

書誌데이터 入力시스템의 인터페이스設計에 관한 研究

— Interface Design for the Bibliographic Data Entry System —

崔 錫 斗*

□ 目 次 □

- | | |
|-----------------|--------------|
| 1. 머리말 | 7. 既存데이터의 利用 |
| 2. 入力計劃 | 8. 文 字 |
| 3. 포 맷 | 9. 카드目錄 |
| 4. 畫 面 | 10. 典據統制 |
| 5. 액세스포인트의 自動生成 | 11. 맺는말 |
| 6. 에러의 체크와 修正 | |

초 록

서지데이터 입력시스템의 인터페이스를 설계할 때 여러 가지 측면과의 관계를 고려해야 할 것이다. 입력계획, 레코드포맷, 화면설계, 액세스포인트의 자동생성, 에러체크, 기존 데이터베이스의 다운로드, 사용문자코드, 카드목록생산, 전거데이터의 이용 등이 그것이다. 본고에서는 이와 같은 사항들을 어떻게 실제의 서지데이터 입력시스템에 적용시킬 것인가에 대해 실례를 들어 논한다.

ABSTRACT

Many factors play a role in designing the interface for bibliographic data entry system in response to the various needs. Among the major variables to be considered are planning, record format, screen design, automatic generation of access points, error checking, download from bibliographic databases, character codes, card catalog production, and use of authority data.

This paper described with examples how these design considerations can be implemented in empirical system.

1. 머리말

書誌데이터의 入力시스템은 도서관자동화과정 중 最初의 實行시스템이 되는 것이 보통이며, 圖書館토탈시스템內 모든 서브시스템의 기본이 되는 데이터를 생산하는 중요한 시스템이다. 그러나 모든 데이터處理시스템에서 가장 어려운 부분이

데이터의 生産作業이며, 書誌데이터처리시스템에 서도 마찬가지이다.

書誌데이터는 대부분의 데이터가 文字列의 自然語形式이며, 길이의 변동폭이 심하며, 項目의 발생 횟수가 일정치 않으며, 레코드에 따라 데이터필드의 構成이 달라진다.” 뿐만 아니라 書誌레코드에는 한글, 한자, 로마자, 일본어의 가나, 특수기호 등이

* 이화여대 도서관학과 교수

混在하고 있으며, 하나의 레코드가 단독으로 존재하는 것이 아니라 레코드間이 밀접하게 연결되어 書誌階層(예: 시리즈- > 제 1 권- > 제 1 부 혹은 雜誌의 誌名變更)의 構成要素가 되기도 한다. 이와 같은 形式과 關係의 多樣性을 고려하여 데이터를 生産하지 않으면 안된다. 이런 관점에서 製作된 書誌데이터베이스가 있다면 그것을 최대한 이용하는 것이 迅速하고 低廉하게 質이 좋은 데이터를 生産하는 방법일 것이다. 그러나 우리나라와 같이 동양서관련 書誌데이터베이스가 빈곤한 상황에서는 당분간 自體의으로 데이터를 生産할 수밖에 없는 실정이다. 自體生産하고자 할 때 西洋書 일부의 다운로드處理를 포함하여 入力시스템 인터페이스를 어떻게 設計하느냐에 따라 入力費用, 데이터의 質, 入力所要期間, 편리함 등에 큰 영향을 미치게 될 것이다.

本稿에서는 이와 같은 問題와 관련하여 單位圖書館에서 自體入力用 書誌데이터 入力시스템의 設計·開發時 入力시스템 인터페이스 向上을 위한 方案을 例와 함께 論하고자 한다.

2. 入力計劃

書誌데이터 入力시스템의 設計는 다른 서브시스템과는 달리 가능한 한 데이터의 量, 人的 資源, 費用, 期間, 機器環境 등의 영향을 받지 않는 것이 좋다. 이들의 制約없이 獨立의으로 設計되고 그 設計에 따라 필요한 모든 것이 具備되어야만 좋은 入力시스템이 구현될 것이기 때문이다. 그러나 현실적으로 環境의 영향을 받지 않을 수는 없다.

入力시스템의 設計와 入力팀의 構成은 다음과 같은 몇 가지의 變數를 기초로 計劃된다. ① 입력 대상데이터의 量과 種類이다. 연간 입수데이터의

量과 溯及入力할 데이터의 量, 그리고 入力할 文字의 種類를 고려해야 한다. ② 入力の 期間과 豫算이다. 입력완료해야 할 계획기간과 예산액에 따라 시스템이 달라진다. 계획기간이 짧다면 外注를 고려해야 하겠지만 데이터의 질은 보증할 수 없을 것이다. 自體入力を 계획한다면 기기예산과 人件費가 늘어나게 될 것이다. 계획하는 입력시스템은 입력작업 자체가 迅速하고 便利하며, 데이터생산에 드는 費用이 低廉해야 한다. ③ 데이터의 質이다. 書誌데이터는 自然言語形式의 복잡한 構造를 갖는 데이터여서 에러의 확률이 높고 利用者의 장래요구까지도 수렴해야 하므로 入力要員에 대하여 상당한 수준의 專門知識을 要한다. 質을 높이기 위해서는 入力시스템의 構築側面에서 예산, 요원, 기간, 기기 등에 대한 여러 각도의 고려가 필요하게 된다. 書誌데이터란 한 번 入力하게 되면 그것을 溯及해서 보완한다는 것은 不可能하다고 생각해야 한다. 따라서 처음부터 質이 좋은 데이터를 生成하는 것이 좋다. 경험이 풍부한 入力業體가 아닌 한 外注보다 自體入력이 바람직하며 既存 데이터가 있다면 그것을 최대한으로 이용하는 것이 좋다. ④ 分擔目錄시스템이 개발되었을 때 會員機關으로서의 참여에 問題가 없어야 한다. 自館을 중심으로 하되 네트워크의 관점에서 他館과의 상관관계도 검토되어야 한다.

이와 같은 사항을 기초로 여러 측면에서의 重複入力を 피하도록 設計되어야 한다. 즉 圖書館間的 重複入力, 部署間的 重複入力(예: 수서부서와 목록부서), 레코드間的 重複入力(예: 시리즈, 부분, 유사자료), 필드間的 重複入力(예: 저자와 저자의 액세스포인트) 등을 排除할 수 있는 현실적이고 실용적인 시스템이 만들어져야 할 것이다.

自體入力を 하는 경우 入力시스템의 팀은 프로그래머그룹, 入力그룹, 그리고 체크그룹으로 構成한다. 入力그룹은 신규입수자료 入力그룹과 소급자료 入力그룹으로 나누고 주로 經驗있는 司書を 참여시킨다. 經驗있는 司書を 중심으로 構成함으로써 印刷形式 워크시트(printed worksheet)가 아니라 직접 畫面워크시트(electronic worksheet)를 使用할 수 있어서 그만큼 데이터의 入力이 빨라진다. 체크그룹은 目錄專門家로 構成되고 入力된 데이터는 2회 이상의 체크과정을 거쳐야 하며 入力시스템의 주도그룹이 되어야 한다. OCLC 會員機關의 入力데이터中에서 修正없이 받아들일 수 있는 正確한 데이터는 60%정도²⁾라는 研究結果에서도 알 수 있듯이 正確한 데이터를 生成한다는 것은 매우 어려운 일이다. 체크그룹의 活動이 데이터의 質과 一貫성을 유지하는데 지대한 영향을 미치는 것이다. 이 체크그룹에서 入力시스템에 대한 상세한 Documentation을 프로그래머그룹과 協力하여 작성하고 新規職員 教育과 필요시의 참조에 供하여야 한다. 프로그래머그룹은 基本設計에 충실하게 시스템을 開發하고 入力作業中 다른 두 그룹에서 나오는 修正要求事項을 충분히 入力시스템에 반영하도록 힘써야 할 것이다.

3. 포 맷

1) 項 目

圖書館에서 書誌데이터 入力프로그램을 開發하고자 할 때 제일 먼저 그리고 가장 심각하게 부딪치는 부분이 포맷(format)이다. 프로그래머와 司書の 생각에는 크게 차이가 있으며, 어느 한쪽의 主張이 완벽하게 다른 한쪽을 說得하지 못하고 있는 것 같다.

