

# 전거제어시스템의 개발

김 태 수\*

## 〈목 차〉

1. 서 론	템 구축
2. 전거제어과정	4. 결 론
3. 패키지구조에 의한 전거시스	참고문헌

## 1. 서 론

전거라는 용어는 말이나 문장의 근거가 된 문헌상의 출처를 의미하는 말로서, 목록에서는 표목의 고유성을 확보하기 위한 방안으로 도입되고 있다(Avram 1984, 331). 전거제어에서 가장 중요한 원칙은 표목의 유일성을 확보하는 것으로, 개인명은 물론 모든 표목이 식별상 고유해야 한다.

현행 전거제어는 저록간의 연결이 아니라 표목간을 연결하는 방식으로서, 저자의 경우 특정 저자의 상이한 이름 중 하나를 표준형식으로 선정하고, 이 표준형식의 이름을 통해 그 저자의 모든 저록과 연결된다. 현행 목록에서는 표목의 형식과 구조가 전반적으로 목록의 구조와 형식을 결정하기 때문에 표목의 제어가 중요시되고 있다.

## 2. 전거제어과정

전거레코드와 서지레코드가 연결되는 것이 가장 이상적인 형태이다. 표목은 전거과일에 한번만 저장되고, 전거파일의 레코드에 변화가 있게 되면 전거레코드만을 수정하고, 서지파일의 표목을 수정할 필요가 없다. 다음은 Mark Twain의 Huckleberry Finn

\* 연세대학교 문헌정보학과 교수

의 2개의 판(New York, 1920; Boston, 1935)의 경우를 예시한 것이다(Burger 1985, 49-50).

**전거레코드**

(레코드 2345)

Twain, Mark

xClemens, Samuel Langhorne

**서지레코드**

(전거레코드 2345)

Adventures of Huckleberry Finn, 1920

(전거레코드 2345)

Adventures of Huckleberry Finn, 1935

상이한 형식의 이름은전거레코드로 해결되고, 서지파일에 있는 자료를 검색하기 위하여전거파일만을 탐색한다는데 그 의미가 있다. 다시 말하면 두개의 파일을 검색하는데 단지 하나의 색인만이 필요하다.

현행 기계가독목록에서도 표준형식의 표목을 선정하고, 이와 상이한 형식의 표목을 연결한 전통적인 제어구조를 도입하고 있으며, 레코드 상호간의 연결방식을 배제하지는 않지만, 주로 표목간을 연결하는 구조이다. 기본적으로 MARC 전거형식에서 제어 방식은 다음과 같다.

100 10 \$aSimmons, Dennis L. \$q(Dennis Lee)

400 10 \$aSimmons, Dennie.

### 3. 패키지구조에 의한 전거시스템 구축

#### 3.1 필요성

표준형식의 표목을 선정하지 않고도 상이한 형식의 표목간을 연결하여, 어떤 형식에 서도 검색이 가능한 새로운 전거시스템을 개발하는 것이 과제이다. 그 이유는 다음과 같다.

- 1) 이용자는 표준형식과 무관하게 자신이 알고 있는 형식으로 목록에 접근하며, 따라서 특정 형식을 기본표목으로 규정할 이유가 없다. 전거제어에서는 상이한 형식의 접근점 중에서 특정 형식을 '선정'하는 기능에서, 이들 상이한 접근점을 상호 '연결' 하는 구조로 전환되어야 할 것이다.
- 2) 앞으로의 환경에서는 서지정보망을 통해 특정 기관의 목록이 외부의 다양한 서지테

이터베이스와 연결되고 통합될 가능성이 있고(Bourdon 1993), 이러한 상황에서 상이한 데이터베이스에 사용된 다양한 형식의 접근점을 연결할 수 있는 전거제어시스템을 고려해야 할 것이다. 그래야만 외부 데이터베이스를 자관의 목록에 수용할 수 있을 것이며, 이런 상황에서 특정 형식의 표목을 유지하는 것은 그 의미를 발견하기 어렵다.

