

MARC 레코드의 構造와 內容標識法에 관한 比較 및 分析

Comparative Analysis on the Structure of the Record and
the Methods of Content Designation of the MARC

吳 東 根*
(Oh, Dong Keun)

抄 錄

MARC포맷의 기본요소 가운데 레코드의 구조와 내용표지법을 UNIMARC와 USMARC, KORMARC, CHINESE MARC, JAPAN MARC 포맷을 대상으로 비교분석하였다. 레코드의 구조는 기본적으로 모든 포맷이 ISO 2709에 준거하고 있으므로 유사한 구조를 가지고 있다. 고정장데이터요소의 표지방법은 문자기호대입법과 유무판별법, 대표지시자 선행방법, 코드표지방법, 수치대입법, 결합지시자표지법이 사용되고 있다. 가변장필드의 데이터요소를 명시적으로 식별하고 추가의 정보를 제공하기 위해 사용되는 내용표지기호로는 표시자와 지시자, 서브필드식별자가 있으며, JAPAN MARC에서는 지시자가 사용되지 않는다.

ABSTRACT

This study intends to compare the structure of the record and the methods of content designation of the MARC formats based on UNIMARC, USMARC, KOMARC, CHINESE MARC and JAPAN MARC.

The structure of the formats takes the similar form to base on ISO 2709. The methods of content designation of the fixed fields is analyzed into 6 categories. And of the content designers, JAPAN MARC does not use indicator.

I. 緒 言

컴퓨터를 사용하여 어떤 데이터를 처리하기 위해서는 이를 機械가 認識할

* 충남대, 전북대 강사

수 있는 機械可讀形式으로 변환시켜야 한다. 이와 같이 變換된 데이터요소들이 관련된 것들끼리 묶여져 어떤 應用프로그램에 의해 하나의 단위로써 다루어지게 될 때, 이를 레코드(record)라 한다. 레코드에는 각 시스템에서 認識하게 되는 어떤 對象이나 단위에 관한 데이터가 收錄되게 된다.

機械可讀書誌 레코드는 ‘어떤 하나의 문헌에 관계된 것으로 자체의 특유한 論理的 構造로서 기계가독형식으로 貯藏되는 情報의 集合’¹⁾이라 定義할 수 있는데, MARC 레코드는 그와 같은 書誌데이터를 收錄하는 대표적인 예에 속하는 것이다.

따라서 MARC 레코드는 도서목록의 한 著錄이나 카드 目錄의 한 세트의 목록 카드에 해당한다고 할 수 있다. 레코드는 探索과 檢索의 절차에 소요되는 시간을 줄이기 위해서 식별이 가능한 필드(field)와 서브필드(subfield)로 구분하여 처리하게 된다.

本稿는 이와 같은 MARC 레코드를 작성하기 위한 기본이 되는 MARC 포맷의 구성요소 가운데 우선 레코드의 구조와 내용표지방법을 구체적으로 검토하기 위해 시도된 것이다. 이를 위하여 본고에서는 기존의 대표적인 포맷인 UNIMARC²⁾와 USMARC³⁾, 동양 3국의 포맷을 직접적인 대상으로 하여 이를 비교분석하고자 한다.

II. 레코드의 構造

데이터를 상호교환하기 위한 수단으로서 채택되고 있는 서지데이터포맷은 대개 물리적 구조(physical structure)와 內容標識記號(content designator), 내용(content)의 세개 기본구성요소로 이루어진다.

이 가운데 레코드의 構造는 ‘레코드의 기본이 되는 기계적인 틀’(machine framework)⁴⁾이라 할 수 있는데, 서지데이터가 컴퓨터의 기억장치에 배열되는

1) *Reference Manual for Machine Readable Bibliographic Descriptions*, 3rd rev. ed., Paris, UNESCO, 1986, 3. 22

2) *UNIMARC Manual*. London, IFLA UBCIM, 1987.

3) *USMARC Format for Bibliographic Data Including Guidelines for Content Designation*. Washington, D.C., 1988.

4) H.D. Avram, *MARC pilot project: Final report on a project sponsored by the Council on Library Resources, Inc., LC, Washington D. C.* 1968. p. 77.

방식을 규정하게 된다. 이것은 데이터가 담겨지는 容器(container)나 운반체에 비유되는 것으로, 데이터는 레코드에 따라 다를 수 있으나 그 운반체는 항상 일정하게 된다.

레코드의 구조에 있어서는 ISO 2709『磁器테이프에 의한 서지정보교환용 포맷』이라는 국제표준이 마련되어 있다. 이것은 磁器테이프에 의하여 서지정보를 상호교환할 수 있도록 하기 위해 개발된 것으로, 그 구조는 레코드레이블(record label)과 디렉토리(directory), 서지데이터필드, 레코드구분기호(record terminator)로 구성되어 있다.

UNIMARC와 USMARC, KORMARC, CHINESE MARC, JAPAN MARC는 모두 ISO 2709를 바탕으로 하고 있으므로, 기본적으로는 동일한 구조를 가지고 있다. 그러나 사용하는 용어에 있어서는 다소 차이를 보이고 있다.

각 MARC별 기본구조는 <圖 1>과 같다.

레코드의 가운데 내용적으로 중요한 부분은 서지데이터를 수록하고 있는 데이터필드이며, 다른 부분은 그 데이터를 틀리지 않고 전달하기 위한 기계적 조건을 조정하는 역할을 한다.⁵⁾

<圖 1> 주요 MARC 레코드의 기본구조 대비표

USMARC 레코드의 기본구조

리더	디렉토리	제어번호필드	제어필드	데이터필드	레코드구분기호
----	------	--------	------	-------	---------

KORMARC 레코드의 기본구조

리더	레코드디렉토리	제어필드	가변장필드	레코드구분기호
----	---------	------	-------	---------

UNIMARC와 CHINESE MARC, JAPAN MARC 레코드의 기본구조

레코그레이블	디렉토리	가변장필드	레코드구분기호
--------	------	-------	---------

1. 레코드레이블

레코드레이블은 USMARC와 KORMARC에서는 리더(leader)에 해당하는 固

5) 石山洋, 「ジャパン マクの今日的意義とその利用へのガイダンス」, 圖書館雜誌 74(6), 1980. 6. p. 270.