單位圖書館의 立場으로 보면 소장자료 데이터베이스의 利用을 中心으로, 東西洋書(주로 우리나라 자료, 서양서, 일본서, 중국서)를 막론하고 入力, 데이터베이스構築, 檢索處理 등을 統合하여 一괄적으로 運營하고 싶은 것이다. 그러나 各國의 MARC는 하나같이 자기나름대로의 特性을 갖고 있으므로 單位圖書館이 어느 MARC포맷을 따라야 하는가가 問題로 대두된다. 또 다른 側面에서는 필드의 細分에 관한 問題이다. 예로 US MARC를 보자. US MARC 포맷에서 태그 001, 100, 110, 111, 240, 245, 248, 250, 260, 503, 600, 611, 700, 710, 711, 745, 9xx의 內容만으로도 일반도서관 利用者나 스텝이 要求하는 情報의 97%를 만족시킬 수 있다고 주장하는 사람도 있다.³⁾ 또한 상기 項目中에서도 情報檢索時 100, 110, 111을 240, 245, 248을, 600, 610, 611을, 700, 710, 711을 각각 區別하고 있는 시스템은 거의 없다. 또한 US MARC 데이터를 다운로드해 보면 '형태사항'에 예 1과 같이 細分하여 入力하고 있다. 資料에서는 어떻게 할 것인가? 어느 정도까지 필드를 細分해야 하는가가 問題로 대두된다.

예 1) 300[^]a[^]x, 292 p. :
300[^]b[^]ill. ;
300[^]c[^]24 cm.

이와 같은 問題를 解決하기 위해서는 分擔目錄포맷을 決定해야 한다. 그것은 既存의 포맷이어도 좋고 새로운 포맷이어도 좋다. 다만 그것이 分擔目錄시스템(shared cataloguing system)을 위한 포맷이라는 것을 모든 圖書館이 인정하는 일이 가장 重要한 일이다. 물론 여기에는 外國語資料의 處理, 로컬情報處理를 위한 시스템構築側面에서의

構想도 포함되어야 한다. 圖書館시스템은 自體所藏데이터와 分擔目錄데이터를 중심으로 운영되어야 하므로 일단 分擔目錄포맷이 결정되면 各國의 MARC를 위시한 世界의 모든 既存 書誌데이터베이스는 所在情報의 역할을 하게 되며 分擔目錄데이터베이스와 함께 다운로드의 對象이 된다.

따라서 현실적으로 심각하게 討議할 가장 중요한 問題는 自館이 처리해야 할 國內外資料의 種類, 利用者의 要求, 서비스計劃, 그리고 分擔目錄시스템의 참여를 前提로 “무엇이 꼭 필요한 항목인가?”이다. 즉 있어야 할 項目이 무엇인가를 決定하는 일이다. 있어야 할 項目이 있고 그 項目을 컴퓨터가 區分할 수 있다면 그 項目이 어디에 어떤 形態로 있는 問題가 되지 않는다.

單位圖書館에서 書誌데이터를 入力할 때 그 데이터의 情報源을 보면 新規入力인 경우에는 資料自體가 되지만 溯及入力인 경우에는 書架目錄이 된다. 溯及入力を 위하여 書架에 있는 모든 자료를 입력부서로 옮겨와 原資料를 기준으로 入力하는 것이 바람직하지만 現實적으로 不可能한 圖書館이 대부분이다. 따라서 入力情報源은 書架目錄이 되며 필요시 原資料를 참조하게 되는 것이 보통이다. 그러므로 目錄카드에 기록된 情報를 중심으로 入力項目을 결정하게 되지만 項目의 결정에 신중을 기해야 한다.

單位圖書館의 自動化時 통합데이터베이스와는 달리 複本の 問題가 있다. 複本の 경우 原本의 레코드에 기록할 것인가 原本과 일치하는 項目을 뺀 나머지 부분만을 별도로 기록하여 原本과 링크시킬 것인가를 결정해야 한다. 長短點이 있으나 原本에 기록하는 경우 다른 시기에 입수된 複本을 處理하기 위해서는 시스템상의 處理過程이 복잡해진다.

2) 資料形態의 網羅性

入力포맷은 東洋書 / 西洋書를 망라하여 處理할 수 있어야 한다. 업무의 분담, 처리절차나 편의를 위하여 구분하여 처리할 수도 있으나 이 때에도 設計思想은 동일해야 한다. 東洋書中에서도 日本書, 中國書 등은 우리나라 資料나 西洋書와는 다른 면이 있으므로 충분히 고려되어야 한다. 예를 들면 日本書의 原書나 翻譯書에서 原書名은 그대로 액세스포인트가 될 수 없으나 西洋書의 原書名은 그대로 액세스포인트가 될 수 있다.

定期刊行物과 單行本은 간행의 형태가 다르다. 특히 定期刊行物의 書誌事項은 변경이 없는 한 한 번만 入力하고 그 후에는 소장사항을 바꾸면 되지만, 單行本은 매회 입력해야 하며 거의 언제나 새로운 레코드가 만들어진다. 定期刊行物을 單行本과 동일한 포맷으로 設計하고 소장사항을 별도로 관리할 수도 있다. 또한 檢索 / 提供用에서는 필요하다면 두 가지를 統合할 수 있으므로 업무내용에 따라 명쾌하게 처리하기 위하여 入力포맷을 달리 할 수도 있다. 이것은 어떤 種類의 資料라 할지라도 마찬가지이다. 다만 여러 가지 形態의 資料를 같은 포맷으로 統合하여 사용하는 경우 같은 필드의 共有에서 오는 曖昧함이 없도록 충분히 고려해야 할 것이다.

畫面上에서 각 필드를 구분하는 코드는 숫자태 그보다 文字레이블이 훨씬 기억하기 쉽다. “서지포맷에서 대부분의 공공도서관 이용자들이 필요로 하지 않는 항목은 없애거나 달리 취급해야 할 것이며, 포맷에서의 태그는 기계가 읽을 수 있을 뿐만 아니라 사람도 읽을 수 있어야 한다”¹⁾라는 Scilken의 말에 주목해야 할 것이다. 共同포맷을 결정하는 팀이 숫자태그를 필요로 한다면 숫자태그와 함께 데이터베이스내에서의 英文필드名도

함께 定義해주는 것이 좋을 것이다. 또한 相異한 形態의 資料에 대하여 동일한 포맷의 동일한 필드라 할지라도 畫面을 달리하여 예 2와 같이 다른 文字레이블을 사용하는 것이 좋다.

예 2) 단행본의 경우 서명 :
출판년 :
정기간행물의 경우 잡지명 :
창간년 :

4. 畫 面

1) 設計 指針^{5, 6)}

人間과 컴퓨터의 對話는 디스플레이畫面을 통하여 이루어지므로 畫面의 설계에 따라 必要情報의 傳達速度가 달라진다. 畫面은 다음의 3가지 사항을 기초로 設計되어야 할 것이다.

첫째, 表現할 情報의 量이다. 利用者에게 필요한 情報만으로 최소화시키는 것이 좋다. 利用者用 畫面領域은 25%를 넘지 않는 것이 좋으며 60%가 限界値라고 연구되고 있다. 동일한 情報라 할지라도 情報의 表現방법에 따라 적은 화면스페이스로 그칠 수 있으며 利用자가 화면상에서 必要정보를 찾는 데 걸리는 時間을 줄일 수 있다. 같은 量의 情報를 적은 화면스페이스로 표현하는 방법으로 다음을 들 수 있다. ① 적당한 略字를 사용한다. 즉 利用자가 익히 알고 있는 用語는 略字로 표현한다. ② 필요 이상으로 詳細한 설명을 피한다. ③ 利用자의 理解水準을 고려하여 간결한 用語를 사용한다. ④ 알고 있는 데이터포맷을 사용한다. 예를 들면 카드포맷은 레이블이 없어도 각 項目를 구분할 수 있다. ⑤ 컬럼헤딩을 불인 表示式을 사용한다.