- 3) 특정 언어뿐만 아니라 어떤 언어의 표목도 수용할 수 있는 제어수단이 요구되며, 이를 통해 실질적으로 상이한 형식의 접근점을 연결하는 효과를 가지게 될 것이다.
- 4) UNIMARC이나 CCF와 같은 일부 기계가독목록에서는 기본표목을 규정하지 않고도 제어기능을 수행할 수 있다. 따라서 현재 서지레코드의 기본표목필드(1XX)는 부출표목필드(700-730)와 통합하여(한국문화예술진흥원 1994, 199-209), 이 통합된 필드를 전거제어용으로 사용하는 것으로, 이미 주제명은 부출표목필드(600-651)를 통해 제어되고 있다.

## 3.2 인명전거시스템의 구축

### 3.2.1 적용원칙

여러 상이한 형식 중 어떤 형식의 이름으로도 접근하기 위해 다음과 같은 원칙을 적용한다(김태수, 최석두, 남영준 1997, 98-112).

- 1) 동일 저자의 상이한 이름형식을 서지레코드로부터 추출하여 전거파일에 저장하되, 특정 형식의 표목을 대표형식으로 고려하지 않는다.
- 2) 기술대상자료에 기재된 형식대로 입력하고, 모든 형식의 이름을 접근점으로 사용한다.
- 3) 특정 저자의 모든 형식의 이름이 입력된 패키지에 대표부호(AuthorID)를 부여하고, 서지데이터베이스의 제어번호와 인명전거데이터베이스의 대표부호를 연결한다.

서지레코드로부터 기본표목의 여부와 무관하게 인명과 관련된 데이터를 모두 추출하여 고유성을 식별한 후, 패키지를 구축하기 때문에 비록 현재 기본표목필드(1XX)에 사용된 인명도 여러 이형 중 하나로 취급한다. 이때 서지레코드에 수록된 식별정보[세계(世系)나 칭호, 기타 명칭, 생몰년, 역조 등의 데이터]도 함께 추출하여 패키지에 포함한다. 아울러 서지레코드와 전거레코드의 연결은 서지레코드의 제어번호를 사용한다.

### 3.2.2 서지레코드의 구조분석

- 1) 제어번호

파일단위로 입수되는 서지레코드를 고유하게 식별하는 기본 값으로 서지레코드의 제어번호(012필드)를 사용하여 전거레코드와 연결한다.

## 2) 서지레코드의 인명관련데이터 분석

100필드(개인명기본표목)의 하위필드 중 역할어(\$e)와 언어(\$l)는 식별데이터가 아니기 때문에 식별정보에서 제외하며, 그 외 데이터는 식별기호를 제거한 후 모두 주기필드에 입력하여 식별데이터로 사용한다. \$e역할어는 특정 저작에서의 저자의 역할정보로 사용한다.

100 0 \$a고종, \$f고려, \$g제23대왕, \$d1192-1259 에서 \$a고종 이외의 하위필드는 식별기호를 제거한 후 주기필드에 저장하여 동명이인의 식별정보로 사용한다. 따라서 인명전거데이터베이스에는 저자명과 식별정보가 다음과 같이 분리되어 저장된다.

고종

고려 제23대왕 1192-1259

245필드에서 \$d와 \$e의 데이터는 각각 첫 번째 저자, 두 번째 이하의 저자로서, 표제면의 형식대로 사용한다.

245 10\$a..., \$dGordon Larsson 지음; \$e배현주 역 에서 \$d의 데이터는 100필드의 \$a의 데이터와 비교하고, \$e의 데이터는 700필드의 \$a의 데이터와 비교하여 처리한다.

507필드(원저자, 원서명에 관한 주기)는 번역자료에 대한 원저자명과 원서명을 기술하는 필드로 \$a에 원저자명 형식을 사용한다.

507 10\$aMailer, Norman

600필드(주제명부출표목-개인명)에서 인명식별데이터로 사용되는 식별기호는 100필드의 식별기호의 경우와 동일하다.