定長필드 부분으로, 시스템에 대하여 그 레코드의 유형을 식별해주는 데이터와 레코드를 처리하기 위해 필요한 정보를 코드화하여 수록하게 된다. 이것은 레코드 전체를 이용하기에 앞서 필요한 예비적인 정보를 식별하는데 도움이 될 수 있도록, '그 레코드의 조작처리상 가장 기본이 되는 요건'⁶⁾을 일차적으로 알려주는 부분이다. 따라서 이 부분은 목록의 정보를 검색하기 위한 것이라기 보다는 정보의 검색을 가능하게 하기 위한 導入部分으로 프로그래밍 기법상의 필수적 요소'⁷⁾인 것이다.

레코드레이블의 구조는 UNIMARC와 USMARC, KORMARC, CHINESE MARC, JAPAN MARC 모두 24자의 고정장필드로 구성되어 있는데, 그 구조와 내용은 <表 1>과 같다.

<表 1>에서 볼 수 있는 것처럼, 레코드레이블의 구조에 있어서는 다섯 포맷이 서로 차이가 없으나, 그 내용에 있어서는 다소 차이를 보이고 있다.

<表 1> 포맷별 레코드레이블의 구조와 내용 대비표

자리수	정보내용	UNIMARC	USMARC	KORMARC	CHINESE	JAPAN
0-4	레코드의 길이	○	○	○	○	○
5	레코드의 상태	○	○	○	○	○
6-9	시행코드	○	×	○	○	○
(6)	레코드유형	○	○	○	○	○
(7)	서지적레벨	○	○	○	○	○
(8)	서지계층레벨	○	×	×	○	×
(9)	공백	○	○	○	○	○
10	지시자의 길이	○	○	○	○	○
11	서브필드식별자 길이	○	○	○	○	○
12-16	데이터기본번지	○	○	○	○	○
17-19	부가레코드정의	○	×	×	○	○
(17)	입력코드레벨	○	○	○	○	×
(18)	목록기술형식	○	○	○	○	○
(19)	연관레코드요건	×	○	×	×	×
20-23	디렉토리 맵	○	○	○	○	○
(20)	필드길이부길이	○	○	○	○	○
(21)	첫자위치부길이	○	○	○	○	○
(22)	시행정의부길이	○	○	×	×	×
(23)	공백	○	○	○	○	○

6) 田村貴代子, "JAPAN/MARCの概要", 「情報管理」, vol.24, no.6, 1981.9. p. 576.

7) 玄圭燮, 「機械可讀目錄法에 있어서의 固定長필드에 관한 연구」, 중앙대학교 대학원 석사학위논문, 1978. p. 52.

레코드의 길이(record length)는 모든 데이터와 내용표지법을 포함하여 레코드 구분기호에 이르기까지의 전체 문자수를 표시하는 부분이다.

레코드의 狀態(record status)는 신규레코드, 수정레코드, 삭제레코드 등과 같이 파일을 유지하기 위하여 레코드의 상태를 표시하는 요소이다. 이외에도 UNIMARC와 CHINESE MARC에서는 이전에 발행된 상위레벨레코드, 불완전한 채 이전에 발행된 출판전 레코드, USMARC에서는 입력코드수준 상향조정, 출판전 입력코드수준 상향조정 등을 표시할 수 있도록 하고 있다. 따라서 KORMARC의 미완성 레코드와 CIP레코드는 레코드의 레벨이 이전의 상태에서 상향조정되었음을 표시하는 것으로서, 그 명칭을 그에 맞도록 변경해야 할 것이다.

旅行코드(implementation code)는 ISO 2709에 코드에 대한 별도의 규정이 포함되어 있지 않기 때문에 시행코드라는 명칭을 갖게 되었다. 그러나 USMARC에서는 이 명칭을 사용하지 않고 있고, JAPAN MARC에서는 이를 서지적 상황코드라 부르고 있다. 우선 레코드의 유형(type of record)은 ‘자료의 知的 特性과 물리적 표현형태’⁸⁾에 의하여 레코드에서 기술되는 자료의 유형을 표시하는 요소이다. JAPAN MARC에서는 언어자료의 인쇄물과 마이크로형태자료만을 표시하도록 하고 있으나, USMARC와 UNIMARC, KORMARC, CHINESE MARC에서는 그밖에도 공문서 및 필사본, 인쇄본악보, 필사본악보, 인쇄본지도, 필사본지도, 녹화자료, 음악용녹음자료, 음악의 녹음자료, 그래픽자료, 컴퓨터파일, 다매체자료, 입체자료, 특수자료 등을 표시하도록 하고 있고, CHINESE MARC에서는 특히 중국善本도서와 拓本을 표시할 수 있도록 하고 있다. 이 코드는 한자리로서 그 레코드에서 다루고 있는 자료에 관한 정보를 제공하기 위한 것이므로, 단행본도서만을 다루는 포맷의 경우에는 그에 대한 정보에만 국한해도 무방할 것이며, 동양자료의 경우 古書資料를 포함할 경우는 CHINESE MARC에서와 같이 이를 별도로 처리하는 것이 바람직할 것이다.

書誌的 레벨(bibliographic level)은 ‘자료의 출판방식이나 그밖의 레코드 처리범주를 식별’⁹⁾해주는 요소이다. JAPAN MARC에서는 단행본만을 표시하며,

8) John C. Attig, "The concept of a MARC", *Information technology and libraries*, vol.2, no. 1, 1983. 3, p. 11.(吳東根 역, "MARC format의 概念", 「國會圖書館報」, 제24권 제5호, 1987. 9~10, pp.35-48)

9) *Loc. cit.*, P. 11.

