

전자용 MARC 포맷에 관한 연구*

A Study on the MARC Format for Authorities

오 동 근 (Dong-Geun Oh)**

목 차

- | | |
|------------------------|-------------------------|
| 1. 서언 | 3. 2 가변길이필드의 데이터요소 |
| 2. 전자용포맷의 구조와 내용표지법 | 3. 2. 1 문자와 표기법 |
| 2. 1 레코드레이블과 디렉토리 | 3. 2. 2 고유명 전자데이터 |
| 2. 2 데이터필드 | 3. 2. 3 주제명 전자데이터 |
| 2. 3 내용표지법 | 3. 2. 4 통일서명과 총서명 전자데이터 |
| 3. 전자용포맷의 레코드 내용 | 4. 결론 |
| 3. 1 고정길이필드의 코드화 데이터요소 | |

초 록

본고는 전자용 MARC 포맷의 3요소로 일컬어지는 구조와, 내용표지법, 레코드의 내용에 대해 UNIMARC와 USMARC, KORMARC를 중심으로 비교분석하였다. 레코드의 구조와 내용표지법은 ISO 2709를 채택하고 있으므로, 대부분 서지용포맷과 동일하다. 다만 데이터필드의 구조와 블록구분은 USMARC와 KORMARC가 전통적인 목록의 전자카드형식을 따르고 있고, UNIMARC는 GARE의 일반적 구조를 따르고 있어, 차이를 보이고 있다. 고정 길이필드의 레코드의 내용에는 USMARC와 KORMARC의 경우에 표목의 선정과 상태 및 범위, 총서에 관련된 사항이 많이 포함되어 있고, UNIMARC에는 국제적 상호교환에 관련된 요소들이 추가되어 있다. 가변길이필드의 비교분석을 통하여 KORMARC의 한자데이터 처리와 관련하여 별도의 서브필드(▼j)를 설정할 것과, 총서에 대한 청구기호와는 별도의 분류기호필드를 설정할 것을 제안하였다.

ABSTRACT

This article analyzes the record structure, content designation, and the content of UNIMARC, USMARC, and KORMARC formats for authorities, through the comparative investigation. Structure and content designation are almost same with those of the bibliographic formats, being based on those of ISO 2709. The data fields of USMARC and KORMARC are divided into blocks based on the traditional authority card formats, and those of UNIMARC are divided into functional blocks based on the GARE. Record contents of the formers in the fixed-length fields include more elements on the selection, status and scope of the heading, and those related to the series. And those of the later include more elements for the international exchange. Based on the analysis of the variable fields, it is recommended that KORMARC should include an additional subfield, say ▼j for the processing of the Hanja(Chinese character) data and add the separate classification fields for series.

* 이 논문은 1993년도 계명대학교 비사연구비의 지원에 의해 연구되었음.

** 계명대학교 문헌정보학과 조교수

1. 서 언

목록에 있어서 검색의 실마리가 되는 표목 또는 접근점의 통일성을 보장하고 일관성을 유지하기 위해서는 전거통제(authority control)가 필요하게 된다. 그럼에도 불구하고 대부분의 도서관의 경우 이러한 전거통제는 전통적인 목록에서는 말뿐인 서비스(lip service)에 불과했던 것이 사실이다. 그러나 목록의 자동화와 더불어 컴퓨터를 이용한 전거업무의 개선에 대한 기대가 높아지면서 전거통제에 대한 관심도 고조되고 있다.

한편 목록의 자동화에 있어서 전거통제의 기본이 되는 전거레코드를 작성하기 위해서는 서지용포맷과는 별도의 전거용포맷이 필요하게 된다. 본고에서는 이와 같은 전거용 MARC 포맷에 대해 MARC 포맷의 원조가 되고 있는 USMARC와, IFLA에 의해 국제표준포맷을 지향하여 개발된 UNIMARC, 현재는 보고서 단계에 있는 KORMARC를 중심으로, MARC 포맷의 3요소를 이루는 포맷의 구조와 내용표지법, 레코드에 대한 비교 분석해 보고자 한다. 그러나 전거통제에 대한 이론적 배경과 역사에 대해서는 이미 선행연구가 이루어져 있으므로 본고에서는 논외로 한다(오동근, 1994a; 최석두). 아울러 KORMARC 전거용 포맷이 거의 전적으로 USMARC 포맷을 바탕으로 개발되었다는 점에서, 특히 UNIMARC의 특징적인 부분에 대해서도 상세히 분석해 보고자 한다. UNIMARC의 전거용포맷은 국내에 구체적으로 소개된 바가 없다는 점과, KORMARC 서지용포맷에서 이

미 UNIMARC 포맷의 상당부분을 도입하고 있다는 점에서, 이러한 분석이 KORMARC 전거용포맷의 개선에 도움이 되길 기대한다.

2. 전거용포맷의 구조와 내용표지법

기계가독형식의 서지데이터를 하나의 응용 프로그램으로 처리하기 위한 기본단위가 되는 레코드(record)는 도서목록의 한 저록(entry)이나 카드목록의 한 세트의 목록카드에 해당한다고 할 수 있다(오동근, 1991, p.20). 레코드의 물리적 구조(physical structure)는 이러한 레코드의 기본이 되는 기계적인 틀로서, 데이터요소가 담겨지는 용기(container)나 운반체에 비유될 수 있는 부분이다. 따라서 이 구조에 따라 전거데이터가 컴퓨터의 기억장치에 배열됨은 물론이다.

주지하는 바와 같이, 이러한 레코드의 구조에 관한 국제표준으로 ISO 2709 「자기테이프에 의한 서지정보교환용포맷」이 마련되어 있어, 대부분의 MARC 포맷에서 이를 따르고 있다. UNIMARC와 USMARC, KORMARC의 전거용포맷에서도 서지용포맷의 경우와 마찬가지로 모두 이 표준을 준용하고 있으므로, 동일한 구조를 갖고 있다. 다만 그 용어에 있어서는 서로 차이를 보이고 있다. 이하에서는 전거용포맷의 특징적인 부분만을 구체적으로 분석해 보고자 한다.

