所採用之分類，則可將資料庫區分如下：

1. 參考性資料庫 (Reference Databases)

含書目性 (bibliographic) 資料庫與諮詢性 (referral) 資料庫。

2. 來源資料庫 (Source Databases)

含數據性 (numeric) 、文字—數據性 (textual-numeric) 、屬性 (properties) 、全文性 (full-text) 等資料庫。

請參見圖一：資料庫分類系統



圖一 資料庫分類系統

(三) 非書目性資料庫

書目性資料庫指示資訊之文獻來源；而非書目性資料庫則直接提供事實資料或實際資訊，如數據、全文資料，此種資料庫可作為來源資訊或原始資料，另尚有許多同義名稱，如來源資料庫 (Source Databases) 、資料檔 (Data files) 、資料銀行 (Data bank) 、資訊銀行 (Information bank) 、事實銀行 (Facts bank) ，歐洲國家常稱之為 Data bank 。

非書目性資料庫或參考性資料庫所包含資料以數據為主，稱為數據資料庫 (Numeric Databases) 。人類自老祖宗起，即對數字很有興趣，於是有日、月、時間、陰曆、陽曆等記時系統之制訂。近世許多學術發展、工商企業、政府機構，以其研究之需求，愈來愈重視數據數值資料，於是促進數據資

料庫之製作，這亦是本文介紹探究之主題。

二、數據性資料庫定義與 書目性資料庫之比較

(一) 定義

數據性資料庫常為科學與技術研究人員或社會學家直接查尋與使用第一手資料之對象。自 1980 年起愈加受到學術研究界之重視，究其原因有：1. 電腦檔案檢索技術之改良，2. 線上貯存成本降低，3. 微電腦普遍使用，4. 資料收集技術改善，便利及時獲取資料，5. 套裝軟體檢索與分析功能加強，6. 計量資料之需要日增，7. 許多問題叢生之領域，如能源與環境方面，需要以科際性方式解決問題與廣泛收集資料。因之，數據性資料庫逐漸受到歡迎。

「數據性資料庫」之定義為：「凡以電腦機讀方式收集數據為主之資料，組織而成之資料庫。」（註三）「數據資料」包含數值或一種與多種屬性或數值資訊，係經過測量、觀察，或計算所得之資料，例如某種化學複合物之沸點，或特定地點、時間之風速。有人將數據資料分成三類：1. 科技資料，即科技實驗室經過實驗測量或觀察所得數字或量化價值之資料。2. 社會資料，社會資料與科技資料愈來愈關係密切，難以區分，如交通統計與生命統計介乎兩者。其他傳統社會學科之範圍，如人口普查、農業、工業、市政、民意、社會調查、選舉統計、公司與政府會計統計及機關文件記錄等。3. 評估資料，將檢索所得資料評估其符合標準之正確度與可信度，稱為評估資料，常適用於科技資料，而難應用在社會與經濟資料。

根據 1982 年 *Directory of Online Databases* 收錄資料庫統計顯示（註四）所作學科資料庫分佈統計，共有 941 個線上資料庫，其中數據資料庫即有 427 個，占 45%，足證其重要程度；依學科性質分析，數據資料庫 80% 為商業、金融或經濟方面，有 78 個屬物理科技性，21 個社會科學性質。以線上資料庫從 1979 年至 1988 年成長近九倍速度來看，一般預測目前數據資料庫之成長亦不低於此，甚或超出數倍不止。但是有許多資料分散各處而鮮為人知，尤其數據資料庫常限制使用或僅限特定代理商，或非營利性機構所發展，故很難正確掌握現有數目。

表二 學科資料庫分佈統計

學科範圍	數據性	屬性	諮詢	全文性	書目性
物理學與技術 (包括生物學、地區學)	78	30	24	—	154
社會科學、教育與人文學	21	—	16	—	96
商學、經濟、法律	328	—	26	76	61
科際性	—	—	41	—	—
總合	427	30	107	76	331

(二)與書目性資料庫之比較

數據性資料庫提供直接第一手資料，以資料取得十分便利，故深受研究人員歡迎。其與書目性資料庫有許多相異處，為增加認識，將這兩類資料庫作一比較。

1. 資料類型：書目性資料庫提供文獻引文檢索資料，指示使用者要尋找資料之文獻，是間接資料；而數據性資料庫則提供數據資料，是第一手資料，使用者的問題可獲得直接立即的答案。不僅檢索直接數據資料，這些資料還可加運用，依學科範疇各有不同功能需求，大部分具備下列能力：1. 圖表呈現功能關係，分子結構、光譜；2. 統計分析；3. 功能關係之決定；4. 報表依各種需求列印；5. 模式模擬（註五）。

2. 資料之可信度：數據資料庫較書目查詢更重視資料之可信度，有些資料庫在列印資料時，也附有品質索引，以顯示資訊品質之評分。在物理學方面，最重視列述資料之來源，設備類型，設備狀況，以供研究者可重複結果。

3. 資料之新穎性：書目查詢比數據資料檢索較需最新之資料，而數據資料一般不受時間而改變，例如，水之沸點等永不改變。

4. 輸出結果：數據資料庫與書目資料庫均須透過電腦終端機與通訊系統查詢，但前者尤其重視圖表輸出展示，故需選購具圖表檢索功能之終端機。過去十年有許多這方面投資，目前圖表資料之展示與運用已充分用作預測經濟趨勢與工業成長。在科學方面則常用來鑑定三度空間組織與描繪物理資料，這些對研究幫助甚大。另外，二者查詢輸出也不相同，書目查詢結果常是 100 - 200 引文，而數據資料之查詢常只是單一數值或一組資料。

