

無典據시스템에 관한 研究

최 석 두*

목 차

- | | |
|----------------------|--------------------------|
| 1. 머리말 | 3.4 분담목록센터의 전거통제 |
| 2. 전거시스템 | 3.5 서지데이터의 입력 |
| 2.1 전거통제의 정의 | 4. 무전거시스템 |
| 2.2 전거통제의 목적 | 4.1 개요 |
| 2.3 전거데이터의 내용 | 4.2 채택-비채택표목의 등가성 |
| 2.4 전거레코드의 종류 | 4.3 서지데이터 입력의 편리 |
| 2.5 표목의 형태 | 4.4 저자기호 및 검색효율을 위한 그룹코드 |
| 2.6 전거시스템의 종류 | 4.5 서지 및 참조데이터 통합의 용이성 |
| 3. 전거시스템의 문제점 | 4.6 자연언어-통제어휘 혼합시스템의 구축 |
| 3.1 전거데이터 관리비용 | 4.7 참조그룹의 작성과 전거작성의 비교 |
| 3.2 전거의 채택 | 5. 맺음말 |
| 3.3 서지 및 전거데이터의 다운로드 | |

1. 머리말

색인작업이란 색인자 혹은 색인시스템과 이용자간의 약속이다. 「이광수」라고 색인했는데 「춘원」으로 찾는다면 양자의 약속은 이행될 수 없을 것이다. 왜냐하면 컴퓨터는 기호를 비교하며 검색하는 것이지 기호가 표현하는 개념으로 검색하는 것은 아니기 때문이다. 따라서 다양한 표현에 대한 처리 없이는 언제나 전체 정보의 일부분만이 검색되어 재현율이 떨어지는 시스템이 될 수 밖에 없다. 서지정보는 검색의 실마리를 제공하고, 대상이 되는 특정자료를 정확히 식별할 수 있는 정보를

제공하여 이용자를 원자료로 손쉽게 인도하는 것을 목적으로 한다. 현재 도서관에서 사용되고 있는 목록규칙은 바로 이 목적을 달성하기 위하여 도서관학자, 정보처리전문가, 서지학자, 출판관계자들이 장기간 연구한 결과라 할 수 있을 것이다. 도서관에서 사용되고 있는 기준은 국가별 자료의 특성에 따라 내용이 어느 정도 차이가 있지만 대부분 국제적인 표준을 준용하고 있다.

좋은 정보검색시스템이란 이용자가 어느 표목으로 접근하든 해당정보를 모두 제공하는 시스템일 것이다. 그러나 색인된 표목을 이용자가 다르게 접근할 가능성이 있는 경우는 광범위하다. 따라서 인명, 단체명, 서명, 주제명, 지리명, 총서명 등에 대하여 모든 표현을 수집하고 각각의 경우에 따라 대표표현(전거형: authority)과 상이한 표현(변형: variant)을 선정하는 일과 양자를 상호참조로 링크시키는 처리방법이 목록분야를 중심으로 많은 연구가 진행되어 왔으며, 많은 성과를 올리고 있다.

그러나 전거통제시스템에서는 서지데이터가 만들어지기 전이나 만들어지는 시점에 전거가 결정되어 변형과 함께 전거화일에 포함되고, 전거는 액세스포인트로서 서지데이터에 기술되어야 한다. 따라서 정합성과 일관성을 갖지 않는 서지데이터는 불완전한 데이터가 된다는 심각한 문제가 발생할 뿐만 아니라, 서로 다른 상황에서 만들어진 데이터를 사용하거나 새로운 전거를 센터시스템에 통합하는 경우에는 이의 통제가 매우 불편하고 힘이 들며 경우에 따라서는 거의 불가능한 경우까지를 생각할 수 있다.

본 연구에서는 전거시스템의 문제점을 보완하기 위하여 목록시에는 지적인 노력을 적게 들이고 컴퓨터시스템의 자원을 적게 사용하며, 검색시에는 동일한 결과를 가지지만 전거를 갖지 않는 새로운 '무전거(無典據)시스템'을 제안하고자 한다.

2. 전거시스템

2.1 전거통제의 정의

전거작업에 의해 전거데이터(전거레코드, 전거화일)가 만들어지며 이것을 이용

하여 선거통제가 행해진다. 선거레코드, 선거화일, 선거작업, 선거통제에 대하여, ALA(1983), Schmierer(Clark 1990, 2에서 재인용), Burger(1985, 3), Hangler and Simmons(1982, 181), 圖書館情報學ハンドブック編輯委員會(1988, 441) 등의 정의를 종합하여 보면 다음과 같다. 선거레코드란 채택표목과 그 표목의 결정근거가 되는 정보원이나 표목형의 정당성을 나타내는 데이터를 기술하고, 그 채택표목으로 혹은 채택되지 못한 표현으로부터의 보라참조, 도보라참조지시 혹은 설명참조를 기술하며, 표목형의 속성이나 참조의 근거가 되는 정보를 기술한 기록이다. 따라서 선거화일이란 일군의 서지레코드내에서 표목의 형을 관리하기 위하여 채택표목과 참조정보를 가지고 있는 선거레코드의 집합이다. 선거화일의 종류에는 저자명선거화일, 통일서명선거화일, 총서명선거화일, 주제명선거화일 등이 있으며, 선거리스트와 동의어이다. 선거작업이란 기존 선거데이터를 유지보수할 뿐만 아니라, 새로운 선거데이터를 수집·검증하여 선거를 결정하고, 그 결정에 관한 정보를 기록하며, 채택된 형식을 목록에서 액세스포인트로 사용하는 것이다. 선거작업은 다음의 다섯 가지 과정으로 구성된다.

- 선거를 신규로 작성하는 과정(선거레코드와 선거화일의 작성)
- 작성된 선거레코드군을 하나의 선거화일내에 모으는 과정
- 그 선거화일을 서지화일과 연결함으로써 전체로서 선거시스템을 형성하는 과정
- 선거화일과 선거시스템을 유지관리하는 과정
- 선거화일과 선거시스템을 평가하는 과정

선거통제란 서지화일에서 표목으로 사용되고 있는 개인, 단체, 주제 등의 선거형이 일관성있게 사용되고 유지되도록 모든 가능한 표현을 찾아내어 일정한 규칙에 따라 액세스포인트를 결정하고 상호참조를 설정하는 것을 말한다. 필요한 참조와 함께 선거형을 기록한 선거화일과 서지화일(데이터베이스)의 일관성과 정합성을 유지하고 보증하기 위한 모든 메카니즘을 포함한다. 특히 자동선거통제란 컴퓨터에 의하여 선거통제과정의 대부분을 관리하는 것이다.

2. 2 전거통제의 목적

모든 서지정보는 그 정보에 대한 액세스포인트, 그 정보의 대체물인 서지기술(bibliographic description), 그 정보의 소장위치의 3가지 부분으로 구성된다. 목록이란 찾는 정보의 유무와 위치를 알려줄 수 있는 찾는 기능(finding function)과 어떤 저자나 주제에 대한 모든 저작과 판차를 알려줄 수 있는 모으는 기능(gathering, collocation 혹은 assembling function)을 직접적으로 지원한다(Clack 1990, 3-4). 전거통제는 이 중에서 찾는 기능보다 모으는 기능과 더 밀접하게 관련된다. 왜냐하면 현재 쓰고 있는 목록규칙들이 특정 저자명이나 서명 등을 찾는 데 편리하도록 자료의 물리적인 단위를 중심으로 기술되어 왔기 때문이다.

저자로 사용된 이름을 이용자는 알아야 한다. 그러나, 여러 가지 다른 이름을 사용할 수도 있으며, 하나의 이름이 여러 가지 형을 가질 수도 있다. 도서관은 저자명의 모든 형을 사용할 수도 있고 그 중 하나를 채택할 수도 있다. 이 때 이용자는 도서관이 사용한 모든 형의 이름을 알거나 도서관이 채택한 이름을 사용해야 한다.

개인명의 경우 같은 사람을 다른 이름 혹은 다른 표기를 사용하여 색인되거나 검색될 가능성이 있는 경우의 수를 보면 다음과 같다.

