1-1-1975

The Lehigh On-Line Cataloging System.

Preston F. Marshall

The Lehigh On-Line Cataloging System

by

Preston F. Marshall, Jr.

A Thesis

Presented to the Graduate Committee

of Lehigh University

in Candidacy for the Degree of

Master of Science

in

Information Science

Lehigh University

1975

ProQuest Number: EP76066

ProQuest EP76066

This thesis is accepted and approved in partial fulfillment of the requirements for the degree of Master of Science.

October 29, 1975
(date)

_____
Professor Donald J. Hillman
Professor in Charge

_____
Professor Robert F. Barnes
Chairman of the Division of
Information Science

ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to Professor Donald J. Hillman for his helpful and valuable suggestions, Lawrence Davis for his assistance and suggestions, Mrs. Edward Farrell for her typing, and most of all his wife, Betsy for her continuing support.

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

This paper describes the data base construction and
on-line operation of the Lehigh MARC cataloging system,
LEMARC.   This system performs automated cataloging of
MARC like records including data base access, record
modification, and communications to a holdings control
system.   In addition each system user can maintain a
private data base of either selected data base records
or user generated records.   The paper includes the
design, internal and user specifications.

1

I. INTRODUCTION

Lehigh University became involved in automated cataloging as the result of two ongoing university activities. The first of these was a program for library mechanization which had already succeeded in automating the circulation control system and acquisitions process. The second was the continuing study into the theoretical framework of information retrieval begun in 1962 by Donald J. Hillman[1] in the Center for Information Science, and which currently involved implementing the LEADER retrieval system within the MART Science and Technology library.[2]

A prototype system was developed in February 1972 to experiment with different ways of using MARC records in both library technical processing and information retrieval systems. For a period of a year this prototype was used, not only by Lehigh, but also by Bucknell University Library, and the staffs of both government and corporate libraries. Based on observations of the usage of the prototype system, a second cataloging system, LEMARC, was developed.

II. SYSTEM REQUIREMENTS

Experimentation with the prototype system showed that a MARC based system could provide significantly more than technical processing service to the library. Typical activities that the prototype system was used for included pre-order searching, inter-library loan verification, and a current awareness service based on titles and authors; the only content indicating fields available in the prototype. A system designed solely for technical processing could not support these requirements; technical processing systems quite often assume that the system user has the work immediately available and can supply the exact form of the search information, for example the title. Thus there exists a single retrieval key for each searchable attribute, and the user can reasonably be expected to supply it.

In many of the applications mentioned this is not a reasonable assumption, as users often have words misplaced or dropped within a title or main entry when they approach a librarian. If a system is to be useful as a verification tool has to be tolerant of this sort of request fault. A second reason for expanding the retrieval capacity to the extent we did was that Lehigh planned to eventually produce a patron access capacity as part of the automation program. It was desired to minimize the degree of system proliferation, and produce a cataloging system whose data base and operations

would be as compatible with the later stages as possible.

While text element retrieval is very desirable for the technical

processing requirements already mentioned, it is essential to

the later development of the patron access system. Therefore

data elements that are likely to be permuted or have words

dropped are searched in "text" mode, in which the only retrieval

requirement is that words that are present in the search also

be present in the data field, with no consideration of order.

Thus incomplete or out of order title or corporate main entry

searches can still be performed successfully. For example any

of the following would retrieve "Compiler Construction for

Digital Computers";

> "Digital Computer Compiler Construction",
>
> "Compiler Construction",
>
> "Digital Compiler Construction".

This approach has certain drawbacks, primarily that there can

exist multiple, and undesired, records for even a search key

that corresponds exactly to the data field.

To allow a more specific search query, both author and

title searches can be combined, retrieving only records that

satisfy both criteria. Results of actual operation have shown

that the combined author-title search is successful, in almost

all cases, in reducing retrieved record lists to a reasonable

length.

Author name keys can also be partially specified, and all

records that satisfy the partial specification can be retrieved.

For example, if JONES is used as a retrieval key, then all Jones records will be retrieved, JONES, A will retrieve those with A Jones as a name, and JONES, AS only AS Jones records. Up to three initials are allowed. This flexibility comes at a certain cost to the user, for it requires additional input and occasionally results in an extra user interaction to screen output. The added user effort has been found to be minor compared to the flexibility obtained.

III. SYSTEM LAYOUT

A. General Layout

The LEMARC system is shown in Fig. 1. Input records
arrive from two sources, either Library of Congress MARC
tapes, or original cataloging performed on the system.
No on-line update is performed. The UPMARC system verifies
the validity of each input record, and does an immediate
update of all key files, after performing the editing
operations described in Appendix E. Record text is
placed in the text file system. Access from key files to
record text is via a document number that is not directly
related to physical location. This was added to provide
more efficient disk storage control during weekly MARC
update cycles, in which 50% of the records are often replace-
ment or deletion transactions.

