

컴퓨터 소프트웨어 및 화일들을 위한 온라인 목록 개발시 MARC 형식의 적용 방안

문 기 주*

MARC FORMAT Implementation on the Development of An Online Catalog For Machine-Readable Data Files

Gee-Ju Moon*

Abstract

One of the major problems on the design of an online database for machine readable data file is on the implementation of MARC format for communication with the Library of Congress or OCLC. Most of the cataloging data to make manual card catalogs are stored on magnetic tapes based on the MARC format at LC or OCLC and are sent to local libraries. Therefore, local libraries can avoid the expensive process of cataloging for the books they own. Instead, they can retrieve the necessary cataloging information from the tape and print out manual card catalogs. A problem with MARC is that it is not designed for databases, but for portability to be read at any type of computer. Therefore, it is not practical to use the format on the development of an online database as long as the database is developed in conjunction with a commercial powerful database package. In this paper a possible methodology to resolve the conflicts between the objective of DBMS and MARC is discussed. It is to satisfy the requirements from a commercial DBMS while leaving a room for MARC to communicate with LC and OCLC.

1. Introduction

It is natural that many college and university libraries have software or computer file collections for the use of staff and students since the past

few years[9]. As time goes, it is not surprising at all that the size of collection is increasing rapidly. The software and computer file written in electronic form are generally called machine-readable data file(MRDF). Finding a proper way to handle

* 동아대학교 공과대학 산업공학과 조교수

the collections are as important as printed materials such as books, newspapers, journals, etc. since most of the publications expected to be in this format in the near future. One problem with the increase of MRDF collection volume is the necessity of an online catalog for an efficient search and various display of the desired information. It could be said that the role of an online catalog is more important if the size of collection turns out to be large.

Online catalog helps patron in a number of ways. It allows patron to located materials with any of available information in a short processing time. In addition, it makes possible to display or print records in various formats as desired. One thing must be in mind during the process of an online catalog development is the implementation of Machine-Readable Cataloging(MARC) format.

The purpose of MARC format is to provide cataloging information in machine-readable form in such a way that the record can be manipulated by computer. This will make available a particular type of file in a machine or printed form so that a local library can avoid the expensive process of cataloging, classifying and keyboarding. Therefore, using any other format would run the risk of requiring a conversion to interchange the data with other libraries or to use the data on MARC tapes distributed by the Library of Congress(LC) or Online Computer Library Center(OLCL).

A problem on the implementation of MARC is the fact that it is not designed to be used in an online catalog[2]. In addition, all Data Base Management Systems(DBMSs) require to use their own formats. Therefore, it is not practical to follow the MARC format on the design of online

database. In this paper, a possible methodology of MARC implementation on the development of online catalog is discussed with cataloging practices.

2. Conflicts on MARC and DBMS

Development of catalog entries for MRDF is different from the one for printed materials. Several studies have been done for the application of the Anglo-American Cataloging Rule, 2nd edition (AACR2) so far[5, 9, 12]. In addition to the application of AACR2, the information that patron would use to find the software they wanted and the strategy that will be used to let them access the software are another important questions on the development of the online catalog. In this section a brief history and necessity of MARC format, requirements of local data, and conflicts with DBMS are discussed.

2-1. The history of MARC format

The study for the development of MARC format started in June of 1965 with publishing LC the first draft of a format based on standard LC cataloging practice[1]. In June 1979, the LC Network Development Office, in cooperation with its Automated Systems Office, began to work on compiling a MARC format for MRDF[6]. And a MARC format for MRDFs prepared by a working group of the United States and Canadian librarians during 1980 and 1981[2]. Machine-Readable Form of Bibliographic Information(MARBI) and the USMARC advisory group approved the USMARC format(Underlying principles) on October 29, 1982[ibid.]. Fi-

nally, the format integrated into MFBID Update 9 in 1984[ibid.].