600 14\$a이광수 \$x평론, 해석

700필드(개인명부출표목)의 식별기호 중 인명식별데이터로 사용되는 기호는 100필드의 식별기호의 경우와 동일하다.

700 1 \$aJefferson, Thomas, \$d1743-1826

765필드(원저저록)의 \$a에 있는 이름형식을 사용한다.

245 10\$a인간의 굴레/\$d썸머셋 모음  
765 0 \$aMaugham, William Somerset, \$t...

767필드(번역저록)의 \$a에 있는 이름형식을 사용한다.

245 10\$a하안전쟁 /\$d안정효  
767 0 \$aAhn, Jeong Hyo, \$t...

900필드(개인명로컬표목)에서는 표목으로 채택된 개인명과 다르게 부출되는 개인명을 기술하며 제1지시기호의 사용법은 100필드의 경우와 동일하고, 식별기호에서 주기 필드에 저장되는 데이터의 경우도 100필드의 경우와 동일하다.

100 1 \$a미우라 아야꼬  
245 10\$a.../\$d三浦綾子  
900 11\$a삼포능자

### 3) 서지레코드의 식별정보

서지레코드의 데이터필드 중 100필드와 600필드, 700필드, 900필드의 하위필드에 수록된 데이터를 추출하여 식별정보로 삼는다. 그리고 역할데이터는 100필드와 700필드, 900필드의 '\$e'의 역할어를 추출하여 저장한다.

### 4) 서지레코드의 관계형태정보

관계정보는 서지레코드의 내용식별기호를 대상으로 하여 관계를 분석한 후 추출한 관계데이터를 인명선거데이터의 입력시 인터페이스를 통해 제시하는데, 기본 값으로 '본명'을 사용한다. 서지레코드에서 추출한 이름의 유형에는 본명, 묘호, 시호, 봉호, 추호, 칭호(명칭) 등이 포함된다. 여기서 묘호나 시호, 봉호, 추호 등은 일반적인 호나 아호 등과 함께 모두 호로 통일한다. 그 외 별명, 필명, 자 등의 관계형태가 있다. 완전명이나 도치명, 번자명 등은 모두 본명으로 처리한다.

## 3.2.3 시스템 개요 및 데이터베이스 구조

### 1) 시스템의 개요

이 시스템은 데이터베이스 구축과 서지정보 검색의 두 부분으로 구성되어 있다. 데이터베이스 구축시 고려할 사항은 패키지개념을 적용하여 모든 이형을 접근점으로 하고, 패키지내 각 접근점간의 관계정보를 구축하여, 서지레코드 내 기술부의 저자명과 접근점이 상이한 경우 두 형식간의 관계에 대한 정보를 제시한다. 동명이인을 식별하

기 위해 패키지ID와 저자명을 결합한 두 개의 필드를 기본키필드로 하여 해결하고, 저자의 식별정보로 생몰년과 직업, 전공분야, 한자표기 등을 사용한다.

## 2) 데이터베이스 구조

관계형 데이터베이스 설계에서는 복잡한 데이터구조를 단순하고 안정된 데이터구조로 바꾸는 정규화단계를 거친다. 이 정규화단계를 거친 개체-관계(entity-relationship)를 논리적인 데이터베이스설계과정이라고 한다. 이 과정을 통해 얻은 개체-관계구조를 실제 파일이나 테이블과 같은 물리적 저장구조로 매핑시키는 것이 물리적인 데이터베이스 설계과정이다.

서지레코드에서 추출한 저자명과 서명, 저자식별정보, 서지정보, 공동저작의 5개의 개체를 추출하여 인명진거시스템의 물리적 데이터베이스를 설계한다.