그밖의 포맷에서는 그밖에도 分出(analytic) 또는 구성요소(component part), 연속간행물, 전집까지를 표시할 수 있도록 하고 있다.

書誌階層레벨(hierarchical level)은 총서내의 단행본과 같이, 서지적으로 계층 관계에 있는 자료에 대하여 계층구조 내에서 그 레코드가 갖는 상관적 위치를 지시하기 위한 것으로서, UNIMARC와 CHINESE MARC에만 설정되어 있다.

이와 같은 旅行코드들은 '자동화된 처리시스템에 의하여 레코드들을 여러범 주로 분류하기 위한 방법을 제공하기 위하여'¹⁰⁾ 설정되는 것으로서, 레코드의 유형 또는 레코드에서 기술되는 자료의 유형을 식별하고, 레코드의 처리를 보조해주는 기능을 갖게 된다.

指示字의 길이(indicator length)는 지시자의 글자수를 표시하는 요소로서, JAPAN MARC에서는 사용되지 않으므로 항상 '0'이 되고, 다른 포맷에서는 제1지시자와 제2지시자가 사용되므로 항상 '2'가 된다.

識別字의 길이(subfield identifier length)는 서브필드식별자의 글자수를 표시하는 요소로서, 다른 모든 포맷이 경계구분 기호와 별도의 데이터요소식별자의 두글자로 이루어지므로 항상 '2'가 되지만, JAPAN MARC는 여기에 데이터부의 길이를 나타내는 세글자와 데이터부의 문자모드를 나타내는 한글자가 추가되어 항상 '6'이 된다.

데이터의 基本番地(base address of data)는 입력된 레코드의 실질적인 데이터가 시작하는 위치를 표시하는 요소이다. 이 기본번지는 레코드의 맨앞에서부터의 자리수로서, 결국 레코드레이블과 디렉토리, 필드구분기호의 글자수를 합산한 숫자와 동일하게 된다.

附加레코드定義(additional record information)는 레코드의 처리에 필요한 추가사항을 제시하는 것으로서, USMARC와 KOMARC에서는 이 용어를 사용하지 않고 있다.

입력코드레벨(encoding level)은 입력레코드의 완전성 정도를 표시하는 요소로서, KOMARC에서는 완전레벨과 불완전레벨, CHINESE MARC에서는 완전레벨과 하위레벨 1,2; UNIMARC에서는 완전레벨과 하위레벨 1, 2, 3; USMARC DP에서는 완전레벨고 완전레벨(자료미검토), 완전레벨 이하(자료미검토), 부분(예비)레벨, 최소레벨, 출판레벨, 미확인, 미사용 등으로 세분하고 있다.

10) *Ibid.*, p. 10.

目錄記述形式(descriptive cataloging form)은 레코드에서 사용된 목록기술의 형식을 표시하는 요소이다. JAPAN MARC에서는 ISBD형식에 의한 것만을 표시하도록 하고 있고; UNIMARC와 KORMARC, CHINESE MARC에서는 완전ISBD형식, 불완전 ISBD형식, 비ISBD형식으로 구분하여 표시하도록 하고 있고; USMARC에서는 완전ISBD형식, 비ISBD형식, AACR2형식으로 구분하여 표시하도록 하고 있다.

聯關레코드要件(linked record requirement)은 USMARC에만 규정되어 있는 것으로서, 실제의 관련레코드를 접근하지 않고서도 기본적인 식별정보를 수록하고 있는 주기를 연관저록필드(76X-78X)에서 얻을 수 있는지의 여부를 표시하게 된다.

디렉토리맵(directory map)은 각 디렉토리엔트리의 '필드의 길이부의 길이' '첫자위치부의 길이' 등을 표시하는 요소이다. 필드의 길이부의 길이는 디렉토리에서 데이터필드의 길이를 표시하는 부분의 글자수를 나타내는 요소로서, 모든 포맷이 네글자로 되어있기 때문에 항상 '4'가 된다. 첫자위치부의 길이는 디렉토리에서 데이터필드의 첫자 위치를 표시하는 부분의 글자수를 나타내는 요소로서, 모든 포맷이 다섯글자로 되어있기 때문에 항상 '5'가 된다. 旅行定義部の 길이는 각 디렉토리의 시행정의부의 길이를 표시하는 요소로서, UNIMA 77Cdhk USMARC에만 정의되어 있으나 사용은 되고 있지 않다.

한편 레코드의 맨마지막에는 ISO 2709에 따라 필드구분기호(field terminator)를 사용하지 않는다.

이상에서 살펴본 바와 같이, 레코드레이블은 레코드에 대한 특정의 정보를 제공해주는 부분으로서, 정보검색을 위한 도입부분이다. 그 구조에 있어서는 모든 포맷이 동일하지만, 用語에 있어서 다소 차이는 보이고 있을 뿐이다. 대부분의 데이터요소도 모든 포맷에 공통적으로 제시되어 있다. 다만 UNIMARC와 CHINESE MARC에서는 서지적 계층관계에 있는 자료의 상관관계를 나타낼 수 있도록 추가의 배려를 하고 있는 점이 특징이라 할 수 있다.