Many online implementation transform USMARC data into a different structure which is better suited to direct access. In some cases, the data structure must suit the needs of a DBMS. No known commercial DBMS can handle USMARC directly[2]. A good comments on the application of USMARC is given by Crawford[ibid.] :

“USMARC was designed as a communications format. Since it was never designed for online use, transformations for such use are natural and proper.”

2-2. Communication and local data

Another point to be discussed is that the necessity of local data in an online catalog for computer files. A computer file needs to have a machine to make itself usable. The role of computer is like the role of paper in a printed material. Not all universities or organizations have the same computer and different computer requires different instructions. Many of local libraries may need to have their own record about system requirements due to the differences on machine type. The method to allow patron access a software or a computer file could be different place by place, too.

The subject-based classification and call number is not helpful for software and computer file collection. Because they can not be shelved like printed materials at library and the way of acces just depends upon the local strategy. If a software or computer file is in needs to be modified to install on a local system, then the information about version, a manual for local version, and additional authors

for editing could be necessary to be altered.

2-3. Search strategy for MRDF

The search strategies for book and MRDF in online catalog might be different. Patrons used to find books or other printed materials by title, name of author, subject, and so on. These information can be obtained easily from reference, bibliography sections of books, or other printed materials. In the case of softwares, manuals do not include reference section on them. Eventually, patrons possibly would try to find the computer file they want based on function, subject, machine type, title of the software, and so on. The differences point out how to decide the indexed elements on the design of online database for MRDF.

Currently, there are good tools such as SPIRES, Oracle, dBASE III, etc. to develop an online catalog. Commercial DBMSs provide excellent retrieval, updating and display functions. However, they require to enter data into their own formatted database. If data must be structured in the physical format of MARC, no known commercial DBMS can be used. The main purpose of MARC format is “communication” while online catalog is “retrieval and display” as desired.

3. DBMS design for MARC

In this section the importance of maintaining one global database, the elements for MARC, framework of database design are discussed.

3-1. One global database

As discussed before, MARC format is necessary

for communication with LC and OCLC while DBMS is necessary for online applications. One way to achieve both purposes is to keep a copy of records in MARC format and keep another copy in a DBMS format for online use. However, this strategy has a serious problem. Keeping two databases eventually turns out to be two different databases if both databases are not properly updated. Another way to achieve the purpose is that keeping only one database in a local DBMS format and using a function of DBMS to rewrite records in MARC format for communication purpose. Neither of the way would be always the best. However, latter could be better than the other since records about computer file in MARC format may have to be cha-

nged frequently to add local data or altered information for installations.

The development of an online catalog for computer files at Iowa State University (ISU) had taken latter strategy. Only one database is kept in a database named SPIRES format for online used and several format files are developed to re-write records as desired. The DBMS is the one developed at Stanford University and runs on VAX computer at ISU. Several other public information systems are developed using the package and run on the same machine. The framework of the database is shown in Fig. 1. The subfile, interfaces, and formats are parts of the system developed through the project.