2.1 레코드레이블과 디렉토리

우선 레코드레이블은 USMARC와 KOR-

MARC에서는 리더(leader)에 해당하는 고정 길이필드 부분이다. 이는 레코드 전체를 이용하기에 앞서 필요한 예비적인 정보를 식별하는데 도움이 될 수 있는 정보를 일차적으로 알려주는 도입부분으로, 세 포맷 모두 24자의 고정 길이필드로 구성되어 있다. 그 구조와 내용은 서지용포맷과 거의 대동소이하다. 다만 레코드의 유형(type of record)에서 UNIMARC는 GARE(Guidelines for Authority and Reference Entries)의 예에 따라 전거엔트리 레코드(x)와 참조엔트리 레코드(y), 일반해설엔트리 레코드(z)의 세 종류로 구분하고 있으나, USMARC와 KORMARC는 전거데이터(z) 하나로만 표시하고 전거레코드의 구체적인 종류는 008 필드의 09 자리에서 표시하도록 하고 있어 차이를 보일 뿐이다. 아울러 서지용포맷에 설정되어 있는 서지적 레벨과 서지계층레벨, 목록 기술형식은 전거통제와는 무관하므로 공백으로 처리되는 물론이다.

다음으로 '도서의 목차' (Library of Congress, 1974. p.4)에 해당하는 디렉토리(directory)는 세 포맷 모두 서지용포맷과 마찬가지로, 동일한 구조와 내용으로 구성되어 있다.

2. 2. 데이터필드

데이터필드(data field)는 레코드의 핵심을 이루는 전거데이터를 기술하는 부분으로, 그 구조와 내용은 각 포맷의 호환성에 절대적인 영향을 미치게 된다. 세 포맷의 데이터필드는 서지용포맷과 마찬가지로, 표시기호(tag)에 따라 기능별로 구분이 가능한 기능별블록으로 구

분할 수 있다. 우선 USMARC와 KORMARC에서는 가변길이데이터필드를 MARC 특유의 데이터요소들을 0XX 블록에 앞세우고, 이어서 대체로 전통적인 목록의 전거카드형식에 따라 표목에 대한 블록(1XX), 참조와 표목지시에 대한 블록(2XX-5XX), 총서처리에 대한 블록(64X), 주기에 대한 블록(66X-68X)으로 구분하고 있는데, 이를 살펴보면 다음과 같다.

- 0XX 제어정보, 식별번호, 분류번호 등
- 1XX 표목(채택표목 및 비채택표목)
- 2XX 복합보라참조
- 3XX 복합도보라참조
- 4XX 로부터보라표목지시
- 5XX 로부터도보라표목지시
- 6XX 총서처리, 주기 등

USMARC와 KORMARC의 경우는 64X에 총서에 관한 별도의 블록을 설정하고 있는 것이 특징이라 할 수 있다. 이 총서전거데이터의 처리에 대해서는 제2장에서 상세히 살펴보고자 한다. 아울러 USMARC와 KORMARC에서는 표목에 관계되는 1XX, 4XX, 5XX 필드의 표시기호는 서지용포맷과 마찬가지로 서로 대등한 의미를 갖도록 배려하고 있다.

한편 UNIMARC에서는 서지용포맷에서와 마찬가지로, 전거레코드의 필드들을 10개의 기능별블록으로 더욱 상세하게 구분하고 있는데, 이는 소위 GARE에서 제시하고 있는 일반적 구조(오동근, 1994b)에 따른 것으로서, 이를 살펴보면 <표 1>과 같다.

〈표 1〉 UNIMARC의 기능별 블록과 GARE의 일반적 구조(전기엔트리)

UNIMARC의 기능별 블록	GARE의 일반적 구조
0 - - 식별블록	
1 - - 코드화정보블록	
2 - - 표목블록	1 전거표목사항
3 - - 정보주기블록	2 정보주기사항
4 - - 보라참조표목지시블록	3 보라참조표목지시사항
5 - - 도보라참조표목지시블록	4 도보라참조표목지시사항
6 - - 분류번호블록	
7 - - 연관표목블록	
8 - - 정보원정보블록	5 편목자주기사항 / 6 정보원사항
9 - - 국가적 사용블록	7 ISADN

표에서도 알 수 있듯이, UNIMARC에서는 GARE의 지침에 따라 전거데이터를 기능별로 구분하여 필드를 설정하고, MARC의 특징적인 블록인 식별블록과 코드화정보블록, 그리고 전거용포맷의 특징적인 블록인 분류번호블록과 연관표목블록을 추가하고 있다.

2. 3 내용표지법

데이터요소를 명시적으로 식별하고 그에 대해 추가의 특성을 부여해주는 내용표지법(content designation)은 일반적으로 서지용포맷과 마찬가지로 표시기호와 지시기호, 서브필드식별기호 등의 내용표지기호(content designator)를 사용하여 이루어지게 된다. 아울러 고정길이필드 및 고정길이 서브필드의 코드화데이터도 데이터요소들의 상대적인 자리수(relative character position)에 따른 별도의 내용표지법을 사용하고 있다. 가변길이필드의 블록별 구분과 내용표지법은 제2절에서 살펴

본 바와 같다. 따라서 전거용포맷의 경우 데이터필드의 블록구분을 제외하고는, 내용표지법상의 특징은 없다고 할 수 있다.

3. 전거용포맷의 레코드 내용

레코드의 내용(content)은 전거용포맷의 핵심을 이루는 요소로서, 각 필드에 기록되는 데이터요소로 이루어지게 된다. 이 데이터요소들이 필드를 구성하게 되고 이 필드들이 다시 하나의 전거레코드를 구성하게 됨은 물론이다. 이하에서는 이를 데이터필드의 고정길이 코드화 데이터요소와 가변길이필드의 데이터요소로 나누어 살펴보기로 한다. 아울러 이 장의 분석의 예는 기본적으로 전거엔트리를 기준으로 한다.