둘째, 情報의 그룹化이다. 데이터의 관련 항목을 그룹化하면 읽기가 쉬우며 다른 그룹과의 관계가 일목요연해진다. 이 그룹들을 표현하는 방법에는 여러 가지가 있다. ① 스페이스를 이용하는 방법(spacing)이 있다. ② 색을 이용하는 방법이 있다. 색으로 관련 데이터를 구분하면 視覺적으로 효과적이다. 그러나 스페이스와 併用하는 것이 좋다. ③ 區劃線을 이용하는 방법(graphical boundaries)이 있다. 상당히 일반적으로 사용하는 방법이며 그래픽문자를 이용하여 관련 데이터그룹을 線으로 區劃하는 방법이다. ④ 돋보이게 하는 방법(highlighting)이 있다. 돋보이게 하는 방법에는 화면과 글자의 反轉(reverse video), 다른 畫面要素와는 다른 색을 사용하는 방법(color), 밝게 하거나(brightness or boldness), 밑줄을 긋거나(underlining), 깜박이는 방법(flashing : 완전히 켜다 켜다 하는 방법, 밝기를 변하게 하는 방법, 反轉시키는 방법) 등이 있다. 이와 같은 방법을 이용하여 그룹화된 데이터를 視覺적으로 명확하게 구분할 수 있다.

셋째, 情報의 配置와 順序이다. 모든 畫面에서 情報의 위치는 利用자가 가장 잘 찾을 수 있는 위치에 놓여져야 하며, 일련의 處理過程과 관련이 있는 모든 畫面은 一貫性이 있어야 한다. 그래야만 利用자는 必要정보가 어디에 있는지를 추측할 수 있으며 빨리 찾을 수 있다. 命令語, 에러메시지, 入力필드 등은 일정한 위치에 돋보이게 하는 방법으로 디자인하는 것이 좋을 것이다. 응용프로그램에 따라 달라지겠지만 畫面타이틀은 중앙, 畫面識別子는 오른쪽 위, 命令語필드나 機能키레이블은 아래, 상대나 에러메시지는 명령어라인 위 등을 생각할 수 있다. 다만 확실한 것은 情報를 가장 빨리 찾을 수 있는 위치는 왼쪽 위이며 가장

時間이 많이 걸리는 位置는 오른쪽 아래이다. 畫面上에서의 데이터位置와 파일상의 데이터構造와는 아무런 관계가 없다는 것에 유의해야 할 것이다.

畫面에 情報를 配列하는 順序는 다음과 같은 기준위에서 결정되어야 할 것이다. 그 중 공통적인 것을 보면 다음과 같다. ① 使用順序이다. 데이터의 入力順序(화면양식)를 情報의 記錄順序 혹은 카드作成의 順序와 가능한 한 일치시킨다. ② 使用習慣이다. 예를 들어 자료의 타이틀페이지는 書名, 著者順序로 되어 있으나 카드목록의 습관상 著者를 먼저 처리하는 습관이 있으면 그 순서를 따른다. ③ 重要度이다. 중요한 데이터를 먼저 배열한다. ④ 使用의 頻度이다. 비어 있을 확률이 많은 필드는 뒤로 돌린다. 즉 一般的인 要素가 特殊한 要素보다 먼저 와야 한다.

2) 畫面間의 關係

단일 레코드의 入力은 同一畫面에서 완료되도록 畫面을 設計한다. 각 필드의 데이터는 서로 깊은 관련을 갖고 있으므로 同一畫面에서 相互參照가 필요하기 때문이다. 同一畫面에서 可變長데이터를 入力하고 필드의 수정, 삭제, 새로운 필드의 추가 등의 處理를 위하여 입력용 스크린에 데이터를 만들면 편리하다. 각 필드를 최대한으로 길게 하여 固定長으로 두고 레코드에 따라 넘는 필드만 별도의 畫面으로 연장하는 방법도 생각할 수 있다. 이와 같은 방법을 사용하고 있는 시스템으로는 CDS/ISIS⁹⁾를 들 수 있다.

입력처리중에 다른 畫面(예 : 카드형식확인화면, 검색화면, 전거복사화면 혹은 연장필드화면 등)이 필요할 때에는 畫面을 分割하거나 일부 겹쳐 사용하는 것이 좋다. 부득이한 경우에는 畫面의 연결정보나 현재의 화면상태정보를 명확하게

알 수 있는 識別子를 준비해야 할 것이다.

5. 액세스포인트의 自動生成

1) 액세스포인트

書誌데이터의 項目中에는 동일한 내용의 데이터가 出力用 데이터(카드상에서의 바디부분)와 액세스포인트(카드싱에서의 표목부분)에서 서로 일치하지 않는 것이 많다. 따라서 동일한 내용의 데이터를 두 번 入力하지 않도록 하기 위해서는 액세스포인트를 自動生成하면 된다. 액세스포인트의 自動生成은 著者에 관한 사항(저자, 역자, 편자, 삽도가 등), 書名에 관한 사항(서명, 부서명, 대등서명, 총서명 등), 出版年度에 관한 사항(출판년, 刊권년, 연호 등)이 있다. 書名의 예를 예 3~5에, 著者の 예를 예 6, 7에 보인다. 예 3은 한자 데이터를 한글로 바꾸면서 副書名까지 액세스포인트로 만들어 주고 있다. 생성된 내용이 불필요할 때는 지운다. 예 4는 書名中의 特殊文字를 제거한 경우를 보이고 있다. 예 6은 西洋人名에서 성명을 自動倒置시킨 例이다. 단체저자도 같은 필드를 사용할 경우에는 길이나 일정한 距離리미터를 사용하여 처리할 수 있다. 人名인 경우에도 人名 뒤에 붙는 II, Jr., Sir, The 1st, 생몰년 등의 경우를 고려하여 處理한다. 예 7은 著者事項이 복잡한 경우를 보인 것이다. 入力데이터와 自動生成데이터가 일치하지 않을 때는 自動生成데이터를 고친다(예 5 참조). 또한 임의로 추가되는 액세스포인트는 해당 '액세스포인트' 필드에 계속 나열한다.

예 3~7에서 각 항목의 出力은 인쇄카드나 온라인목록에서 檢索結果를 디스플레이 할 때의 형식을 나타내고 있다.

예 3) 서명(입력) : 國家와 政治理論 : 現代資本主義의 國家와 階級

액세스포인트(자동생성) : 국가와 정치이론\현대 자본주의의 국가와 계급

서명(출력) : 國家와 政治理論 : 現代資本主義의 國家와 階級

예 4) 서명(입력) : “조선문법론”에 대한 연구

액세스포인트(자동생성) : 조선문법론에 대한 연구

서명(출력) : “조선문법론”에 대한 연구

예 5) 서명(입력) : 林巨正

액세스포인트(자동생성) : 임거정 → ‘임격정’으로 고침

서명(출력) : 林巨正

예 6) 저자(입력) : 洪吉童, J.F. Kennedy(共著)

액세스포인트(자동생성) : 홍길동\Kennedy, J.F.

저자(출력) : 洪吉童, J.F. Kennedy 共著

예 7) 저자(입력) : (organized by) Gesellschaft

fur Mathematik und

Datenverarbeitung mbH Bonn :

(co-sponsored by)

International Federation for Systems Research([and others]) :

(edited by) Donald G. Fink ; (tr. by) Soong-Hun Lee ;

(introduction by) Ann B. Dobie ;

(illustration by) Jacques

Durand ; (photography by) Leonard

Bloomfield

액세스포인트(자동생성) : Gesellschaft

fur Mathematik und

Datenverarbeitung mbH Bonn\

International Federation for

Systems Research\Fink, Donald

G.\Lee, Soong-Hun\Dobie,

Ann B.\Durand, Jacques\Bloomfield,

Leonard

저자(출력) : organized by Gesellschaft fur Mathematik und

Datenverarbeitung mbH Bonn :

co-sponsored by International

Federation for Systems Research [and others] ; edited by

Donald G. Fink ; tr. by Soong-Hun

Lee ; introduction by Ann

B. Dobie ; illustration by Jacques

Durand ; photography by

Leonard Bloomfield

2) 主要語

온라인목록에서 利用者가 主題語나 키워드로 檢索할 수 있다면 아주 편리할 것이다. 그러나 목록카드를 만들 때 主題語 혹은 키워드를 기록하지 않았던 圖書館이 온라인목록을 위하여 主題名이나 키워드를 入力한다는 것은 불가능한 일은 아니지만 매우 어려운 일이다. 資料를 分析하여 主題를 入力하기 위해서는 모든 도서관자료가 한 번씩은 主題專門家의 손을 거쳐야 하기 때문이다. 入力과 豫算이 있다하더라도 이들을 일관성있게 統制해줄 道具가 없다. 그러나 書誌項目의 검색만으로는 좀 미흡하게 느껴진다. 이를 어느 정도 보완할 수 있는 것이 主要語(keyword)를 抽出하는 방법이다. 主要語는 서명, 부서명, 총서명, 원서