2. 디렉토리

디렉토리(directory)는 '圖書의 目次'¹¹⁾에 비유될 수 있는 부분으로서,

11)

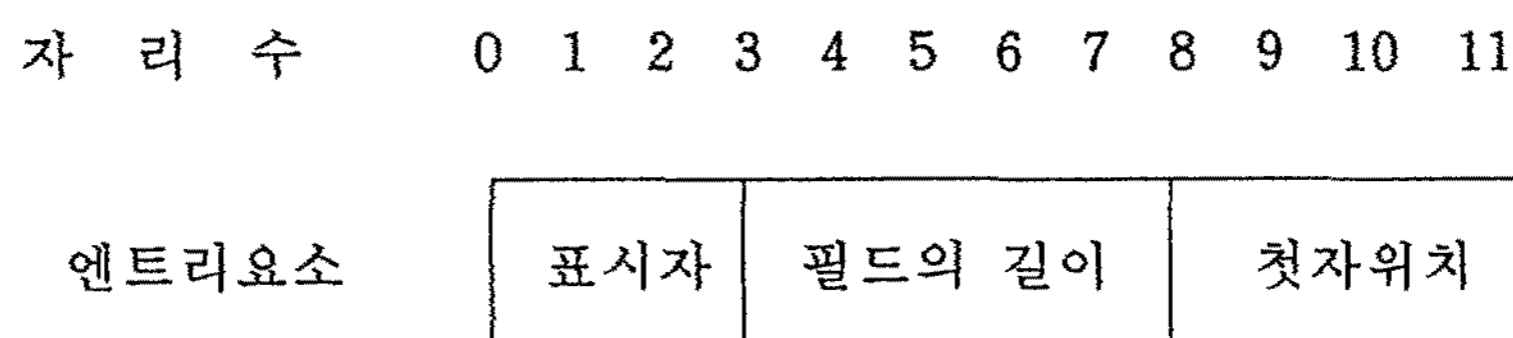
KORMARC에 있어서는 레코드디렉토리에 해당한다.

디렉토리는 레코드를 구성하고 있는 각 필드와 그 필드의 레코드 내에서의 위치를 표시해 줌으로써, 각 데이터요소의 위치를 예상하여 표시자(tag)별로 신속하게 검색할 수 있도록 하기 위해 편성된다. 즉 디렉토리를 설정해줌으로써 情報檢索시에 찾고자 하는 데이터요소에 접근하기 위해서 가변장필드의 모든 데이터 요소를 일일이 접근하지 않고서도, 표시자를 매개로 하여 디렉토리에 제시된 위치로 접근할 수 있기 때문에 정보검색시간을 단축시켜 줄 수 있는 것이다.

디렉토리는 모든 포맷이 동일한 구조와 내용으로 구성되어 있다. 즉 다섯 포맷 모두 디렉토리는 표시자, 필드의 길이(field length), 첫자위치(starting character position)의 세요소로 이루어지는 여러개의 디렉토리엔트리(directoryentry)로 구성되며, 각 디렉토리엔트리는 항상 12자를 한단위로 편성되는 固定長이 된다. 그러나 이 디렉토리엔트들의 집합으로 이루어지는 디렉토리는 표시자의 수에 따라 달라지기 때문에 可變長이 된다, 디렉토리 내에서 각 디렉토리엔트리는 오름차순에 의해 배열되게 된다.

디렉토리엔트리의 구조는 <圖 2>와 같다.

<圖 2> 디렉토리엔트리의 구조



表示字는 레코드의 해당데이터필드를 표시해주는 부호로서, 현재는 다섯포맷 모두 세자의 아라비아숫자로 표시하고 있다. 이 표시자는 디렉토리에만 포함될 뿐이며, 해당필드의 데이터에는 부가되지 않는다. 필드의 길이는 지시자와 서브필드식별자, 데이터요소, 필드구분기호 등을 포함한 각 필드의 전체길이를 표시한다. 첫자위치는 각필드의 시작위치를 표시해주는 요소이다.

3. 데이터 필드

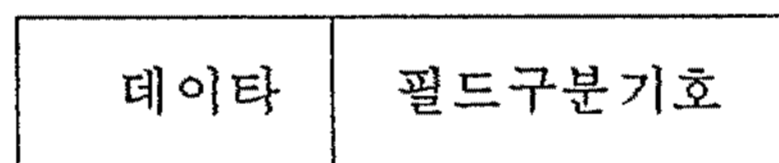
데이터 필드(date field)는 MARC레코드의 핵심을 이루는 부분으로, 그 구조와 내용은 각 포맷의 互換性에 절대적인 영향을 미치게 된다. 데이터 필드는 크게 제어필드(control field)와 기타 가변장 데이터 필드로 구분할 수 있다.

(1) 制 御 필 드

제어필드는 레코드를 처리하기 위해 필요한 정보를 코드화하여 제공하는 고정장으로 이루어지는 부분이다. 이 필드는 UNIMARC와 CHINESE MARC, JAPAN MARC의 경우는 레코드 識別番號필드에 해당한다. 제어필드는 '00'으로 시작되는 표시자를 갖게 된다. 또한 데이터 자체와 필드구분기호만으로 구성되며, 이 필드에서는 지시자나 서브필드식별자가 사용되지 않는다. 제어필드는 디렉토리의 다음에 오게 된다. USMARC와 KORMARC에서는 제어번호필드 외에도, 부호화정보필드 또는 고정장데이터요소(008 필드)를 여기에 포함하고 있으나, 그밖의 포맷에서는 이를 코드화정보블록으로서 가변장필드에 포함시키고 있다. 그러나 이들 포맷의 경우는 코드화정보블록의 서브필드는 고정장으로 이루어지게 된다.

제어필드의 구조는 <圖 3>과 같다.

<圖 3> 제어필드(00X)의 구조



(2) 其他 可變長 데이터 필드

가변장 데이터 필드는 서지데이터나 목록데이터를 수록하게 되는 필드로서, 제어필드의 다음에 오게 된다. 기본적으로 서지레코드는 그 가운데 상당수의 데이터가 가끔씩만 나타나고, 경우에 따라서는 레코드에서 서로 다른 다양한 횟수로 반복되기도 하고, 어떤 데이터요소는 각 레코드에 따라서 길이가 달라질 수도 있기 때문에, 자연히 고정장으로는 적합치 않다.¹²⁾ 그리하여 대부분의 포맷에서는 데이터 필드를 내용표지기호를 부여하여 가변장으로 사용하고 있

12) Ellen Gredley and Alan Hopkinson, *op. cit.*, pp. 49-50.