- 완전한 개인명(부모로부터 받았거나 법적으로 등록되어 있는 이름) : 李光洙, Mary Roverta Smith
- 칭호 : 세종대왕, The Marchioness of Bucktooth: Prime Minister of Lower Slobbovia
- 약어(상기 이름의 다양한 약어표기) : 세종, Mary R. Smith, M. Roverta Smith, M.R. Smith, Lady Bucktooth, The P.M.
- 별명(많이 알려진) : Bobbie, Buckie, Cuddles
- 호(필명, 자, 시호, 아호 등) : 春園, Agent 009, M.R.S., Jane Doe
- 공식적인 이름(상기 이름들에 대하여) : Mrs. John Young, The Duchess of Worcester
- 한자의 한글음 : 이광수(李光洙), 삼도유기부(三島由紀夫), 빙우란(憑友蘭), 모택동(毛澤東)

- 외국인명의 한글표기 : 마크 트웨인(Mark Twain), 미시마 유키오(みしまゆきお - 三島由紀夫), 셰익스피어(William Shakespeare)
- 외국인명의 다른 한글표기 : 마-크 트웨인(Mark Twain), 미시마 유키오(みしまゆきお - 三島由紀夫), 셰익스피어(William Shakespeare)
- 한글표기 인명의 저자에 의한 로마자표기 : Too Young Lee(李斗榮)
- 한글표기 인명의 표준번자표에 의한 로마자표기 : Du Yung i(李斗榮)
- 비로마자 인명의 자국어 표현 : みしまゆきお(三島由紀夫), А.П. Чехов (Anton Cechov)
- 비로마자 인명의 자국음에 대한 로마자번자 : Yukio Mishima(三島由紀夫), Yu-lan Feng(憑友蘭), Anton Cechov(А.П. Чехов)
- 비로마자 인명의 다른 로마자 표현 : Anton Cechov, Anton Tchekhoff, Anton Tchekhov, Anton Cexov, ... (А.П. Чехов)
- 도치명(姓/名으로의 도치) : Mao, Tse-tung; Mao Tse-tung(毛澤東), Lee, Too Young; Lee Too Young(李斗榮)

예를 들어 인명인 경우 'Dostoevski, Fedor Mikhailovich'와 'Maugham, William Somerset'라는 저자명이 우리나라의 번역작품에 실제로 나타난 예를 나열하면 다음과 같다.

Dostoevski, Fedor Mikhailovich, 1821-1881

Dostoevski, F.M.	도스토예브스끼
Dostoyevsky, Fyodor	도스토엠프스끼
도스토예브스끼, 웨너 미카이로비치	도스토에프스끼
도스토엠프스끼	도스토에프스끼이
도스토에프스끼	또스토이에프스끼

Maugham, William Somerset, 1874-1965

모음	모음, 윌리엄 서머셀
모음, 서머셀	모음, 윌리엄 서머셀
모음, 서머셀	모음
모음, 씨머셀	

또 다른 예로 「The Cherry Orchard」의 저자명표기(Cyril 문자로는 А.П. Чехов)가 각국의 여러 번역물에서 다양하게 번자된 예를 보면 다음과 같다(Hagler and Simmons 1982, 221).

Anton Cechov [English, German, Italian]	Anton Tchekhoff [English]
Anton Cechow [German]	Anton Tchekhov [English, French]
Anton Cexov [English]	Anton Tchekhov [French]
Anton Chehov [English]	Anton Tchekoff [English]
Anton Chekhov [English]	Anton Tschechov [German]
Anton Czechow [Polish]	Anton Tschechow [German]
Anton Tchegov [English]	Anton Tsjechov [Dutch]

또한 주제의 경우 주제어의 변형도 매우 다양하다. 예를 들어 “정보사부지사기 사건”에 대한 신문기사의 사용용어를 보면 다음과 같이 20여 종류로 다양함을 알 수 있다.

정보사땅사건	정보사부지사기
정보사땅사기	정보사부지사기극
정보사땅사기극	정보사부지사기사건
정보사땅사기사건	정보사사건
정보사부지매각사건	정보사사기사건
정보사부지매각사기사건	정보사터매각사기사건
정보사부지매매사기	정보사터매매의혹사건
정보사부지불하사건	정보사터사기
정보사부지불하사기사건	정보사터사기사건
정보사부지사건	

이용자의 탐색을 위하여 이 중에서 어느 것을, 어디까지를 액세스포인트로 채택할 것인가를 결정해야 할 것이며, 이름의 다른 변형에 대한 링크가 필수적이다. 즉 선거통제의 기본개념은 이와 같은 변형중 액세스포인트를 표준화하여 다른 표현과 구분되는 유일한 표현을 찾고, 그 유일한 표현을 다른 표현과 링크시키는 데 있다.

정보의 탐색시 여러 가지 다른 표현중의 하나를 사용하는 경우 그렇게 기술된 정보만이 탐색될 것이다. 그러나 선거통제화일을 유지하고 탐색시 전자화일을 이용한다면 이용자가 여러 표현중 어느 표현으로 탐색하더라도 시스템은 자동으로 채택표목과 연결시켜 해당 문헌 전체를 망라적으로 탐색할 것이다. 왜냐하면 상기 여러 가지 표현중 하나를 전거로 채택하여 표목(색인어)으로 삼아 색인하였으므로 나머지 표현은 직접 해당 표목 아래 모아져 있기 때문이다. 즉 서지데이터에는 전거인 표목만을 기술하고 해당 표목의 변형은 전자화일에 기록유지한다. 색인어는 표목들이 되지만 탐색시 언제나 전자화일을 참조하여 전거를 확인하고 그 전거로 탐색하게 된다. 새로운 변형이 출현했을 때는 서지데이터베이스를 갱신할 필요 없이 전자화일의 채택표목 아래에만 모으면 끝날 수 있다.

2.3 선거데이터의 내용

전거표목이란 현재 사용되고 있는 목록규칙 또는 각 도서관의 목록상황에 의해 규정되어 있는 액세스포인트의 형식을 말한다. 전거레코드는 표목에 관한 정보와 함께 표목자체를 기록하기 위하여 작성된다. 따라서, 이 표목이 목록이나 기계가독 서지화일 내에서 통일되어 사용된다. 다시 말하면, 전거레코드는 출판물내에 있는 어떤 형식에서 서지레코드내에 있는 액세스포인트로 안내해주는 도구이다. 따라서, 표목이 되는 개인명, 단체명, 회의명, 주제, 지명, 통일서명에 대한 전거표목이 만들어져야 할 것이다. 다만 '주제'에 대해서는 '시소러스'에서 해결할 수도 있을 것이다.

개인명은 대부분 출생시에 부여되지만 이 때의 이름은 본인이 아니라 다른 사람이 붙인 이름이다. 목록작성자에 의해 고유의 이름을 선택했을 때 서지의 세계에서 과연 고유한 이름인가를 확인할 필요가 있다. 고유하지 못하면 생물년을 추가하거나 완전형으로 기술하거나 저자가 전공하고 있는 분야의 주제를 추가하거나 하여

고유한 것으로 만들 수 있다.

단체명은 어떤 특정 활동을 위하여 조직된 개인의 집합체인 모든 단체에 적용된다. 단체명은 개인명과 같이 단체 그 자체가 만들어진 날부터 존재한다. 그러나 개인명의 경우와 비교하여 단체명의 변화는 보다 급속하며 이유가 다르므로 목록규칙에서는 다르게 취급한다. 또한 단체는 무한히 존속할 수 있으며 개인보다 수명이 길다. 또한 단체명은 내부에 하위표목 단체명이 존재할 수 있다.

회의명은 단체명으로 취급하고 있다. 그러나 회의명을 고유한 것으로 만들기 위한 한정방법이 다르다. 회의명에 추가가능한 정보는 세 종류가 있으며 그것은 번호가 붙어 있는 회차, 개최일자, 개최장소이다.

지리명은 목록의 세계에서 많은 혼란을 일으켜 왔다. 왜냐하면 지리명은 정부나 관할구의 공식명칭이므로 본질적으로도 실제상으로도 단체로 볼 수 있기 때문이다. 물론 지리명은 지역이나 지방의 명칭기도 하다. 개인명이나 단체명과 같이 지리명도 변화한다. 지리명이 지시하는 실체가 분할되거나 다른 지역과 합병되기도 하며 새로운 지역을 형성하기도 한다.

통일서명은 단일의 저작이 복수의 서명에 의해 식별될 때 사용된다. 통일서명은 저작을 식별하는 표준적인 서명이며 그 통일서명에 대하여 다른 서명이 참조된다. 다른 저작이 동일한 통일서명일 때는 창작년이나 형식세목을 추가하여 유일한 것으로 만든다.

한 건의 전거레코드에는 표목, 참조, 주기, 적용이나 취급의 지시, 각 도서관의 필수정보(목록작성자의 두문자, 목록의 대상이 되는 저작의 청구기호 등) 등의 여러 가지 종류의 전거데이터를 포함된다.

