

全國第一套 當代文學史料影像全文系統

莊健國 中央圖書館閱覽組編輯

一、前言

本館曾於民國 74 年 5 月間舉辦「當代文學史料展」，展示 30 餘年來臺灣現代文學之成果，並以人工建立約 6 百位作家之資料卷計 8 百餘冊。輯錄資料範圍含括作者生平、手稿、相片、著作年表、作品評論等，由於所收資料內容豐富，且具獨特性，故廣為學術界、藝文界、出版界與讀者使用，且深獲好評。因原資料卷建檔至今已有 10 年，且使用率過高，加以新資料不斷產生，為考慮資料之長久保存，及擴增其使用彈性，爰採用儲存量大、具永久保存及高速存取能力，並可處理圖形、影像等特性之全文影像處理系統製作。在本館閱覽組歷任兩位主任的規劃下，本系統始得以順利完成，並榮獲 1994 年十大傑出中文資訊產品獎，前主任張錦郎先生對文學史料之徵集獨具慧眼；現任主任宋建成先生對自動化之推動頗有前瞻性。本文擬簡述該系統之開發經過及其功能。

二、收錄範圍

本系統收錄 40 年來臺灣地區約 6 百餘位當代文學作家之基本資料及其生平傳記、手稿、照片、著作年表、作品目錄、評論文獻、翻譯文獻、名句及歷屆文學獎得獎紀錄，總計約 5 萬餘頁資料，以全文及影像方式輸入電腦，並公開於臺灣學術網路上，亦可透過國際電腦網路將本系統傳播至世界各地，嘉惠全球愛好中國文學人士。

本系統收錄之資料共有九種類型，茲簡述如下：

1. 基本資料：記錄作家之本名(其他名)、性別、生卒年、籍貫、學經歷、主要寫作風格。
2. 照片：將作家提供給本館之照片掃描入電腦，使

讀者可在螢幕上一窺其真面目。

3. 手稿：將作家之親筆手稿輸入電腦，使得以目睹其真蹟。

4. 作品目錄：將作家之作品書目鍵入電腦，方便讀者用書名任意字查檢到該作品，即所謂之全文檢索功能。

5. 評論文獻：將別人評論該作家作品之文章全文掃描入電腦，讀者可在電腦上閱讀整篇文章。

6. 傳記文獻：將有關作家生平傳記之文章全文掃描入電腦，使讀者可在電腦上閱讀整篇文章。

7. 翻譯文獻：將別人翻譯該作家作品之文章全文掃描入電腦，讀者可在電腦上閱讀整篇文章。

8. 名句：名句乃將作家作品中有名之文句輯錄，使讀者可快速欣賞其作品之精華。例如：隱地在其「理想繼續燃燒」一文中，有此名句「晚上洗臉，是一天的結束；早晨洗臉，是一天的開始……洗著、洗著……人就老了！」

9. 文學獎：輯錄臺灣地區歷屆文學獎得獎紀錄及各獎項敘獎說明。

三、資料編輯

(一)資料蒐集方式

1. 報紙：地毯式逐日翻閱國內各大報副刊，蒐集作家相關資料。
2. 期刊：逐期翻閱國內重要文學雜誌，並利用本館開發之中華民國期刊論文索引光碟系統，增補資料。
3. 書籍：從各種文學專書、學位論文、學報中查檢作家相關資料，進而由其參考書目追蹤其他資料。
4. 網路：利用電腦網路查詢作家最新著作及相關文獻。

5. 作家：部份資料直接向作家本人徵詢。

(二)資料審核

本館當初建立作家資料卷時，並未考慮今日會以電腦處理，因此有部分資料書目不全，如缺漏刊名、卷期、頁碼，甚或漏頁等，均須設法補正，費力不少。但缺漏之處在所難免，且資料蒐集乃一無止盡之工作，今後仍將繼續增補資料，以維持書目之新穎與完整。