다. 따라서 가변장 데이터필드는 데이터와 필드구분기호는 물론 지시자와 識別字를 포함하게 된다. 다만 JAPAN MARC의 경우는 ‘일본어 서지정보 처리상 특히 필요가 없었기 때문에’¹³⁾ 지시자를 사용하지 않고 있다.

가변장 데이터필드의 구조는 〈圖 4〉와 같다.

지시자는 JAPAN MARC를 제외한 모든 포맷에서 제1지시자와 제2지시자가 사용되고 있다. 서브필드식별자는 JAPAN MARC를 제외한 다른 포맷에서는 境界區分記號(delimiter)와 데이터 要素識別字(data element identifier)를 사용한다. JAPAN MARC에서는 이 두부호 다음에 데이터부의 길이를 나타내는 가변장 데이터필드의 구조는 〈圖 4〉와 같다.

〈圖 4〉 포맷별 가변장 데이터필드의 구조 대비표

UNIMARC, USMARC, KORMARC, CHINESE MARC의 경우

지시자	서브 필드 식별자	데이터	...	필드구분기호
-----	-----------	-----	-----	--------

JAPAN MARC의 경우

서브 필드 식별자	데이터	...	필드구분기호
-----------	-----	-----	--------

지시자는 JAPAN MARC를 제외한 모든 포맷에서 제1지시자와 제2지시자가 사용되고 있다. 서브필드식별자는 JAPAN MARC를 제외한 다른 포맷에서는 境界區分記號(delimiter)와 데이터 要素識別字(data element identifier)를 사용한다. JAPAN MARC에서는 이 두부호 다음에 데이터부의 길이를 나타내는 3자의 숫자와 데이터부의 모드(mode)를 나타내는 1자의 숫자를 추가하고 있다. 모든 포맷에서는 각 필드내에서 서브필드를 필요에 따라 반복하여 사용할 수 있도록 한다. 이와 같은 각각의 데이터필드는 디렉토리에 수록되어 있는 표시자에 의해 식별되게 된다.

한편 모든 필드의 맨마지막에는 항상 필드구분기호가 부가되며, 레코드의 마지막에는 레코드區分記號(record terminator)가 부가된다.

13) 「JAPAN/MARC マニュアル」, 第3版, 東京, 國立國會圖書館, 1988, p. 5.

Ⅲ. 內容標識方法

내용표지법(content designator)은 '어떤 레코드에 수록되어 있는 데이터요소를 明示적으로 식별하고 추가의 특성을 부여하는 동시에 그 데이터의 조작을 돕기 위해 설정된 여러가지 코드와 관행'¹⁴⁾을 말한다. 따라서 이론상으로 말한다면 전통적인 카드목록에서 사용하는 ISBD의 구뚫점도 이와 같은 내용표지를 위한기호라 할 수 있다.

일반적으로 MARC에서는 앞서 살펴본 표시자와 지시자, 서브필드식별자를 사용하여 내용표지를 하게 된다. 그러나 본고에서는 고정장필드 및 고정장 서브필드의 코드화데이터에 대해서도 그 표지방법을 함께 고찰해보고자 한다.

1. 固定長데이터要素의 內容標識방법

레코드레이블을 포함한 고정장필드 또는 서브필드의 데이터요소들은 상대적인 자리수(relative character position)에 의해 식별되며, 대개는 그 부호값(coded value)에 의해 내용을 나타내게 된다.

이와 같은 고정장 데이터요소를 표지하는 방법은 문자기호대입법과 수치대입법, 유무판별법, 대표지시자선행법, 코드표지방법, 결합지시자표지법 등으로 구분할 수 있다.¹⁵⁾

文字記號代入法은 규정된 문자기호를 해당하는 위치에 표지하여 그 내용을 판별할 수 있도록 하는 방법으로, 레코드레이블에서 레코드의 상태를 표시하는 데있어서 신규레코드는 'n', 개정레코드는 'c', 삭제레코드는 'd' 등으로 기호를 부여하고 이 가운데 해당하는 문자기호를 선택하여 입력하도록 하는 것이 그 예이다. 그 밖에도 레코드레이블의 시행코드의 입력코드레벨, 목록기술형식과, 고정장 데이터요소의 자료형태코드와 목록전거코드, 문학형식코드, 전기코드 등에서 이 방식을 사용하고 있다.

數値代入法은 해당자리에 일정한 수치를 입력하는 방법으로서, 레코드레이블에서 데이터의 기본번지에 해당자리수를 입력하도록 하는 것이 그 예이다.

14) *MARC for library use, 2nd ed: understanding integrated USMARC, op. cit.* pp. 331-332.

15) 玄圭燮, 「機械可讀目錄法에 있어서의 고정장필드에 관한 연구」, *op. cit.*, pp. 59-63.

그밖에도 레코드레이블의 레코드의 길이와 지시자의 길이, 서브필드식별자의 길이, 디렉토리맵과, 고정장 데이터요소의 입력일자 등에서 이 방식을 사용하고 있다.

有無判別法은 검색하고자 하는 정보가 대상자료에 포함되어 있는지의 여부를 판별할 수 있도록 有無관계기호만을 표지하는 방법으로, 고정장 데이터필드의 색인지시자와 회의간행물지시자, 기념논문집지시자 등에서 해당하는 경우는 '1'로 표시하고 해당하지 않는 경우는 '0'으로 표시하는 방법이 여기에 속한다.