서지레코드에서 인명관련필드의 데이터를 추출하여 저자명테이블을 구축하며, 공동저작테이블은 저자명테이블과 서명테이블의 다:다[N:N] 관계에서 파생된 테이블로 저자의 저작에 대한 역할어에 대한 정보를 담게 된다. 식별정보테이블은 각각의 저자명 패키지에 대해 하나의 레코드를 가지게 된다. 서지정보테이블의 경우 완전한 서지정보를 포함하는 대신 서지레코드 파일에 대한 포인터를 저장하고 서명테이블과 연결하여 검색시 적합한 서지레코드를 제시하도록 한다. 다음 그림 1은 테이블의 구조와 관계를 도식한 것이다(정영하 1997, 44-59).

- 가) 저자명테이블(AuthorName)에서는 AuthorID와 AuthorName이 기본 키가 된다. 특정 저자가 여러 형식의 이름을 사용한 경우 동일한 AuthorID를 갖게 되며, 따라서 동일한 AuthorID는 여러 이형의 이름을 묶는 패키지 개념을 구현하는 도구이다.
- 나) 서명테이블(WorkTitle)의 WorkID 역시 특정 저작이 상이한 서명을 가지는 경우, 이형의 서명에 동일한 WorkID를 부여하여 패키지개념을 구현한다. 이 시스템에서는 레코드의 제어번호를 WorkID로 한다.
- 다) 저자명테이블과 연결되어 있는 식별정보테이블(AuthorNotes)에서는 저자명테이블의 AuthorID를 외래 키로 연결하여 특정 저자명패키지에 대한 식별정보를 저장하며, 필요한 경우 이 식별정보를 통하여 고유성을 확인할 수 있다.
- 라) 서지정보테이블(BibInformation)은 서명테이블의 WorkID를 외래 키로 한다. 하나의 파일 속에 하나의 서지레코드를 저장해 두고, 검색결과에 따라 적합한 서지레코드를 출력하도록 한다. 서지정보테이블의 FileID는 서지레코드의 파일이름을 저장한 필드이다.
- 마) 공동저작테이블(CoWork)은 저자명테이블과 서명테이블의 다:다[N:N]관계에서 나온 테이블로 저작에 대한 저자의 역할정보를 수록한 테이블이다. 따라서 공동저작 테이블에서는 저자명테이블의 AuthorID와 서명테이블의 WorkID를 외래 키로 참조해 준다.

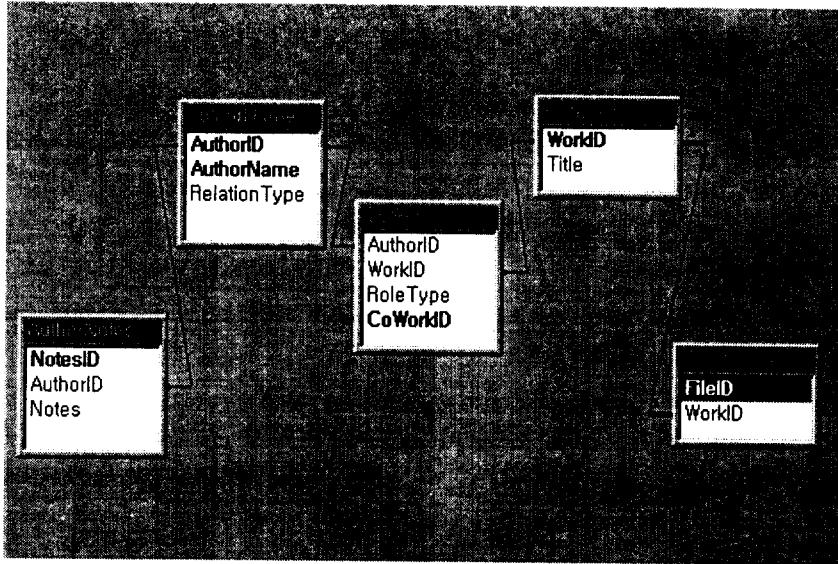


그림 1: 테이블의 구조와 관계

### 3.2.4 시스템설계

서지레코드를 패키지구조의 RDBMS에 저장하는 DB구축프로그램과, 패키지구조로 구축된 인명전자데이터에 접근하여 검색하는 서지정보 검색 프로그램으로 나누어진다.

