

## e-化世界

### 《中國期刊網》已廣為人知，可您知道《中國期刊網》 是 CNKI 工程的源資料庫之一嗎？

#### CNKI--高度共建共用的知識資訊資源與網路化互動服務平臺

清華同方光碟股份有限公司知識網路事業部副總經理 張義民

中國知識基礎設施(China National Knowledge Infrastructure)工程的概念，源於世界銀行 1999 年初發表的《1998 年度世界發展報告》提出的 NKI。該報告認為，在知識經濟到來、經濟出現全球化的形勢下，發展中國家整體提高國家的知識創新能力和競爭能力成為國際生存與發展的關鍵；作為一種積極、長久的應對策略，應當充分利用現代資訊基礎設施，將知識生產、傳播、擴散的活動聯繫並組織起來，形成全社會可充分利用、互動的、網路化的知識基礎設施，建立高效的交流合作環境與學習環境。

CNKI 的目標是建造中國高度社會化共建共用知識資訊資源的基礎核心資源與網路化互動服務平臺：建成一個基於在「知識元」水平上全文關聯的，資料結構標準化的，人與資源、人通過資源可以即時以多媒體資訊交互的知識服務平臺，並通過標準化將各種動態更新的專業知識資訊資料庫關聯起來。無論資料庫的物理位置何在，無論資料庫專業性質和具體結構如何，都可以使其在知識資訊的內容上跨越資料庫界限而渾然成為一體。人們在這個平臺上可以用平臺的各種個性化服務手段，從所有資料庫中任意獲取所需要的知識資訊，並可以在人與人的交流互動過程中得到新的知識。其未來的定位是，將 CNKI 的資源、服務平臺，以技術優勢、市場優勢與全球同行的資源、平臺進行知識內容鏈結和電子商務鏈結，成為所有知識資訊資源社會化服務的主要門戶；通過數位化、資訊化、知識化技術的推廣，使 CNKI 的技術平臺成為各單位和個人的知識管理平臺；通過各種網路會議平臺、網上採編平臺、作者資料庫等技術與設施的應用，使 CNKI 成為知識生產的互動服務平臺。

CNKI 目前的主要建設內容--CNKI 源資料庫是按照知識資訊的不同載體形式分別開發的集成化全文資料庫和多媒體資料庫，是 CNKI 知識平臺

最重要的資訊來源。目前已開發的源資料庫有「中國期刊網」,「中國重要報紙資料庫」、「中國博碩士論文全文資料庫」,「中國重要會議論文全文資料庫」等等。開發源資料庫的目的是為了及時、完備地採集各種顯性資訊資源素材,為各種形式的資訊與知識服務打造數位化資源基礎。源資料庫大多加工成為純文本和軟體形式,可以利用不斷發展的資訊處理技術對其進行各種深度的再開發。如引文關聯、知識元關聯、語音識別、自動標引、自動聚類、視覺化處理等等。這樣才能充分地利用知識資訊資源的價值,充分滿足知識資訊資源高度共用的各種需求。

當源資料庫等資料庫開發完成後,所有資料庫將通過知識元鏈結、引文鏈結等被關聯成爲一個整體化的「知識網路」,既 CNKI 知識網路--就是利用知識內容之間的關係將知識資訊關聯起來的知識庫集合。知識之間的關係有很多種,如邏輯關係、繼承發展關係、交叉關係、滲透關係、轉移關係、拓展關係、借鑒關係等等。在傳統觀念看來,一個學科形成之後,本學科的知識必然遵循人們認識客觀事物存在與變化的本質規律的方法,而被整理成自己的邏輯體系。所以,傳統的圖書館必須有一個分類法,把知識納入一個分類體系當中。

用現代資訊學和認知理論看來,分析與綜合無時不在交叉進行,沒有什麼「道理」是孤立存在於某種學科範圍內的,知識之間的關係表現出很大的不確定性。傳統的知識分類已經不能滿足人們利用一切已有的知識發現新知識的需要。因此,人們希望有另一種查找知識資訊的方法,即不經過學科分類、概念分類這種先驗的、邏輯的思考,而直接找到知識之間的聯繫,而通過這種聯繫找到相關的知識資訊。

建立這種方法的途徑目前有兩種:

(1)引文鏈結

文獻之間的引證關係說明了知識資訊之間存在著某種關聯,而這種關聯則指明了知識發現的基礎和途徑。運用引文鏈結將文獻關聯起來,不僅可以找到知識之間的聯繫,而且是深度開發利用知識資訊資源極爲有效的