명 등의 서명에 관한사항에서 自動生成할 수 있다. 생성된 主要語는 書名을 완전하게 기억하지 못하거나 主要語의 순서가 바뀌어도 검색할 수 있으며, 어느 정도 主題語의 역할을 기대할 수 있다. 다만 文學類인 경우에는 書名類에서 추출된 용어가 主題語로서 큰 意味가 없으므로 처리여부의 고려대상이 될 것이다. 主要語의 自動生成時 불용어 제거(동/양서), 조사 및 접미사 제거(국어), 어미처리(영어) 기능을 둔다. 主要語의 質을 높이기 위하여 不用語파일은 目錄部署에서 관리하며, 입력작업중 언제라도 削除, 追加가 가능토록 한다. 완전한 不用語리스트, 조사/접미사리스트, 조사제거를 위한 숙어리스트를 단번에 만든다는 것은 불가능한 일이기 때문이다.

우리말의 助詞의 제거를 위하여 우리말 역순사전⁹⁾과 國語辭典을 이용, 조사/접미사의 형태가 포함된 숙어사전을 만들어 참조한다. 이 辭典이 없으면 예8과 같은 單語도 일부가 잘리어 '가, 가, 늙은, 도서관학,...'등이 되어 버리기 때문이다. 예9에 처리의 결과를 보인다. 目錄者가 主要語라고 생각되지 않는 것은 지운다.

- 예 8) 가로
가을
늙은이
도서관학과
민주주의
보은
애국가

- 예 9) 서 명: 도전적 여성을 위하여
주요어: 도전, 여성

英語의 語尾는 명사의 복수형, 동사의 3인칭단

수 현재의 s, 과거/과거분사의 ed, 현재분사의 ing, 형용사의 비교급, 최상급의 er, est, 부사의 ly, 줄임말 they're, I'm 등, 소유격 s' 및 's 등을 40여종의 規則으로 처리할 수 있다. if~then~의 형식으로 표현한 規則의 예를 예10에, 處理結果는 예11에 보인다. 다만 西洋書에서 복수형으로 쓰는 용어의 's'나 복합명사인 경우에는 目錄者가 점검하여 고치거나 用語辭典 파일을 이용하여 處理할 수 있다. 예11과 같이 動名詞의 'ing'를 특별히 處理하지 않을 수도 있다.

예 10) 동사의 과거 ed의 처리

- 1) 단어끝 부분의 구성이 "- - 2 1 1 e d"이며,
1="t, p"중 하나,
2="a, e, i, o, u" 중 하나이면,
어간은 "- - 2 1"이다.
(예: stopped)
- 2) 단어끝 부분의 구성이 "- - 1 p e d"이며,
1="a, e, i, o, u" 중 하나이면,
어간은 "- - 1 p"이다.
(예: developed).

- 예 11) 서 명: Theoretical foundations of programming methodology
주요어: theoretical, foundation, programming, methodology

국어 복합명사의 分割/組合方法은 國語辭典으로 가능하며(예: 정보검색시스템-> 정보, 검색, 시스템, 정보검색, 정보시스템, 검색시스템, 정보검색시스템), 人名이나 團體名을 자르고 싶지 않을 때(예: 와인버그-> 와인, 버그: 산업경제연구원->

산업, 경제, 연구원)는 人名 및 團體名辭典을 만들면 된다. 生成된 主要語는 目錄者가 점검하여 최종 결정한다.

學術誌나 學位論文의 抄錄을 入力하는 경우 主要語를 自動生成하는 연구결과가 많이 발표되고 있으나 文章의 意味解析까지가 필요함으로 實用的인 시스템으로는 아직 미흡하다. 따라서 상기 主要語의 추출방법을 기본으로 하여 統計的技法과 온라인 시소러스를 병행하여 사용하게 되면 상당한 효과를 얻을 수 있을 것이다.

상기 예3, 6, 7에서와 같이 하나의 필드를 몇 개의 서브필드로 구분할 필요가 생기게 된다. 이 구분을 위해 사용하는 딜리미터는 가능한 한 記憶하고 區別하기 쉬우며, 論理的으로 단순한 것 중 한번의 터치로서 入力할 수 있는 것은 채택한다. 물론 데이터와는 구분할 수 있는 特殊文字여야 한다. 入力者의 入力코드는 단순히 하고 프로그램으로 對應코드로 變換시킬 수도 있다. 예를 들면 入力者가 형태사항을 "x, 292 p. : ill. : 24 cm"로 入力하더라도 파일이 만들어질 때는 필요하다면 "\$ax, 292 p : \$bill. : \$c 24 cm"와 같이 生成되도록 하여 入력을 간단히 할 수도 있다. 請求番號에서의 분류번호, 저자기호, 권차표시간의 딜리미터는 스페이스로 가능하다. 상기 예3, 6, 7에서의 모든 액세스포인트間的 딜리미터를 역슬래쉬(\)로, 主要語는 쉼표를 사용하고 있다. 데이터의 내용상 이와 같은 딜리미터가 데이터와 구분되지 않을 때는 두 자의 文字를 겹쳐서 사용하거나 그것을 自動生成할 수 있다.

6. 에러의 체크와 修正

1) 카드形式 確認機能

오래동안 카드形式의 데이터를 취급해온 司書가 入力화면에 익숙하려면 상당한 시일이 요구된다. 화면에 익숙해지더라도 入力데이터에 대한 形式上의 에러는 카드형식에서 더 간단히 발견된다. 카드형식에서는 바디(body)에 관한 사항, 부출지시(엑세스포인트)에 관한 사항, 복본에 관한 사항이 명확하게 들여짜기로 구분되기 때문이다. 入力完了 데이터의 誤謬를 확인하는 데 入力화면에서 카드형식으로 확인할 수 있다면 入力誤謬 復元에 크게 도움이 될 것은 자명한 일이다. 入力畫面에서 登錄番號를 入力함으로써 완전한 카드형식의 入力內容을 볼 수 있어야 할 것이다. 카드형식확인 화면은 별개의 화면을 사용하는 것보다 같은 화면을 겹쳐서 사용하는 것이 좋다. 이 형식으로 종이에 출력하면 카드형식 에러체크리스트가 된다. 畫面上의 카드형식 디스플레이의 예를 그림 1에 보인다.

2) 에러체크리스트

카드형식의 에러체크리스트에 主要語, 複本の 請求番號 등 카드상에 出力할 수 없는 데이터가 있으며 빈 줄이 많아 체크리스트의 量이 많아진다. 따라서 入力데이터 전체를 볼 수 있는 라인모드 에러체크리스트 出力機能을 둔다. 入力된 모든 필드를 보여주며, 複본인 경우에는 등록번호, 서지 제어번호 및 청구번호를 출력한다. 각 항목에 文字레이블을 붙이는 것이 좋으며 데이터가 入力되어 있는 필드만 出力하는 것이 좋다. 그림2는 출력의 예이다. 이 예는 複本을 原本레코드에 포함하지 않는 형이며 複本の 書誌事項은 없다. 따라서 原本만을 修正함으로써 에러修正은 끝난다.

등록번호 20202780_ 원본번호 20000000_ 유사번호 20000000_ 소장표시 _____ 청구번호 _____

서 명 _____

서명입기 _____

저 자 001.5 R19g A guide to commercial artificial intelligence : fundamentals and real-world applications / Wendy B. Rauch-Hindin. - Englewood Cliffs, N.J. : Prentice Hall, c1988. xxvi, 523 p. : ill. 판 포 시

저자입기 _____

원 서 명 _____

출판사항 20202780 Bibliography: p. 497-508. 출판년 _____

총 서 명 20202781-c.2 ISBN 0133687708 언 어 _____

I S B N 20202782-c.3

주 기 1. Artificial intelligence. 2. Expert systems (Computer science). I. Title. II. Rauch-Hindin, Wendy B.