#### 1) 패키지데이터베이스 구축

인명전자데이터베이스를 구축하기 위해서는 서지레코드의 인명관련필드를 분석한 후, 레코드에 포함된 모든 형식의 이름을 데이터베이스의 저자명필드의 데이터와 비교한다. 이때 데이터베이스에 없는 형식에 대해서는 패키지개념을 적용하여 새로운 패키지를 만들고 그 구조 속에 이형들과 함께 저장한다. 만약 기존의 데이터베이스에서 동일형식이 발견되면 동명이인인지의 여부를 확인하여 동명이인인 경우에는 새로운 패키지에 저장하고, 동일인인 경우에는 인명데이터를 해당 패키지 내의 이형들과 비교하여 동일형식이 없을 경우 이 이형을 기존의 패키지에 추가한다(정영하 1997, 35).

#### (가) 작업흐름도

서명패키지의 서지정보테이블과 서명테이블에 적절한 정보를 파일형태의 서지레코드로부터 추출하여 저장한다. 그 다음에는 서지레코드로부터 혹은 입력에 의해 적절한 정보를 인명패키지의 저자명테이블과 식별정보테이블에 저장한다. 마지막 단계로 공동저작 테이블에 저자패키지ID(AuthorID)와 저작패키지ID(WorkID)를 저장하고 저작에 대한 역할정보를 선택하여 입력한다.