3. 1 고정길이필드의 코드화 데이터요소

데이터필드의 고정길이 코드화 데이터요소는 UNIMARC의 코드화정보블록(1 - -)의

필드의 고정길이서브필드와, USMARC와 KORMARC의 고정길이데이터요소필드(008)의 데이터요소들이 여기에 해당한다고 할 수 있다. UNIMARC의 경우는 이 요소들을 모든 전거레코드에 적용할 수 있는 기본적인 코드화데이터를 담고 있는 일반처리데이터(100), 2 - - 표목필드의 표목이 단체명이나 회의명, 지명 등 고유명일 경우에 그에 관한 추가의 정보를 제공하기 위해 사용되는 고유명에 대한 코드화데이터필드(150), 사용된 편목규칙과 주제명시스템을 나타내주는 규칙(152), 2 - - 표목필드의 표목이 통일서명일 경우에 사용되는 통일서명에 대한 코드화데이터필드(154), 표목과 관련된 지역이 있을 경우에 이를 나타내기 위한 지역코드(160)로 구분하여 설정하고 있다. 고정길이 코드화데이터요소에는 전거레코드 전반에 관한 일반적인 코드화정보 또는 1XX 표목이나 4XX/5XX 표목지시필드의 특수한 측면에 관한 코드화정보들이 고정길이로 편성되어 있는데, 이 데이터요소들은 검

색과 데이터관리를 위해 도움이 될 수 있는 요소들이다.

이들 요소에 있어서는 USMARC와 KORMARC가 거의 동일한 반면, UNIMARC는 상당한 차이를 보이고 있다. 우선 세 포맷의 공통적인 데이터요소들을 살펴보면 <표 2>와 같다.

표에서 볼 수 있는 것처럼, 세 포맷은 모두 파일입력일을 제외하고는, 표목의 처리나 선택, 상태에 관한 요소들을 공통적으로 설정하고 있다. 전거표목상태코드나 채택수준코드는 용어만 다를 뿐 동일한 요소이며, UNIMARC의 경우는 다양한 문자를 수용해야 하므로 번자코드를 두고 있는 반면 다른 두 포맷에서는 로마자화표코드만을 두고 있는 것이다. 물론 각 포맷의 구체적인 코드에 있어서는 다소 차이가 있음은 물론이다. 한편 KORMARC에서는 정부기관유형은 필요없는 부분(윤구호 등, 1994, p.13)으로 간주하여 삭제하고, 그 자리에 서지용포맷과 일치시키기 위하여 수정레코드라

<표 2> 고정길이 코드화데이터요소의 공통적 요소

UNIMARC 1-- 블록		UNIMARC와 KORMARC (008필드)	
필드 및 자리수	내 용	자리수	내 용
100/\$a/0-7	파일입력일	00-05	파일입력일
100/\$a/8	전거표목상태코드	33	채택수준코드
100/\$a/12	번자코드	07	로마자화표코드
150/\$a/0	정부기관유형	28	정부기관유형*
152/\$a	편목규칙	10	목록기술규칙코드
152/\$b	주제명시스템	11	주제명표목시스템/시소러스코드
154/\$a/0	총서유형코드	12	총서유형코드
160/\$a	지역코드	필드043	지역코드

* 정부기관유형은 USMARC에만 설정되어 있음.

는 요소를 설정하고 있다. 그러나 KORMARC의 경우도 정부기관에 관련된 표목이나 전거 데이터를 용이하게 모을 수 있도록 하기 위해서는 서지용포맷의 경우 만큼 상세할 필요는 없지만, 이 요소를 살려두는 것이 바람직할 것이다. 이를 위해서는 현재 공백으로 처리되고 있는 008/38를 사용하는 것이 서지용포맷과도 일치한다는 점에서 -- 서지용포맷의 경우는 008/38-39가 한국정부기관부호로 되어 있음 -- 적합하다고 생각한다.

한편 UNIMARC에서는 공통적인 요소 외에 편목언어(100/\$a/9-11), 문자세트(100/\$a/13-16), 부가문자세트(100/\$a/17-20), 편목문자(100/\$a/21-22)와 같은 데이터요소들이 추가로 설정되어 있다. 이러한 예들은 UNIMARC가 국제적 상호교환을 목적으로 하고 있음을 보여주는 예라 할 수 있다.

또한 USMARC와 KORMARC에는 앞서 살펴본 공통적인 데이터요소 이외에 많은 데이터요소들이 포함되어 있다. USMARC와 KORMARC의 데이터요소를 대략 구분해보면, 레코드처리관련사항과, 언어 및 문자관련사항, 표목의 상태, 표목의 범위, 표목의 선정 등 표목관련사항, 총서관련사항 등으로 나눌 수 있을 것이다. 레코드처리관련사항으로는 파일입력일과 편목정보원코드(cataloging source code)를 들 수 있다. 표목관련사항은 당연히 표목이나 접근점의 처리를 위한 포맷인 전거용포맷의 핵심을 이루는 데이터요소가 될 것이다. 표목의 상태에 관련된 요소로는 레코드 종류코드, 표목사용코드들, 참조평가코드, 레코드갱신중, 동명이인, 채택수준코드 등을 들

수 있을 것이다. 표목의 범위와 관련된 요소로는 직접/간접지리세구분코드, 주제세구분유형코드(USMARC의 경우)를 들 수 있을 것이다. 표목의 선정에 관련된 요소는 앞서 제시한 목록기술규칙코드(descriptive cataloging rules code)와 주제명표목시스템/시소러스코드(subject heading system/ thesaurus code)를 들 수 있다. 총서에 관련된 요소로는 총서유형코드와 번호부여/미부여총서코드를 들 수 있다. 한편 KORMARC에서는 서지용포맷과 일치시키기 위해 일부 요소의 자리수를 변경하고 있으며, 한국내 자료를 처리하기 위한 코드를 추가하고 있다(윤구호 등, 1994, pp.11-13).