B. Data Base Structure

The LEMARC data base structure consists of four sets of
data files, controlled through entries in the system control
file, MARCSKL. This file includes both the physical and
logical attributes of the files in the data base, plus the
system definition of each file as a File Definition Block
(FDB). This control file describes a logical combination of
files, and their relationships. All attach and access
information is contained in MARCSKL so no data base character-
istics need be included in the access programs.

INFORMATION FLOW IN LEMARC
Fig. 1
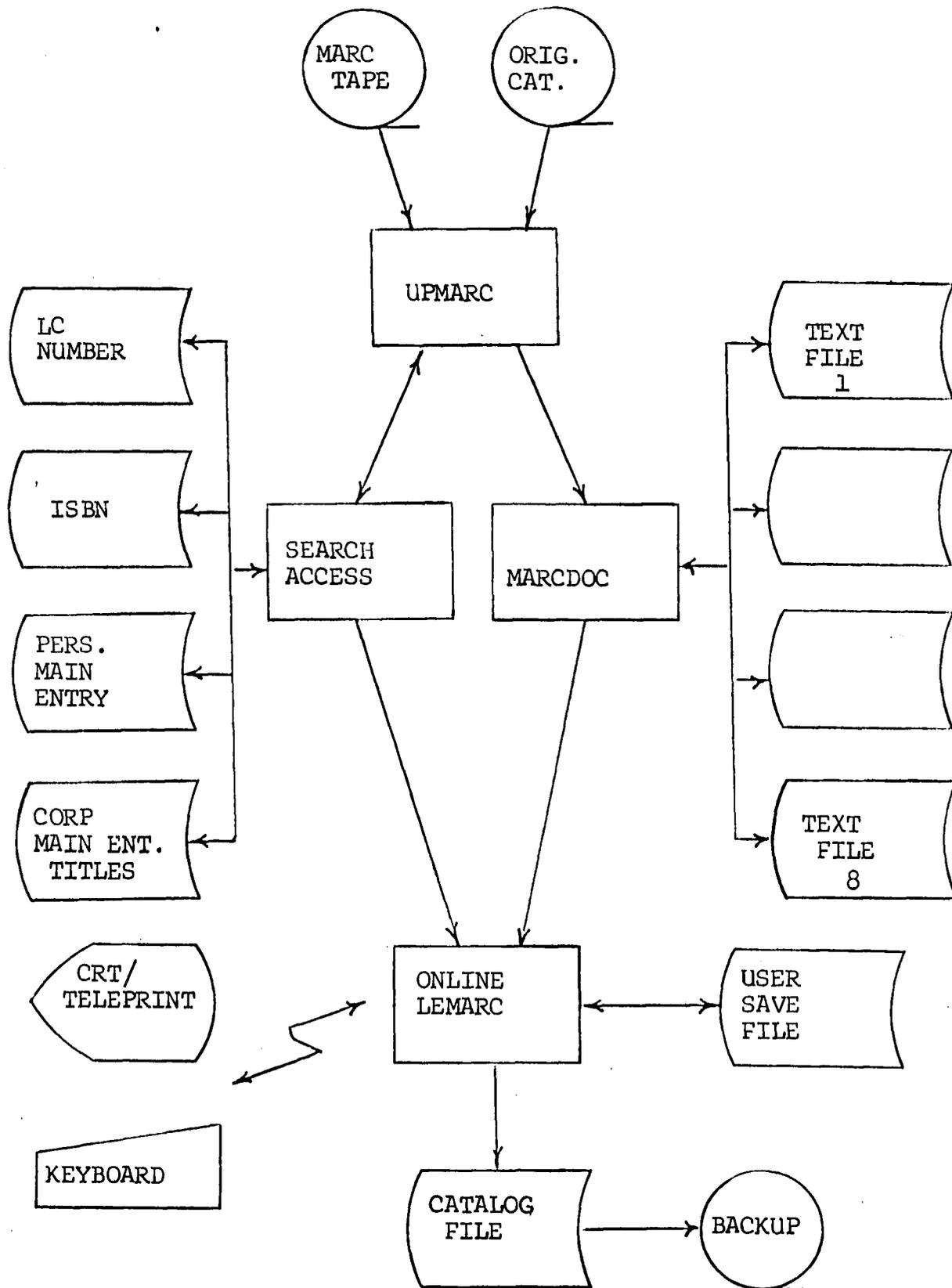
1. Text File System

The text file system contains from two to nine files. The first, MARCDØC, is a word addressable files. Entries in this file are 1 word long and point to the text file and word location of each record in the system. By restricting access to the text system to logical record number only it is possible to reorganize or restructure the text system without updating the key files. The text files, MRCTXT1 - MRCTXT8 are also word addressable files. The records are stored in LEMARC internal format.