5. 成本：數據資料庫之成本各系統均有不同。與書目資料庫成本之易於預估計算比較，數據資料庫則依資料複雜與分析程度，其費用多寡相差甚大。

6. 檢索者：書目檢索與數據資料庫均需檢索者具備相當知識，但有許多不同處。數據檢索較重視學科或主題專長，尤其在資料運用與分析時。所以數據資料庫服務系統平均提供 1 ~ 5 資料檔，而書目系統則相對可有 40 ~ 100 個資料庫。這表示在同樣學科範圍，數據資料庫檢索者需要知道更多電腦語言、學科專長與增多解釋資料之工作。因此，數據資料庫之使用者多為具備學科專長之「資訊需求者」(end-users) 或中介者。

三、數據資料庫之製作者 ，服務系統與使用者

(一) 資料庫製作者

數據資料庫之製作者常常是各種機構單位。有些資料庫製作者同時也是出版品或報告之出版者；有些則屬營利性質，尚提供所製作資料庫之研究、顧問與諮詢服務；有些則是政府機構，不僅同時製作書目資料庫與數據資料庫，且負傳播相關學科資料與資訊之職。也有些製作者將政府之資料加以組織製作成資料庫，在組織過程中，製作者不僅整合各種來源之資料，且添加一些資料以提高其價值，如預測性資料等。

世界上數據性資料庫以美國製作的居多，而美國聯邦政府扮演極重要的角色，國際間也有許多機構單位從事製作，使用者本身也有許多研究單位貢獻心力，因此將分三方面：美國、國際活動與使用者專門研究團體，分別介紹數據性資料庫之製作者。

1. 美國：美國聯邦政府尤其致力於數據性資料庫之製作，有許多獨立單位或多單位合作計畫。*Encyclopedia of Information Systems and Services* 一書即介紹許多美國聯邦資料中心提供電腦化數據性資料庫系統。科技資料方面之製作者主要有：National Standard Reference Data System 與 Numerical Data Advisory Board 等。統計資料方面重要的有：Statistical Policy Division of the U.S. Office of Management and Budget, Bureau of the Census, Bureau of Economic Analysis, U.S. Department of Transportation 等。未來趨勢是聯邦政府將繼續支持數據性資料庫與系統之發展。

2. 國際間活動：國際上亦有許多單位從事製作，如：CODATA、ADD、UNISIST、ICSU World Data Centers、IAEA 與 COGEODATA 等。社會科學方面有 IASSIST (Int'l Association for Social Science Information Service and Technology)、INTER-UNIVERSITY CONSORTIUM FOR POLITICAL AND SOCIAL RESEARCH。

3. 使用者專門研究團體：民間有許多研究單位或學術團體為其學術研究需要而從事製作，如：IWGDE (ERDA Interlaboratory Working Group for Data Exchange)，社會科學方面有 APDUC (The Assn. of Public Data Users)，美國圖書館學會亦成立 Government Documents Round Table：從事研究社會科學數據資料庫。科技方面有美國電腦學會 (Association for Computing Machinery) 在其專門研究小組設立 SIGSOC (Special Interest Group on Social and Behavioral Science Computing) 製作並出版 *SIGSOC Bulletin*。

(二) 資料庫服務系統

提供數據性資料庫之系統稱為「數據性資料庫服務系統」(Numeric Database System)。一般將之定義為：「結合一種或多種數據資料庫與檢索數值資料之查尋系統，並且提供資料處理之功能。」(註六)一般須具備兩科功能：1. 直接查尋與檢索數據資料，2. 將檢得資料予以運算與使用。簡言之，即可直接查尋，檢索並可運用與分析數據資料之系統，以其結合數據資料庫與查尋系統，使得數據資料庫之檢索與利用更有效率。

有許多數據資料庫系統並不公開發行，僅限於特定使用團體，例如美國 ERDA 社區或聯邦政府。而美國聯邦政府支援許多數據資料庫與系統，其中有 National Standard Reference Data System of the U.S. National Bureau of Standards 與 Statistical Policy Division of the U.S. Office of Management and Budget 協助之聯邦統計系統，均是著名之計畫。國際間 CODATA 也極力推動數據資料之產生、組織評估與傳播；EURONET 也大力推動這類服務。

由於數據資料庫常有其特定主題、學科與使用者，故亦局限其市場。但是也有一些例外，例如 Lockheed's 公司之 Predicasts 檢索／統計服務，Data Resources Inc's (DRI) 公司之經濟資料庫與模擬系統，或是 Man-Labs Inc's 公司之熱力學資料庫，均為極成功之營利性數據資料庫系統。科技數據資料最大的系統 Chemical Information System (CIS) 也早已不屬於美國聯邦政府。另一方面，也有許多系統提供多功能、使用方便之套裝軟體，以利數

據資料之檢索與運用，如 Battelle's BASIS、Informatics MARK IV、MRI System's SYSTEM 2000 與 Cincom System's TOTAL 等。此外，由於數據資料庫無法與系統之查尋、檢索、處理、分析資料功能分離，故許多數據資料庫事實上亦是資料庫系統。