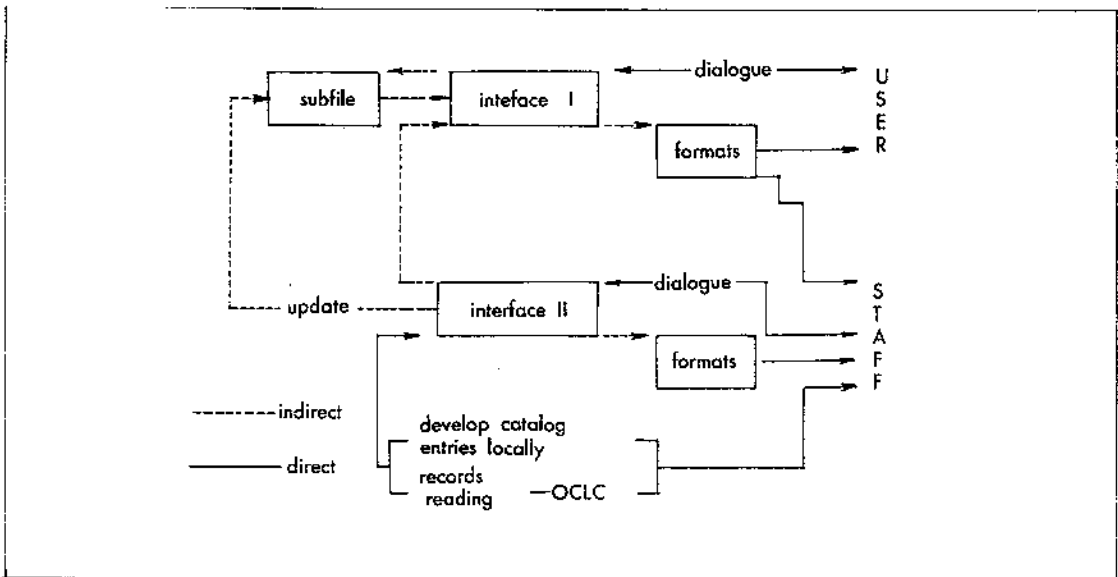


Fig. 1. Framework of database design.

3-2. The elements for search

The indexed elements in an online catalog are provided for search purpose, however, they also

need to be in MARC format. The indexed elements include some unnecessary characters on the search point of view. For instance, the entry in an online catalog "00\$aWord Star" and "Word Star" could

Table 1. Elements in subfile SOFTCAT

Name of elements	Remarks	Name of elements	Remarks
001, CALL_NUMBER, RECORD_ID	used as ID since it is unique index	500, GENERAL_NOTE	general note
000, LEADER	leader	. INDICATOR	indicator
005, TRANSAC	transactions	. NOTE	note
. ADDED_DATE	added date of a software	. REST	other information except above
. ADDED_TIME	added time of a software	TITLE_SRC, SOURCE	source of title
008	008	520, SUMMARY	summary, abstract, or annotation note
245, TITLE_STMT	Title statement	. INDICATOR	indicator
. INDICATOR	indicator	. BRIEF	brief summary
. TITLE	software title	. REST	other information except above
. REST	other information	538, TECH_DETAIL	technical details note
. RESPON_STMT	responsible statement	. INDICATOR	indicator
020, ISBN	ISBN	. REQUIREMENTS	minimum requirements
STD_NUMBER, STD	standard number other than IBSN	. REST	other information except above
040, CAT_SOURCE	cataloging source	590, LOCAL_NOTE	local note
STACK0	tag between 001 and 099 except above	. INDICATOR	indicator
100, MAIN_ENTRY	main entry	. NOTE	note
STACK1	tag between 100 and 199 except above	. REST	other information except above
250, EDITION_STMT	edition statement	STACK5	tag between 500 and 599 except above
. INDICATOR	indicator	690, LOC_HEAD	local subject added entry, topical
. EDITION	edition	. INDICATOR	indicator
. REMAINDER	remainder	. SUBJECT	subject
. REST	other information except above	. REST	other information except above
260, IMPRINT	imprint	653, KEYWORD	subject added entry, uncontrolled
. INDICATOR	indicator	650, TOP_HEAD	subject added entry, topical
. PLACE	place	. INDICATOR	indicator
. PUBLISHER	publisher	. HEADING	heading
. PUBLISHED_DATE	date	. REST	other information except above
. REST	other information except above	STACK6	tag between 600 and 699 except above
STACK2	tag between 200 and 299 except above	700, NAME_ADD	added entry, personal name
300, DESCRIPTION	physical description	710, CORP_ADD	added entry, corporate name
. INDICATOR	indicator	740, TRACED_TIT	added entry, title traced differently
. EXTENT	extent	. INDICATOR	indicator
. DETAILS	other physical details	. KNOWN_AS	also known as
. DIMENSIONS	dimensions	. REST	other information except above
. ACCOMPANY	accompanying material	STACK7	tag between 700 and 799 except above
. REST	other information except above	STACK8	tag between 800 and 899 except above
STACK3	tag between 300 and 399 except above	STACK9	tag between 900 and 999 except above
STACK4	tag between 400 and 499 except above		

be quite different. Only the entry "Word Star" will be matched if a user requests a search for exact "Word Star." Wild card search with the request "Word Star" will retrieve "00\$aWord Star," but it will also retrieve "Word Star Kit" that is not wanted. A way to ignore the indicators, subfields and tags those are important only in MARC could be developed. However, it will be a troublesome with the use of a commercial DBMS.