2. Key System

The key system contains three files for three different entry types. The block and record size for each key type is contained in MARCSKL. Access to the key files is by a common randomizing algorithm. The output range is controlled by the MARCSKL file. For each file there is allocated within it prime storage area and an overflow area. The prime area is equal to the number of buckets times the length of prime buckets. The overflow area can be established with a different record length depending on the expected statistical character- istics of the randomizer output. For unique keys it is reasonable to assume that any entry can be found in one read 95% of the time with a file configured to fill 90% of the prime storage area.

3. Update Control Files

The update control requires two files. The first,
MARGAR holds location-length pairs for free space in the
text files. Typically this space is made available when
a record is deleted. This file is kept in core during the
update process. The second file is MARCMEI. This file
contains entries destined for the MARCME file. Since access
to these keys is not required during the update process, it
is more efficient in large volume updates to handle these
keys in a sequential manner. Direct sequential update is
a update run option.

4. On-Line Control Files

These files interface between the various on-line
activities and the off-line processes. MARCUSERFLxx holds
user stored records that may or may not be complete MARC
records. This file is initialized batch and utilized by the
on-line system only. It may be reorganized in batch mode.
The other communication file is MARCSAVEFLx, which holds
cataloging output from on-line activities. It is read and
reset by the batch card production system.

IV. ON-LINE PROGRAM ORGANIZATION

A. Requirements

The on-line component of LEMARC was written to meet
the following objectives:

1. Minimize Memory Requirements. Since the Lehigh computing
environment is centered around a single CDC 6400, minimal
memory requirements are essential to receiving reasonable
service from the time sharing system. This requirement
was felt to be more important to system performance than
the amount of central processor or channel time required.

2. Operate on TTY compatible terminals. Although cataloging
operations must be performed on a full ASCII terminal
(ex. A Beehive Super Bee) inquiry and verification should
be performable on any TTY like device. This precluded
micro-programmed or intelligent terminals.

3. Potential for Reentrant Operation. This first version of
LEMARC was not reentrant, since our performance studies
showed that at least four active terminals were required
to make reentrant operation break even. However, later
plans, such as patron access would easily meet this level,
so the programming developed for LEMARC was set up to allow
relatively simple conversion to multi-user use.

4. Operate with Different Mass Storage Configurations. One of
the major costs of running an on-line system is the cost of
disk mass storage. It is also true that very little of

this resource is actually used at any one time. In a library
system, the major storage requirement is for record text.
Therefore we set LEMARC up so that the system could operate
with different sets of records available. Currently this
division is by receiving date,but subject area is an equally
reasonable division criteria. The key files contain keys to
all active records so a user can be informed when a record
that is not available is retrieved. Different sets of records
can be scheduled depending on the needs of the user group.
For an inconvenience in accessing 5% or so of the search
requests, up to 50% of the storage costs of the system can be
eliminated. This is by no means a very desirable solution,
but it does allow a greater degree of control over the system
configuration and fixed costs.

B. Implementation

The implementation of LEMARC is very modular, not only for
the obvious programming advantages modularity yields, but also
to simplify the use of LEMARC routines in other systems. All
input is channeled through a single routine, CØMMAND. Each
command type is handled by a separate command processing routine.
Since all input processing is contained in COMMAND, various
terminal characteristics can all be dealt with in one location.
COMMAND, after verifying the command, then transfers control
to the appropriate service routine. No user dependent storage
is contained within each service routine, so as long as two users

11

are not scheduled into the same command the system can easily be made reentrant.

The command service routines, and the routines they require, all reside on a single library. To configure a LEMARC system, a set of macro calls cause the loading and linking of the desired service routines. LEMARC generation MACROs include:

1. CØMMAND text, routine

>   This allows version of LEMARC to be assembled to support different activities. Some systems may not include features like cataloging or private store area. Since each user on LEMARC has his own copy, the more compact each version is the better the entire system will perform. Also this allows multiple input commands to be made synonymous, which is convenient if users are used to different systems. It also allows single keys, such as the program function keys to be made equivalent to text commands such as LCNUMBER or TITLE.

2. TERMINAL type

>   Certain versions of LEMARC are generated to assume, and require, that the terminal they are active at is a full ASCII terminal, and transmit both upper and lower case characters. Others display only a restricted character set and either display the shift characters, or ignore them.

3. MAXDOCL number

Different usages of LEMARC have differing

requirements in terms of maximum number of

documents that can be displayed as the

result of a single search. For cataloging

versions up to 512 documents can be displayed;

for inquiry versions only 128 documents are

displayable in each search pass.