代表指示字先行法은 먼저 대표되는 지시자 한자로서 그다음 부속되어 있는 데이터요소의 내용을 표지하고, 그 다음에는 그 내용을 세분시켜 입력하는 방법으로서, 고정장 데이터요소 가운데 출판년을 표시하는 경우에 이방법을 사용하고 있다. 즉 KORMARC의 경우 출판년유형에서 이를 단일년도(s)와 판권연도(c) 등으로 세분하고, 그에 따라서 출판년1과 출판년2의 내용을 결정하여 표시하도록 하는 것이 그 예이다.

코드標識方法은 미리 설정된 코드표에 따라 해당하는 코드를 선정하여 입력하는 방법으로, 고정장 데이터요소 가운데 출판국명코드와 언어코드 등이 여기에 해당한다.

結合指示字標識法은 미리 설정된 다수의 지시자를 결합하여 표지하는 방법으로, 고정장 데이터요소의 삼도부호와 내용형식부호가 여기에 해당한다.

2. 內容標識記號의 分析

내용표지기호(content designator)는 '여러 데이터요소를 식별하고 어떤 데이터요소에 관한 추가의 정보를 제공하기 위한 코드'¹⁶⁾로서, 레코드에 수록되어 있는 서지정보의 유형을 명시적으로 식별하기 위한 문자난 코드¹⁷⁾라고 할 수 있다. 목록카드와 같이 일정한 紙面에 기록하거나 단위카드와 같이 著錄단위로 목록이 구성되어 있는 경우에는, 개개의 데이터요소를 개별화하거나 식별하기 위한 별도의 코드를 부여할 필요가 없었다. 즉 목록카드에 있어서는 목록정보가 默示的으로 표현되어 있는 경우라 하더라도 인간이 이를 눈으로 직

16) UNIMARC manual, IFLA UBCIM, p. 2.

17) R.Renaud, Rasolving conflicts in MARC exchange: the structure and impact of local options, *Information technology and libraries* 3(3)(Sep. 1984), p. 265.

접속하여 각각의 데이터요소의 내용을 식별할 수 있을 뿐만 아니라, 각각의 데이터요소가 시작되고 끝나는 위치를 구별할 수 있는 것이다.

그러나 컴퓨터의 경우는 이와 같은 명시적인 식별기능이 불가능하기 때문에, 각각의 데이터요소를 분명히 구분시켜 주고 각 요소가 시작되고 끝나는 위치를 상세하게 결정해줄 수 있는 수단을 제공함으로써, 각각의 데이터요소가 明示的으로 식별되도록 해야 하는 것이다.¹⁸⁾ 결국 내용표지기호는 이와 같은 명시성을 부여하기 위한 수단이라 할 수 있다.

앞서 살펴본 바와 같이 내용표지기호는 표시자와 지시자, 서브필드식별자로 구성되는데, 다음에는 그 각각에 대하여 구체적으로 살펴보고자 한다.

(1) 表 示 字

표시자는 해당필드에 대한 레이블로서 사용되는 것으로,¹⁹⁾ 다섯 포맷 모두 세자리의 숫자코드르 사용하고 있는데, 필드 자체에는 수록되지 않고 디렉토리엔트리에 슬고된다. 이 표시자에 의해 식별되는 필드는 대개 전통적인 목록의 데이터요소의 집합으로서 구성되게 된다.²⁰⁾

USMARC와 KORMARC에 있어서는 표시자의 첫번째 문자는 전통적인 목록레코드에서의 데이터의 기능, 즉 基本標目과 부출표목, 주제명부출표목 등을 식별하며, 두번째 문자는 개인명과 서명, 단체명 등과 같은 필드의 데이터의 종별을 식별하고, 세번째 문자는 이를 더욱 세분하고 있다. 따라서 이 두 포맷의 가변필드는 표시자의 첫번째 문자에 의하여 블록화될 수 있을 것이다. 다만 KORMARC의 關聯著錄필드의 표시자(451-473)는 760-789로 변경되어야 할 것이다.²¹⁾ 그 이유는 한국문헌자동화목록법 연속간행물용 실험용포맷에서도 이를 760-787필드에 설정하고 있고, USMARC에서도 이를 760-787필드에 설정하고 있을 뿐만 아니라, 현재의 451-473필드는 총서표시블록의 중간에 삽입되어 있어, 부출되지 않거나 변형되어 부출되는 총서명이 연관저록에 포함되게 되는 불합리성을 갖게 되기 때문이다. 이것은 포맷간의 일관성유

18) *Information System Office, LC. MARC manuals used by the Library of Congress, 2nd ed., vol.2, Chicago, ALA, 1970,p.2.*

19) *UNIMARC manual., op. cit, P.4.*

20) 데이터요소의 구성과 표시자와의 관계, 목록규칙과의 관계에 대해서는, 吳東根, "MARA와 목록규칙", 「도서관」, 제46권 제3호, 1991. 5/6. pp.3-23 참조.

21) 吳東根, "MARC 포맷에 관한 비교연구(II)", 「情報管理學會誌」, 제3권 제2호 1986. 12,p. 97.

지는 물론 장차의 통합된 단일포맷을 작성할 경우에 대비한다는 점에서도 정당화될수 있을 것이다.²²⁾

한편 USMARC와 CHINESE MARC, JAPAN MARC에서는 엄격히 구분된 기능별블록에 따라 표시자를 부여하고 있다. 즉 이들 세 포맷은 우선 데이터를 기능별블록으로 나누어 첫번째 문자를 부여하고, 두번째 이후의 문자로 데이터내용의 유형 등을 규정하고 있는 것이다.²³⁾

JAPAN MARC도 블록의 구분은 식별블록과 코드화정보블록, 기술블록, 접근블록으로만 구분하고 있으나, 표시자는 UNIMARC에 따라 부여하고 있다, 다만 0XX필드와 1XX필드를 제외한 시자에는 UNIMARC의 표시자에 50이나 60을 더한 숫자를 사용하고 있다. 이것은 JAPAN MARC의 경우 서명과 저자 표시가 漢字로 표시되어 있는 것이 많고, 이것을 그대로 배열할 수 이가 없기 때문에, ‘質的 差異를 고려하여 다른 기호로 표시하였기 때문이다.’²⁴⁾결과적으로 표시자의 의미가 UNIMARC와 완전히 동일하지는 않다는 점에서 JAPAN MARC의 설계자들은 표시자를 변경시켜 사용하도록 하였던 것이다. 또한 JAPAN MARC에서는 기술블록과 접근점블록의 데이터를 연결시키기 위해, 기술블록의 서명저자사항필드와 접근점블록의 서명의 讀音과 저자명의 독음필드와 표시자 말미 두글자를 동일한 번호로 표시하도록 하고 있다.