키 워 드 _____

주 제 명 _____

입력 Insert 수정 Update 삭제 Delete Confirm <I,U,D,E,Y,N> c

<그림 1> 카드형식 디스플레이화면의 예 (서양서)

등록번호 : 10292835
 청구번호 : 792 R182b
 서 명 : 새론 俳優藝術
 서명입기 : |새론 배우예술|배우예술|배우의 수련|스타니슬랍스키의 연기론
 |배우예술론|L'Art du comedien|Stanislavsky on the art of the stage
 저 자 :
 저자입기 : Rappaport, S. A., 1863-?|Magarshack, David|꼬끄랑
 역 자 : 李光來, 金興雨, 李淑(공역)
 역자입기 : 이광래| 김흥우| 이숙
 출판사항 : 인천 : 우성문화사, 1985. 출 판 년 : 1985
 형태사항 : 289 p. : 삽도.
 총 서 명 : 극단배우극단연극신서 ; 1
 총서입기 : 극단배우극단연극신서 ; 1
 주 기 : 목차 : 제1부: 俳優의 修練/르파보르트; 李光來 譯. 제2부: 스타니슬랍스키의 演技論(Stanislavsky on art of the stage)/데이빗 마가르샤크; 金興雨 譯. 제3부: 俳優藝術論(L'Art du comedien)/꼬끄랑; 李淑 譯.
 키 워 드 : 새론, 배우예술, stage, 스타니슬랍스키, 연기론, art, comedien, stanislavsky
 부출지시 : 792 St24s한|792 꼬18a한

등록번호 : 10292836 원본번호 : 10292835
 청구번호 : 792 R182b c.2

<그림 2> 에러체크리스트의 예 (레이저빔 프린터)

체크리스트의 出力은 파일 전체의 出力, 등록번호의 일련번호나 임의선택번호에 의한 出力, 분류번호선택에 의한 出力이 가능해야 한다. 分類番號를 선택하여 출력하는 경우 100대, 200대, 110대, 211대 등으로 선택할 수 있고, 分類番號順으로 정렬(sort)하여 출력할 수 있으면 書架目錄과 쉽게 대조할 수 있어 편리하다. 또한 出力時 도트매트릭스프린터, 레이저비프린터, 라인프린터 등의 모드 선택이 가능하다면 용도에 따라 선택할 수 있어 좋다.

3) 修正과 削除

에러체크리스트의 체크결과나 作業中 발견한 에러는 수시로 修正하거나 削除할 수 있어야 한다. 에러의 修正이 불편하면 入力시스템 전체가 불편한 시스템이 된다. 또한 原本을 修正함으로써, 관련 데이터가 모두 修正(propagation updating) 되도록 한다. 入力화면에서 '입력, 수정, 삭제' 어느 모드를 선택하더라도 레코드 ID를 넣으면 그 데이터가 畫面에 디스플레이 되도록 하면 편리하다. 書誌레코드뿐만 아니라 불용어, 조사 및 조사의 음이 포함된 用語辭典의 修正, 追加, 削除機能도 入力화면과 겹치게 하거나 처리루틴이나 메뉴화면을 가까운 곳에 두어 쉽게 이용할 수 있도록 하는 것이 좋다. 다만 위험하거나 실행 후 회복할 수 없는 처리에는 언제나 확인메시지를 내야 할 것이다.

데이터를 入力하게 되면 불가피하게 몇 종류의 파일그룹으로 나누어야만 한다. 예를 들면 필요에 따라서 新規入力 데이터파일, 遡及入力 데이터파일, 온라인目錄用 데이터파일, 基本圖書 데이터파일, 指定圖書 데이터파일, 機關이라면 해당기관출판 데이터파일, 大學이라면 교직원 저작파일 등이

臨時파일이나 二重파일로 존재할 수 있다. 같은 종류의 파일이라도 入力일자, 入力자, 처리내용 등에 따라 다시 몇 개의 서브파일로 分割될 수 있다. 어떤 레코드를 修正하는 경우 그 레코드가 어느 파일에 있는지를 明確하게 구분할 수 없는 경우가 많으므로 修正時 파일을 지정하지 않더라도 찾아낼 수 있도록 하여야 한다. 찾아낸 레코드의 修正時 같은 필드내에서의 삽입모드(insert mode)와 겹쳐쓰기모드(overwrite mode)의 양쪽이 가능하면 修正作業의 能率이 높아지게 된다.

4) 에러의 防止

각 필드의 데이터형이나 길이를 보아 다르다면 에러메시지를 내어 에러를 防止할 수 있다. 예를 들어 숫자만의 필드에 문자가 들어오거나 길이가 다르다면 경고음과 에러메시지를 내는 방법이다. 또한 에러를 에러로 處理하지 않을 수도 있다. 예를 들면 각 필드의 마지막 句讀点(예: 서명끝/출판사항끝의 마침표)은 入力, 비입력, 혹은 이중 入力에 상관없이 目錄規則에 따라 프로그램이 체크함으로써 人間の 知的 努力을 경감할 수 있다.

같은 성격의 데이터群이 집중적으로 入力될 때는 필요한 항목이나 항목의 일부분에 디폴트值(default value)를 사용함으로써 에러를 防止하고 入力の 效率化를 기할 수 있다. 예를 들면 어떤 集書를 집중적으로 入力할 때 집서기호항목, 등록번호의 일정부분, 분류번호의 일정부분, 언어항목, 入力일자 등에 대하여 디폴트值를 사용할 수 있다. 또한 모든 숫자의 入力시 일정한 부분을 固定시키며, 뒤에서 채우기(right justify) 방법으로 선행하는 "0"을 入力하지 않도록 하면 入력이 쉽고 에러를 줄일 수 있다. 예를 들면 圖書館의

登錄番號가 동서 10000000, 양서 20000000 등으로 시작되는 경우 “동서 123”면이면 “123”만을 入力함으로써 “10000123”이 되어 선행하는 “10000”을 入力하지 않아도 좋다.

7. 既存 데이터의 利用

1) 다운로드

데이터를 컴퓨터로 처리하기 위해서는 우선 機械可能型으로 變換시킬 필요가 있다. 그러나 單位圖書館은 中央集中式 典據統制와 함께 網羅的인, 質 좋은, 그리고 一貫性 있는 데이터를 최소의 資源負擔으로 실현시키고 싶은 것이다. 이를 위하여 카드를 OCR로 入力시키고 자동포맷인식(automatic format recognition) 기술을 이용하여 機械可讀型으로 變換하는 방법을 論하기도 한다. 그러나 가장 좋은 방법은 모든 데이터를 포괄하고 있는 데이터베이스에 터미널을 연결하여 찾는 방법이지만 그런 것은 존재하지 않는다. 차선택으로 分擔目錄시스템을 이용하는 방법이다. 이 방법도 우리나라에는 分擔目錄시스템이 아직 構築되어 있지 않으므로 불가능하다.

현재 쓸 수 있는 방법은 既存의 權威있는 書誌 데이터베이스를 호스트컴퓨터에 로드시키거나 CD-ROM을 도입하여 特定目錄으로 檢索하고 그 데이터를 다운로드시키는 정도이다. 실제의 다운로드時에는 확장문자세트에서 發生되는 問題, 프린터와의 問題, 포맷上의 問題, 目錄規則 또는 觀點의 問題, 날권 또는 統合目錄의 問題 등 여러 가지 問題點이 있다. 예를 들면 독어, 불어 등의 데이터를 다운로드시켰을 때 擴張文字에 해당하는 자리가 비어 있거나 이상한 文字가 나오게 되므로 原本을 보고 고쳐야 하는 不便이 있