(三)使用者

儘管數據資料庫成長快速，但大部分仍針對某些團體或特定使用者而設計，如許多營利性數據資料庫包括經濟、金融、公司與工業資料及非營利性數據資料庫，是專為「資訊需求者」所發展。因此數據資料庫傾向「資訊需求者」之使用。據 1979 年美國學者 Hardy 調查結果顯示，資訊需求者之使用占 80%，館員與資訊專家中介者之使用不到 20%。然而隨著微電腦技術之普及，未來館員與資訊專家被期望扮演更積極之角色。

四、數據資料庫之資料內容

營利性線上查尋系統如果同時提供書目性資料庫與數據性資料庫，通常二者查尋方式相同，且較易使用，例如 Dialog 系統。倘若由個別數據性資料庫系統單獨提供如 I. P. SHARP，則常有不同之程式結構且使用者之輸入方式亦較複雜，常需要特別訓練使用者熟悉系統。然而這種複雜設計將有助於有效控制查尋效果，使檢索更具彈性且可製作各類格式報表，使檢索更精確。

儘管各數據性資料庫之資料會因學科不同而各有特色，大不相同，然資料呈現方式仍有共通處，首先認識這些資料 (Data)、變數 (variables) 與相關之術語，將有助於數據資料庫之檢索與利用。

(一)資料

首先應先了解資料庫資料之性質與格式，如同了解書目性資料庫知其為書目性，故資料庫之記錄以書目方式呈現，有書名、作者、主題、出版者、出版年等。而數據性資料庫常因資料庫之性質，資料特性各有不同，如化學數據性資料庫之資料有化學結構、複合物、分子公式、屬性等，而人口統計數據性資料庫之資料可能提供的是時間、地區與人口數量等。因此需要認識各數據資料庫之資料特性、變數、敘述語 (descriptors)、有關名詞 (term) 之定義與統計學通用名詞。

(二) 變數

統計數字中常以變數來敘述資料，如書目性資料庫之變數有作者 (author)、題名 (title)、出版日期 (publication dates)、資料來源 (sources)，而數據性資料庫常有下列四種變數：

1. 資料類型 (type)：標示資料之類型，如利率資料庫之資料類型將包括：抵押利率、原價利率、商業貸款利率等。

2. 時間 (time)：包括時間頻率與時間長度。時間頻率是指提供資料之頻率，常有每週、每月、每季、或每年。時間長度指資料庫資料涵蓋之時間範圍，例如 I. P. SHARP 資料庫的 Consumer Price Index 時間長度是從 1914 年起，而在 Dialog 的卻僅有二十年。

3. 單位 (denomination)：資料表達之單位，例如元 (dollars)、百分比、數量等，通常同樣主題有相同單位，如利率常用百分比。

4. 詳細層次 (level of detail)：資料常依詳細程度分成若干層次，如地理地區、工業、年齡、性別與種族等常有相同主題，不同層次。

(三) 其他相關名詞

1. 基數 (index No.)：將某時間之資料與基數時間之資料相比較，以觀察其變化，如消費物價指數 $1967 = 100$ ， $1968 = 105$ ，相當於增加 5 %。

2. 時間序列 (time series)：收集某段時間之統計數字，以觀察其規則性與變化，例如以時間序列來觀察消費物價指數。

3. 參見 (cross-section)：顯示同時間內各種不同之資料，有參考比較之功能，如 1978 年 8 月之薪水與工作時數。

4. 季節調整 (seasonal adjustment)：為避免資料受到尖峰或淡季時段影響而將原始資料作調整，例如修改 2 月與 12 月工作時間，以免受到 2 月時間過短，12 月假期較多之影響。

5. 迴歸分析：應用統計方法，同時顯示兩種以上變數之關係，可用作預測。

本文附錄一些實例，具體說明數據資料庫之資料內容，以加深印象，請參見附錄各例（註七）。一般為幫助了解專門變數與其應用於數據資料庫之方式，應仔細閱讀各資料庫使用手冊與相關文件，並輔以經濟學、統計學及各學科相關字典，將可使數據資料之檢索與運用得心應手。

五、各學科數據資料庫簡介

Fried 與 Kovac's 預測，自 1980 年代起數據資料庫會持續成長，僅就 1982 年 Cuadra Associates 所列 941 種營利性線上資料庫有 427 種數據性資料庫。由於許多科技數據資料庫係由研究單位個別自製，並限制使用，因此，據說尚有為數千種，甚至萬種的數據資料庫不為外界所知。據 Berger 與 Wanger 調查，大多數數據資料庫（約有 90%）屬於商業範圍，如經濟、金融、公司、工業專門資料等；約 10% 為科技數據資料庫。由於數據性資料庫已開發者數量相當多，限於篇幅無法一一敘述，僅以科學技術、社會科學、經濟與金融三方面簡要敘述其資料庫之發展狀況。

(一)科學與技術

科學方面的數據資料庫常是實驗室在界定條件與良好控制下各種實驗測量之數值記錄。各科技學科，因主題不同，資料類型各自相異。以物理化學為主之科學資料，其定義清楚，係經多次實驗與測量程序而產生；工程與材料科學相關之資料則不易重複實驗或界定範圍方式獲得；生命科學、生物科學與公共衛生在本質上不易分析量化，而較多敘述性「非數據」資料。地球科學、環境科學、大氣科學、氣象學與太空科學之資料偏向觀測結構，常由測量各獨立事件，連續記錄有系統方式收集而得。

使用此類資料，首要確定資料之有效性、正確性、可靠性與資料適用之條件。須詳細了解測量情形，包括測量工具、畫刻度、使用標準、數學計算或統計方式，以免錯誤使用資料庫而導致錯誤結論。