결국 레코드의 상호교환의 측면에서 볼 때, 고정길이 데이터요소는 일부를 제외하고는 본질적인 차이점을 갖고 있는 것은 아니다. 따라서 포맷 상호간의 호환성을 증진하기 위해서는, 각 코드의 표준화가 필수적이라 하겠다.

3. 2 가변길이필드의 데이터요소

이 절에서는 전거용포맷의 중심이 되는 데이터필드의 가변길이 데이터요소에 대해 살펴보고자 한다. 이 절에서 특히 한국의 전거데이터의 처리와 관련하여 문자와 표기법의 처리 문제를 구체적으로 분석해 보고 아울러 고유명과 주제명, 총서명 전거데이터의 처리에 대해서도 고찰해 보고자 한다.

3. 2. 1 문자와 표기법

문자와 표기법의 처리 문제는 전통적인 목록과 서지용포맷은 물론 전거용포맷에서도 반

드시 고려해야 할 중요한 문제이다. 특히 이 문제는 전통적으로 한자를 사용하는 동양자료의 처리에 있어서는 반드시 고려해야 할 문제 이면서도 항상 충분한 배려를 받지 못하는 문제이기도 하다. 서지용포맷의 경우에는 같은 동양삼국의 경우에도 서로 차이를 보이고 있을 뿐만 아니라, KORMARC의 경우에도 충분한 배려가 부족한 것이 사실이다(Jeong Pilmo and Oh Dong-geun, 1992). 특히 표목은 한글로 표기한다는 한국목록규칙의 원칙에 따라 서지용포맷에서 표목을 한글로만 표기할 경우 수많은 동명이인이나 동의어들이 나타나게 된다. 따라서 이러한 문제점을 해결해야 할 전거용포맷의 경우에는 한자처리의 문제를 더욱 신중하게 고려해야 할 것이다.

3. 2. 1. 1 USMARC의 별도의 그래픽에 의한 표현방법

전거용포맷의 경우, USMARC에서는 원칙적으로 로마자화된 형식의 표목을 채택하게 된다. 따라서 앞서 살펴본 것처럼 로마자화에 사용한 방식을 나타내기 위한 코드(008/07)를 고정길이필드에 마련하고 있는 것이다.

그러나 현재 USMARC의 전거용포맷에서는 서지용포맷에만 설정되어 있던 소위 880 필드, 즉 'Alternate Graphic Representation' 필드를 새로이 도입하여, 별도의 문자를 그래픽으로 나타낼 수 있도록 하고 있다(USMARC Authority, May 1993. pp.880-p.1-880-p.3). 이와 관련하여 066 필드, 'Character Set Present' 에 그래픽 문자가 포함되어 있는지의 여부가 표시됨은 물론이다. 그 구조는 서지용

포맷과 동일한데(서지용포맷에 대해서는 오동근, 1991. pp.142-143 참조), '김소월'의 예를 들어 도해하면 다음과 같다:

```
066 bb †CS1
100 1b †6880-01 †aKim, So-Wol
400 1b †6880-02 †aKim, Jeong-Shik
880 1b †6100-01/ †a김소월
880 1b †6400-02/ †a김정식
```

여기서 서브필드 †6는 두 필드를 서로 연결시키기 위한 서브필드이며, 표시기호 다음에 나타나는 '-01'과 '-02'와 같은 번호는 출현 번호(occurrence number)이다. 아울러 이 때는 다음과 같은 두 개의 상호참조가 작성될 것이다(USMARC. 880-p.3):

김소월

search under: Kim, So-Wol

김정식

search under: Kim, So-Wol

3. 2. 1. 2 UNIMARC의 연관표목필드

국제적인 이용을 전제로 하여 작성된 UNIMARC의 경우는 고정길이필드는 물론 가변길이 필드에도 언어나 문자를 처리하기 위해 상당부분의 배려를 하고 있다. 즉 UNIMARC에서는 로마자화를 포함한 번자(transliteration) 방식을 나타내기 위한 번자코드(100/\$a/12)와, 편목에 사용된 언어를 표시하기 위한 편목언어(100/\$a/9-11), 사용된 문자를 나타내기 위한 편목문자(100/\$a/21-22)의 고정길이 데

이터요소를 설정하고 있고, 아울러 대등하게 사용되는 다른 언어나, 문자 또는 표기법을 가변길이필드에서 처리하기 위한 지침을 별도로 마련하고 있다.

우선 다른 언어로 된 형식을 바탕으로 하는 대등표목(parallel heading) -- 예를 들면 불어로 된 표목 "Bibliothèque nationale du Canada"와 영어로 된 표목 "National Library of Canada"가 있는데, 국내의 경우에도 한글 표목과 영어표목의 문제가 이에 해당한다고 할 수 있을 것이다 -- 의 처리는 두가지 방법을 제시하고 있다. 첫째는 서로 다른 언어로 된 대등표목을 각각 표목으로 하는 별도의 전거엔트리를 작성하는 방법이다(UNIMARC, pp.22-23). 즉 불어표목과 영어표목을 표목으로 하는 별도의 전거엔트리를 작성하는 것이다. 이 때 두 표목은 7 - - 연관표목블록(linking heading block)을 통하여 연결시킬 수 있다. 연관표목블록은 서지용포맷의 연관저록필드와 유사한 성격을 갖는 블록으로서, 이 블록에는 표목블록의 표목과는 다른 언어나 문자로 된 형식들이 수록된다(UNIMARC, pp.69-73). 대등표목의 예를 들어 도해하면 다음과 같다:

레코드 1 (한국어 레코드)

001 11111
 210 02\$a국립중앙도서관
 710 02\$322222\$8eng\$aThe National
 Library of Korea

레코드 2 (영어 레코드)

001 22222
 210 02\$aThe National Library of Korea

710 02\$3111111\$8kor\$a국립중앙도서관

여기서 \$3는 관련레코드의 레코드번호를 식별하기 위한 서브필드이며, \$8은 해당표목이 2 - - 레코드의 표목과 다를 때 그 편목언어를 식별하기 위한 서브필드이다.