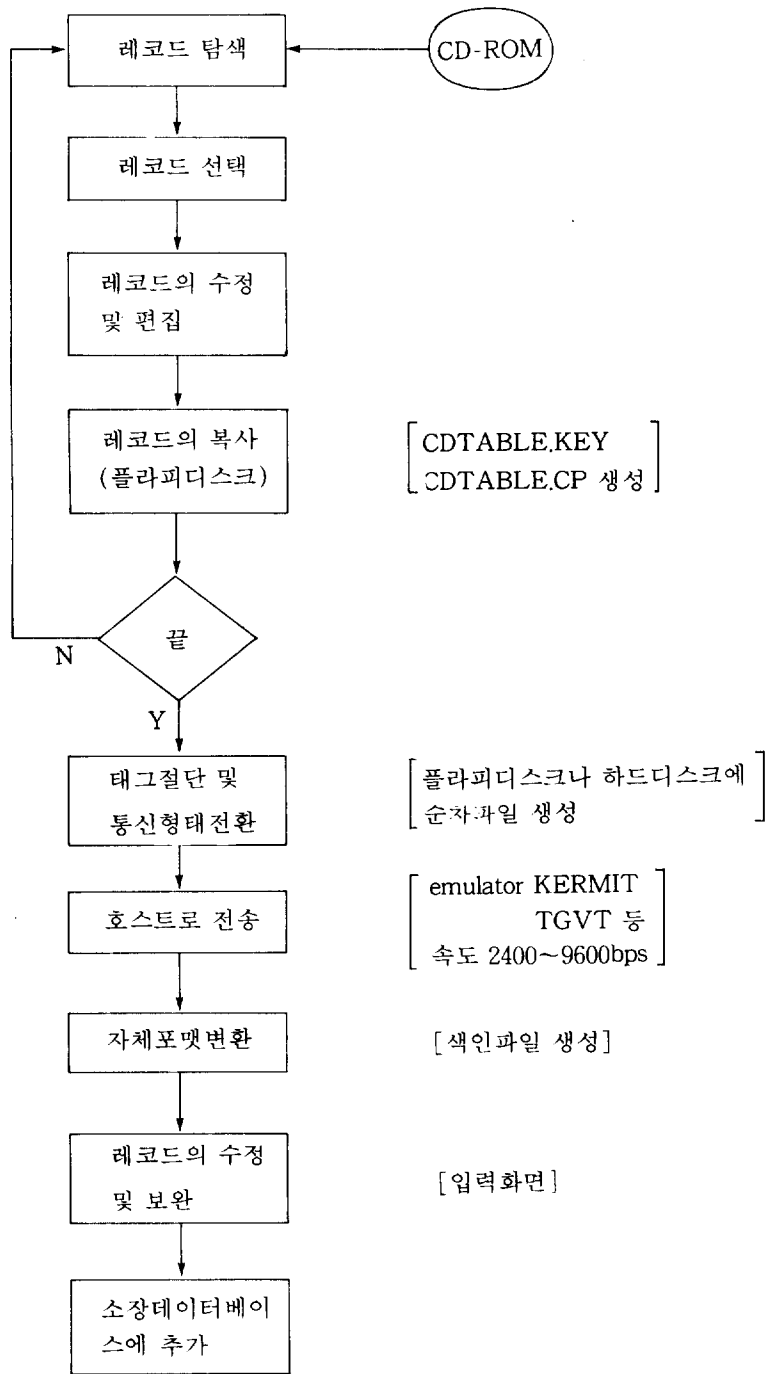
다. 또한 目錄作成의 觀點이 달라지면 著者나 書名 등의 入力데이터 내용이 쓸모없는 것이 되기도 한다. 그러나 이 方法은 지금도 많은 外國의 圖書館이 사용하고 있으며 司書의 업무량을 현저하게 줄여주고, 신속하게 質이 좋은 데이터를 얻고 있다.⁹⁾

우리나라에서도 최근에 韓國標準研究所를 선두로 다운로드방식에 의한 데이터베이스 構築이 활발하게 진행되고 있다.¹⁰⁾ 현재 널리 사용되고 있는 既存目錄데이터로서는 Biblio File(LC MARC CD-ROM)과 OCLC CAT CD450(A Cataloging System on CD-ROM from OCLC)이 있다. 다운로드의 過程을 그림 3에 보인다. 그림 3은 金昌根의 그림¹¹⁾을 일부 變更한 것이며 전술한 두 데이터베이스의 다운로드 概念은 동일하다. 그림에서 “레코드의 탐색”부터 “태그절단…”까지는 個人用컴퓨터에서 處理되며, “자체포맷변환”이후는 호스트컴퓨터에서 處理된다.

2) 檢索機能 및 複寫機能

기존데이터를 다운로드하는 것과 같은 概念으로 自館이 이미 入力한 데이터를 檢索하고 複寫하여 再利用하는 방법이다. 모든 데이터를 레코드단위로 入力하는 경우 복분/시리즈를 혹은 闡發만 다른 자료 등 비슷한 書誌事項을 갖는 데이터가 있게 된다. 이와 같은 비슷한 데이터를 다시 入力한다는 것은 浪費이므로 이미 入力된 데이터를 複寫하여 상이한 부분만 고칠 수 있는 기능을 두어야 한다.

複寫해 올 레코드의 ID번호를 알고 있을 때는 ID번호로 複寫가 가능해야 한다. 書架目錄으로 遡及變換을 하는 경우 ID번호를 언제나 알 수 있으므로 편리하고 손쉽게 비슷한 데이터를 入力



<그림 3> 서지데이터의 다운로드과정

할 수 있다. ID번호를 모르는 경우에는 여러 가지 檢索機能에 의한 檢索하고 해당 레코드를 찾아 人力畫面으로 複寫한다.

書名에 의한 검색화면의 예를 그림 4에 보인다. 그림4는 '노동'으로 시작하는 書名을 우측절단 검색방법으로 검색한 것으로 간략한 書誌事項을 보여주고 있으며, 긴 것은 잘라버리고 있다. '36'은 檢索件數를 나타낸다. 또한 複寫는 해당레코드에 커서를 옮겨 "c(copy)"를 넣도록 한다. 검색 화면은 입력화면의 일부를 겹쳐 사용하며 화면단

위, 라인단위의 상/하스크롤(up-scroll / down-scroll) 을 가능케 하여야 편리하다.

新規入力데이터에는 청구번호란이 비어있을 수도 있으며 이것은 分類作業보다 入力作業이 선행된 것으로 차후 入力한다. 原本番號項目에서 검색하고 복사하면 原本番號를 자동적으로 가지고 오며 類似番號에서 검색하고 복사하면 데이터만 複寫되도록 하여 類似데이터 入力を 손쉽게 할 수 있다.

등서 자료 입력

등록번호 _10000001_ 원본번호 _10000000_ 유사번호 _10000000_ 소장표시 _____

청구번호 _____

저 자 _____

저자입기 _____

역 자 _____

역자입기 _____

서 명 _____

연번	등록번호	청구번호	저 자	서 명	출 판 사 항	출판년도
1	10164939	331 조76L	趙容範 (편)	노동경제학	서울 : 풀빛, 1983.	1983
2	10164940	331 조76L c.2	趙容範 (편)	노동경제학	서울 : 풀빛, 1983.	1983
3	10164941	331 조76L c.3	趙容範 (편)	노동경제학	서울 : 풀빛, 1983.	1983
4	10298704	331.09 K952r한	(지은이:)위르겐	노동계급 등장의 역사	서울 : 푸른산, 1989.	1989
5	10298705	331.09 K952r한 c.2	(지은이:)위르겐	노동계급 등장의 역사	서울 : 푸른산, 1989.	1989
6	10298706	301.15 노3157	이보형, 주장현(노동과 곳 : 일하는 사람들의 삶과 세계관	서울 : 학민사, 1989.	1989
7	10298707	301.15 노3157 c.2	이보형, 주장현(노동과 곳 : 일하는 사람들의 삶과 세계관	서울 : 학민사, 1989.	1989
8	10288565	331.02 문94L 1987	文亨南	勞動關係法 實務便覽 : 書式作成 要領	서울 : 중앙경제사, 1	1987
9	10207084	331.02 신79L	辛仁鎔	노동기본권 연구 : 한국노동자의 노동삼권 보장	서울 : 이화여자대학	1985
10	10207085	331.02 신79L c.2	辛仁鎔	노동기본권 연구 : 한국노동자의 노동삼권 보장	서울 : 이화여자대학	1985
11	10288218	335.4 M369Z강좌	강좌전집위원회(노동쟁쟁과 농민문제	서울 : 학민사, 1989.	1989
서 명 노동				Hit Count:	36	