LEMARC PROGRAM ORGANIZATION
Fig. 2

KEYBOARD

DISPLAY

INPUT
TERMINAL
CONTROL

OUTPUT
TERMINAL
CONTROL

COMMAND

COMMAND
SERVICE
ROUTINES

USER
AREA

DATA
MANAGER

LEMARC
DATA BASE
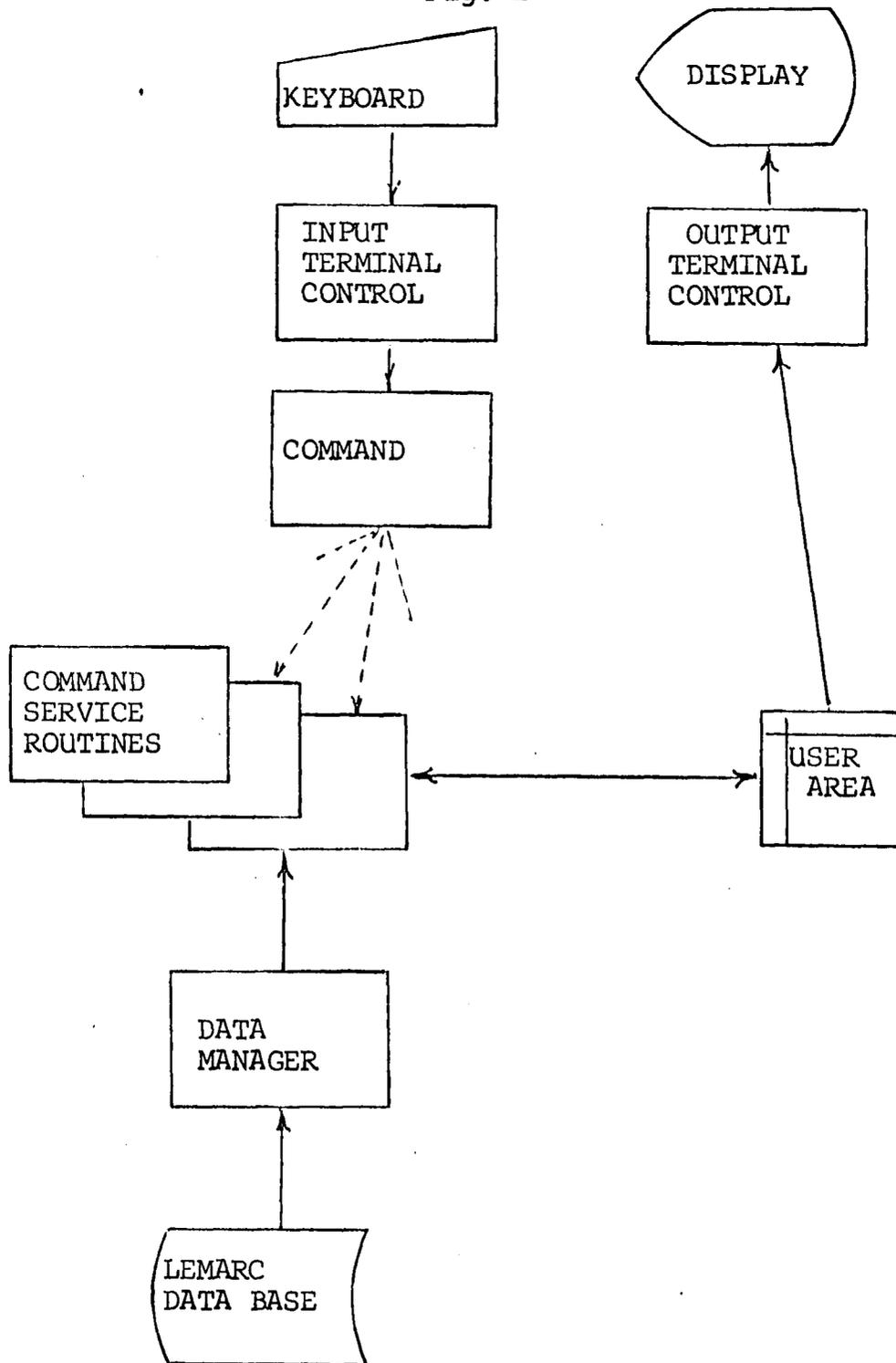
V. ADDED FEATURES

Although the basic functions of a cataloging system have been covered, there are a number of additional features that can add to the utility and ease of use of a system, although they are not essential.

A. Private Store Area

In our initial study of library requirements, a need for user oriented, temporary storage capacity was significant. This resulted for several reasons. The first was a desire on the part of the library cataloging supervisor to be able to verify the validity of cataloged records, on a complete, or partial basis. There was also a desire to put a record aside for later use, or pass a record to another user.

This need was met through the provision of a private store area for each cataloging user. This file stores up to one hundred records that may or may not be complete (i.e. main entry, title, call number, etc.) and so are only accessible through a number assigned to each record. The user can either request to

store the current record in number nn, or

store the current record in the next available location.

In the latter case the system would inform the user of the location that was used. Users can move records in and out

of the private store from the current record area.