The elements to be displayed in a special format also make a trouble if indicators, subfields and tags are not separated properly. The "00\$aWord Star" will be troublesome to remove the "00\$a" from the real title part when it is necessary to be displayed in a specific format except MARC. Good examples can be found in the system named SOFTCAT (SOFTware CATalog). The online catalog for MRDF is developed using SPIRES(Stanford Public Information RETrieval System) as mentioned be-

fore. Among the elements in the system, indexed elements and the elements to be displayed in a specific format have separated indicators and subfields.

Explanations are given in Table 1. Element name 'STACKS' are the elements not to be used neither for search nor for display purpose. These are the records only for MARC. Therefore, they do not have separated indicators and subfields.

3-3. The elements for MARC

The elements in the database can be classified into LC or OCLC, common, and local data group as shown in Fig. 2. LC or OCLC group includes the elements which do not need locally but necessary to keep for regeneration of the data in MARC format. Common group includes the elements to be obtained through MARC tape and also useful

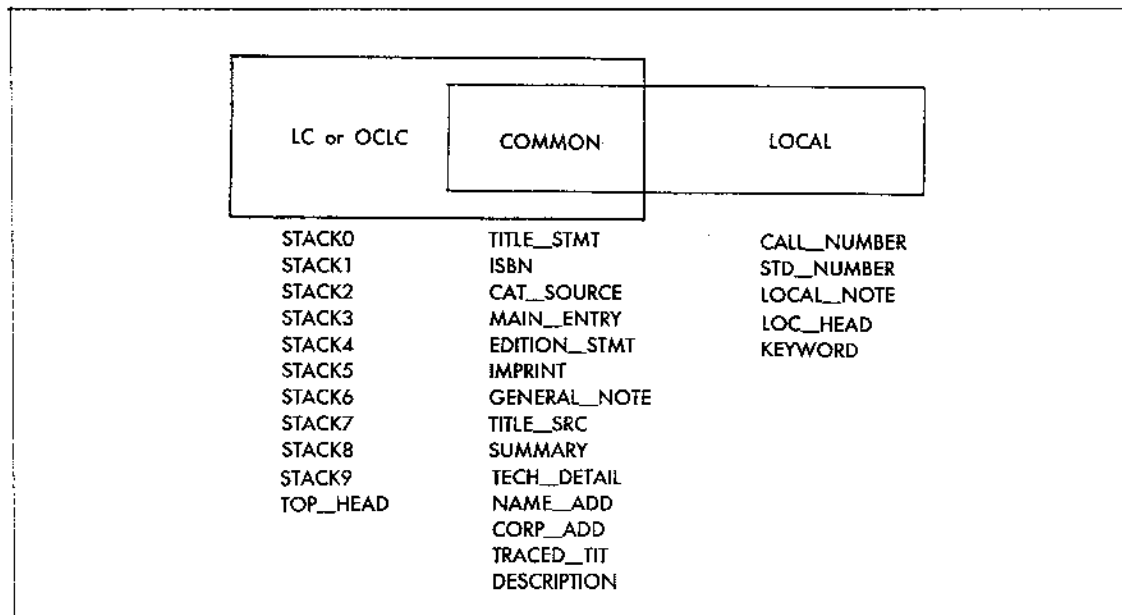


Fig. 2. Elements classification.

```

000=*****nmm 22***** a 4500
001=IPP187 P78.1
005=19871213133240.0
008=870906s1986 waun ua 00010 N/A d
040= $aGXM
100=0 $aBob Wallace
245=00$aPC-WRITE$hcomputer file/$cdeveloped by Bob Wallace.
250= $aVersion 2.70
260= $aSeattle, WA : $bQuicksoft,$cc1986.
500= $aRestrictions on use : Prior permission required.
      Contact MSE dept. Can be copied by anyone to use.
520= $aPC-WRITE is a word processor that lets you enter,
      edit, format and print anything written with letters,
      numbers, and other characters.
538= $aIBM PC or IBM PC compativle ; DOS 2.0 or higher.
590= $aLocation : 229 Engineering Annex
690= $a Productivity Word Processing
653=0 $aText editing printing letters numbers writing word processor
    
```

Fig. 3. An example display with the format MARC.

on the database. Local group contains the elements of those only necessary for local use. As shown in Fig. 2, LC or OCLC records are going to be stored in the elements named STACK. Tags form the beginning upto 099 will be stored in STACK0, 100 upto 199 will be in STACK1, 200 upto 299 will be in STACK2, and so on. This strategy is taken to keep records in the original sequence if they are to be regenerated in the MARC format.