手段。《中國期刊網》從 2001 年 10 月實現文獻之間的引文鏈結後，用戶即可以通過文獻之間的引證關係，直接從資料庫中得到引證線索和引用文獻的全文，使文獻調研的速度和效果達到了前所未有的境界。目前，CNKI 資料庫間的引文鏈結工程已全面展開。今後，還將與高校圖書館等以及其他資料庫等建立引文鏈結。讀者能夠通過文獻之間交錯複雜的引用途徑在整個知識網路中自由地遨遊，從而最大限度地、系統地獲得相關知識資訊。

## (2)知識元鏈結

作為人們對事物存在特性與發展規律的認識，知識本身可以用文字給出一個簡明的描述，這種描述在資料庫中可以作為一種知識資訊元素，即「知識元」，它不僅可以直接表述知識的內容，也可以通過全文資料庫的索引技術將使用這一知識元的文獻關聯起來。另外，由於知識描述本身必將引用與其相關的知識元，所以利用對一個知識元的描述，將可以建立知識元之間的内容關聯。將大量不斷更新，不斷出現的知識元建成知識元資料庫，將會產生各種知識內容之間的內在網路化關聯，體現出個學科專業交叉滲透的關係特徵。將知識元資料庫嵌入到各種資料庫當中，它將把資料庫內部和各資料庫之間的知識資訊從内容上緊密關聯起來，並將通過文獻對知識發現的認識過程，使所有資料庫在知識元資料庫這一神經網路的支配下成為一個内容高度關聯的有機的整體。

以上兩種方法，給出的是知識之間必然存在的關係，比傳統的關鍵字、主題詞對文獻的描述更為具體、準確、全面、深刻。用這兩種方法將資料庫中大量的知識資訊直接鏈結起來，就形成了「知識網路」。

用這種方法逐步將國內外的各種資料庫鏈結起來，就將知識網路擴展到了世界範圍。資料庫還可以和圖書館、情報所等一切有非資料庫的資源鏈結起來，通過手工方式提供服務。只要用戶需要知識，就可以從鏈結中找到；只要資料庫是内容鏈結的，就有人從中攝取知識。

有了知識網路，還需要知識服務網路，後者為開展全面的知識服務提供堅實的基礎。無論資料庫是誰的、什麼結構、什麼軟體、什麼語言、放

在那裏，只要在 INTERNET 上授權使用，「清華同方分佈異構統一檢索平臺」都可以對其檢索。只要資料庫開發商、服務商願意把它的資料庫納入知識網路，就無須去購買它的整個資料庫，無須去學會它的檢索方法，無須從一般的檢索方法得到的大量無關的資訊中挑選有用的東西，人們就會「不惜代價」地買回它的知識資訊。但是，網路之複雜、資料庫版權之珍貴，把所有的網上服務都集中到一個物理點上是不可能的，也是行不通的，所以需要很多服務網點。使用同樣的軟體平臺和價格標準連鎖服務，就可以很快建立一個很大的市場。

資訊學將資訊一般分為顯性資訊與隱性資訊兩類，前者為以某種媒體方式顯示表達出來的資訊，後者為人腦中的意識。有的人將包含在大量顯性資訊中但有待分析、綜合等發現過程方可為人得知的資訊，也稱為隱性資訊。

我們所做的工作在傳統意義上稱之為資訊服務，就是通過某種服務方式向人們提供資訊，比如資料庫、新聞網，一般都指的是顯性資訊的提供。CNKI 的資訊服務不止於此，一方面，它需要通過人們互動的過程將人腦中活的隱性資訊顯性化，加以開發利用；另一方面，它將不僅提供給人們資訊，還將提供給人們解決問題需要的方案，即知識，稱之為知識服務。這種知識服務是人們最期望的，也是最增值的。

知識網路和知識服務網路共同構成了 CNKI 知識服務平臺，而 CNKI 知識服務最終的實現是依靠 CNKI 知識服務平臺來提供的。該平臺為人們獲取資訊和知識提供資訊資源保障、技術保障、服務保障。

#### (1) 資訊資源保障

CNKI 在建設自身源資料庫的同時，把國內外各學科專業知識資訊的核心資料庫、印刷館藏等各種資源，通過電子商務等商業機制和基於知識內容關聯的資訊組織法則組織起來，形成全球化的知識網路，使人們可以在更廣泛的知識網路中容易的找到所需要的資訊。