참조는 전거레코드에서 두번째로 주요한 부분이다. 일반적으로 말하면 참조 또는 참조부출지시는 전거표목의 2자택일 또는 다른 형식이며, 전거표목에 대하여 어떤 특수한 관계를 갖는 관련표목인 경우도 있다. 따라서 어떤 하나의 표목은 참조를 갖는 않는 경우도 있으며 많은 참조를 갖는 경우도 있다.

주기는 전거레코드에 여러 가지 사항을 기록하는 데 사용된다. 예를 들면 전거레코드를 작성하는 시점에 점검한 자료나 탐색결과를 기록할 수도 있으며, 표목과 다른 표목과의 관계를 명확히 하기도 하고, 그 표목의 배경이나 정보를 주거나 그 표목의 형성과정에 대한 변천을 기록할 수도 있다. 적용이나 취급의 지시에는 대부분

총서에 대한 표목을 어떻게 적용하여 왔는가에 대한 정보를 기록하고 있다.

2. 4 전거레코드의 종류

전거레코드는 액세스포인트만이 레코드가 되는 것은 아니다. 레코드를 작성할 때 레코드의 종류를 구분하여 작성하여야 할 것이다. USMARC에서의 008/09(레코드 종류)에서 정의하고 있는 레코드의 종류를 보면 다음과 같다(Library of Congress 1987).

- 1) 채택표목레코드(established heading record) - 1XX 필드에 채택표목이 기술된 전거레코드를 말한다. 이 레코드에는 보라참조필드, 관련표목, 표목선정 관련 정보 등에 대한 주기를 가질 수 있다.

따라서 채택표목레코드는 일반적으로 다음과 같은 가변장필드를 가지게 된다.

0XX	제어정보, 식별 및 분류기호 등
1XX	채택표목
360	복합도보라참조
4XX	보라참조
5XX	도보라참조
6XX	처리정보, 주기 등

- 2) 참조레코드(reference record) - 1XX 필드에 비채택표목이 기술된 전거레코드를 말한다. 이 레코드에는 이용자에게 채택표목으로 안내하기 위하여 260(복합 보라참조-주제), 664(복합보라참조-이름) 혹은 666(일반설명참조-이름)를 가질 수 있다. 이 레코드의 1XX 표목이 다른 채택표목레코드의 4XX에 출현하는지(traced: 보라참조있음) 안하는지(untraced: 보라참조없음)를 008/09에서 구별하여 정의하고 있다.

- 3) 세목레코드(subdivision record) - 1XX 필드에 채택되지 않은 부분표목이 기술

된 전거레코드를 말한다. 즉 채택표목의 주제세목부분으로 사용된다는 것을 의미한다.

4) 채택표목/세목레코드(established heading and subdivision record) - 1XX 필드에 다른 채택표목의 주제세목으로도 사용할 수 있는 채택표목이 기술된 전거레코드를 말한다. 기관에 따라 채택표목레코드와 세목레코드로 분리하여 작성할 수도 있다.

5) 참조/세목레코드(reference and subdivision record) - 1XX 필드에 참조나 채택표목의 주제세목으로 사용할 수 있는 비채택표목이 기술된 전거레코드를 말한다. 기관에 따라 참조레코드와 세목레코드로 분리하여 작성할 수도 있다.

6) 노드레이블레코드(node label record) - 1XX 필드에 시소러스에서 분류의 논리적인 기준을 나타내기 위해 사용되는 표준용어를 기술한 전거레코드를 말한다. 이 용어는 비채택표목이며 서지레코드에서 사용하지 않는다.

2. 5 표목의 형태

MARC 전거레코드에서 표목은 1XX, 4XX, 5XX 필드의 내용이며, 두 가지 형태의 표목이 있다.

1) 채택표목(established heading) - 서지레코드에서 기본표목(1XX), 주제명부출표목(600-651), 부출표목(700-730) 혹은 총서명 부출표목(440 혹은 800-830)으로 사용할 수 있는 표목이다. 전거레코드에서 채택표목은 채택표목레코드의 1XX와 5XX에 나타난다.

2) 비채택표목(unestablished heading) - 전거화일내에서 참조하기 위한 표목이다. 비채택표목은 채택표목레코드의 4XX에 나타나거나 참조레코드의 1XX에 기술된다.

2.6 전거시스템의 종류

기계화 이전의 레코드 연결구조는 관련을 갖는 모든 레코드를 한 곳에 모아 배열하거나, 목록레코드상의 참조나 주기에 다른 레코드와 연결되어 있다는 내용을 명시하는 방법밖에 없었다. 그러나 기계화시스템에서는 레코드간의 연결은 물론이고 표목을 연결할 수 있는 장점이 있다. 표목의 연결은 양쪽 레코드에 공통으로 존재하는 표목의 연결로 달성된다.

전거화일은 목록레코드상에서 사용된 표목의 형식에 대한 기록이다. 이 기록으로 전거화일과 서지화일이 서로 논리적으로 연결되게 되지만 이 연결방법은 시스템의 구축방법에 따라 달라진다. 지금까지 전거화일과 서지화일간의 관계는 다음의 세 종류가 있다(Burger 1985, 49-50).

- 1) 서지화일과는 완전히 독립된 전거화일: 아래의 예와 같이 전거레코드가 채택된 표목으로서 'Twain, Mark'를 사용하도록 규정하고 있음에도 불구하고 서지레코드의 표목은 다르다. 어떤 도서관은 1920년판을 Twain 밑에 기입하고 또 어떤 도서관은 1935년판을 Clemens 밑에 기입하도록 결정하고 있다. Twain을 규정하는 전거화일이 있음에도 불구하고 서로가 다르며 Twain과 Clemens 양쪽을 다 찾지 않는 한 동일저자인 두 저작은 아마 검색되지 않을 것이다.

전거레코드	서지레코드
Twain, Mark	Twain, Mark
x Clemens, Samuel Langhorne	Adventures of Huckleberry Finn
	1920
	Clemens, Samuel Langhorne
	Adventures of Huckleberry Finn
	1935

- 2) 서지화일과 동격의 전거화일: 전거화일과 서지화일과의 관계는 전거화일에서 규정된 이름의 형을 서지레코드에서 사용하도록 규정하고 있다. 채택된 형이 양쪽에 사용되고 있으므로 양쪽 레코드 모두 검색될 것이다. 다만 이름의 적절

한 형을 확인하기 위하여 전거화일을 탐색하지 않으면 안될지도 모른다.

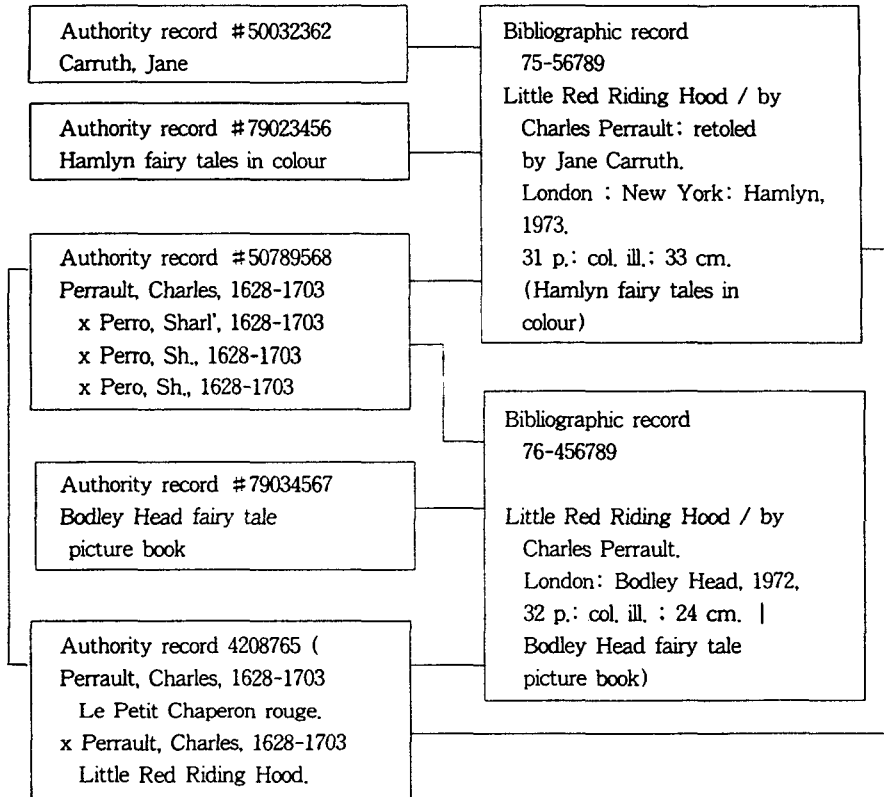
전거레코드	서지레코드
Twain, Mark	Twain, Mark
x Clemens, Samuel Langhorne	Adventures of Huckleberry Finn
	1920
	Twain, Mark
	Adventures of Huckleberry Finn
	1935

3)서지화일의 일부로서의 전거화일: 이름의 변형은 전거레코드에 의해 결정되고 있지만 서지화일에서 레코드를 탐색하기 위하여 전거화일만을 탐색하는 능력도 구비하고 있다. 다시 말하면 두 가지 화일을 탐색하기 위해 필요한 색인은 하나뿐인 것이다. 전거화일중의 표목이 서지화일에 사용되면 포인터에 의하여 연결되며, 표목은 전거화일중에 단 한번 출현하게 된다. 따라서 표목이 변경되는 경우 서지데이터의 변경없이 전거화일만의 변경으로 끝나게 된다.