둘째는 2 - - 표목의 대등형식을 다른 이형(異形)이나 참조와 마찬가지로 단순한 이형이나 참조로 취급하는 방법이다(UNIMARC, pp.23-24). 앞서의 예를 이에 따라 도해하면 다음과 같다:

210 02\$a국립중앙도서관
 410 02\$aThe National Library of Korea

이 때 전거표목의 결정은 채택기관의 관례에 따르게 되는데, 한글표목을 원칙으로 하는 한국목록규칙에 따를 경우에는 당연히 한글표목이 전거표목이 될 것이다.

다음으로 별도의 문자로 된 표목(alternative script heading)의 처리문제가 있다. 즉 번자의 경우나 한국이나 일본과 같이 별도의 표기법을 가지고 있는 경우에 이들을 처리해야 하는 문제가 생기게 된다. UNIMARC에서는 이 경우에도 두가지 방식을 제안하고 있다. 첫째는 100필드에 표시된 문자의 형식으로 된 전거표목(2 - -)에 대한 별도의 표기법을 2 - - 필드를 반복해서 사용하여 추가의 표목필드를 설정하여 기록하고, 100필드에 표시된 문자와 다른 문자는 서브필드 \$7을 사용하여 표시하도록 하는 방법(UNIMARC, p.46)이다. 이를 일본어 데이터를 예로 들어 설명하면 다음과 같다:

100 \$a/21-22 = dc
 210 b0\$akouzu, harushige
 210 b0\$7db\$a高津春繁

001 44444
 200 b0\$akouzu, harushige
 700 b0\$333333\$7dc\$a高津春繁

이 예에서 “dc”는 일본어가 가나형으로 기록됨을 표시하며, \$7의 “db”는 일본어가 한자형으로 기록됨을 표시하는 것이다.

둘째는 대등표목의 경우와 마찬가지로 별도의 표기법을 별도의 레코드에 기록하고, 7 - - 연관표목블록을 이용하여 연결시키는 방법이다. 이 때 서로 다른 표기법은 7 - - 블록의 서브필드 \$7을 사용하여 표시하게 된다. 이를 예를 들어 설명하면 다음과 같다:

레코드 1 (한자레코드)

001 33333
 200 b0\$a高津春繁
 700 b0\$34444\$7db\$akouzu, harushige

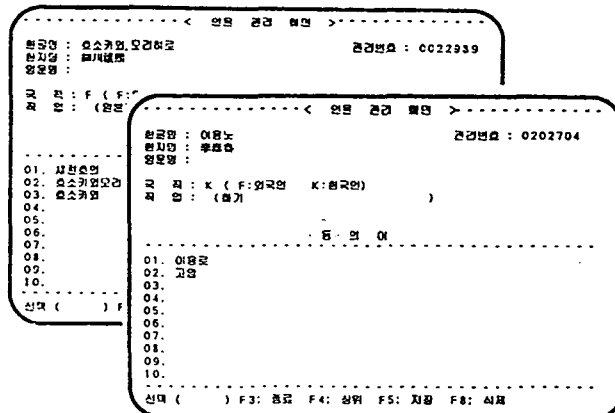
레코드 2 (가나레코드)

3.2.1.3 KORMARC의 처리방법과 개선방안

한편 KORMARC 전거용포맷의 경우는 이와 같은 언어나 문자, 표기법에 대한 배려를 볼 수 없다. 다만 개인명표목(100)의 예에 한자데이터의 처리예가 나타날 뿐이다(윤구호 등, 1993. p.137). 따라서 KORMARC의 경우는 특히 한자표목과 영어표목의 처리에 대한 부분을 보완해야 할 것이다.

우선 한자표기와 관련해서는 다양한 예들이 발견되고 있다. 우선 UNIMARC의 異文字 처리의 첫번째 예와 같이 한글표기와 한자표기를 대등한 표목으로 설정하는 것으로, 시소러스의 예이기는 하지만 중앙일보사의 디스플레이된 예(중앙일보사, p.61)를 들 수 있을 것이다. 이를 예로 들면 <그림 1>과 같다.

<그림 1> 중앙일보 시소러스 인명파일 디스플레이의 예



두번째로는 한자표기를 별도의 언어에 대한 취급의 경우와 마찬가지로, 단순한 이형이나 참조로 취급하는 방식으로, KORMARC의 예(윤구호 등, 1993, p.137)가 이에 해당하는데, 이는 다음과 같다:

100 1b▼a김동리
400 1b▼a金東里

중앙일보사의 방법은 목록의 경우에도 특히 책자형이나 한글과 한자의 자동변환이 가능한 자동화된 시스템의 경우 등에서는 한자형이 그대로 표목으로 사용되는 예를 자주 볼 수 있다는 점에서 타당성을 갖는다고 본다. KORMARC의 예는 한자형으로 목록에 접근하는 경우가 많다는 점에서 역시 타당성을 가질 수 있으리라 본다. 그러나 전거데이터로서의 한자데이터를 처리하는 데 있어서는 다소 미흡한 것으로 생각된다. 즉 한자데이터가 별도의 데이터로 다루어져야 한다는 데는 이견이 없겠지만, 한국에서 사용하는 전거데이터에 나타나는 한자데이터는 별도의 전거데이터라기보다는, 특히 인명의 경우는 생몰년과 함께 전거표목을 한정하는 수식어의 성격이 강하기 때문이다(오동근, 1994b, pp.80-81). 따라서 이와 같은 수식어적 성격을 고려한다면, 별도의 필드보다는 전거표목필드의 서브필드에 한자데이터를 나타낼 수 있는 별도의 서브필드를 설정하는 것이 바람직할 것이다. 이를 위해서는 미사용의 서브필드 ▼j를 동양인명의 한자표기를 나타낼 수 있는 별도의 서브필드를 설정할 수 있으리라 생각한다. 아울러 수식어적

성격을 나타낼 수 있도록 괄호로 묶어 표시하는 것이 바람직할 것이다. 이를 예를 들어 설명하면 다음과 같다:

100 1b▼a김소월▼j(金素月)▼d1902-1934
400 1b▼a김정식▼j(金廷湜)▼d1902-1934

한편 이 때 한자표목을 그대로 사용하는 경우나 한자데이터로 접근하게 되는 경우는 한자표목에 대하여 별도의 파일을 유지함으로써 이를 해결할 수 있을 것이다.