(三)資料撰寫

目前各作家之基本資料由館員根據蒐集之報章、雜誌、書籍及相關資料撰寫而成，雖然辛苦，但頗具意義。

四、系統開發與特色

本系統委託工業技術研究院電腦與通訊工業研究所，自民國 82 年 4 月至 83 年 3 月止，歷經 12 個月完成軟體開發工程。

開發本系統所使用之硬體設備包括：1. SUN SPARC 系統主機一套(主記憶體 64M bytes，硬碟 4.7G bytes)；2. 工作站五套(主記憶體 32M bytes，硬碟 424M bytes)；3. PC 工作站一台(主記憶體 64M bytes，硬碟 525M bytes)；4. 影像掃描機一台；5. 雷射印表機一台，噴墨式印表機二台；6. 備份磁帶機一台。

系統具有下列幾項特色：

(一)整合性系統

本系統是一個整合系統(Integrated system)，從資料庫的建立，到資料的修改、刪除、查詢、顯示、列印、排序，以至於系統管理等種種功能皆囊括在內。(請參閱當代文學史料影像全文系統架構圖，見次頁)

(二)生產線式的文件輸入方式

本系統將作家之傳記、評論、翻譯等文獻利用影像掃描機，輸入電腦，再利用影像處理技術做最佳之處理後，以中文 OCR 技術，將其轉換為中文文字檔，再應用中文文字校正技術整合之。使用中文 OCR 技術，可節省打字建檔時間，並方便進行遠地傳輸。系統設計採用工廠生產線原理，將其分為文件輸入、影像處理、中文辨識、文字校正、全文排版等階段，使文件輸入作業一貫化，以增加其生產效率。而輸入

之文字檔並可與系統之索引資料庫自動結合，提供查詢時之影像及全文顯示功能。

此外，本系統以自動化管理理念處理文件輸入方式，除了單筆文件建檔功能外，更採用批次建檔設計，結合人力規劃，在工作時間及程序上，發揮軟硬體配合的效率，達到最高的生產力。

(三)全文檢索系統

本系統以改良式逐字索引方式建立全文索引檔，並以資料庫管理系統(Informix)管理之。可快速建立新的全文索引資料，並不會因資料量之增加而減低系統之功能，同時亦提供快速之中文檢索服務。

(四)親和周備的檢索與查詢

本系統提供 19 種檢索點，方便使用者從各種角度查詢本資料庫。

- | | | | |
|---------|---------|----------|---------|
| 1. 作家 | 2. 評論者 | 3. 翻譯者 | 4. 傳記作者 |
| 5. 作品 | 6. 評論文獻 | 7. 翻譯作品 | 8. 生平傳記 |
| 9. 作品類別 | 10. 出版年 | 11. 出版者 | 12. 報刊名 |
| 13. 名句 | 14. 語言 | 15. 性別 | 16. 生卒年 |
| 17. 籍貫 | 18. 文學獎 | 19. 寫作風格 | |

畫面設計採親和力較高之選項式查詢模式，讓使用者能以最少之操作指令，查獲其所需之資料；其或無需輸入中文字，只要會運用瀏覽功能，也可進行資料檢索。

本系統還提供再次查詢功能，讓使用者可以任意往來於各檢索點，進行資料檢索，而不必每次都要跳回主畫面，再進入其他查詢子畫面。

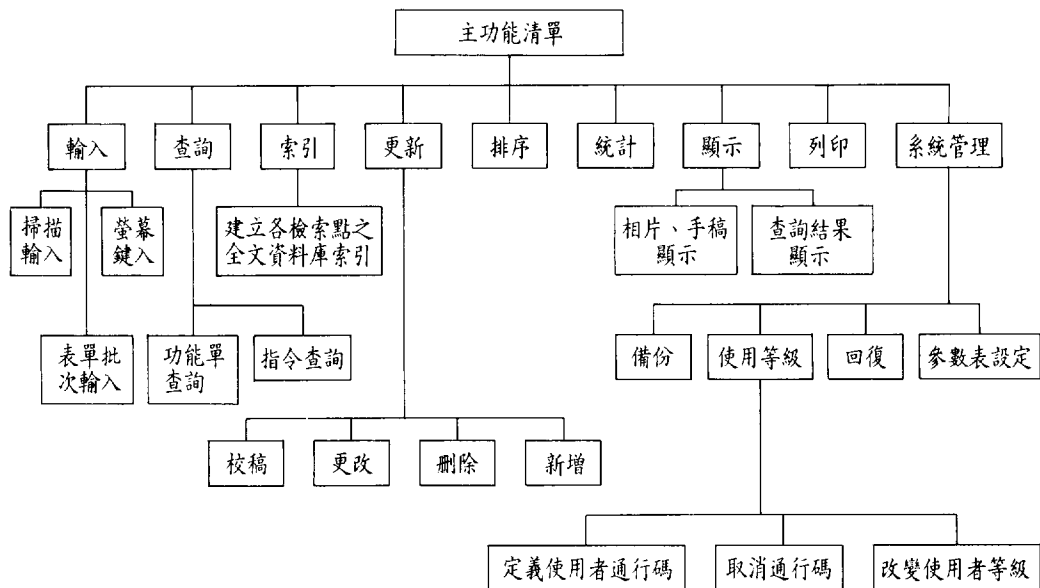
系統亦具備作家權威控制功能，針對擁有許多筆名之作家，只要檢索一次，即可查獲其所有作品，而不必對其所有筆名，逐一進行檢索。