Most importantly, users can access other users private store if it is not being used, and if they know the password. In that way a supervisor can examine the completed cataloging output, or a record that has caused a problem and leave comments or solutions in the private store. There is one more advantage to the private store, particularly for installations that have a reasonably small usage rate. Since the private store area can exist on a separate device from the main data base, private store operations, including cataloging, can be performed even if the data base itself is unavailable. This technique can reduce the total resources required to maintain a cataloging system if it is acceptable to the user group, and the unused resources can be used by another system.

B. Message Switching

Experience with operating the LEADERMART system showed users (or the system itself) sometimes required the ability to communicate directly with a human. In the LEADERMART system a COMMENT command would display the user's comment on a Teletype at the Center for Information Science, to be noted or acted on by a staff member. The LEMARC system displays a message on a specified terminal (if available) in the following situations:

1. Every time a user enters or leaves the system,

2. Every time a system problem develops, ex. a file

is found to be missing or out of space,

3. At a user request.

A staff member can then respond by sending a message to
the terminal or rectifying the problem situation. Users
can also communicate directly with other system users.
Since we have very little operational experience with
this feature it is not possible to determine its values
to the user community.

VI. CONCLUSION

Since we have had very little operational experience with the system it is not possible to determine the acceptance of the system at this point. The system has several advantages:

1. It can be configured to fit a wide variety of requirements and resource availabilities.

2. It includes many capabilities beyond those required for technical processing, and so can be used to discover reaction to features such as message switching, information retrieval and patron access.

3. It is compatible with the requirements for bibliographic information retrieval systems.

4. It is easily extendable to additional requirements with a minimum of additional programming and data base changes.

REFERENCES

1. Hillman, D.J. and Kasarda, A.K., "The LEADER Retrieval System", Proceedings 1969 AFIPS Spring Joint Computer Conference, Boston, Mass. (1969).

2. Kasarda, A.K. and Hillman, D.J., "The LEADERMART System and Service", Proceedings of the ACM 1972 Annual Conference", Boston, Mass. (1972).

3. Kilgour, F.G. et.al., "Title-Only Entries Retrieved by Use of Truncated Search Keys," Journal of Library Automation V4: p.207-210,(Dec.1971).

4. Long, P.L., and Kilgour, F.G., "Name-Title Entry Retrieval from a MARC File," Journal of Library Automation, V4: p. 211-212,(Dec. 1971).

5. Kilgour, F.G., et.al., "The Shared Catalog System of the Ohio College Library Center," Journal of Library Automation, V5: p.157-183, (Sept. 1972).

APPENDIX A

LEMARC USERS MANUAL

This appendix is a reproduction of the current LEMARC USERS manual. It consists of chapter I-VIII.

In addition five Exhibits are included

Introduction

The Lehigh MARC System (LEMARC) is designed to aid library technical staff members working with the Library of Congress MARC II data base. The objective of the system is to aid the user in obtaining a machine readable cataloging record either from the data base or by original entry. Once the suitable record has been found the system provides the user with the capability to retrieve, modify and store the record for off-line catalog card production.

This document describes the basic characteristics of the system with both narrative and examples. Additional exhibits are included to provide background information for system users.

I. Initiating LEMARC

   The LEMARC system is entered by first logging into the
Lehigh computer with a valid user name and password and valid
user description

   Type

   MARC.

II. Initial Messages

When LEMARC is first initiated it will inform you of several important facts about the data base, including the date of the last update and the last MARC issue.

Normally LEMARC will display
READY.

when it is ready for your first command. It is possible that it may have trouble collecting the MARC data base. Discussion of this is in Chapter VIII.

LEMARC is now ready. The LEMARC system has the capacity to perform many different operations on the MARC data base and on particular MARC records. The next chapters describe the various operations that can be performed and discuss how to combine them to perform various purposes.

Command Entry

This section will discuss the entry of commands into the LEMARC system. The particular commands will be described later.

The LEMARC system performs particular functions for the user. It is the responsibility of the user to combine these functions together. All commands entered are of this form:

Command, parameter 1, parameter 2, lll

Command is a string of characters (described later) that refer particular functions such as DISPLAY, GØ or LCNUMBER.

The parameters that follow depend on the particular function, they may consist of an LC number, text or a line number.

Each line begins in column 1. Parameters are separated from each other and the command by commas. The last parameter (or the command itself if there are none) can be terminated by a blank. When a line is ready for transmission, type carriage return, send or enter depending on your terminal.

LEMARC requires that only enough of the command be typed to make it unique from any other possible commands. It is acceptable to type more than this; for example

LC, 68-4648

LCN, 68-4648 and

LCNUMBER, 68-4648

will all be recognized. Generally the first two characters are sufficient to guarantee uniqueness. The command may contain characters in both upper and lower case.