<그림 4> 입력화면에서의 서명검색

8. 文 字

1) 使用코드

사용하는 文字세트는 한자 4888자, 한글 2350자, 특수문자 1117자의 KS 2바이트 完成型(KS C5601)을 사용하는 것이 좋다. 여기에는 그리스 문자, 일본문자, 러시아문자, 그래픽문자 등이

포함되어 있으나 안타깝게도 독일문자 및 프랑스 문자가 없다. 한편으로는 있다하더라도 액세스포인트로는 사용할 수가 없다. 利用者의 입장에서 보면 入力方法이 달라 이와 같은 特殊文字들을 入力할 수 있는 사람은 많지 않을 것이기 때문이다. KS C5601을 사용하지 않는 컴퓨터는 自體코드를 사용하게 된다. 이 때 圖書館間 데이터互換性

의 문제가 생긴다. 바람직한 것은 自體코드가 KS C5601과 1:1로 對應되는 것이지만, 양쪽이 1:1로 對應되지 않으면 어느 한 쪽이 다른 한 쪽을 완전히 포함한다 하더라도 데이터의 互換性에는 무리가 있다. 또한 필요에 의해 自體코드를 擴張해 가거나 KS C5601이 擴張되면 코드의 互換性은 더욱 낮아지게 된다. 어느 한 쪽에 특정 한자코드가 없는 경우에는 한글로 바꾸는 방법을 생각할 수 있으나 한글(古字 포함)이나 특수문자의 일부가 對應되지 않는 경우가 생길 수 있으므로 주의해야 한다.

현재 KS C5601 문자세트로서는 不足한 文字가 많다. 그러나 이 코드體系를 사용함으로써 타 기관 데이터와 완전한 互換性을 유지할 수 있을 것이다. 왜냐하면 극히 일부 기종을 제외한 대부분의 컴퓨터시스템이 이 KS 문자세트를 지원할 것이기 때문이다. 없는 한자가 있을 경우, 固有名詞에 가까운 것은 한자/한글을 혼합하여, 一般用語는 모두 한글로 표기하는 정책을 채택할 수도 있다. 한자가 없으면 절대 入力할 수 없는 데이터는 處理를 후일로 미룰 수도 있으며, 單位圖書館에서 필요한 코드와 폰트(font)를 만들어 쓸 수도 있다. 國家標準코드란 모든 응용분야의 처리내용을 만족시킬 수는 없으며, 필요하다면 單位圖書館에서 코드를 추가할 수도 있는 것이다. 다만 KS 문자세트의 擴張이 具體化되고 있으므로 KS C5601에서 정해둔 利用者 定義域(한글 94자, 한자 94자, 특수문자 94자, 계 282자)의 범위를 넘어설 때는 여러 가지 面에서 신중할 필요가 있다.

2) 2바이트 데이터의 入力

한자입력을 편리하게 하기 위하여 漢字의 音節 變換方法和 利用者辭典 및 學習機能을 갖는 시스

템을 構築할 필요가 있다. 또한 日本語의 入力を 키보드입력이나 특수문자 선택형식으로 入力하는 것보다는 로마자 입력방법이 편리하다. 입력속도가 빠르며 영문키보드만 알면 入력이 가능하기 때문이다.¹²⁾ 다만 利用者가 入力하기 어려운 特殊文字는 미리 標準表記方法이나 일반적인 표기로 바꾸어 入力하는 것이 좋다. 特殊文字를 입력하더라도 액세스포인트만은 標準表記方法으로 바꾸는 것이 좋다. 예를 들면 우물라우트(umlaut)를 그대로 檢索하게 하면 利用者가 入力하기 어려우므로 'e'를 부기하는 방법이 좋다. 또한 'Σ, Ω' 등을 書誌事項에는 그대로 入力하고 액세스포인트에서는 '시그마, 오메가' 등으로 生成할 수도 있다.

9. 카드目錄

書誌데이터를 入力하여 카드目錄을 생산할 것인가? 현재 온라인목록시스템이 없다면 당연히 카드目錄을 생산해야 한다. 온라인목록시스템이 있는 경우에는 ① 생산하지 않는 방법, ② 書架目錄만 생산하는 방법, ③ 閱覽目錄까지 생산하는 방법의 3가지를 생각할 수 있다. ②와 ③의 방법은 遡及變換이 끝나지 않은 圖書館에서의 과도기적인 방법으로 遡及變換이 끝나면 결국 카드를 생산하지 않게 될 것이다. ②의 방법은 카드비용, 카드정리업무가 줄고 利用者가 신규자료부터는 온라인목록을 이용해야 하므로 利用教育을 여유있게 체계적으로 시킬 수 있으며, 遡及變換이 끝날 때까지 목록부서에서 書架目錄만으로 업무를 처리할 수 있는 利點이 있다. 다만 利用者가 신규자료를 찾기 위해 온라인목록을, 기존자료를 찾기 위해 카드목록을 검색해야 하는 不便이 남는다. ③의 방법은 카드목록으로 全體資料를 檢索할 수 있다는 長點

이 있으나 카드目錄의 生産, 保管, 整理에 費用이 많이 든다는 短點이 있다. 온라인목록시스템의 구축 후 서가목록만을 생산하다가 遡及變換이 끝나면 생산을 중단하는 방법이 좋을 것이다. 다만 카드를 프린터로 생산할 때는 부출카드까지 생산하는 것이 좋으며, 다시 複寫하여 부출카드를 만든다거나 헤딩을 매뉴얼로 기입 또는 修正하는 방법은 止揚하는 것이 좋다.

카드목록을 생산할 때는 하나의 숫자群(예: 출판년도)이 두줄에 걸쳐 인쇄되거나 英語單語의 音節이 잘못 절단되는 경우, 句讀點(예: :, ;, ', ,)이나 특수문자가 라인의 선두로 오는 경우, 부출

지시사항의 일부가 다음 카드로 연결되는 경우 등을 防止하도록 노력해야 할 것이다. 또한 複本이나 시리즈의 등록번호가 많아 복수의 카드가 되는 경우도 무리없이 處理해야 할 것이다. 기기에 따라 기성제품용지를 사용할 수도 있으며 보유기에 맞는 카드용지를 設計하여 注文할 수도 있다. 레이저 프린터에 의한 카드出力의 例를 그림 5에 보인다(그림 2 에러체크리스트의 出力例와 비교하여 참조). 그림 5에서 ㄱ이 副出指示事項의 일부가 다음 카드에 계속되면 앞카드에 스페이스가 있더라도 副出指示事項을 모두 다음 카드에 印刷하는 것이 좋다.

792
R182 ㅂ 새론 俳優藝術 / 李光來, 金興雨, 李淑 공역. - 인천 :
우성문화사, 1985.
289 p. : 삽도. - (극단배우극장연극신서 ; 1)

10292835 목차: 제1부: 俳優의 修練/르바쁘르트; 李光來 譯. 제
10292836-c.2 2부: 스타니슬랍스키의 演技論(Stanislavsky on art of
the stage)/데이빗 마가르샤크; 金興雨 譯. 제3부: 俳優
藝術論(L'Art du comedien)/꼬끄랑; 李淑 譯.

○

(계속)

(카드 2)

792
R182 ㅂ 1. 792 St24s 한. 2. 792 ㄱ18a 한. I. 새론 배우예술.
II. 배우예술. III. 배우의 수련. IV. 스타니슬랍스키의
연기론. V. 배우예술론. VI. L'Art du comedien.
VII. Stanislavsky on the art of the stage.
VIII. Rappaport, S. A., 1863-?. IX. Magarshack, David.
X. 꼬끄랑. XI. 이광래. XII. 김홍우. XIII. 이숙.
XIV. 송서명.

<그림 5> 카드출력의 예

○

10. 典據統制

Cutter가 말한대로 圖書館目錄은 특정저자, 특정서명, 특정주체의 자료를 찾을 수 있어야 하며, 그 자료의 소장여부를 알 수 있어야 하며, 자료의 판차나 자료의 성질로 特定資料를 선택하는데 도움을 주어야 한다.¹³⁾ 이것은 온라인목록에서도 맞는 말이다. 이를 위하여 수집된 情報를 액세스할 수 있는 액세스포인트를 결정하고 그 결정된 事項과 變形(variant, cross reference 혹은 see)을 기록하여 統一되고 一貫性있는 형태로 표현하기 위한 일련의 처리를 하게 된다. 이것을 典據統制(authority control)라 한다.