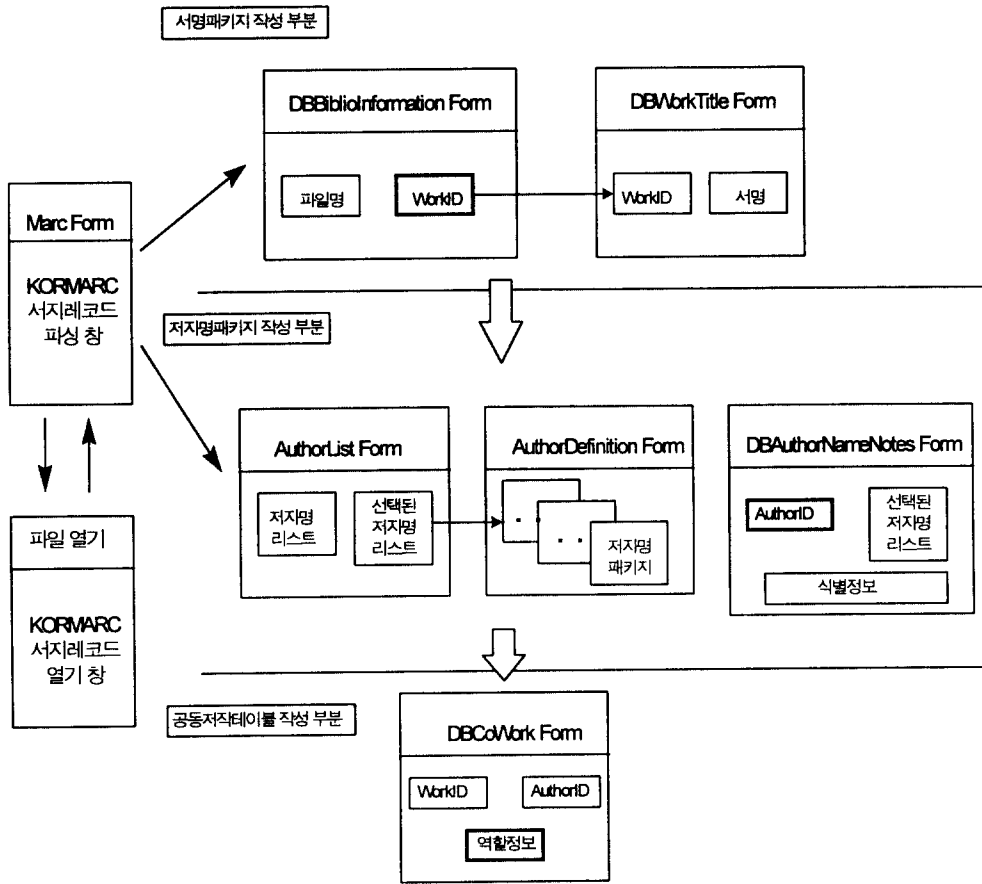


그림 2: 데이터베이스 작업흐름도

그림2에서와 같이 초기화면(Marc Form)에서 서지레코드를 열고(FileOpen Form), 서지정보 테이블(BibInformation)에 파일명(FileID)과 저작키(WorkID)를 입력하기 위한 창(DBBibInformation Form)에 제어번호를 저작 키로 하여 입력과정을 확인한다. 서지정보테이블에 데이터를 추가한 후 서명데이터(WorkTitle)에 해당 저작키에 대한 서명을 입력하게 되는데(DBWorkTitle Form), 이 과정은 서지레코드의 서명저자필드의 서명데이터를 추출하여 자동으로 처리한다.

서명패키지부분을 처리하면, 다음 단계로 인명패키지부분을 처리하게 되는데, 인명 데이터를 추출하는 창(AuthorList Form)을 호출하여, 이형의 이름들을 선택한 후 패키지 작성하는데, 이때 선택된 모든 형식의 이름으로 데이터베이스를 검색하여 동명이인의 여부를 판단하게 되는데 동명이인인 경우(A)와 동일인의 경우(B)로 분기된다.

먼저 동명이인인 경우(A), 동명이인 식별창(AuthorIdentification Form)에서 동일한



이름형식들이 패키지 단위로 각 패키지에 대한 식별정보와 함께 출력된다. 여기서 동명이인을 판단한 후, 동명이인일 경우에는 후자의 처리단계(B)를 거친다. 동일인일 경우, 새로 입력된 레코드에 대해 기존의 패키지에서 새로운 이형이 나타났을 경우, 혹은 식별정보를 추가하고자 할 때도 역시 (B)에서 패키지데이터를 수정하며, 새로운 데이터가 없을 경우에는 인명패키지부분의 처리를 종료하게 된다. 동일인인 경우(B)에는 인명패키지 데이터 입력창(DBAuthorNameNotes Form)에서 인명패키지번호(AuthorID)를 입력받고 식별정보를 추가하여 이전 폼(AuthorList Form)에서 넘어온 인명데이터를 식별정보와 함께 저자명테이블(AuthorName)과 식별정보테이블(AuthorNotes)에 저장하게 된다. 서명패키지와 인명패키지의 처리가 끝나면 마지막으로 공동저작테이블(CoWork)을 처리하기 위해, 공동저작 데이터 입력창(DBCoWork Form)에서 저자의 역할정보를 입력받아 서명패키지 기본 키(WorkID)와 인명패키지 기본 키(AuthorID)와 함께 저장한다.

#### (나) 시스템 구성

##### ① 데이터베이스 구축대상 파일 선택

데이터베이스 구축대상 파일은 인명관련데이터의 분석창을 이용하여 파일에 접근할 수 있다.

##### ② 서지레코드 파싱

해당 파일을 열면, 원서지정보로부터 인명관련필드의 데이터를 추출하여 해당 컨트롤에 보여준다.

##### ③ 서명테이블과 서지정보입력테이블

인명관련데이터 분석창에서 파일명과 서명, 제어번호를 파싱하여 WorkTitle테이블과 BibInformation테이블에 서지정보를 입력한다.

##### ④ 인명데이터 추출

인명관련데이터 분석창에서 인명관련데이터를 바탕으로 하나의 인명패키지를 구성할 인명데이터를 모두 선택하여 인명패키지컨트롤에 추가한다.

##### ⑤ 동명이인 식별

동명이인을 식별하기 위해 추출한 인명데이터를 탐색어로 하여 데이터베이스를 검색하고 같은 이름형식의 데이터를 보여준다. 추출된 모든 인명데이터가 탐색어가 될 수 있다.