(2) 指 示 字

지시자는 ‘필드의 내용이나, 레코드에 수록되어 있는 필드사이의 관계, 어떤 특정데이터의 처리과정에서 필요한 조치에 관한 추가의 정보를 제공하는’²⁵⁾ 기호로서, JAPAN MARC를 제외한 다른포맷에서 모두 사용되고 있다.

지시자는 제1지시자와 제2지시자가 사용되며, 아라비아숫자로 표시된다. 지시자는 각 필드와는 관계없이 독립적으로 규정되며, 그 의미도 독립적으로 해석된다.

지시자의 기능은 크게 데이터처리에 대한 지시기능과 데이터내용의 세분 및 수식기능으로 구분할 수 있다. 데이터처리에 대한 지시는 出力指定에 관한 지

22) *Loc. cit.*

23) UNIMARC에 있어서의 기능별 블록을 포함한 MARC와 목록규칙의 관련성에 관한 포괄적인 고찰은 吳東根, “MARC와 目錄規則”, 「도서관」, 제46권 제3호, 1991. 5-6, pp.3-23 참조.

24) 石山洋, 「ジャパン マク仕様書の概要」, *op. cit.*, p.11.

25) *UNIMARC manual*, *op. cit.*, p. 3.

시와 排列指示로 세분할 수 있고, 내용의 세분 및 수식은 이용가능성표시와 언어지시, 데이터의 품질에 관한 지시, 데이터의 修飾세분 등으로 세분할 수 있다.²⁶⁾

출력지정에 관한 지시는 데이터의 부출여부나 주기에 있어서 標出語의 표시 여부 등을 지시하는 것으로서, KORMARC에서 서명저자표시필드(245)의 제1지시자를 사용하여 서명의 부출여부를 지시하는 것이 그 예이다.²⁷⁾

배열지시는 배열의 순서나 배열에 있어서 무시되는 문자수 등을 지시하는 것으로서, KORMARC에서 일반주기필드(500)의 제2지시자를 사용하여 서명의 첫머리에서 제외되는 문자수를 지시하는 것이 그 예이다.

이용가능성의 표시는 KORMARC에서 국립중앙도서관 청구시호필드(052)의 제1지시자를 사용하여 자료의 소장사항을 지시하는 것이 그 예이다.

言語指示는 자료의 언어나 표기에 사용된 언어를 지시하는 것으로서, KORMARC에서 기본목표의 각 필드의 제2지시자를 사용하여 데이터를 표기하는 데 사용된 언어를 지시하는 것이 그 예이다.

데이터 품질에 관한 지시는 기록의 완전성 등에 대한 지시로서, KORMARC에서 내용주기필드(505)의 제1지시자를 사용하여 내용주기의 완전성을 표시하는 것이 그 예이다.

데이터의 수식세분은 데이터의 내용을 세분하는 것으로서, KORMARC에서 개인명에 관한 각 필드의 제1지시자를 사용하여 개인명을 단일성과 복합성 등으로 세분하는 것이 그 예이다.

한편 JAPAN MARC에서 지시자를 사용하지 않음으로써 포맷을 복잡하게 하는하나의 원인이 되고 있다. 그 대표적인 예로서 서명의 첫머리에서 배열시에 무시되는 冠詞 등이 있을 때에는, 이를 표시하기 위하여 서명저자표시사항(251)에서는 관사를 포함시켜 표시하고, 서명의 讀音필드에서는 관사를 생략하여 표시하고 있는 것이다. 그러므로 JAPAN MARC에서도 앞서 살펴본 기능을 효과적으로 수행할 수 있도록 하기 위해서는 지시자의 사용이 고려되어야 할 것이다.²⁸⁾

26) 丸山昭二郎, “主要MARCフォーマットと目録法-その比較と分析”, 「圖書館學會年報」, vol.27, no. 3. 1981, pp. 106-107.

27) 본절에서는 KORMARC의 대표적인 예만을 제시하고자 한다.

28) JAPAN MARC에 있어서 지시자의 필요성에 대해서는 小田泰正, *et. al.*

“JAPAN/MARCの移用システムとその問題点”, 「圖書館界」, vol.34, no.4, 1983. 1, p. 318; 丸山昭二郎, *op. cit.*, p.109

(3) 識 別 字

서브필드식별자는 가변장필드의 각각의 서브필드를 식별하기 위한 코드로서, 필드내의 데이터를 세분하고 이를 식별하며 부가적인 정보를 부여한다.

서브필드식별자로는 JAPAN MARC를 제외한 다른 포맷에서는 두자의 코드를 사용하며, JAPAN MARC에서는 여기에 서브필드의 데이터부의 길이를 표시하는 세자의 숫자와 데이터부의 모드(mode)를 표시하는 한자의 숫자를 추가한 여섯자의 코드를 사용하고 있다. JAPAN MARC의 데이터부의 길이는 서브필드데이터의 검색을 용이하게 하기 위한 것이며, 데이터부의 모드는 사용된 문자세트가 英數字(alphanumeric)인지 漢字인지를 표시하기 위한 것이다. 특히 JAPAN MARC의 경우에는 최소의 서브필드만을 사용하고 있고, 또한 서브필드의 갯수나 동일항목의 반복횟수, 길이 등이 일정치 않으며, 문자의 표시에 있어서도 영수자모드와 한자모드를 함께 사용하기 때문에, 그 처리를 용이하게 하기 위해 설정된 것이다.