As discussed before, indexed elements are separated from their indicators and other subfield records for good search results. The elements name CALL_NUMBER, ADDED_DATE, KNOWN_AS, TITLE, ISBN, STD_NUMBER, PUBLI-

SHER, RESPON_STMT, PUBLISHED_DATE, REQUIREMENTS, HEADING, and SUBJECT are indexed. Rest of the elements which have separated indicators and subfields are for display purpose with the formats other than MARC.

3-4. Update and display

The function to add records taken form MARC tape into the system directly is not provided yet. The main reason is that no MARC tape for computer software is provided at LC or OCLC ; so it can not be subscribed. Secondly, some of the local

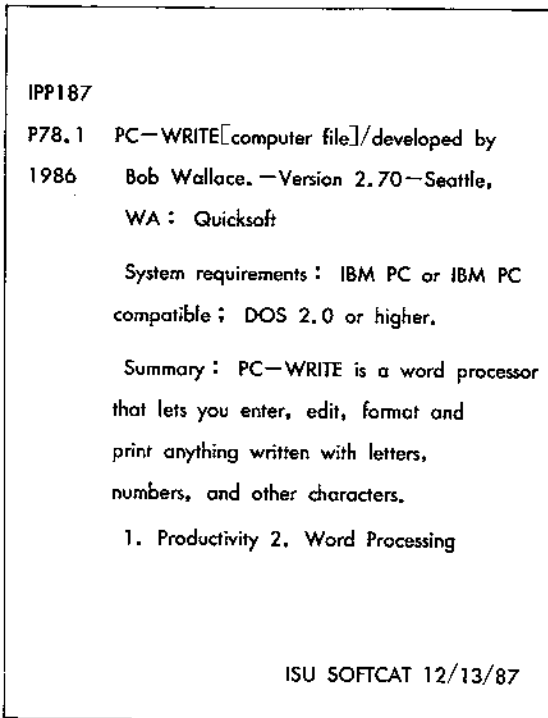


Fig. 4. An example display with format CARD.

information are too important to ignore for search and display purposes in comparison with printed materials. This means that some additions or modifications of the records in MARC format are going to be necessary for local installation. Therefore, no serious attention is given to develop the function reading records from MARC tape and adding them to online catalog directly yet. Currently, a standard format for SPIRES is provided as a worksheet and cataloging information is typed or copied from MRDF to the worksheet. And then information in this worksheet is entered to the online database system directly.

A record stored in the SPIRES format can be displayed in various formats. Two examples are given in Fig. 3 and Fig. 4 for demonstration purpose. The one in Fig. 3 is displayed with the for-

mat file name MARC for MARC tagged display. This is only a tagged display of MARC records, not the original format of MARC as it was on magnetic tape physically. For this, another program is developed in BASIC to read the tagged display record and create data in the original MARC format as on magnetic tape. It could be sent to LC, OCLC, or other local libraries. Another example in Fig. 4 is displayed with the format file named CARD. This format is designed for manual card catalog printing to be used at library if electric power is out and no computer is available.

