



館務簡訊

Unicode工作小組93年度工作成果

為因應 Unicode 字碼日益普遍應用於圖書館界之趨勢，使多國語文的資料能正常查詢及顯示，臺灣地區 INNOPAC 系統使用者聯盟於民國 93 年 4 月組成 Unicode 工作小組，由聯盟成員館自願報名加入工作小組，參與的單位包括中央研究院計算中心及文哲所圖書館、國立臺灣大學圖書館、國立政治大學圖書館、逢甲大學圖書館、朝陽科技大學圖書館、臺北醫學大學圖書館及國家圖書館等七所圖書館。該小組由國家圖書館書目資訊中心負責召集，自 93 年 6 月至 12 月間共計召開 8 次會議，由各館指派代表共同討論，以解決圖書館自動化系統中文字碼轉換為 Unicode 可能遭遇的問題。與會人員來自各館編目組、資訊組、系統資訊組、採錄組、閱覽組及書目資訊中心等單位，結合系統管理者、建檔者及使用者，會議期間亦邀請字碼專家宋玉先生與會指導並提供諮詢。

在宋玉先生及 Innovative Interfaces Inc. 提供的字碼對照表基礎上，該小組 93 年的工作重點在於查核彙整不同來源的對照表，以修訂出 CCCII 及 Unicode 雙向轉換對照表，作為圖書館界共同遵循的字碼轉換標準。工作小組討論過程中發現問題較大的是多碼對一字形的 CCCII 與 Unicode 互轉的過程，從多碼對一字形的 CCCII 轉換為 Unicode 較無對應問題，但從 Unicode 轉換回 CCCII 時，必須在多個 CCCII 碼中選定一個對應的內碼。由於對應表的查核相當耗費人力時間，為及早將對應表提供各界利用，小組成員決議優先整理相同一個字形，其 Unicode 對應至多個 CCCII 的問題。各館代表依據共同訂定的選碼原則，以分工方式逐字依據下列工具查核對照表，以求較完整蒐集目前應用於各資料庫的 CCCII 碼：

1. 行政院文化建設委員會資訊應用國字整理小組編，張仲陶主編，《中文資訊交換碼》第 2 冊第 3 版（臺北市：行政院文化建設委員會，1985）1992 重印。

2. 行政院文化建設委員會資訊應用國字整理小組編，張仲陶主編，《中文資訊交換碼》第 3 冊（臺北市：行政院文化建設委員會，1987）。
3. 行政院文化建設委員會資訊應用國字整理小組編，《中文資訊交換碼異體字表》第一次綜合修訂稿（臺北市：行政院文建會，1992）。
4. The Unicode standard, version 4.0 : the Unicode consortium, edited by Joan Aliprand [et al.]. Boston, MA : Addison-Wesley, c2003.
5. Unihan Database, <http://www.unicode.org/charts/unihansearch.html>.
6. CNS11643 中文標準交換碼全字庫，<http://www.cns11643.gov.tw>.
7. 運用 PN300 查核昌泰中文內碼。

該小組於 93 年度查核修訂完成 1,325 組多碼 (CCCII) 對一字 (Unicode) 的對應表，該表總計蒐集 2,906 個 CCCII 碼。此外，並將各個來源的對照表匯集為可對應 Unicode 之 CCCII 碼計 27,436 個；可對應 CCCII 之 Unicode 碼計有 25,868 個。94 年度的工作計畫除了上述對照表之推廣應用外，亦將包括持續進行新增 Unicode 之增補及對應，Unicode 簡繁字對照表之整理等相關事項。字碼整理過程相當繁瑣，期望有更多熱心的同道加入整理工作，以便早日讓各圖書館順暢步入 Unicode 環境。（許靜芬）

新近出版品

電子出版品

《國家圖書館遠距學園93年度課程》

國家圖書館附設資訊圖書館製作 / 國家圖書館 / 9312 / CD-ROM / 3 片光碟 / 584MB / ISBN 957-678-413-1 / 盒裝 / 450 元 / GPN 4309303749

國家圖書館遠距學園 93 年度開設新課程「國家圖書館網路資源——鄉土篇」、「國家圖書館網路資源——財經篇」、「文件探勘 (Text Mining) 的原理、工具與實務應用」，歡迎上網學習，網址為 <http://cu.ncl.edu.tw>，並配合製作本課程光碟，以