General Flow

There are some important concepts that make LEMARC usage more easily understood. Essentially the LEMARC user performs three major procedures

1) Request Entry

2) Inspection ans Selection

3) Modification and Storage

Request entry (described in Chap. IV) yields a list of documents from 0 on up. This is the current document list. If it contains no items, it means that either the MARC II data base does not contain what you are looking for, or your key was not formed properly (see Exhibit B). There are tools in LEMARC for easily dealing with this list by reducing its content.

The next phase is selection of the desired record. The PAGE and DØCUMENT commands aid in selecting records from the current list. Any LC or ISBN search, or a DØCUMENT command yield a single current record.

This current record can then be modified and saved.

III. Record Display

1. DISPLAY, line 1, line 2

This command displays the current document, from line number 1 to line number 2. If the parameters are missing, the entire record is displayed. If no record has been retrieved, the display will be aborted.

2. PAGE, docnum

This command starts the page display beginning with document < docnum >.   This command is appropriate after a main entry search.   The page display shows only the first line of the title and main entry, and a single page is displayed. For a 25 line terminal, 11 documents will be summarized. If docnum is not specified, the display will be continued from where it was left off, or at the beginning, document 1, following a search.

3. DØCUMENT, docnum

This command displays a particular document from the list resulting from a main entry or title search. Typically it is used to examine a particular document viewed in the page display. If docnum is not specified, the next document in line, or the first following a search will be displayed. This document then becomes the current document, and can be modified, saved or stored (described later).

4. GØ

If the display for a LEMARC record will continue
beyond the bottom of the current display screen, the
LEMARC will request that the user enter a GØ to continue
the display. This process will continue until the record
has been completely displayed (or until the last line has
been displayed if a limit was set in a DISPLAY command),
when a READY will be displayed. Even though the system is
requesting a GØ, any valid LEMARC command may be entered
at any point and the display continued with a GØ at a
later point.

IV. Record Retrieval

1. LCNUMBER, lcno

This command retrieves and displays the record for lcno. The hyphen can be included, and zero's following the first two digits are not required.

LC, 68-4648

LC, 68-04648

LC, 68-004648

are all equivalent.

2. ISBN, isbn

This command retrieves and displays the record for isbn. The hyphen's can be included. An SBN can be converted to an ISBN by adding the country code (a zero for the U.S.) onto the left end of the SBN. The resulting number is nine digits plus a tenth which serves as a check digit and can be left off.

3. PMENTRY , name

This command retrieves the record based on the author's name. The name consists of the author's last name, plus any desired initials. It is desireable to specify as many initials as possible, as this will minimize the number of recrods retrieved. However if the MARC record does not contain the initials you have specified it will not be found, and it will be necessary to try the search with one less initial. The following search requests will all return the desired document,

although, the first requests will yield fewer undesired

documents

        PM, J∅NES AS

        PM, J∅NES A

        PM, J∅NES

4. CMENTRY, corporate main entry

This searches for records based on the corporate main entry. The text of the main entry should be entered in full, except all punctuation should be left out. In order for a record to be retrieved, its main entry must contain all the words in the request, although order or any other words that may be present are not considered.

5. TITLE, title

This functions in the same manner as the CMENTRY search except that the title is examined.

6. ANDD∅CS

Often both the title and main entry searches will yield too many documents to easily inspect. To minimize the number of records that must be inspected, it is possible to specify both the title and main entry. The procedure is to logically 'and' the results of two searches together, yielding only records present in both lists.

After one search has been completed, and before the second search enter:

ANDD∅CS   29

For example, to search U S CØNGRESS; NØMINATIØN HEARINGS, enter

CM, U-S CØNGRESS

AND

Tl, NØMINATIØN HEARINGS

7. NSETM

On some searches, the internal record list of LEMARC may be exceeded. LEMARC will return as many records as it can, and will indicate that there is another set of documents. The retrieved list can be dealt with as with any document list. When it is exhausted, enter;

NSETM

to get the next set. It is even possible that no documents were retrieved, although there may be documents in a subsequent set. If one of these sets is 'AND'ed, the original request, followed by an immediate NSET should be entered. The 'AND' can then be performed against this set. It is important to realize that 'AND' operates on the current list, so it must be performed on each set individually. It is therefore important to do the search which is likely to result in the smallest list first.

V. CATALØGING + RECØRD MØDIFICATIØN

1. WSPACE

This command sets up a blank workspace for original cataloging. Data can be entered as alter statements (shown below).

2. ALTER

This command alters, deletes, or adds data to the current MARC record. The form is:

-, line no, tag, data

    line no. This specifies the data element being referenced, and is of the form n.m. Initially line numbers are assigned to the data elements in the record, incrementing n by 1 and with m=0, i.e. 1.0, 2.0, 3.0. To insert a new item between element 2 and 3, line no should be 2.5. If the line number is the same as an existing one, it replaces the existing one. If the line number is blank, the entry will be added to the end of the record.

    tag - This specifies the tag to be stored for the new field. If it is a new entry, it must be present. If the entry replaces or adds to an existing entry, the tag field can be left out, and the old tag value will be used.

data - This is the text of the data to be stored.
All data from the first non-blank character
following the comma after tag through the last
non-blank character of the line are stored.
No leading or trailing blanks can be stored.