4. Conclusions and comments

The methodology implemented here for the MARC format implementation on the development of an online database works out well. However, it is not quite clear yet what kind of information is going to be used by patrons to find the MRDF. Those are probably different one from the other. It can be guessed that several elements different from the printed materials are going to be used. This helpful information in the design of online database could be obtained in a function for the statistics like tally sheet is provided in the interface between the database and patrons.

Currently, there is no specialist in this MRDF cataloging field. This is located between librarians and computer workers. The job requires knowledge in the both fields. MRDF cataloging, especially for subject classification and key words development could be a good expert system application. This also indicates that it could be one of the interesting areas for industrial engineers in modern society.

Cataloging MRDF does not have a stable guideli-

nes yet. In addition there are many of unanswered questions on the development of an online catalog for the files. The characteristic of existing in electronic form makes handling the files for cataloging real hard. It allows copy or modification ease. Another important characteristic is that it requires to have physical medium to be stored, machine to be executed, and printed manual to know how to use. These factors will give a big trouble to know where to find proper information for cataloging if all three sources give different cataloging information each other.

The DBMS application as shown here could be one of the possible answers. The DBMS allows to display records in any formats as desired and allows ease to update. This will make practical to follow up frequent computer file updates. Also it could be one of the best ways to give various information to user how to access the file. As discussed before, the way to access depends upon the type of machine and local implementation. Using call number to let a patron access the material would not be possible before any management level decision is made.

References

[1] Avram, H.D., J.F. Knapp, and L.J. Rather. *The MARC II FORMAT—A Communications Format For Bibliographic Data*, Library of Congress, Washington, D.C., 1968.

[2] Crawford, W., *MARC for Library Use : Understanding the USMARC Formats*, Knowledge Industry Publication, Inc., White Plains, NY, 1984.

[3] Crawford, W. et al., *Bibliographic Displays in the Online Catalog*, Knowledge Industry Publica-

tion, Inc., White Plains, NY, 1986.

[4] Demas, S., "Microcomputer Software Collections," *Special Libraries*, Vol. 76, No. 1, pp. 17-23, 1985.

[5] Dodd, S. *Cataloging Machine—Readable Data Files—An Interpretive Manual*, American Library Association, Chicago, 1982.

[6] Dodd, S. and A. Sandberg-Fox, *Cataloging Microcomputer Files—A Manual of Interpretation for AACR2*, American Library Association, Chicago, 1985.

[7] Faust, J., "Microcomputers as On-Line Catalogs in Special Libraries," *Special Libraries*, Vol. 79, No. 3, pp. 133-139, 1986.

[8] Gorman, M., "Microcomputers and Online Catalogs," *Drexel Library Quarterly*, Vol. 20, No. 4, pp. 25-33, 1984.

[9] Gorman, M. (Editor), *The Joint Steering Committee for Revision of AACR. Anglo-American Cataloguing Rules, Second Edition. Chapter 9, Computer Files(draft revision)*, The American Library Association, Cfhicago, 1987.

[10] Piele, L. et al., "Teaching Microcomputer Literacy : New Roles for Academic Librarians," *College & Research Libraries*, Vol. 47, No. 4, pp. 374-377, 1986.

[11] Stanek, D., "Videotapes, Computer Programs, and the Library," *Information Technology and Libraries*, Vol. 5, No. 1, 1986.

[12] Templeton, R. and A. Witten, *Study of Cataloguing Computer Software : Applying AACR2 to Microcomputer Programs. Library and Information Research Report 28*, The British Library Wetherby, West Yorkshire, Great Britain, 1984.

[13] Troutner, J. *The Media Specialist, the Micro Computer, and the Curriculum*, Libraries Unlimited, Inc., Littleton, Colorado, 1983.