다음으로 영어표목과 관련해서는, 중앙일보사의 예에서는 한자표기와 마찬가지로 한글표기와 대등한 표목으로 설정하고 있고, KORMARC에서는 해당되는 예를 찾아볼 수가 없다. 그러나 영어표목의 경우는 UNIMARC의 두번째 방식, 즉 한글표기를 전거형(1XX)으로 삼고 영문표기를 단순한 이형이나 참조(4XX)로 취급하는 방식을 채택하는 것이 바람직하다고 본다. 그 이유는 국내에서는 동양인명이나 주제명, 서명에 대해 영어표목을 채택하고 있는 도서관이 없을 뿐만 아니라, 기존의 포맷에 어떠한 변경을 가하지 않고서도 가능하기 때문이다.

3. 2. 2 고유명 전거데이터

기본적으로 서지레코드에 대한 표목이나 접근점으로 사용되는 모든 고유명 또는 이름(names)은 전거통제가 필요하게 되며, 따라서 이들에 대한 전거레코드가 작성되어야 한다. 이와 같은 전거레코드가 작성되는 대상을 편목규칙에서는 저자(author)나 책임성(respon-

sibility)이라는 용어로 표현하게 되는데, 이는 일반적으로 어떤 저작의 지적 예술적 내용에 대해 책임을 가지고 있는 개인이나 단체를 가리킨다.

UNIMARC에서는 이를 개인명(X00)과 단체명(X10), 지리명(X15), 가족명(X20) 등의 필드로 구분하고 있고, USMARC와 KOR-MARC에서는 개인명(X00)과 단체명(X10), 회의명(X11), 지리명(X51) 등으로 구분하고 있다. 특히 UNIMARC에서는 고유명이 서명에 포함되어 있을 경우에 대하여, 어떤 저작의 다양한 판들과 총서의 자료들에 대한 레코드를 함께 모으기 위한 고유명/서명필드(X40)와, 다작의 작가의 다양한 저작에 대한 레코드를 함께 모으기 위한 고유명/종합통일서명필드(X45)를 별도로 설정하고 있는 것이 특징이다. 주지하는 바와 같이, 이와 같은 경우 USMARC와 KORMARC의 경우에는 각각 고유명 블록의 고유명 서브필드와 서브필드 †t와 ▼t를 결합하여 나타내게 된다.

지시기호와 식별기호는 서지용포맷과 대동소이하다. 따라서 KORMARC의 경우는 개인명의 경우 묘호와 시호 등을 나타내기 위한 지시기호와, 역조, 한국과 중국의 세계를 나타내기 위한 식별기호와 같이, 동양인명의 특징을 반영할 수 있는 요소들이 추가되어 있다. 이와 같은 요소는 동양인명의 전거통제에 유익한 요소로서, 국제적인 포맷에도 반영되어야 할 것이다.

3. 2. 3 주제명 전거데이터

일반적으로 주제를 통한 검색을 가능하도록

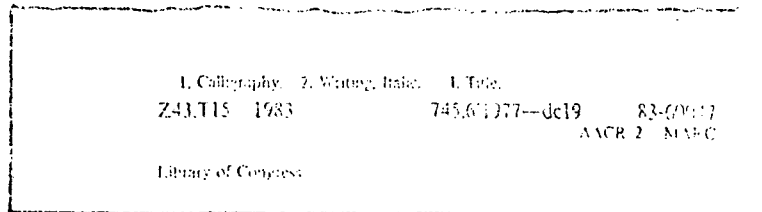
하는 주제목록(subject catalog)은 주제를 나타내는 단어에 따라 알파벳순으로 배열하는 알파벳순 주제명목록과 이를 문헌분류법의 기호법에 따라 분류하는 분류순목록으로 구분된다(DDC 20. p.xxv). 여기서 분류표에 관련된 분류데이터에 대한 전거레코드는 별도로 다루어지는 것이 바람직하며(윤구호등, 1994. p.9), USMARC의 경우는 이미 분류데이터를 처리하기 위한 별도의 포맷이 마련되어 있다.

이렇게 보면 주제목록 가운데 전거용포맷에서 다루게 되는 기본대상은 주제명 데이터가 된다. 이론상으로 주제명전거통제의 기능은 이용자가 사용하는 어휘와 문헌에서 사용하는 어휘를 정확하게 연결시켜 주는 일관성있는 어휘를 사용하도록 하는 것이다(Clack, p.5). 따라서 주제명표목을 통제하기 위해서는 목록에서 실제로 사용하는 어휘와 참조구조가 일관성을 갖도록 해야 한다. 이와 같은 일관성을 보장하기 위한 도구들이 바로 주제명표목표와 시소러스인 것이다. <그림 2>는 주제명표목의 전거통제와 관련하여 서지레코드의 데이터와, 미국의회도서관주제명표목표(LCSH)의 데이터, 전거데이터의 형식을 제시한 것이다.