CODATA(The Committee on Data for Science and Technology of the Int'l Council of Scientific Unions)出版 *The CODATA Series of Dictionaries of Data Sources for Science and Technology* 介紹科技數據資料庫之處理方式，並分別就 Crystallography、Hydrology、Astronomy、Zoology、Seismology、Chemical Kinetics、Nuclear and Elementary Particle Physics、Molecular Spectroscopy、Geodesy 等敘述各科之數據資料庫。

以下僅就重點學科簡單介紹資料特性，並輔以著名相關之數據資料庫說明。

1. 物理／化學資料

此方面資料常經由實驗室實驗測量所得之數據，測量可不斷重複，在不

同時間、地方，由各個研究人員來作，數值結果可作比較，例如物理或化學屬性可不斷重複，光速在真空之測量與某些壓力之水沸點可重複與比較。有關此類資料已有許多可靠之數據資料庫，如核子資料 (nuclear data)、原分子熱力學與熱力化學資料 (atomic thermodynamic 與 thermochemical data)、固態物理屬性 (solid state physical properties) 等。對於學術研究界幫助甚大，然在資料傳播、查尋與使用上仍有很大困難，應請資料庫開發者努力改善，使用方便並幫助科學教育界與專業團體認識。

CIS (Chemical Information System) 係提供科學資料庫最大服務公司，由U.S. National Institute of Health(NIH)與U.S. Environmental Protection Agency (EPA)聯合提供之 NIH-EPA Chemical Information System 是一結合數據資料庫與書目資料庫系統，提供查尋二十個化學資料庫，包括物理與化學屬性、光學資料、生物資料、毒物資料、環境資料等。

有關次原子物理 (sub-atomic physics)，有 National Neutron Cross Section Center Brookhaven National Laboratory 收集與傳播有關中子誘導核子反應之實驗與評估資料，貯存在 Experimental Neutron Data Library，可透過 Cross Section Information Storage and Retrieval System 查尋。

有關熱力與其他屬性 (thermodynamic and other properties)，由 Man-Labs Inc's 公司提供之 Man Labs-NPL Materials Data Bank 可檢索複合物、金屬與合金系統，並計算個別平衡狀態與相位圖表等資料。

2 工程／材料資料

這類資料雖與物理化學相似，但有二點相異：(1)其測量可再製，但須依照原來準備之條件與處理方式；(2)以其使用者大多位在私人公司，故與市場產品息息相關，依賴這些資料可避免一般工程結構如建築、汽車、飛機、橋樑、電子工廠等之失敗。工程資料有關工程設計使用材料之強度與耐力，包括彎曲強度、拉張強度、易碎程度。這些特性無關於礦物之來源，而在於其被應用方式。是以相同材料在不同溫度與大氣中有不同之強度與效果。這意謂每一工程數據應詳述準備之過程與條件環境等。因之，工程資料庫不易產生，評估與維護，而從工程師需求觀點更不易獲得所需之資料庫。

¹Westbrook 與 Desai 所著之 “Data Sources for Materials Scientists and Engineers” 即列舉許多相關資料庫，可供參考。

BASIS (Battelle Automated Search Information System) 提供二個相關資料

庫；一為 Metals and Ceramics Information Center 資料庫，MCIC 係於 1971 年合併美國 Defense Metals Information Center 而成立，以收集軍事系統相關資料與傳播應用為服務目標；一為 Mechanical Properties Data Center 於 1960 年在美國 Department of Defense 成立，其宗旨在為增加科學家、工程師從事 Department of Defense 科技計畫之生產力。資料範圍包含金屬與應用在結構、國防之合金屬性。

3. 生命科學／生物科學／公共衛生資料

這三類資料關於強調質之觀察重於量之測量，故提並論。生命科學數據資料庫常指非數值資料，但不含書目資料庫且資料庫極其分散。Foundation of American Societies for Experimental Biology 主導生物學方面資料庫之編輯。

公共衛生科學的許多研究對人口有很大影響，例如流行病學 (epidemiology)。這類學科常針對大型取樣人口與控制人口，在某些主題，條件進行比較研究，例如比較某種藥物必須小心而正確地選擇取樣團體，於是其取樣方式、資料統計方式、資料測量程序與記錄極其重要。

美國國立醫學圖書館 (NLM) 的線上服務系統 MEDLINE、TOXLINE 與 CHEMLINE 提供 Toxicology Data Bank (TDB)，這是 NLM 的第一個數據資料庫，提供個別物質之資料，包含其化學物理屬性、毒物學反應、藥理學的、緊急治療、環境、職業等資料。

4. 地球科學／太空／大氣／氣象學／環境科學觀察資料

這些資料大部是觀察所測得之資料，常應用在非重複性的自然界事件上，且資料之收集與貯存規模龐大，遠超過其他學科。其設備常設置於人造衛星或地球表面上進行連續測量，並將資料送回觀測站，以磁帶收集。因此，為短時間收集資料磁帶，須有快速處理資料之技術，以配合系統處理大量資料。因之，觀察資料常有找尋磁帶位置與上磁帶之困難，所以在設計磁帶位置、索引製作、資料識別與保存需充分考慮資料量大，以更密集方式處理資料貯存（例如現在研究應用光碟技術處理 U. S. Geological Survey），並改良資料檔文獻製作，發展自動查尋相關資料之方式，以增加查尋效果。如 United States, Department of Energy Information Administration 正發展電腦化 Data Resources Directory 以查尋與敘述自身資料檔。這些困難與問題亦期待資料庫製作者與供應商發展新的資料處理技術，並傾向以使用者為導向設計。