전거레코드	서지레코드
(레코드번호 2345)	(전거레코드번호 2345)
Twain, Mark	Twain, Mark
x Clemens, Samuel Langhorne	Adventures of Huckleberry Finn
	1920
	(전거레코드번호 2345)
	Twain, Mark
	Adventures of Huckleberry Finn
	1935

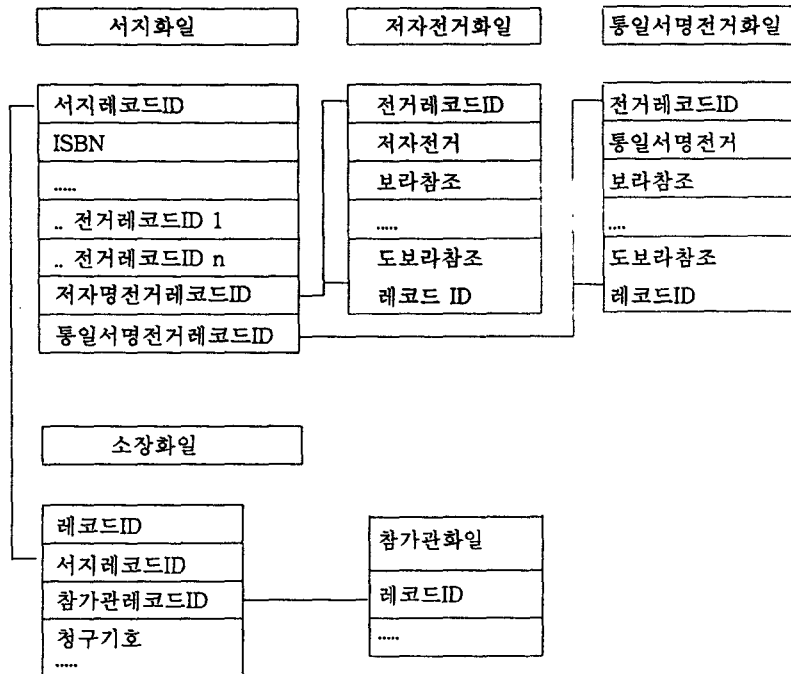
시스템내부에서의 서지레코드와 전거레코드간의 표목에 의한 연결관계 개념을 예로 보면 <그림 1>과 같다(Burger 1985, 8).

〈그림 1〉 서지데이터와 전거데이터의 연결



분담목록시스템에서의 전거시스템의 구조를 설계해 보면 다음 〈그림 2〉와 같다. 전거화일은 저자전거화일과 통일서명전거화일의 예를 든다. 서지화일은 단행본서지화일 혹은 연속간행물서지화일 등으로 구분할 수도 있다. 그림에서와 같이 서지화일이 각종 전거레코드 ID를 명시적으로 가질 수도 있다. 다만 양쪽 화일의 입력과 갱신이 어려우므로 전거레코드 ID를 갖지 않고 표목에 의한 암묵적인 연결도 가능하다.

〈그림 2〉 서지화일과 선거화일의 연결



3. 선거시스템의 문제점

3.1 선거데이터 관리비용

외국에서는 인명선거 혹은 주제명선거 등을 상당수 보유하고 있어서 정보검색의 기본 도구로서 중요한 역할을 담당하고 있다. 원래 선거데이터란 자국의 국어와 관련된 문제로 우리의 것은 우리가 제작하는 방법외에는 없으나 선거작업시 외국에서 이미 만들어진 선거화일을 이용하면 해당국가의 선거데이터 구축시 많은 참조

가 된다. 그러나 참조가 될 뿐 그것으로 충분하지는 않다. 예를 들어 외국의 선거 화일에서 'Mozart, Wolfgang Amadeus, 1756-1791'이란 이름에는 'Mozart, Johann Chrysostom Wolfgang Amadeus'라는 다른 표기가 있다는 것은 도움을 받을 수 있지만, 「모짜르트, 볼프강 아마데우스: 모차르트, 볼프강 아마데우스」등의 표기는 우리나라에서 만들어야 하며, 만드는 관점에 따라 표목이 바뀔 수도 있는 것이다.

불행하게도 현재까지 한국어의 다양한 표현중 어느 것을 대표(전거)로 삼아야 하며 변형은 어떤 것이 있는가에 대한 조사나 연구가 일천하여 전거데이터에 대한 표준 포맷도 범용성이 있는 기본 전거데이터도 만들어져 있지 못한 실정이다. 다만 일부 기관이 자관에서 정한 규칙에 따라 카드식 전거데이터를 만들어 사용하고 있을 뿐이다.

전거데이터의 작성 및 전거통제과정은 전문적인 조사와 판단이 필요하고 노동집약적이므로 비용이 많이 드는 일이다(Hagler and Simmons 1982, 181). 예를 들어, 텍사스대학이 1977-78 회계년도의 1년동안 이름/서명 전거통제에 14만5천달러를 사용하여 자료당 1.25달러가 들었으며, ARL도서관들이 1979년 한해 동안 전거통제에 든 비용이 무려 5백만달러였다(Taylor 1989, 33). 뿐만 아니라 컴퓨터를 이용하여 자동전거시스템을 구축하는 경우, 서지데이터의 표목을 관리하고 정합성을 유지하기 위하여 전거통제와 관련되어 사용되는 컴퓨터의 자원은 막대하다. 예를 들어 MARC포맷을 사용하는 경우를 보자. 우선 입력되는 모든 표목에 대하여 전거통제를 받아야 한다. 또한 새로운 1XX, 4XX, 5XX중의 하나가 추가되었을 때, 전체 전거데이터를 대상으로 서로를 점검해야 할 뿐만 아니라 표목이 바뀌면 서지데이터까지 점검하고 고쳐야 한다. 자관용, 분담목록용을 막론하고 저렴한 비용으로 전거시스템을 구축하고 전거데이터를 관리할 수 있어야 할 것이다.

3. 2 전거의 채택

목록자는 기관의 규칙이 있겠지만 여러 가지 표현중에서 전거를 채택해야 한다. 전거를 채택하기 위해서는 여러 가지 정보원을 이용하여 모든 표현을 수집하여야 한다. 'Mark Train'을 예로 들어보자.

Twain, Mark
 Clemens, Samuel Langhorne
 트웨인, 마크
 트웨인, 마-크
 클레멘스, 새뮤엘 랑혼
 클레멘스, 사무엘 랑혼

상기 예에서 어느 것을 전거로 채택할 것인가? '영문표현을 채택한다' 혹은 '한글표현을 채택한다'는 것이 원칙적으로 정해져 있더라도 문제는 남아 있다. 모든 표현을 수집하지 못해 전거로 채택해야 할 표현이 채택된 후에 나타나는 경우도 있을 것이다. 그리고 전술한 대로 이 결정이 서지데이터를 작성하는 시점에 이루어져야 하는 것이다.

3. 3 서지 및 전거데이터의 다운로드

많은 도서관들이 기존의 서지유틸리티를 이용하여 서지데이터나 전거데이터를 다운로드할 수 있는 기회가 많아졌다. 이 때 타관의 서지데이터를 손쉽게 이용하기 위해서는 서지사항 뿐만 아니라 표목까지도 일치하여야 한다. 해당 표목에 해당하는 전거데이터도 함께 다운로드하여 참조할 수 있기 때문이다. 그러나 전거의 채택 기준이 다르면 표목이 달라진다. 항상 서지데이터를 다운로드할 때마다 자관의 전거기준과 맞는지를 점검해야 하고, 다르면 서지데이터를 수정해야 하는 번거로움이 따르게 된다.