If the first two characters of the data
field are $$, the data in the line is entered
as a continuation. If blanks follow the $$,
they are included in the record.

For example:

To make an original cataloging entry:

-,1,82, $a 301.6

-,2,100, $a Horowitz, David

-,3,245, $a Countr Culture and Revolution

To correct line 3

-,3,, $a Counterculture and Revolution

3. SAVE

This command uses the current MARC record to make a
permanent cataloging entry. The exact actions taken will
depend on the user's Library Description File (see Exhibit C).
The standard library values can be overriden by parameters
following the SAVE. For example if a user normally has CAT=1,
PRINT=0 (card production), the user can cause the document to
be printed only by

SAVE, CAT=0, PRINT=1

If the users VERI flag (see Exhibit C) is 1, then then save will be rejected. The user can perform a SAVE only by adding

       VERI=0

to the parameter list.

VI. PRIVATE STORE AREA

Each cataloging user has 100 positions to temporarily store MARC records. These records are not part of the MARC data base and are retrievable only through the private store area commands.

1. GETREC,n

   This command makes the record stored in position n the current document. If n is not specified the next record in sequence is returned, any unused records being skipped. The document list is not affected.

2. STØREC,n

   This command stores the current record in position n. If n is not specified, the next unused position is utilized, and the user is informed of the number. The document list is not affected. This command will be aborted if no open spots can be found if n is not specified. If n is specified any document in that position is overwritten. It is possible that the store will be aborted even if a position is open if no space is available. It may be necessary to purge a record to have two contiguous positions open.

3. PURGER, n

   This purges record n from the private store. If n=ALL, the entire private store is cleared.

4. USE, ID, PASSW

   This command allows one user to examine another user's private store area. It is necessary to specify both the ID

(2 characters) plus a password (PASSW) associated with

that user. The original user cataloging description

still applies, but the new users private store area is

now accessible. If the USE attempt fails, the user will

have no private store area. To regain the proper store

area enter:

USE, HERE

VII. Miscellaneous Commands

1. DØNE

This ends the session with LEMARC. The user can
select another LEADERMART option.

2. MESSAGE, ID, message

This sends 'message' to terminal ID. This allows
users to communicate among themselves. If ID is absent,
the message is sent to the C.I.S. LEADERMART monitor
terminal. When using this feature it is helpful if you
identify yourself in the message, for ex.

M,, WHEN WILL MARC LC BE UP...SHARØN

3. SETP, type, val

This command allows the user to override normal defaults
during a LEMARC session. Type specifies the default value
to be overriden, and val specifies the new value.

type default (see Exhibit C)

1 NLINES
2 NCHARS

for example, if

SE,1,19

is entered, the terminal display will be adjusted to a
19 line terminal.

VIII. ERRØRS AND PRØBLEMS

A. Initiation

During program initiation (ie. before the first READY-appears) several problems can be detected by the system. Both you and the Center for Information Science command terminal will be informed of any of these problems.

The LEMARC system consists of a set of data files. It is possible that for one reason or another one or more of these files may not be available to the LEMARC program. In that case one of the three messages below will be displayed.

1. MARCTXT file missing, some documents may not be retrievable

For any MARCTXT file that is missing, approximately 1/8 of the MARC records will not be available. References to them will be retrieved but at the point of display a message will indicate that the document is unavailable.

2. _____ command not available

In this case the data file needed for the command is not available. That command will not be recognized if it is entered. For example

LCNUMBER command not available

If these restrictions are not tolerable for your activity do an immediate DONE after LEMARC displays READY-.

3. MARC data base is not available

The MARC data base is not available so LEMARC will do a DØNE itself. The next question you will get will be a chance to select a LEADERMART option.

B. During normal activity

Occasionally errors in MARC usage will occur. Most of these errors are user induced, but a few are errors detected by the system. An example of this later type would be error 1, which alerts you to a problem in the system. Most of the error messages are self explanatory and require only that you examine the last line entered and correct the mistake, for example mistyping a command. Others, indicating a record is unavailable. are not user errors but indicate an error condition in the MARC data base. The only action the user can perform is to skip over the document.

Exhibit A

ERRØR MESSAGES

1. DOCUMENT FILE MISSING, RECORD UNRETRIEVABLE.

When LEMARC was initiated, one or more
MARCTXT files was missing.

2. COMMAND NOT RECOGNIZED.

User entry was not a valid in EMARC entry.
This could be because the command was mistyped,
or not enough characters were entered. It is
also possible that a file being missing caused
a normally valid command to be unavailable.