相關之數據資料庫，如美國 Water Resources Division 之 Water Resources and Retrieval System 提供數以萬計美國領土地上與地下水觀測站資料。結合 Daily Values File Peak Flow File Water Quality File 成為 WATSTORE 資料庫提供水流、貯水容積、水溫度、沈澱物、地下水層次、河流階級、物理、化學、生物、放射性化學、地質、水文、水井結構與產品資料。

有關環境品質方面，U. S. Environmental Protection Agency 發展有關環境品質數據資料庫。著名者有二，一為 National Water Quality Information System (STORET) 可提供查尋水質、污染來源、污染造成魚死，與其他環境品質資料。一為 Aerometric and Emissions Reporting System (AEROS) 提供污染之處理與報告資料。

(二)社會科學

社會科學是第一個以電腦化處理資料之對象，早溯自 1890 年人口普查以 Hollerith 孔卡與 1951 年 U. S. Bureau of the Census 用 UNIVAC I 來處理人口統計資料。雖其早期即已發展，然目前機讀線上查尋社會科學資料庫之數量仍不及科技數據資料庫。美國聯邦政府從事許多社會資料收集活動。The Statistical Policy Division of the U. S. Office of Management and Budget 並且摘要許多聯邦政府之統計資料，提供大眾公開查尋聯邦統計資料。而 National Technical Information Service、National Archives and Records Service、Statistical Reporting Service of the Department of Agriculture 與 the Bureau of Labor Statistics of the Department of Labour 等機構提供許多統計資料。

1. 教育

National Center for Education Statistics 提供 Educational Statistics Information Access Service，包含美國學校資料，與全國 Higher Education General Information Survey、The Post Secondary Vocational Education Directory 與 Elementary and Secondary Education General Information Survey 等統計。

2. 人口統計學

SITE II 市場分析資料庫系統提供使用者依美國地區查尋人口統計、住家，與其他資料庫，如 Census of Population and Housing、The Time Series City and County Data Book、OBERS、County Business Patterns、Census of Manufacturing、Transportation 與 Agriculture 等資料。

(三)商業與經濟學

商業與經濟數據資料庫主要提供商品產品價格趨勢及產品之成長、成長率或成長費用等之數據資訊。近年來此方面資料庫成長迅速，其涵蓋範圍有經濟學、財務統計、貨幣匯率、公司與金融機構之財務資料、保險與商品資料、物價、利率、股票、職業、薪資與工業特別資料等。

資料來源有多種：政府機關、國際機關、私人研究機關與個人研究等。可能來自一種資料來源，也可能來自多種資料來源。這類資料庫之製作者可能重新計算處理資料才在線上供應，如 National Planning Assn. 重新計算從 U. S. Bureau of Census 收集資料，另加文件說明，以幫助資料檢索。另一種附加價值特色是允許使用者可從不同資料庫建立個人資料檔，以提供各種分析使用。這類資料庫之軟體具有轉換資料、檢索、印製報表、建立模式並執行許多統計處理之功能。

這類資料庫大部分是營利性的，Data Resources Inc. (DRI)、Chase Econometric Associates Inc.、Warton Econometric Forecasting Associates Inc. 等是主要提供經濟預測資料之單位。

DRI 是美國最大經濟與商業資料公司，主要提供經濟資料庫、經濟計量模式、預測模式、資料處理服務與諮詢。其客戶來自零售商、製造商、銀行、政府與學術界。DRI 資料包括巨型經濟變數，如全國總產量、利潤、利率、投資、消費資料與美國經濟工業統計等資料。

Chase Econometric Associates Inc. 提供美國與國際性大型經濟、金融、工業、農業、能源、人口普查之資料及預測模式。可透過 Interactive Data Corporation's 分時網路查尋有關的二十個資料庫系統。

National Bureau of Economic Research Inc. 維持近三千種經濟時間系列與其經濟、管理研究活動。時間系列包括全國收入與生產帳戶、金融、美國國際事務、製造業、商業、人口、職責、收入與預測資料。客戶可透過十六個商業分時網路來訂購 NBER 之資料庫。