국가적인 수준의 전거데이터가 만들어져 있다고 하더라도 로컬이 그 데이터를 그대로 사용하지 않는 한 로컬의 요구에 맞도록 로컬시스템용으로 변환되어야 한다 (Clack 1990, 7)는 것은 당연한 일이다. 그러나 그 데이터를 그대로 사용할 수 있다면 업무의 효율성이 상당히 제고될 것은 틀림없다.

3. 4 분담목록센터의 전거통제

많은 회원기관에서 전송되는 서지데이터 및 전거데이터를 관리하고 통제하는 분

담록센터에서는 전거데이터의 관리가 가장 문제가 된다. 특히 각 기관의 데이터가 한꺼번에 오는 초기화 단계에서는 많은 혼란이 예상된다. 물론 회원기관이 입력하여 전송한 서지데이터의 질을 높이기 위하여 센터시스템은 채택 및 비채택표목에 대한 검색 및 참조기능을 강화하여야 할 것이다.

센터시스템은 로컬시스템에게 새로운 전거의 채택이나 기존 전거의 변경을 허락하지는 않을 것이다. '보라참조'는 채택이나 변경이 가능할 수도 있다. 문제는 각 회원기관에서 전송된 데이터에 대한 표목을 분석하고 결정하며 전달하는 전거통제의 과정이 센터의 많은 자원을 소비하게 될 것이라는 데 있다.

3.5 서지데이터의 입력

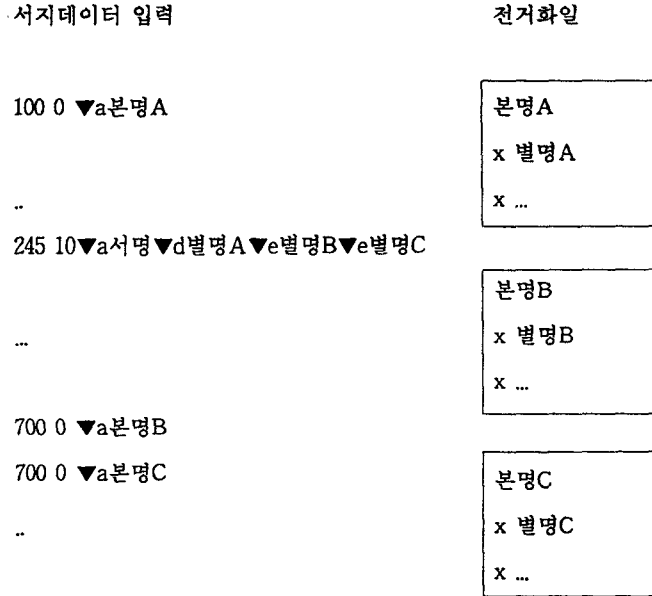
전거시스템에서 서지데이터를 MARC형식으로 입력하는 경우 상당히 번거로운 과정을 거치게 된다. 번거로운 과정은 '이용자 인터페이스'로 어느 정도 해결이 되겠지만 근본적인 것은 마찬가지이다. 가장 편리한 방법을 사용하는 서지데이터 입력시스템의 전거통제 과정을 보자.

다음 <그림 3>은 3인의 개인명 공저자가 있다고 가정하고, KORMARC(국립중앙도서관 1993)형식에 따라 입력한 것이다. 표제지의 저자는 모두 전거화일의 '보라참조'에 해당하는 '별명A, 별명B, 별명C'로 표시되어 있으며, 그들의 전거는 각각 '본명A, 본명B, 본명C'라 가정한다.

먼저 필드 100에서 별명A를 입력하고 그 자리에서 전거통제를 받는다. 전거화일에서 전거가 본명A이므로 본명A를 필드 100에 넣고 필드 100에 입력된 별명A는 245 ▼a▼d별명A로 옮겨진다. 다음에 필드 245에서 ▼a서명을 입력하고 별명A는 있으므로 ▼d별명A 다음에 ▼e별명B를 입력한다. 이 자리에서 별명B가 전거통제를 받는다. 별명B의 전거가 본명B이므로 별명B는 그 자리에 두고 본명B를 필드 700에 옮긴다. 다음에 별명C를 입력하고 전거통제를 받는다. 본명C가 전거이므로 별명C는 그 자리에 두고 본명C를 필드 700을 반복하여 옮긴다. 표목은 필드 100과 700이 된다.

이와 같은 입력과정은 시간이 걸릴 뿐만 아니라, 전술한 바와 같이 전거와 변형의 관계를 시스템이 자동으로 처리해주어도 복잡하다. 자동으로 옮겨주지 않으면

〈그림 3〉 서지데이터의 입력과 전거통제



전거와 변형을 모두 입력해야 하며 오류가 발생할 여지가 많다. 또한 전거화일에 해당 변형 혹은 전거가 없으면 새로운 전거레코드를 만들어야 한다.

4. 무전거시스템

4.1 개요

전거화일과 서지데이터를 연결하기 위한 장치는 다음과 같은 조건을 갖출 필요가 있다(Lauro and Sly 1985, 4-6). 첫째, 기억장소를 절감할 수 있어야 한다. 기관명이나 회의명과 같은 긴 이름은 전거화일에 한번만 기억되어야 한다. 긴 표목 대

신에 코드를 사용할 수도 있다. 둘째, 시간을 절약할 수 있어야 한다. 즉 긴 이름 대신 코드의 입력, 새로운 레코드의 생성시 단 한번의 검증, 전거레코드의 수정으로 서지데이터까지 수정 등으로 사서의 시간을 절약할 수 있을 것이다. 셋째, 서지데이터입력시 입력오류를 최소화 하여야 한다. 즉 코드를 입력함으로써 오류를 줄일 수 있으며, 전거화일과의 자동체크와 함께 여러 가지 안내 메시지를 낼 수 있을 것이다. 넷째, 정보검색이 효율적이어야 한다. 즉 전거화일이 도치화일의 역할을 할 수도 있으며, 어느 변형으로 찾더라도 채택표목으로 검색하여야 하며, 상호참조를 출력할 수도 있을 것이다.

그러나 전거시스템의 중심이 되고 있는 전거의 채택이 사실상 많은 문제점을 야기하고 있다. 기본적인 전제는 전거를 채택하지 않으면 많은 문제가 해결될 것이라는 데 있다(김태수 1993).

다만 전거를 채택하지 않았을 때에도 전거를 채택할 때와 같이 다음 세 가지 기능을 가져야 한다. 즉, 첫째, 이용자는 여러 가지 표현중 어느 표현을 사용하더라도 모두 망라하여 검색할 수 있어야 한다. 둘째, 저자기호의 채택에 문제가 없어야 한다. 셋째, 검색시간이 짧아야 한다. 아울러 전거를 사용하지 않으므로 서지데이터의 입력이 편리하며, 기억장소를 절감할 수 있고, 비용이 적게 들고, 서지데이터의 통합이 효율적이라면 더욱 바람직할 것이다.

이와 같은 관점에서 전거를 채택하지 않는 시스템을 '무전거(無典據)시스템'이라 하고, 상기 목적을 달성하기 위하여 다음과 같은 세 가지 원칙을 세운다.

■ 원칙 1 : 참조그룹을 작성한다

무전거시스템은 전거를 채택하지 않고 단지 '참조그룹'만을 만드는 것이다. 예를 들면 다음과 같은 표현이 모두 '마크 트웨인'을 가리킨다면 대표표현 없이 모두를 한 그룹으로만 모아 두는 것이다.

Twain, Mark
 Clemens, Samuel Langhorne
 트웨인, 마크
 트웨인, 마-크

클레멘스, 새뮤얼 랑혼

클레멘스, 새뮤얼 랑혼

....

■ 원칙 2 : 표목은 원자료에 따른다

서지데이터의 입력시 전거가 없으므로 표목은 원자료에 기술된 저자의 표기대로 사용한다. 사용된 표목은 해당 참조그룹에 포함되어야 할 것이다.

■ 원칙 3 : 참조그룹에는 그룹코드를 부여한다

각 그룹에 고유한 코드를 부여한다. 이 그룹코드는 두 가지 역할을 한다. 하나는 저자기호용이며, 또 하나는 검색의 효율성을 위한 것이다.

4.2 채택-비채택표목의 등가성

MARC 전거레코드에서 뿐만 아니라 모든 전거통제시스템에서는 채택표목과 비채택표목이란 두 가지 형태의 표목으로 나누며, MARC 전거레코드에서는 각 표목이 1XX, 4XX, 5XX 필드에 기술된다. 선택여부에 따라 채택표목레코드가 생성되거나 기술위치가 달라지게 된다.