本系統為方便使用者再次利用先前已查過之資料，保留最後 50 組之查詢結果。

(五)多重鍵值查詢與重點提示功能

在查詢系統方面，本系統多方考量使用者的觀點，並充分發揮電腦科技的效益，以應驗電腦實現人類之夢想。舉例來說，假如要查知民國 80 年以後是否有人撰寫比較王禎和與黃春明小說差異性之文章，此時正是利用其多重鍵值查詢特定主題之好時機，只要在適當欄位鍵入王禎和、黃春明、小說、80 年四個鍵值，就可立刻進行檢索，這種功能在傳統的目錄索

當代文學史料影像全文系統架構圖



引上，是很難做到的。

在顯示書目資料畫面時，本系統也做了貼心的安排，此乃在該畫面顯示前，系統會自動查核該筆資料是否得過文學獎，如是，則會在該筆書目資料右側標示“Ψ”符號，以提醒使用者注意該筆資料，而不必跳離該檢索點，再進入文學獎查詢子系統。此即重點提示功能。

五、結語

當今從事學術研究工作者，在其研究某一主題之歷程中，最感迫切需要的是該研究主題之完整性，期能盡可能地避免遺漏該主題之重要文獻。然而在資訊爆炸的時代，利用傳統書本式目錄或卡片目錄來蒐集資料，是很難求其完備的，更何況有些媒體(例如報紙)沒有編製索引，可知研究人員蒐集資料之不易。

有鑑於此，又基於文學乃研究一個民族的重要史料，為適時保存當代國家文化資產，並為文學研究人士提供一個迅速、便捷、完整的查詢環境，本館不惜動員有限人力及財力，竭盡所能蒐集當代文學史料相關文獻，努力建立一個全國最完整的文學資料庫。

在建立此套資料庫系統時，乃企圖採取國外大型購物中心之經營理念，滿足顧客的多樣購物需求，只

要到該中心，即可將所需之貨品一次採購完畢(one stop shopping)。也就是說，本系統希望使用者可以在這邊找到他所需要的所有文獻，而不必分別查詢各類索引，因此本系統儘量提供作家之全部資料。舉例而言，過去研究臺灣文學史料者，可能對作家之作品尚可勉力蒐集，但卻不易掌握作家之評論或傳記資料，如今只要在電腦前面，輕輕敲下幾個鍵，作家相關資料即可呈現在眼前。因此本系統之開發，成功的為文學界開創一個應用電腦科技保存文化素材，並提供迅速、便捷、多人同時享用的查詢環境。

「當代文學史料影像全文系統」堪稱全國第一套，其系統軟體於民國 83 年 3 月完成，同年 4 月及 10 月分別在本館閱覽大廳展示，11 月榮獲 1994 年十大傑出中文資訊產品獎，12 月復參加 83 年資訊月展覽，於世貿大樓展出 9 天，皆深獲好評。日後將公開給國內外文學界人士及愛好文學之社會大眾免費使用，使用者可利用一般個人電腦透過臺灣學術網路或以電話撥接方式，進入本系統。

目前本系統雖已完成，但工作人員仍將繼續研發系統之升級，以符合時代之潮流，而文學史料之蒐集更是永無止盡，本館仍將本著服務讀者的精神，在有限的人力與經費下，力求資料之完整性。