⑥ 동일인 식별

동일인을 식별하기 위해 저자명테이블에서 선택한 레코드에 대해 패키지와 식별정보를 출력해 준다. 이 정보를 기준으로 동일인이며 데이터의 수정이 필요한 경우에는 수정한다.

⑦ 기존패키지 수정

추출된 인명데이터를 참고로 적절하게 수정한 후 기존의 저자명테이블을 갱신한다. 이 창을 통해 저자명테이블(AuthorName)과 식별정보테이블(AuthorNotes)이 갱신된다.

⑧ 공동저작테이블 입력

저자명과 식별정보에 변경이 생기면 패키지ID(AuthorID)와 제어번호(WorkID), 역할 데이터가 공동저작테이블에 입력된다.

⑨ 인명패키지 작성

동명이인을 식별하기 위해 저자명ID를 입력하면 추출된 인명데이터와 관계정보를 확인할 수 있고 이때 적절한 수정을 한 후 입력하면 저자명테이블이 갱신된다. 식별정보는 새로이 입력받아 식별정보테이블에 저장된다.

2) 서지정보 검색

(가) 작업흐름도

서지정보검색은 그림 3과 같이, 먼저 저자명 검색창(NameSearch Form)에서 저자명을 입력하고, 이 저자명으로 데이터베이스 내 저자명 테이블(AuthorName)을 검색하여 식별정보와 함께 저자식별창(Identification Form)에 출력하여, 적합한 저자명을 선택한다. 여기서는 이용자가 선택한 저자명에 대해 상세한 인명정보를 제시하고자 할 때 저자상세정보창(NameAuthority Form)을 통해 패키지 내의 모든 이름형식과 이름간의 관계정보, 그리고 패키지의 식별정보를 알 수 있다.

간략서지창(AuthorWorks Form)에서는 저자식별창에서 선택한 저자에 대한 자료를 모두 제시해 주며, 그 중 상세서지정보를 원하는 저작을 선택하면 상세서지창(BibInformation Form)이 활성화된다. 이것으로 하나의 검색과정이 끝나게 되고, 또 다른 접근점으로 검색하거나 검색작업을 종료하게 된다.

(나) 시스템 구성

① 정보검색 - 저자명

접근점이 되는 저자명을 입력받아 인명패키지데이터베이스를 검색하며, 이때 부분일치탐색이 가능하다.

② 저자명 탐색결과

입력된 저자명을 기준으로 데이터베이스 내에서 동일한 형식의 저자명을 패키지번호와 관계정보와 함께 출력해 준다. 이 중 한 레코드를 선택하여 인명패키지를 선택하거나 새로운 접근점으로 검색을 다시 시도할 수 있다.

③ 인명패키지 정보

저자명 탐색결과에서 인명패키지를 선택한 경우, 선택된 각각의 레코드에 대한 식별정보를 패키지 내 모든 저자명데이터와 함께 출력해 준다.

④ 저작사항 - 간략서지

저자명 탐색결과나 인명패키지에서 저작사항을 선택하면 해당 저자의 모든 저작물을 간략형식으로 제시한다. 선택한 레코드에 대한 상세정보를 원할 경우에는 상세정보를 선택한다.