그러나 무전거시스템에서는 원칙 1과 2에 따라 이들의 구분이 없어지고 동등해진다. 즉 어느 것을 전거로 선택할 것인가 노력할 필요도 없으며, 참조그룹에 속하는 표현은 모두 표목으로 사용할 수 있다. 또한 전거시스템에서처럼 1XX와 4XX의 정합성을 검증할 필요가 없어진다.

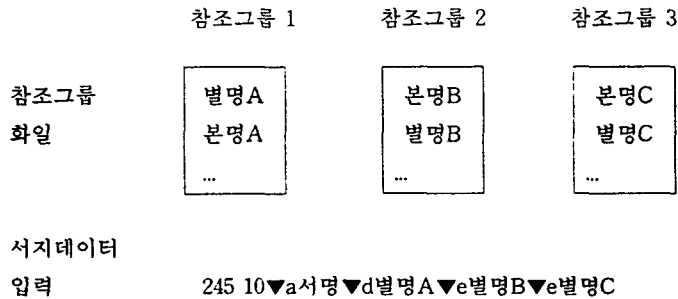
USMARC 전거포맷이 정의하고 있는 세목(subdivision) 및 노드레이블(node label)에 해당되는 표현을 별도로 처리하고 싶을 때는 얼마든지 다른 방법으로 구별할 수 있을 것이다.

4.3 서지데이터 입력의 편리

전거를 표목으로 사용하지 않으면 서지데이터 입력형식에서 표목부가 없어도 좋

다. 예를 들어 개인저자명인 경우 <그림 3>의 입력사항이 줄어 <그림 4>와 같이 필드 245의 입력으로 완료되며, 필드 100과 700은 필요없게 된다. 포맷이 단순해질 뿐만 아니라 입력량과 처리량이 대폭 줄어들게 되어 인력과 컴퓨터 자원을 절감할 수 있다. 표목은 전거가 없으므로 원자료에 있는 대로 '별명A, 별명B, 별명C'가 된다.

<그림 4> 서지데이터입력과 참조그룹



다만 필요에 따라서는 전거데이터가 갖는 주기, 적용이나 취급의 지시, 각 도서관의 필수정보 등 여러 가지 정보를 참조그룹에 추기할 수 있을 것이다.

4. 4 저자기호 및 검색효율을 위한 그룹코드

각 그룹에 고유한 코드를 부여한다. 이 그룹코드는 두 가지 역할을 한다. 하나는 저자기호용이다. 표목을 통제하지 않고 원자료에 준하게 되면 동일한 저자의 저작이 저자기호의 상이로 떨어져서 배열될 수도 있으나 이 그룹코드를 이용하면 한 곳에 모을 수 있다(정준민 1993). 그룹코드는 편의에 따라 임의의 문자열을 이용하여 만들 수 있다. 저자에 종속하는 저자기호를 사용하지 않고 연대순저자기호를 사용할 수도 있을 것이다.

또 하나는 검색의 효율성을 위한 것이다. 참조그룹의 모든 구성원에 대한 문헌번호를 그룹코드에 링크시킨다. 따라서 이용자가 참조그룹내에 있는 임의의 구성원

으로 검색했을 때, 하나 하나 모두를 검색하여 논리합으로 연산할 필요 없이 한번의 검색으로 전거시스템에서와 같은 효과를 얻을 수 있다. 그룹코드는 검색시 전거시스템에서의 전거와 같은 역할을 한다.

〈그림 5〉는 참조그룹과 색인화일간의 관계를 단순하게 나타낸 것이다. 색인화일에는 원래의 문자열이 아니라 코드를 갖기 때문에 색인화일의 관리가 단순해지고 기억영역을 크게 줄일 수 있다.

4.5 서지 및 참조데이터 통합의 용이성

전거의 선정문제가 발생하지 않으므로 각 기관은 원자료에 표기된 사항에 충실하게만 기술하면 서지데이터의 통합은 상당히 간단해진다. 표목부가 단순해지므로 동일저작의 검증이 쉬워지며, 기관마다 가지고 있는 전거의 선정규칙이나 개인적인 관점에 대하여 독립적이므로 오류의 발생여지가 줄어들게 된다.

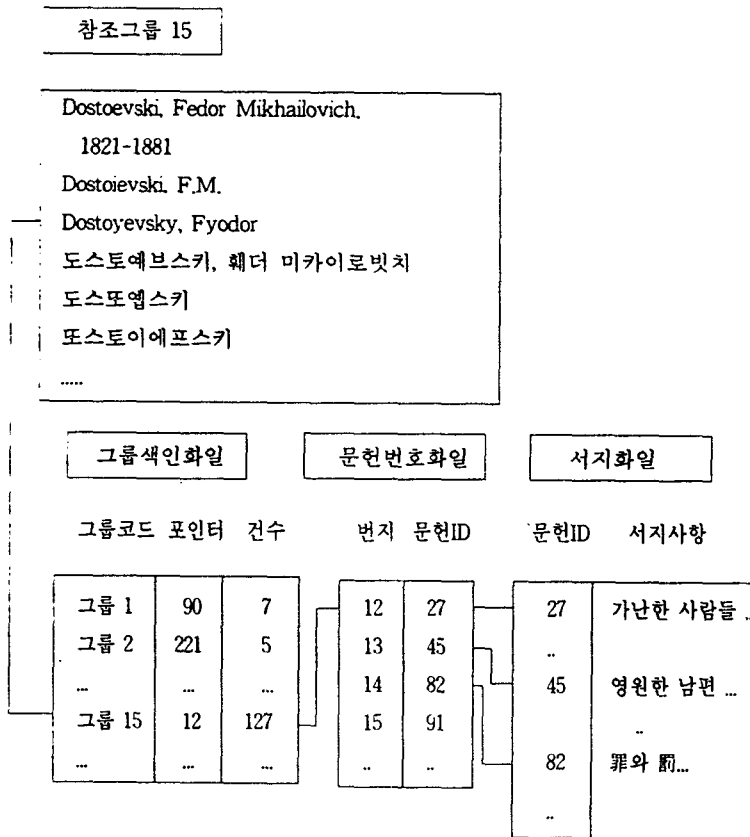
따라서 분담목록센터에서는 서지사항 기술의 정확성과 표목에 대한 참조그룹의 관리만을 하게 되므로 시간, 인력, 비용의 부담이 훨씬 줄어들게 된다. 여러 나라의 MARC데이터를 다운로드하여 사용하는 기관도 채택된 표목이 원자료에 표기된 것인가 아닌가만 확인하여 다른 표기일 경우에만 원자료의 표기를 사용하고, 표목으로 채택된 것은 참조그룹에 추가시키면 되는 것이다.

분담목록시스템이 무전거시스템을 채택한다면 로컬기관에서는 그 서지데이터 뿐만 아니라 참조데이터도 다운로드하여 그대로 사용할 수 있다. 센터에서는 전거통제시스템을 사용하고 로컬기관에서만 이 방법을 사용한다면 원자료의 표기대로 사용되었는가만 확인하여 다른 표기일 경우에는 원자료의 표기를 사용하고 센터의 전거데이터를 다운로드하여 1XX와 4XX를 합하여 자관의 참조그룹에 추가시키면 될 것이다.

원자료에 기술된 표기인가 아닌가의 검증과 수정도 이용자 인터페이스를 구축하여 시스템이 자동검증과 수정이 가능하도록 할 수 있다. 즉 서지데이터의 표목부를 시스템이 서지기술부(원자료의 표기)와 대조하여 상이한 경우 서지기술부의 문자열로 대체할 수 있으므로 로컬기관이나 분담목록센터에서 사서의 할일이 줄어들게 된다. Miller(1982, 43-44)가 제안한 대로 단위도서관의 전거통제란 너무나 어렵고

복잡한 일이기 때문에 선거통제시스템을 네트워크상에 구축하고 도서관들이 공유하는 방법도 있을 것이다. 이 때 현재의 선거시스템으로도 가능한 하겠지만 선거의 관리라는 어려운 문제는 여전히 남는다. 무선거시스템을 채택한다면 많은 짐을 덜 수 있을 것이다.