서브필드식별자의 첫번째 코드는 경계구분기호(delimiter)라 하는데, USMARC에서는 ‘#’를 사용하고 다른 포맷에서는 ‘\$’를 사용한다. 한편 데이터 요소 식별자(data element identifier)로 불리는 두번째 코드로는 UNIMARC와 USMARC, CHINESE MARC에서는 알파벳소문자와 아라비아숫자, KORMARC에서는 알파벳소문자, JAPAN MARC에서는 알파벳소문자를 각각 사용한다.

서브필드식별자는 각 필드와 독립적으로 정의되며, 기본적으로 데이터요소를 식별하기 위해 정의되는 것이며 각 요소의 배열순서를 결정하기 위한 것은 아니다. 따라서 데이터요소의 배열순서는 각 포맷이 준거하고 있는 목록규칙에 의해 결정되어야 한다.

(4) 內容標識記號의 互換性

각 포맷에서는 경우에 따라 동일한 기능을 수행하는 데 있어서 서로 다른 내용표지기호를 사용하기도 한다. 이와 같은 호환성이 생기게 되는 근본적인 이유는 기본적으로 목록규칙에 있어서 규정상의 차이나 데이터 세분에 있어서의 차이 등에서 생기는 것이지만, 각 내용표지기호, 즉 표시자와 지시자, 서브필드식별자에 대한 해석상의 차이에도 그 원인이 있다고 할 수 있는데, 그 대표적인 예를 살펴보면 다음과 같다.

우선 표시자와 서브필드식별자가 호환성을 갖는 경우로서, KORMARC에서

는 350필드에 별도로 표시하고 있는 價格을 JAPAN MARC에서는 360필드의 \$B 서브필드에 표시하도록 하고 있는 것이 그 예라 할 수 있다.

두번째는 표시자와 지시자가 호환성을 갖는 경우로서, UNIMARC에서는 602필드에 별도로 표시하도록 하고 있는 家族名을 KORMARC에서는 100필드의 제1지시자를 사용하여 표시하도록 하는 것이 그 예이다.

세번째는 서브필드식별자와 부호가 호환성을 갖는 경우로서, UNIMARC에서는 210필드의 서브필드식별자 \$e와 \$g를 사용하여 표시하도록 하고 있는 인쇄지와 인쇄처를 KORMARC에서는 260필드의 서브필드 \$h에서 ‘:’를 사용하여 구분하도록 하고 있는 것이 그 예이다. 이와 같은 예는 특히 최소의 서브필드만을 규정하고 있는 JAPAN MARC에서 다른 포맷의 서브필드식별자와 동일한 기능을 특정의 부호로서 처리하게 되는 경우가 많다.

이와 같은 차이로 인하여 해당하는 데이터가 없어서 소정의 내용표지기호를 완전히 부여할 수 없을 경우에 대비하여 UNIMARC에서는 소위 補充文字(fillcharacter)를 사용하도록 하고 있다. 그러나 이와 같은 문제를 근본적으로 해결하기 위해서는 내용표지기호에 대한 정의와 사용법에 대한 통일적인 규정이 명확히 설정되어야 할 것이다.

IV. 結 言

이상에서는 MARC포맷의 구성요소 가운데 레코드의 구조와 내용표지법에 대하여 고찰한 바, 그 내용을 요약하면 다음과 같다.

레코드의 구조는 UNIMARC와 USMARC, KORMARC, CHINESE MARC, JAPAN MARC가 모두 기본적으로 국제표준으로서 제정된 ISO 2709『磁器테이프에 의한 情報交換用 포맷』에 준거하고 있기 때문에, 레코드레이블과 디렉토리, 데이터필드, 레코드구분기호를 기본으로 하는 동일한 구조를 갖게 된다. 다만 사용하는 용어와 시행코드의 일부데이터에 있어서 차이를 보일 뿐이다.

상대적인 자리수에 의해 식별되고 대개 그 부호값에 의해 내용을 표지하게 되는고정장데이터요소의 표지방법으로는 문자기호대입법과 유무판별법, 대표지시자선행방법, 코드표지방법, 수치대입법, 결합지시자표지법 등의 방법이 사

용되고 있다.

한편 가변장필드의 데이터요소를 명시적으로 식별하고 그에 대해 추가의 정보를 제공하기 위한 코드로서 사용되는 내용표지기호는 표시자와 지시자, 서브필드식별자가 있으며, JAPAN MARC에서는 지시자를 사용하지 않고 있다. 이와 같은 내용표지 기호의 사용에 있어서는 목록규칙상의 규정의 차이나 데이터 세분상의 차이 등으로 인하여 호환성이 발생하는 경우가 있음을 알 수 있다. 따라서 이를 위해서는 내용표지기호의 정의와 사용에 대한 통일적인 규정을 마련해야 할 것이다.

결과적으로 레코드의 구조와 내용표지법은 서지데이터의 효율적인 처리를 위한 보조적인 부분임을 이상의 분석에서 분명히 알 수 있다. 따라서 MARC에 있어서 레코드의 처리를 위해 가장 핵심이 되는 부분은 레코드의 내용을 이루는 각각의 데이터요소인 것이다. 레코드의 내용은 주로 기록대상서지자료의 성격과 레코드의 용도, 레코드에서 사용하게 되는 각종의 표준에 따라 결정된다.

이와 같은 데이터요소를 처리하기 위한 대표적인 규칙이 목록규칙이라는 점에서, 결국 논의의 초점은 다시 目錄規則으로 돌아갈 수 밖에 없을 것이다.