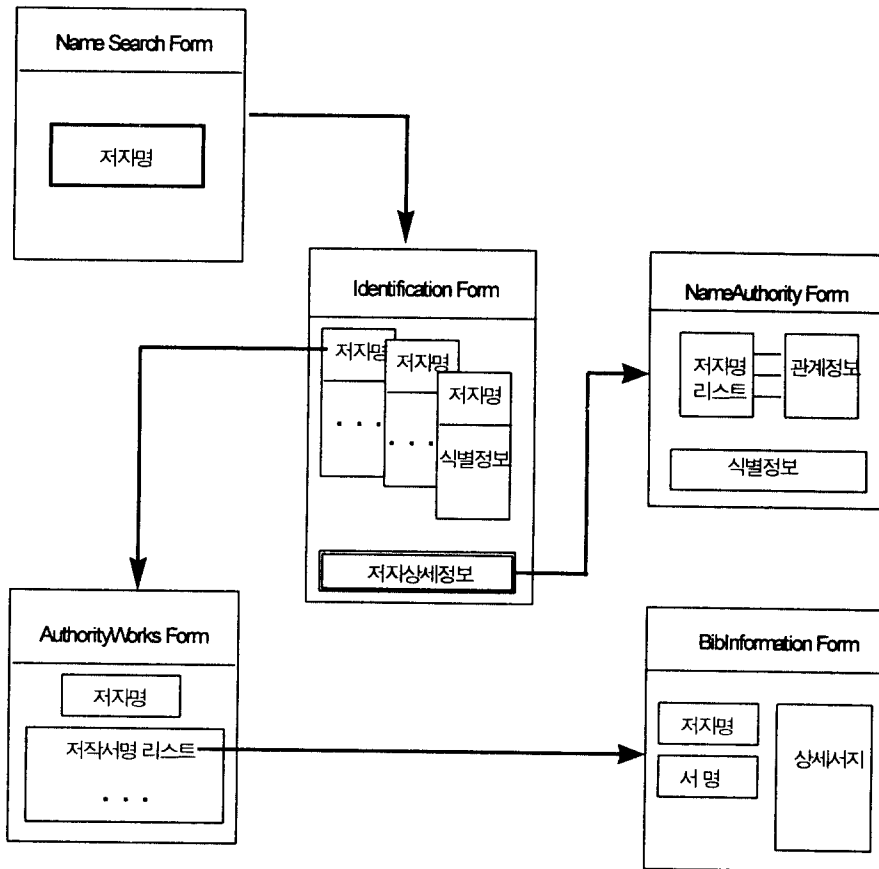


그림 3: 서지정보검색 작업흐름도

⑤ 상세서지 - 서지정보

상세서지-서지정보에서는 하나의 패키지에 대한 모든 저자명데이터와 식별정보, 패키지번호(저자ID)를 제시해 주며, 저작사항에서는 제어번호와 서명과 함께 서지레코드 자체를 출력해 준다. 새로운 검색을 시작하거나 적절히 선택할 수 있다.

## 4. 결 론

이 전거제어시스템은 서지레코드를 원시데이터로 하여 인명과 관련된 모든 필드를 분석하여 데이터를 추출한 후, 동명이인의 식별과정을 거쳐 패키지에 저장하되, 문헌에 사용된 형식을 그대로 입력할 수 있고, 시스템의 구축이 용이하여 다음과 같은 효과를 기대할 수 있다(김태수, 최석두, 남영준 1977, 116).

- 1) 표목의 표준형식을 결정할 필요 없이 자료에 표기된 형식대로 입력함으로 시간과 비용을 절약할 수 있다.
- 2) 검색시 자연어나 제어어휘와 같은 어떤 형식으로도 원하는 자료에 접근이 가능하다.
- 3) 전거레코드의 형식이 단순하여 입력이 용이하다.
- 4) 상이한 서지데이터베이스의 통합 시 그 유용성이 기대된다.

## 참 고 문 헌

1. 김태수, 최석두, 남영준. 1997. 디지털도서관용 시소러스 및 전거데이터 구축에 대한 연구. [서울]: 오름정보.
2. 정영하. 1997. 패키지구조를 이용한 인명전거시스템 구축에 관한 연구. 석사학위논문. [서울]: 연세대학교 문헌정보학과.
3. 한국문화예술진흥원. 1994. 예술자료관 보유데이터 변환을 위한 연구. [서울]: 동진진흥원.
4. Avram, Henriette D. 1984. "Authority control and its place", *Journal of Academic Librarianship*, vol. 9: 331-335.
5. Bourdon, Françoise. 1993. *International cooperation in the field of authority data: An analytical study with recommendations*. Translated from the French by Ruth Webb. München: Saur.
6. Burger, Robert H. 1985. *Authority work: The creation, use, maintenance, and evaluation of authority records and files*. Littleton, Col.: Libraries Unlimited.