六、數據資料庫之問題與未來趨勢

(一)現況問題

數據資料庫經過十多年之蓬勃發展，但仍有許多現存問題，如缺乏統合

服務系統、標準問題、資料品質、資料評估與成本等，茲分別陳述於後：

1. 缺乏統合服務系統：書目性資料庫有許多著名成功的統合系統如：Dialog、ORBIT、BRS等，可提供多種資料庫選擇。而數據資料庫大部分僅有專門服務系統，提供一種或某主題多種資料庫，專供特定客戶，如Data Resources Inc. 是世界聞名之線上資料庫公司，僅提供一種數據資料庫。因之，使用客戶須訂購多種線上服務系統以應所需。並且須學習各系統之功能，然系統常變，十分不便，費時費錢，致使使用者常感挫折與困難。加以有關各數據資料庫之廣告與介紹資料極少，很難找尋這類服務系統。
2. 缺乏標準：各數據資料庫系統有其不同之語言與說明書。資料庫之檔案結構與形式尤其複雜，並且詞彙之使用、定義、測量單位、代碼、索引法與查尋語言亦各有不同。資料庫與系統之間彼此無法適用，常引起資料交換與系統溝通之困難，因此亟須發展統一數據資料庫之標準格式。
3. 使用者兼備學科知識與查尋技巧：為有效查尋，使用者應同時具備相關資料庫主題之學科知識、電腦資訊觀念及系統操作之技巧。然而資訊需求者雖長於學科知識，卻缺乏電腦資訊觀念與系統技巧；資訊學家與圖書館員中介者除需有系統查尋技術且要了解相關學科背景，因此愈形困難，尤其有些數據資料庫之查尋技巧與專門學科知識更是複雜。
4. 資料評估與品質控制困難：數據資料庫之正確與可信十分重要，然以資料有許多變動參數，很難控制品質。而資料評估須對主題作深入了解，費時費錢，但若不作好控制，則線上資訊服務價值全失，不僅浪費金錢，且會使用者失望，中介者尷尬。尤其經濟與金融資料庫對國計民生影響甚大，一旦資料錯誤，將引起巨大危機。
5. 成本複雜：各數據資料庫成本均不相同，包括線上時間與印製費。主要費用是與電腦主機連線處理時間，與檢索之複雜程度相關；若檢索簡單，即使印製結果費時，成本仍很便宜，但倘須機器複雜計算，則成本昂貴。因之，成本計算較為複雜。

(二) 未來趨勢

預測未來數據資料庫將有下列發展趨勢：

1. 成長迅速、愈來愈多：隨著資料庫管理系統、人工智慧、關係資料式、微電腦、通訊網路等電腦科技更臻成熟，數據資料庫會大量增加，尤其在商業、家庭消費市場方面。在資訊中心與服務使用者，均會長足進步，許多

資訊中心會像書目資料庫一樣普遍與自然發展。

2. 改善檢索方式：數據資料庫成功之路較書目性資料庫障礙重重，如設計複雜，輸出報表需多功能，資料控制很難完整與有效，且資料輸入十分費時等。然而數據資料庫之製作者與系統將致力改善其檢索方式，加強系統之模擬與分析功能，並將結合 CAD / CAM 電腦輔助設計技術，支援未來發展活動。

3. 考慮使用者需求：未來數據資料庫將配合使用者之需求，朝向使用便利、查尋有效、輸出多功能等目標努力。

4. 發展光碟型數據資料庫：近年來已有將數據資料庫資料儲存在光碟上，商業發行光碟數據資料庫。其優點有：可免於分時系統連線成本，提供更多資料運用之附加軟體，可購置多份，方便多位使用者等。市面已有發行如 Micro/Scan Disclosure 將 Disclosure II 資料庫之數據資料轉移儲存於光碟；Wharton World Economic Service Data 包含 1960 年美國等 41 國經濟與財務統計資料；TRINET's Survey of Commercial Buying Power 收錄有關經濟統計調查資料。儘管光碟產品崛起僅有幾年光景，但已提供另一種有效查尋資料之方法，其發展趨勢值得重視。

七、圖書館的角色

有關圖書館對於數據資料庫所應扮演之角色，早在 1965 年美國一項名為 INTREX Planning Conference Report 計畫中，學者 Pool 即認定圖書館應發揮傳播機讀式社會科學資料之功能。近年美國即有三所大學 University of California at Los Angeles (UCLA)、University of Princeton、University of Florida 實施大學圖書館人口普查資料計畫，由圖書館之參考服務主動提供人口普查資料。

1984 年美國 ARL 協會進行一項學術圖書館使用數據資料庫調查（註八），針對 ARL 協會之 101 學術圖書館調查圖書館使用數據資料庫之情形。各館反映經常使用之數據資料庫有 18 種：Foundation Directory、Books in Print、Electronic Yellow Pages、GPO Publication Reference File、CHEMNAME、CHEMSEARCH、Harfax Industry Data Sources、GAS Online、PTS (Predicats) U. S. Time Series、CENDATA、American Men and Women of Science、D&B

Principal Int'l Business or Million Dollar Directory、EIS Industrial Plants、PTS (Predicats) U. S. Forecasts、CHEMZERO、PATDATA(DATS)、CHEMSIS、Federal Research in Progress 等。

調查結果同時顯示：雖然數據資料庫在 1975 年才開始萌芽，但此次調查已知有許多學術圖書館開始使用。但僅限於同時提供書目性資料庫之營利性服務系統，以其系統易於使用，查尋者熟悉系統之程式語言、查尋策略與資料檔功能，故使用不困難。檢索者 95% 是館員中介者，以連線費用昂貴，需由曾受訓練人員操作較有效率。調查之圖書館並建議圖書館學校應將數據資料庫列入線上查詢課程，訓練學生使用，以克服抗拒使用之心理。

由這項調查得知，美國學術圖書館已開始重視數據資料庫，但一般使用仍限於小部分營利性資料庫服務系統，並未如同書目性資料庫一般興趣與普遍接受。究其原因有：(1)大部分資訊需求者直接使用，並不透過圖書館館員；(2)館員不熟悉數據資料庫，不易使用資料庫且難提供解釋與分析資料之服務。如果館員繼續排斥數據資料庫於其服務之外，則未來資料庫製作者與系統將直接對資訊需求者發行，不再依賴館員中介者，於是圖書館將逐漸被摒棄與淘汰。為確保圖書館地位，發揮資訊傳遞功能，館員應努力投入數據資料庫之研究，了解科技、社會科學、經濟與商業各學科數據資料之特色，學習查尋技巧與策略，以充分掌握各數據資料庫之主題、範圍、資料內容、成本、技術、文件與品質等。圖書館學與資訊科學學系也應提供相關課程，以訓練「資訊館員」順應此一資訊時代之需求。