<그림 5> 검색시스템의 화일 구성



데이터베이스의 종류가 주제별로 다양해짐에 따라 이용자의 편의성을 위하여 LC의 MUMS(multiple use MARC system)과 같은 시스템이 출현할 수 밖에 없게 되

며, 모든 표목은 하나의 색인으로 통합하게 된다. 따라서 전거통제시스템은 복수 MARC시스템과 같이 복수전거통제시스템을 사용하여야 한다(Petersen 1987, 23-29). 무전거시스템을 사용한다면 다른 데이터베이스용 전거데이터를 통합하는 것이 쉬워지게 되므로 복수전거시스템을 채택할 필요가 없어지게 된다.

4.6 자연언어-통제어휘 혼합시스템의 구축

1950년대부터 통계, 구문, 의미분석 등의 방법에 의한 자연언어 자동색인이 계속 연구되어 왔으며 성과를 올리고 있다. 자동색인의 목적은 '빨리, 저렴한 비용으로, 통제어휘와 동등한 혹은 그 이상의 효율성을 갖는 방법은 없을까'에 있다. 다만 그렇게 많은 연구와 자연언어의 효율성을 증명하는 많은 결과가 있었지만, 현재 도서관이나 정보센터에서 탐색용으로 쓰고 있는 상용 대단위 데이터베이스를 대상으로 한 연구는 별로 없었다. 뿐만 아니라 대부분의 연구는 과학기술분야 데이터베이스를 대상으로 하고 있었다. 그러나 확실한 것은 '현재까지의 연구로 나타난 결과로서는 자연언어가 좋기 때문에 통제어휘를 버릴 수 있다거나 버려야 한다는 아무런 근거를 발견할 수 없다'는 것이다(Harter 1986, 57). 그렇다면 어느 방법이 가장 이상적이며 최선의 방법인가?

우선 자연언어와 통제어휘의 특성을 비교하면 <표 1>과 같다(Harter 1986, 54).

자연언어의 장점을 색인측면에서 보면 최신용어를 사용하여 용어의 최신성을 유지할 수 있으며, 저자가 사용한 어휘이므로 색인작성자에 의한 오해의 소지가 없으며, 융통성과 표현력이 좋으며, 입력비용이 적게 들며, 데이터베이스의 교환이 용이하며, 사고시 색인의 재생성이 용이하다는 점이다. 검색측면에서 보면 특정성이 높아 정도율이 높으며, 탐색자는 자연언어상의 어휘를 사용할 수 있다는 점이다. 자연언어의 단점을 보면 동의어와 유사동의어의 존재를 처리하기 어려우며, 동음이의어를 처리하기 어려우며, 문맥상의 애매함을 처리하기 어려우며, 개념간의 계층 관계를 정의할 수 없다는 점을 들 수 있다.

통제어휘의 장단점을 보면 자연언어와 반대로 자연언어의 장점이 통제어휘의 단점이 되고 자연언어의 단점이 통제어휘의 장점이 되고 있다.

사용하는 검색시스템 및 대상정보에 따라 달라지겠지만 현재 최상의 시스템은 자

연언어와 통제어휘를 병립시키는 혼합시스템일 것이다. 자연언어를 사용하면 좋은 경우와 그 반대의 경우가 있기 때문이다. 즉 자연언어를 사용하는 것이 좋은 경우로는 의미의 뉘앙스가 중요할 때, 통제어휘로는 그 개념을 표현하는 적당한 용어가 없을 때, 주제분야가 부드러운 분야(수학쪽보다 예술쪽에 가까운 것)일 때이며, 통제어휘를 사용하는 것이 좋은 경우로는 통제어휘를 사용함으로써 개념이 정확하고 애매하지 않을 때, 망라검색을 하려고 할 때, 주제분야가 딱딱한 분야(예술쪽보다 수학쪽에 가까운 것)일 때일 것이다.

전거시스템은 통제어휘의 개념과 동일하며, 무전거시스템은 자연언어와 통제어

〈표 1〉 자연어와 통제어휘의 특성

자연언어	통제어휘
표현력이 좋다	표현력이 나쁘다
표현이 자유롭다	표현이 엄격히 제한되어 자유롭지 못하다
망라탐색이 어렵다	망라탐색이 비교적 쉽다
액세스포인트가 다양하다	액세스포인트가 한정되어 있다
동의어의 문제가 있다	동의어를 통제한다
동형이의어의 문제가 있다	동형이의어를 통제한다
false drop의 문제가 있다	전조함으로 false drop을 막는다
융통성이 있다	융통성이 없다
다양한 관점을 표현할 수 있다	하나의 관점을 표현한다
훈련이 필요없다	훈련이 필요하다
새로운 개념을 표현하기 쉽다	새로운 개념을 표현하기 어렵다
복잡한 개념을 표현하기 쉽다	복잡한 개념을 표현하기 어렵다
애매하다	명쾌하고 정확하다
표준화되어 있지 않다	표준화되어 있다
색인작업이 필요없다	일관성있는 색인작업이 필요하다
탐색어는 이용자가 생각해야 한다	상호참조에 의해 탐색어를 제시할 수 있다
그렇게 간결하지 않다	매우 간결하다
실재를 그대로 대표한다	실재를 그대로 대표하는 것은 아니다

휘의 혼합개념과 동일하다. 따라서 무전거시스템을 이용하여 혼합시스템을 구축할 수 있다. 즉 그룹화일을 통하여 접근하면 통제어휘시스템이며 용어색인화일을 바로 접근하면 자연언어시스템이 된다. 이 개념을 그림으로 나타내면 <그림 6>과 같다.

특히 최근에는 서지적으로 불완전하거나 도서와는 다른 성질을 가진 데이터의 비정규화 용어 혹은 자유키워드로 검색할 수 있는 기능과 도서데이터와의 통합 및 통합검색 요구가 증가되고 있다(Bregzis 1978, 124). 무전거시스템에 의한 혼합시스템은 이와 같은 요구를 충분히 충족할 수 있을 것이다.

4.7 참조그룹의 작성과 전거작성의 비교

1) 전거의 작성

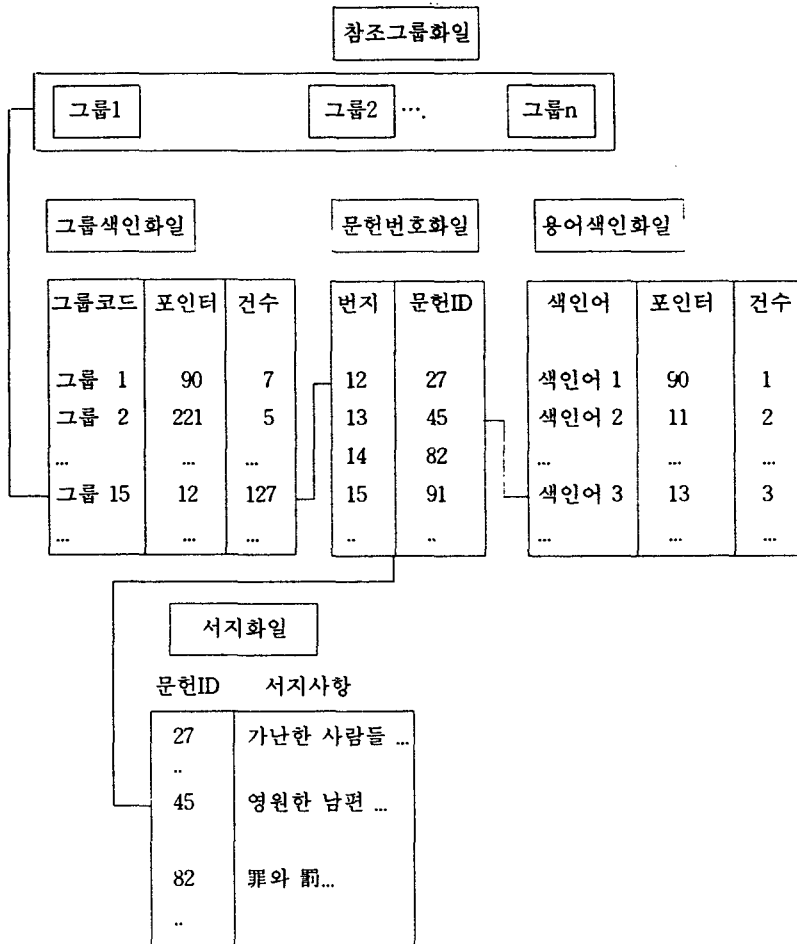
전거레코드의 작성시에는 두 가지 방법이 있다. 하나는 모든 표목에 대하여 전거레코드를 작성하는 방법이다. 즉 변형을 갖지 않는 표목도 전거레코드를 갖는다. 이것은 '보라참조'는 없지만 '도보라참조'가 있을 수 있으며, 변형이 생기거나 후에 두 개의 유사한 표목이 생길 것에 대비하여 미리 상이점을 정의함으로써 업무의 복잡성을 들 수 있는 잇점이 있다. 물론 이 전거가 다른 전거레코드의 '보라참조'에 포함되어 있는지를 점검하여야 한다. 이 방법은 시스템이 전거레코드를 자동으로 생성하게 할 수도 있다. 다만 전거가 바뀌면 레코드를 재작성해야 하는 불편과 전거화일이 방대해지는 단점이 있다.