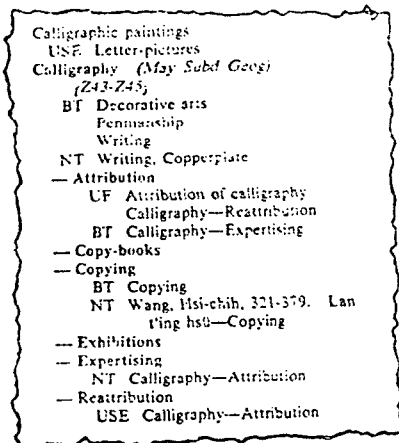
주지하는 바와 같이, 전거용포맷에서는 이와 같은 전거데이터를 나타낼 수 있도록 하기 위하여, 표목과 보라참조표목지시, 도보라참조표목지시와 같은 필드를 설정하여, 통제어와 보조적인 참조를 연결시킬 수 있도록 하고 있다. 아울러 전거용포맷에는 표목과 표목지시 사이의 관계를 나타내주는 “NT”나 “BT” 등의 기호를 생성할 수 있도록 하기 위한 별도의 제어서브필드도 설정되어 있다. 세 포맷 모두 주

〈그림 2〉 주제명전거의 사용예 (Clack, p.173)

(가) 서지레코드



(나) LCSH



(다) 전거데이터

Calligraphy
 BT Penmanship
 Writing

제명표목과 관련되는 4XX 필드의 데이터들은 모두 “UF”로 생성할 수 있을 것이다. “NT”와 “BT”를 생성할 수 있도록 하기 위해서는, USMARC의 경우는 #w의 첫번째 자리, USMARC의 경우는 \$5의 첫번째 자리의 코

드를 이용하면 되는데, 두 포맷 모두 “NT”에 대해서는 “h”, “BT”에 대해서는 “g”의 코드 값을 부여하고 있다. 각각의 경우에 별도의 표시상수(display constant)를 표시할 수 있음은 물론이다.

한편 UNIMARC와 USMARC에서는 주제명표목과 분류번호를 연결시키기 위한 별도의 필드를 설정하고 있다. 분류기호는 이용자를 다른 주제명으로 안내해주지는 않는다는 점에서 보면 상호참조와는 다르다. 그러나 분류기호는 어떤 자료나 주제를 찾을 수 있는 서가상의 위치를 이용자에게 안내해주는 것이다 (Drabenstott and Vizine-Goetz, pp.68-69). 그러므로 전거용포맷의 분류번호필드에 기록되는 분류번호는 주지하는 바와 같이, 서지용포맷의 분류번호는 물론 분류데이터용포맷의 분류번호와도 그 성격이 다른 것이다. 그러나 KORMARC에서는 USMARC에 설정되어 있는 분류기호필드들을 삭제하고 있다. 다만 관련논문(윤구호 등, 1994, p.15)에서는 주제명에 대하여 필드 052(국립중앙도서관 청구기호)에서 분류기호만을 추출할 수 있다고 밝히고 있는 것으로 보아, 청구기호 관련필드를 총서뿐만 아니라 주제명표목에 대해서도 사용하도록 하고 있는 것으로도 추정된다. 그러나 포맷의 설명이나 예에서는 그와 같은 내용이 포함되어 있지 않다. 따라서 KORMARC에도 주제명표목과 해당분류번호를 연결시키기 위한 필드가 추가되어야 할 것이다.

3. 2. 4 통일서명과 총서명 전거데이터

어떤 저작이 둘 이상의 이름으로 나타날 때는 그 저작에 대한 모든 레코드를 한 곳에 모을 수 있도록 하기 위하여 전거통제가 필수적으로 따르게 되는데, 이러한 서명 가운데 전거통제의 대상이 되는 것은 통일서명(uniform title)과 총서명(series)이다.

통일서명의 경우는 UNIMARC와 USMARC 및 KORMARC가 그 취급에 있어서 서로 차이를 보이고 있다. 즉 UNIMARC의 경우는 이를 통일서명(X30)과 종합통일서명(X35), 고유명/서명(X40), 고유명/통일서명(X45) 등의 필드로 세분하고 있으나, USMARC와 KORMARC에서는 고유명과 결합된 통일서명이나 총서명은 해당고유명의 필드에 표시하도록 하고, 그 밖의 통일서명만을 통일서명필드(X30)에 표시하도록 하고 있다.

한편 총서명의 경우는 UNIMARC에서는 고정길이필드의 총서유형코드(154)를 제외하고는 포맷 자체에서 총서명을 처리하기 위한 별도의 배려를 하고 있지 않다. 그러나 이와는 달리 USMARC와 KORMARC에서는 고정길이필드(008)는 물론 가변길이필드(64X)에서도 총서명 전거데이터를 처리하기 위한 상당부분의 필드를 설정하고 있다. 특히 두 포맷에는 총서가 하나의 전집으로 다루어질 경우에 이를 나타낼 수 있도록 하기 위한 별도의 청구기호필드들(USMARC: 050, 060, 070, 090; KORMARC: 050, 052, 09X)이 설정되어 있어, 특정주제의 총서들을 쉽게 찾을 수 있도록 배려하고 있는 것이 특징이다. UNIMARC의 경우는 총서를 통일서명필드에서 다루게 되며, 청구기호에 대한 배려는 나타나지 않는다.

4. 결 언

이상에서는 UNIMARC와 USMARC, KORMARC의 전거용포맷에 대해 그 특징을

중심으로 분석해 보았다. 그 내용을 요약하면 다음과 같다:

전자용포맷의 구조와 내용표지법은 서지용 포맷과 대체로 유사하다. 데이터필드의 구조와 블록구분은 USMARC와 KORMARC가 전통적인 목록의 전자카드형식을 따르고 있고 UNIMARC는 GARE의 일반적 구조를 따르고 있다.

고정길이 코드화 데이터요소에서는 USMARC와 KORMARC의 경우가 UNIMARC에 비해 표목의 선정과 상태 및 범위, 총서에 관련된 사항이 많이 설정되어 있고, UNIMARC에는 국제적 상호교환에 관련된 요소들이 추가되어 있다.