因此，館員今後在數據資料庫所扮演之角色應趨於主動與積極，須充分掌握查尋技巧與主題專長，並儘量協助資料庫系統之設計，務求易於使用，以充分發揮中介者之功能。

八、結論與建議

線上資訊系統從 1970 年起即開始風行世界，一般大眾普遍使用書目性資料庫，卻忽視了另一寶藏——數據資料庫。它宛若一處尚未發掘、有待琢磨之鑽石寶礦：以其直接提供第一手資訊、各學科之數值資料，使用者可獲得立即答案，並且附加資料運用、分析、模式模擬、圖表程式、資料評估等多功能，說它是價值珍貴之寶藏一點也不過實。然而以數據資料庫各有其專

門主題與學科範圍，故常限制使用的對象，或不公開發行，因此雖云數據資料庫問世的數量已有許多，但是很難找尋。而資料庫各有其電腦程式語言與查尋方式，且缺乏標準與統合之資料庫系統，致使用者不易，故說它是未經發掘，有待仔細雕琢之鑽石，未能顯露光芒，十分可惜。

以數據資料庫之學科專長特色，目前使用者大部分為資訊需求者，資訊專家與館員中介者之使用僅占少數。隨著電腦科技進步，數據資料庫將會持續成長，如果圖書館館員不重視此一資訊寶藏，資料庫系統將直接針對資訊需求者發展，而圖書館之功能也會逐漸沒落。因此，身為此資訊時代之圖書館員應主動扮演積極的中介者角色，深入了解數據資料庫，確實掌握查尋技巧，廣泛學習相關主題之知識，並協助資料製作者與服務系統針對使用者需求設計。並從事各項引進、介紹、推廣與宣傳之工作，使數據資料庫迅速、容易、完整的為資訊需求者接受與使用。

台灣地區近年來積極推行科技整體發展，也主動引進國外若干著名科技資料庫與系統，並有計畫進行製作國內資料庫。先後引進 Dialog 、 ORBIT 、 BRS 等系統，民國 75 年又開始籌畫實施「科技性全國資訊網路」(註九)，第一階段引進國外資料庫如： BIOSIS 、 PREVIEWS 、 CA SEARCH 、 COMPENDEX 、 ERIC 、 INSPEC 、 MEDLINE 、 NTIS 等書目性資料庫。而國內科技資料庫之製作仍以科技文獻、人才資料等書目性資料庫為主，尚未見及有關數據資料庫之引進與製作，實在是一大缺憾。盼未來「科技性全國資訊網路」能評估數據資料庫之價值，並逐漸引進推廣，以免國內研究發展有遺珠之憾。

附 註

- 註 一 Cuadra/Elsevier, *Directory of Online Databases*, 9 (January, 1988): vii.
- 註 二 Cuadra, P.V.
- 註 三 James A. Luedke, "Numeric Databases and Systems," *Annual Review of Information & Technology*, 12 (1977): 120.
- 註 四 S. V. Meschel, "Numeric Databases in the Sciences," *Online Review*, 8 (February 1984): 79.
- 註 五 S. V. Meschel, p. 79.
- 註 六 James A. Luedke, pp. 120-21.
- 註 七 Patricia Suozzi, "By the Numbers: An Introduction to Numeric Databases," *Databases*, 10 (February 1987): 20-22.
- 註 八 Trudi Bellardo and Judy Stephenson, "The Use of Online Numeric Databases in

Academic Libraries: A Report of a Survey," *The Journal of Academic Librarianship*, 12 (July 1986): 152-55.

註 九 科學技術資料中心簡訊，第4卷(民78年1月)，頁3—4。

參考書目

- Bartkus, Edward P. "Use of Numeric Databases in Reference and Information Services." *Drexel Library Quarterly*, 18 (Summer-Fall 1982): 205-219.
- Bellardo, Trudi & Stephenson, Judy. "The Use of Online Numeric Databases in Academic Libraries: A Report of A Survey." *Journal of Academic Librarianship*, 12 (July 1986): 152-157.
- Blake, Monica. "Aspects of Electronic Archives." *Electronic Publishing Review*, 6:151-67.
- Bourne, Charles P. "Online Systems: History, Technology and Economics." *Journal of American Society for Information Science*, 31 (May 1980): 157.
- Chen, Ching-chih and Herson, Peter. *Numeric Databases*. Norwood, NJ: Ablex, 1984.
- Fried, John B. "Numeric Databases in the 80s." *Drexel Library Quarterly*, 18 (Summer-Fall 1982): 2-10.
- . "Online Numeric Databases." *Bulletin of the American Society for Information Science*, 1 (1975): 17-18.
- . "Online Numeric Database System." *Online*, 1 (1977): 70-72.
- Gallina, Paul. "On-line: Changing Products, Technologies and Roles." *Canadian Library Journal*, 43 (June 1986): 169-72.
- Hardy Nancy F. "The World of Non-bibliographic Information." *Information Manager*, 1 (July/Aug. 1979): 20-22.
- Introduction to Non-bibliographic Online Database Services*. Cuadra Associates, Inc., 1979.
- Jarvelin, Kalervo. "A Straightforward Method for Advance Estimation of User Charges for Information in Numeric Databases." *Journal of Documentation*, 42 (June 1986): 65-83.
- Lee, Lucy Te-Chu. "Non-bibliographic Databases." *Journal of Library and Information Science*, 9 (April 1983): 74-93.
- Luedke, James A. "Numeric Data Bases Online." *Online Review*, 1 (Sept. 1977): 213.
- . "Numeric Databases and Systems." *Annual Review of Information & Technology*. Washington: ASIS, 1977.
- Meschel, S. V. "Numeric Databases in the Sciences." *Online Review*, 8 (Feb. 1984): 77-103.
- Suozi, Patricia. "By the Numbers: An Introduction to Numeric Databases." *Database*, 10 (Feb. 1987): 15-22.