目前，CNKI 工程研製的「數位化圖書館檢索管理平臺」採用最新國

際標準，可將分佈存儲在各個數位化圖書館中的資料庫納入平臺中，而各資料庫可以是分佈異構的，通過「知識元資料庫」、「引文鏈結索引庫」和其他元資料庫鏈結索引庫耦合，使得用戶在該平臺上將所有資料庫視為同一「知識網路」來應用。

(2) 資訊與知識 資訊與知識服務技術保障

A. 可跨地區、跨平臺、跨資料庫檢索的超平臺搜索、檢索工具……分佈異構資料庫超級引擎，使全球各地的各種非規範資料庫可以無縫統一檢索。

B. 基於知識服務的檢索功能，如智慧聚類、自動摘要、知識視覺化等知識挖掘技術以及自動傳送等個性化服務技術。

C. 多媒體點播、檢索、加工處理技術。

D. 多媒體網路會議技術，用於直接互動交流。

E. 網上協同工作平臺，用於隨時隨地的工作研究。

F. 網路教育技術，用於社會化的各類教育活動。

G. 各種專業學習工具，用於處理整個學習過程的資訊流。

(3) 服務保障

資訊服務業的規模應當通過業界的緊密合作而發展，才能使已建成和正在建設的資訊資源以最快的速度在全社會共用。借鑒國際成功經驗，CNKI 工程在自身建設的同時，與資料庫提供商、圖書館、情報服務部門、網路運營商結成行業戰略聯盟，建立共用資源與行業標準，同時，形成以 CNKI 資料庫為內容服務門戶的知識服務網路體系。

以知識服務為理念的 CNKI 工程，對資訊資源分享方式的設計與構想超越了一般數位化圖書館。目前的數位化圖書館，通過數位集成與網路資訊技術，大大提高了人們獲取與傳播知識資訊的能力。但是，僅僅通過標題、關鍵字、作者等「元資料」進行關聯檢索或全文檢索等對「求知者」而言，並無法直接準確的提供全面針對的知識資訊，加之目前的資料庫結構、資訊加工標準與質量上的差異，使得資料庫之間的資訊關聯、內容關

聯更加難以實現。從這個意義上說，在現有數位化圖書館建設目標下開發的知識資訊資源，本質上還是一本本「堆積」起來的數位化圖書與期刊。除了信息量大、檢索快等方便之外，人們獲得資訊的手段與方式，依然類似於傳統圖書館「電腦書目檢索」與「開架借書」。而 CNKI 與數位化圖書館的服務模式的不同在於，人們從 CNKI 知識資訊資源中獲得的不僅僅是一條一條的資訊，而且是一個一個針對特定問題的解決方案，即知識，讀者可以進行「知識挖掘」，享受「個性化服務」。

CNKI 工程與數位圖書館的差別概括起來有以下幾點：

(1)總體目標不完全相同

數位圖書館的目標是顯性資訊資源的共用。CNKI 的目標是互動的知識生產、傳播、擴散、利用，顯性資訊資源的共用只是其內容之一。

(2)CNKI 資訊資源開發的範圍大於數位圖書館

CNKI 包括對隱性資訊、隱性知識的開發。

(3)CNKI 更強調實現資訊互動、協同工作

資訊化社會，互動是創造的源泉。數位圖書館只是靜態的資訊服務。

(4)CNKI 更注重知識服務

資訊服務回答事物存在與否的問題，知識服務則除此之外回答「是什麼」、「為什麼」、「怎麼做」的問題。

(5)CNKI 工程建設涉及全社會

數位圖書館則主要是它自身的事情。比如，做期刊資料庫，要同時搞期刊出版資訊化；搞博碩士論文資料庫，要同時辦博碩士論文數位化出版和博碩士論文評估；開發圖書資料庫，要同時辦各出版社的圖書網路出版等等。因為「雞」和「卵」必須共存共榮。

根據 CNKI 的理念，CNKI 數位圖書館不僅僅是清華同方自己開發的資料庫，而且要把國內外不斷生長的各種知識資訊資源組織、調動起來，供人們長期利用。因此，CNKI 是一個全社會獲取知識資訊的集成化、開放式、動態更新的超級資源平臺，是一種基於知識網路和知識服務網路的數位圖書館。