문자와 표기법의 처리에서, 세 포맷은 각기 다른 표시방법을 채택하고 있는데, KORMARC의 경우에는 한자데이터의 처리와 관련하여, 수식어적인 성격을 나타낼 수 있도록 별도의 서브필드(▼j)를 설정하여 괄호로 묶어 표시하도록 하는 것이 바람직할 것이다. 아울러 주제명 전자데이터의 처리와 관련하여, 총서를 표시하기 위한 청구기호필드와는 별도로 분류기호필드를 별도로 설정하도록 하는 것이 바람직할 것이다.

전자통제는 기본적으로 목록의 표목이나 접근점으로 사용되는 개인명이나 단체명 등을 포함한 고유명과, 주제명, 통일서명 및 총서명의 통일성과 일관성을 유지시킴으로써, 목록의 검색기능을 개선시키기 위한 것이다. 전자용 MARC포맷은 이와 같은 전자통제의 결과를 기록하기 위한 형식인 것이다. 따라서 MARC는 전통적인 목록의 요건을 바탕으로 개발되

었고, 컴퓨터를 이용한 전자업무의 처리에는 더 정확한 업무분석과 이해가 필요하다는 점에서 보면, 전통적인 전자업무의 관행이 정립되지 못했고, 표목과 접근점의 선정과 그 형식, 전자통제에 대한 연구가 부족한 우리의 실정에서는 그에 대한 연구와 기반구조의 구축이 선결과제가 될 것이다.

인용문헌

- 國立中央圖書館. 1986. 中國機讀編目權威記錄格式初稿. 臺北, 國立中央圖書館.
- 권수미. 1989. 수작업전자제어와 자동전자제어의 비교연구: 저자명전거를 중심으로. 석사학위논문. 이화여대.
- 도태현. 1991. 자동화목록에서의 전자통제. 도서관학논집 제18집. pp.217- 243.
- 신민정. 1992. 자동전자통제시스템에서의 전자데이터 작성과 이용에 관한 연구. 석사학위논문. 이화여대.
- 오동근. 1991. 편목규칙과 MARC 포맷에 있어서 동양자료의 서지적 기술에 관한 비교분석. 중앙대학교 박사학위논문.
- 오동근. 1994a. 전자통제에 관한 편목이론적 고찰. 도서관 제49권 제3호 pp.31-43.
- 오동근. 1994b. 전자레코드의 작성에 관한 연구. 한국문헌정보학회지 제27집. pp.75-97.
- 윤구호 등. 1993. 전자통제용 KORMARC형식 개발에 관한 연구. 서울, 국립중앙도서관.
- 윤구호 등. 1994. 전자통제용 KORMARC형

- 식의 개발에 관한 연구. 정보관리학회지 제11권 제1호 pp.3-16.
- 이병기. 1990. KORMARC를 위한 자동전자화일의 기초적 설계. 청량정필모박사 화갑기념논문집. pp.681-717.
- 조옥현. 1994. 전자화일의 구조에 관한 연구. 석사학위논문. 청주대.
- 중앙일보사. 1994. 시소러스 사전구축: 제조업 경쟁력강화사업 최종연구보고서 (3차년도). 서울, 체신부.
- 최석두. 1993. 무전자시스템에 관한 연구. 한국 문헌정보학회지 제25집. pp.233-264.
- 현규섭. 1986. 자동화목록법서설. 서울, 아세아 문화사.
- Authority control: The key to tomorrow's catalog: Proceedings of the 1979 Library and Information Technology Association Institutes. 1982. Edited by Mary W. Ghikas. Phoenix, Oriz, Oryx Press.
- Authority control symposium: Papers presented during the 14th Annual ARLIS/NA Conference, New York, N.Y. February 10, 1986. 1987. Edited by Karen Muller. Occasional paper no.6. Tucson, Ariz, Art Libraries Society of North America.
- Bourdon, F. 1993. International cooperation in the field of authority data. Munchen. K.G. Saur.
- Burger, Robert H. 1985. Authority work: The creation, use, maintenance, and evaluation of authority records and files. Littleton, Colorado.
- Clack, Doris Hargrett. 1990. Authority control: Principles, applications, and instructions. Chicago, ALA.
- Drabenstott, K.M. and Vizine-Goetz. 1994. Using subject heading for online retrieval. San Diego. Academic Press.
- Gredley, Ellen and Hopkinson, Allan. 1990. Exchanging bibliographic data. London, LA.(한국어판: 정필모, 현규섭, 오동근 공역. 1993. 서지정보의 상호교류. 서울, 아세아문화사.)
- IFLA. 1984. Guidelines for authority and reference entries. IFLA International Programme for UBC.
- IFLA. 1987. UNIMARC manual. IFLA UBCIM Programme.
- IFLA. 1991. UNIMARC/Authorities: Universal format for authorities. London, K.G. Saur.
- ISO. 1981. Documentation: format for bibliographic information interchange on magnetic tape. ISO 2709-1981. Geneva. ISO.
- Jeong, Pil-Mo and Oh, Dong-Geun. 1991. The influence of pronunciation and scripts in East-Asian MARC formats. In International Cataloguing and Bibliographic Control. Vol.20. No.4. pp.56-60.

Library of Congress. 1974. Information on the MARC system. 4th ed. Washington.

Library of Congress. 1987. USMARC specifications for record structure, character sets, tapes. Washington.

Library of Congress. 1988. USMARC format for bibliographic data including guidelines for content designation. Washington.

Library of Congress. 1991. USMARC format for authority data including guidelines for content designation. Washington.

Oh Dong-Geun. 1995. Comparative analysis of the MARC in Korea, Taiwan and Japan. In Program Vol. 29, No.2. 123-134.

Truitt, Marc. USMARC to UNIMARC / Authorities: A qualitative evaluation of USMARC data elements. In LRTS Vol.36, No.1. pp. 37-58.