예로 著者名 典據統制를 하게 되면 저자기호가 같아지므로 표기는 다르지만 같은 사람 같은 主題의 著作이 書架에 나란히 꽂히게 된다. 또한 온라

인목록에서 著者名 典據파일을 이용하면 어느 이름(변형 혹은 참조명)으로 入力하더라도 全體 著作을 모두 檢索할 수 있는 利點이 있다.

여러 가지 典據파일이 만들어지면 이를 書誌데이터의 入力에서 사용할 수 있다. 이 典據파일은 목록부서에서 만들고 필요시 참조하지만 入力時에 複寫함으로써 신속하고 정확한 入力이 가능해진다. 다만 目錄業務에서 典據파일을 만들 때 典據採擇의 사유, 안내, 제한, 범위 등을 기록하기 위한 典據注記項目이 필수적으로 준비되어야 할 것이다. 典據데이터의 複寫를 위한 檢索時에는 절단검색이 필수적이다. 입력시 典據데이터 複寫를 위한 檢索畫面의 예를 그림 6에 보인다. 複寫畫面에서는 典據레코드가 길더라도 한 줄만 보여주지만, 複寫時에는 완전한 典據레코드를 갖고 오게 된다.

등서 자료 입력

등록번호 _10000001_ 원본번호 _10000000_ 유사번호 _10000000_ 소장표시 _____

청구번호 _891.73 D742b한C _____

저 자 _마그리타 아브루제세 _____

저자입기 _Abbruzzese, Margherita _____

역 자 _____

역자입기 _____

서 명 _____

Authority Control	
Entry	Dostoevski, Fedor Mikhailovich, 1821-1881
Ref. 1	도스토예브스키, 페더 미카일로비치
Ref. 2	도스토예프스키
Ref. 3	도스토예프스키
Ref. 4	도스토예브스키
Ref. 5	도스토예프스키
Ref. 6	도스토예프스키
Ref. 7	도스토예프스키어
Ref. 8	도스토이예프스키
Note	
Name(Series)	Dos

<그림 6> 전거데이터 복사화면의 예

典據데이터를 典據파일에서 複寫할 때 入力데이터와 典據데이터와의 關係를 잘 파악할 필요가

있다. 예로 著者名에서의 關係를 표 1에 보인다.

〈表 1〉 저자명 전거데이터와 관계

전 거	선 택	저자명	엑세스포인트
한글(한자)	전 거	한 자	한 글
	진 거	한 글	한 글
	변 형	변 형	한 글
한 글	진 거	한 글	한 글
	변 형	변 형	한 글
영 문	전 거	영 문	영 문
	변 형	변 형	영 문

표1에서 '전거' 항목은 典據(authority 혹은 heading)의 형태를 의미하며, 한글과 한자를 併記한 것, 한글만인 것, 영문인 것의 3종류로 나눌 수 있다. '선택'은 複寫를 위해 典據를 선택했는가 變形을 선택했는가를 나타내고 있다. '저자명'과 '엑세스포인트' 항목은 이 선택결과가 變形을 선택했다면 저자명항목에는 선택된 變形(한글)이 복사되며 엑세스포인트에는 典據(英文)가 복사된다는 것을 의미한다. 만약 典據를 선택했다면 저자명항목에는 倒置된 典據(영문 혹은 외국인의 한글표기)를 바로 잡아 複寫되며, 엑세스포인트에는 典據가 그대로 複寫된다. 또한 典據가 한글/한자인 경우, 典據를 선택하면 '저자명'에 한자가, '엑세스포인트'에 한글이 복사되며, 變形을 선택하면 '저자명'에 한자가, '엑세스포인트'에는 變形이 복사되어야 한다. 다만 典據는 한글/한자가 併記되어 있으나 자료가 한글로만 되어 있는 경우에는 '저자명'의 한자를 한글로 바꾸는 處理가 필요할 것이다.

複數著者이면 각각의 典據데이터를 複寫해야

하는 경우가 생기게 되므로 反復하여 複寫할 수 있도록 해야 할 것이다. 그림6과 표1은 著者名의 경우이며 그림6은 두번째 저자를 복사하려는 화면을 나타내고 있다. 만약 'Ref. 1'을 複寫하게 되면 入力畫面에는 다음과 같이 나타날 것이다.

저 자 : 마그리타 아브루제세, 웨더 미카
이로빗치 도스토예브스키
엑세스포인트 : Abbruzzese, Margherita\Dos
toevski, Fedor
Mikhailovich, 1821-1881

11. 맺는말

人間과 컴퓨터와의 對話는 특별한 경우를 제외하고 모두가 디스플레이畫面을 통해서 이루어진다. 지금은 대부분 CRT를 사용하고 있지만 대형 액정디스플레이(large-scale liquid-crystal display)나 발광패널(electroluminescent panel) 사용이

급속하게 늘고 있다. 현재까지 데이터入力에서 가장 널리 사용되고 있는 방식은 메뉴방식과 樣式을 채우는 방식의 인터페이스(그림 4 참조)가 主題를 이루고 있다. 최근에는 高解像度 디스플레이를 이용한 멀티윈도우(multiple display window), 풀다운 메뉴(pull-down menu), 아이콘(icon) 등에 의해 動的인 畫面이 폭넓게 사용되고 있으며, 컬러화면, 좀 더 넓은 화면, 그래픽 터미널과 그래픽기술을 이용한 멀티데이터베이스用 하이퍼텍스트(hypertext, hypermedia, hyperdocument) 등도 급속도로 보급되고 있다. 또한 Davies¹가 論하는 바와 같이 目錄業務는 여러 가지 特性上 專家시스템을 도입하기에 적당한 領域이다. 書誌데이터 入力시스템의 인터페이스 向上을 위하여 새로운 機器를 이용하거나 이들의 새로운 概念을 시급히 導入해야 할 것이다.

참 고 문 헌

- 1) 사공철 등저, 「최신정보검색론」(서울 : 구미무역, 1990), pp.165-166.
- 2) E.T. O'Neill, Diane Vizine-Geotz, *Quality Control in Online Databases*. In : Martha E. Williams ed. *Annual Review of Information Science and Technology, volume 23*(Amsterdam : Elsevier Science Pub., 1988). pp.125-156.
- 3) Lucy A. Tedd, *An introduction to computer-based library systems. 2nd ed.* (New York : John Wiley, 1984). p.104.
- 4) Marvin H. Scilken, *The Catalog as a public service tool*. In : Maurice J. Freedman, S. Michael Malinconico ed. *The Nature and Future of the Catalog : Proc. of the ALA's Information Science and Automation Division's 1975 and 1977 Institutes on the Catalog* (Phoenix : The Oryx Press, 1979). pp.89-101.
- 5) Martin Helander, *Handbook of Human-Computer Interaction* (Amsterdam : North-Holland, 1988), pp.377-412.
- 6) Charles R. Hildreth, *Online Public Access Catalogs : The User Interface* (Dublin : OCLC, 1982). pp.144-158.
- 7) Unesco, *Mini-Micro CDS/ISIS Reference Manual, Version 2.3* (Paris : Unesco, 1989). pp.87-94.
- 8) 유재원, 「우리말 역순사전」(서울 : 정음사, 1985).
- 9) Barry B. Baker, Lynne D. Lysiak, *From Tape to Product : Some Practical Considerations on the Use of OCLC-MARC Tapes* (Ann Arbor : The Pierian Press, 1985).
- 10) 이상현, "CD-ROM 기술을 이용한 도서관 편독시스템 개발에 관한 연구," 연세대학교 도서관학과 석사학위논문, 1988.
- 11) 金昌根, 申龍珠, "CD-ROM을 利用한 書誌情報데이터베이스의 信賴性 向上 技法 研究," 情報管理研究, Vol.29, No.2(1990). pp.69-99.
- 12) 崔錫斗, "圖書館시스템의 맨 머신 인터페이스에 관한 研究," 도서관, Vol.44, No.3(1989). pp.5-24.
- 13) Roy Davies, *Cataloguing as a Domain for an Expert System*. In : Roy Davies ed. *Intelligent Information Systems : Progress and Prospects* (Chichester : Ellis Horwood Ltd., 1986). p.3에서 재인용.
- 14) Roy Davies, pp.54-77.