附 錄

SETTIMER
 CLEAR
 RESETOPTIONS
 AVERAGES
 AUTOLABEL
 TITLE 'CPI-PITTSBURGH (1967=100)'
 QUARTERLY, DATED 1/78 TO 4/78
 PUT USCPI 'PIT/UUO,UU81100,UU82100'
 TABLE ABOVE (Command for TABLE, displayed below)

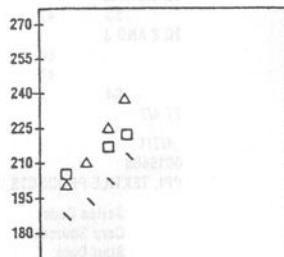
	CPI—PITTSBURGH (1967 = 100)		
	ALL ITEMS	FOOD	SHELTER
1ST/78	184.90	203.00	200.70
2ND/78	192.00	211.30	209.70
3RD/78	199.90	217.00	224.20
4TH/78	203.95	219.25	235.95
YEAR/78	195.19	212.64	217.54

PLOT ABOVE (Command for PLOT, displayed below)

ALL ITEMS :
 FOOD :
 SHELTER :
 CPI—PITTSBURGH (1967 = 100)

QUARTERS @
1/78 3/78

+ — + —



例一 U.S. Consumer Price Index Numeric System (I.P. SHARP)

PRODUCER PRICE INDEX

	1979	1980	1981	1982
U00000000 ALL COMMODITIES	235.57	268.84	293.39	299.32
U03 TEXTILE PRODUCTS AND APPAR	168.71	183.52	199.72	204.33
U0382 TEXTILE HOUSEFURNISHINGS	190.39	206.94	226.71	239.95
U057 PETROLEUM PRODUCTS REFIN	444.75	674.68	805.94	761.45

	1ST/81	2ND/81	3RD/81	4TH/81	YEAR/81
U00000000 ALL COMMODITIES	287.57	294.10	296.10	295.80	293.39
U03 TEXTILE PRODUCTS AND APPAR	194.07	198.97	202.20	203.57	199.72
U0382 TEXTILE HOUSEFURNISHINGS	220.80	223.33	229.37	233.33	226.71
U057 PETROLEUM PRODUCTS REFIN	777.33	834.77	811.93	799.73	805.94

例二 不同時段之資料檢索結果 (檢索 I.P. SHARP 資料庫系統)

File 565: Chase Econometrics - 86/Oct
 Copr. Chase Econometrics 1986

Set	Items	Description	PROTIV	
7S PRODUCER(W)PRICE(W)INDEX				
0	0001	PRODUCER	RA010	
164	1111	PRICE	RA010	
371	3333	INDEX	RA010	
S1	0	PRODUCER(W)PRICE(W)INDEX	RA010	
7S PPI				
S2	60	PPI	RA010	
7S TEXTILE				
S3	47	TEXTILE	RA010	
?C 2 AND 3				
60	2			
47	3			
S4	3 AND 3	2 AND 3	RA010	
?? 4/7				
4/7/1				
0019608				
PPI, TEXTILE PRODUCTS, UNITED STATES				
Series Code:	PPITXAU			
Corp Source:	BLS : PRODUCER INDEX RELEASE			
Start Date:	JANUARY 1948 (4801)			
Frequency:	MONTHLY			
Units:	INDEX 1967 = 100, NOT SEASONALLY ADJUSTED			
1986	JAN	215.4	FEB 215.3	MAR 215.6
	APR	214.9	MAY 214.9	JUN 215
	JUL	215.1	AUG 214.7	
1985	JAN	216.6	FEB 216.7	MAR 216.5
	APR	216.5	MAY 216.1	JUN 215.4
	JUL	215.1	AUG 215	SEP 215
	OCT	214.5	NOV 215.4	DEC 215.2
1984	JAN	214.9	FEB 217.1	MAR 217.5
	APR	217.5	MAY 218.1	JUN 218
	JUL	217.9	AUG 217.6	SEP 217.7
	OCT	217.4	NOV 216.8	DEC 216.5
1983	JAN	209.6	FEB 208.9	MAR 209.4
	APR	209.9	MAY 210.8	JUN 210.3
	JUL	211	AUG 212	SEP 212.4
	OCT	213.8	NOV 214.6	DEC 215.1
1982	JAN	215.1	FEB 215.8	MAR 214.5
1981	APR	214.7	MAY 214.4	JUN 213.3
1980	JUL	211.8	AUG 211.7	SEP 211.7
1979	OCT	211	NOV 210.9	DEC 210.5

例三 數據資料之關鍵字查詢(Dialog)

FIELD	ITEM	ITEM	ITEM	ITEM
PROD	0001	1111	3333	RA010
PRIC	1111	2222	3333	RA010
INDX	3333	4444	5555	RA010
PERI	4444	5555	6666	RA010

(參見資料庫資料夾之資料庫名稱)