또 하나는 변형이 서지레코드에 발생할 때만 전거레코드를 만드는 방법이다. 이때는 대상이 전거이든 변형이든 두 가지 형을 포함하는 새로운 전거레코드를 만들어야 하는 단점이 있는 반면 전거레코드의 재작성에 따른 오류가 줄어들며 전거화일이 작아지는 잇점이 있다. 어느 방법을 사용하든 전거레코드의 유지보수에는 막대한 시간과 비용을 요하게 된다.

2) 참조그룹의 작성

서지데이터의 어떤 항목이 표목으로 사용되었을 때 참조그룹을 만드는 방법은 전거의 작성과 같이 두 가지 방법이 있다. 하나는 표목이 한 종류만 있는 경우에는 참

〈그림 6〉 자연언어-통제어휘 혼합시스템



조그룹을 만들지 않고, 두 종류 이상의 표목이 사용되었을 경우에만 참조그룹을 만드는 방법이다. 또 하나는 표목으로 사용된 것은 모두 참조그룹을 만드는 방법이다.

전자의 방법을 채택하는 시스템에서는 다음과 같은 세 가지의 경우가 생기게 된다. 첫째, 참조그룹이 있으며, 그 표목이 구성요소로 이미 존재하는 경우이다. 이 때

는 확인으로 끝난다. 둘째, 참조그룹은 있으나 그 표목이 포함되어 있지 않은 경우이다. 이 때는 참조그룹에 포함시킴으로써 처리는 완료된다. 셋째, 포함시킬 참조그룹이 없는 경우이다. 이 경우에는 하나의 구성요소를 갖는 참조그룹을 만든다. 후자의 방법을 채택하는 시스템에서는 첫째, 둘째의 과정을 거치면 된다.

따라서 서지데이터의 입력시 '4.3 서지데이터 입력의 편리'에서 논한 바와 같이 '245 10▼a서명▼d별명A▼e별명B▼e별명C'만의 입력으로 모든 것이 해결되는 것은 아니다. 다만 선거에서와 같이 채택표목과 비채택표목의 정합성을 검토할 필요 없이 같은 그룹인지만 확인하면 되므로 업무의 내용이 선거작성보다는 사서와 시스템 양측면에서 대단히 단순해 진다.

5. 맺음말

선거시스템의 문제점은 1) 선거데이터작성 및 통제과정은 전문적인 조사와 판단이 필요하고 노동집약적이므로 비용이 많이 들며, 2) 선거의 결정이 어려우며, 3) 로컬기관에서 기존의 서지데이터 다운로드시 서지데이터를 수정하거나 선거데이터를 그대로 사용할 수 없으며, 4) 분담목록센터에서는 회원기관의 선거데이터를 관리하기 어려우며, 5) 서지데이터포맷 및 선거포맷이 복잡하여 양쪽 데이터의 입력이 어렵고 번거롭다는 데 있다.

이를 해결하기 위해서는 선거를 채택하지 않는 방법이 최선이라 사료된다. 선거를 사용하지 않는 방법을 '무선거시스템'이라 명명하고, 선거의 단점을 보완하기 위하여 선거를 사용하지 않는 대신 1) 참조그룹을 작성하고, 2) 표목은 원자료에 따르고, 3) 참조그룹에는 그룹코드를 부여하는 원칙과 응용하는 방법을 제안하였다.

무선거시스템을 사용하면 선거시스템의 기능을 모두 가지면서 다음과 같은 잇점을 생각할 수 있다.

- 1) 선거의 채택과 정합성의 검증에 드는 비용을 대폭적으로 줄일 수 있다.
- 2) 서지 및 선거포맷이 단순하게 되어 입력이 편리하다.
- 3) 색인화일의 관리가 단순해지고 기억영역을 크게 줄일 수 있으며, 컴퓨터 자원

을 절감할 수 있다.

- 4) 각종 서지데이터의 통합이 용이하다.
- 5) 자연언어-통제어휘 혼합시스템의 구축이 쉽다.

본 연구에서 제안한 무전거시스템은 단위기관과 분담목록센터관의 입장에서 많은 보완연구와 실험결과가 있어야 할 것이다. 다만 무전거시스템을 기능, 효율성, 시스템 등의 측면에서 현재까지 검토한 바로는 무리가 없으며, 전거시스템보다는 효율적이라 판단된다. 전거데이터를 생산하는 국가서지기관에서는 전거시스템을 사용하는 기관을 위하여 전거데이터를 생산할 수도 있으나, 이를 사용하는 로컬기관에서는 이를 무전거시스템의 참조데이터로 변환하여 사용하는 것이 효율적이라 사료된다.

참고문헌

- 국립중앙도서관(1993). 한국문헌자동화목록형식 - 단행본용 -. KS C 5867-1993.
- 김태수(1993). 한남대학교 문헌정보학과 김태수 교수와 1993년 6월 23일 속초에서 면담.
- 圖書館情報學ハンドブック編輯委員會(1988). 圖書館情報學ハンドブック. 東京 : 丸善(株).
- 정준민(1993). 전남대학교 문헌정보학과 정준민 교수와 1993년 7월 16일 대전에서 면담.
- ALA(1983). *The ALA Glossary of Library and Information Science*. Chicago : ALA.
- Bregzis, Ritvars(1978). "Integrated Access to Information and the Economic Reality." In *What's in a Name? Control of Catalogue Records Through Automated Authority Files*. Proceedings of the Workshops Sponsored by National Library Canada, Vancouver May 25-26, 1978.
- Burger, Robert H.(1985). *Authority Work: The Creation, Use, Maintenance, and Evaluation of Authority Records and Files*. Libraries Unlimited, Inc.
- Clack, Doris Hargrett(1990). *Authority Control: Principles, Applications, and Instructions*. Chicago : American Library Association.
- Hagler, Ronald and Peter Simmons(1982). *The Bibliographic Record and Information Technology*. Chicago : American Library Association.
- Harter, Stephen P.(1986). *Online Information Retrieval*. New York : Academic Press.
- Lauro, Anne Di and Maureen Sly(1985). *Recommended Methods for Development-Information Systems*, Volume 2: Guidelines for the Building of Authority Files in Development-Information Systems, IDRC-TS52e.
- Library of Congress(1987). *USMARC Format for Authority Data Including Guidelines for Content Designation*. Washington : Library of Congress.
- Miller, R. Bruce(1982). "Authority Control in the Network Environment." In (Mary

-
- W. Ghikas ed.) *Authority Control: The Key to Tomorrow's Catalog*. Proceedings of the 1979 Library and Information Technology Association Institutes. Phoenix : The Oryx Press.
- Petersen, Toni(1987). 'Multiple Authorities in Library Systems.' In (Karen Muller ed.) *Authority Control Symposium*. Occasional Paper No.6. Papers Presented During the 14th Annual ARLIS/NA Conference, New York, Feb. 10, 1986. Tucson : Art Libraries Society of North America.
- Taylor, Arlene G.(1989). 'Research and Theoretical Considerations in Authority Control. " In (Barbara B. Tillett ed.) *Authority Control in the Online Environment: Considerations and Practices*. New York : The Haworth Press.

ABSTRACT

A Study on the Non-Authority Control System

Suk-Doo Choi*

The authorized form is called the heading must be unique and consistent. Without both uniqueness and consistency, retrieval from the bibliographic database will not be optimal and in some cases may even be impossible. But the authority control is extremely expensive, because it is labour-intensive and dependent on professional investigation and judgement. In fact, it is by far the most expensive function in all of bibliographic control.

The processes used to determine the form of access points, maintain, use, and evaluate very accurate authority records reduce the efficiency to generate or merge bibliographic items without modification from any national level bibliographic utilities as well as individual library.

In this paper non-authority control system was suggested to solve the disadvantages of authority work. The apparent advantages of the non-authority control system are it is easy to determine the heading, input the bibliographic data, integrate the access points into an existing file, and generate and merge bibliographic data without modification. In addition, non-authority control system can be transmitted into online information retrieval system conducted with a mixture of controlled vocabularies and natural language.

* Assistant Professor, Department of Library and Information Science, Ewha Womans University