3. SYSTEM DOCUMENT NUMBER INVALID.

If following a SYSDØC command, the document
number was too large. If it occurs at any
other point, it is an internal error to
LEMARC, please notify C.I.S. of the situation.

4. DOCUMENT HAS BEEN DELETED.

This MARC record has been deleted by the Library
of Congress. This will occur infrequently.
Ignore it and keep going.

5. YOU HAVE NO CATALOGING CAPACITY.

You are not authorized to catalog, or else
system was unable to find space to catalog
your records.

6. YOU HAVE NO PRIVATE STORE CAPACITY.

You do not have a private store area.

7. YOUR PRIVATE STORE IS FILLED.

There is no more space in your private store
area, purge some or all records.

8. NO RECORD STORED IN THAT POSITION.

No record has been stored in that position
of your private store area.

9. SETP - VALUE NOT RECOGNIZED.

> The set parameter value in a SETP command
> is out of range.

10. ILLEGAL LC KEY.

> The LC number specified is illegal. Most
> likely it is greater than eight characters
> (not counting hyphen); is less than three;
> or contains alp abetic characters, in other
> than trailing position.

11. ILLEGAL ISBN KEY.

> The ISBN Key is illegal, typically for
> reasons similar to (10) above.

12. MARC RECORD BAD.

> The LC MARC record is unusable, skip it.

13. LEMARC INTERNAL ERROR.

> The LEMARC program has detected an internal
> error. You will lose your current search,
> although anything you saved will not be lost.
> C.I.S. will automatically be informed of the
> situation. If this problem persists, do a
> DØNE.

14. PASSWORD INVALID.

> The password specified in the USE command is
> not correct, try again with the correct password
> or return to your own user area.

15. SPECIAL INSTRUCTIONS INVALID.

> User options in SAVE command are not correct,
> see Exhibit C for description.

16. ACCOUNTING FILE ERROR.

> Accounting for this usage may be inexact,
> please notify C.I.S. to insure an exact record.

17. NO MARC RECORD HAS BEEN RETRIEVED.

> An attempt has been made to STØRE, DISPLAY or
> ALTER when there is no current record.

18. NEXT/PAGE BEYOND DOCUMENT LIST.

> A NEXT or PAGE request has been issued
> that would exceed the number of records
> retrieved.

19. BAD NUMERIC PARAMETER.

> A parameter that should contain only digits
> contains other characters.

20. LINE NUMBER BAD.

> ALTER request would require existence of
> a line number that does not exist.

21. ILLEGAL USER ID.

> User ID is not a valid user ID for the LEMARC
> system.

22. COMMAND NOT YET IMPLEMENTED.

> This feature of LEMARC has not yet been
> completed. Please be patient.

23. USER FILE UNAVAILABLE.

> The user file requested in an USE command is
> in use by another user.

24. USER NOT KNOWN TO LEMARC SYSTEM.

> The user ID specified in a USE command is not
> known to the LEMARC system.

25. ERROR in ALTER STATEMENT.

> ALTER statement is improperly composed.
> Look for missing or extra commas.

26. NO ROOM TO STORE SAVED RECORD.

> All LEMARC cataloging save files are filled.
> Notify C.I.S.

Exhibit C

Save Parameters and Library Description

        In order to execute any SAVE functions under LEMARC it is
necessary that for each cataloging user, a User definition be
completed. The parameter values, and a sample form are in this
exhibit.

        Part of the User description is the specification of the
standard catalogs which entries from that user will be sent to.
For each catalog listed, a Catalog Description Form should also
be completed. (In most cases one Catalog Description Form will
serve multiple users). These define the existing card catalogs
in the library system. They may also define archiving files.
A special Catalog, PRINT, is defined for all users (if desired).
Records sent here will be printed and mailed to the "SHIP TØ"
address in the User's Definition Form.

Exhibit D

USER Definition Values

PASSW   Password required to access private store are.
               If O, is no access is possible.

VERI   I = SAVE allowed
               O = SAVE not allowed. SAVE request will be
                   rejected unless VERI=1 is listed as a
                   special parameter.

HØLD   Holding library number. If no record is to be kept,
               this should be O. (This value is originally
               assigned by C.I.S.

HØLDT   Holding library text. This text will appear in
               holding library area of card.

NCAT   Number of catalogs. Normally this will be the
               number of catalogs defined under LCAT. It is
               possible to define more LCAT entries than are
               reflected in the NCAT value. The extra catalogs
               will only have an effect if the NCAT value is
               set at SAVE issuance.

NLINES   Number of lines displayed on terminal screen.

NCHARS   Number of characters across one line.

Catalog Description

LABEL	Catalog Label	10 chars
DISP	Disposition code[1]	
CTYPE	Catalog type[2]	0-7
CØNTS	Continuation[3], default=1	0-7
CALLT	Call no. type, data element	
CALLF	Call Format[4]	
LINEPC	Line position; primary call, d=3	
ØVS1	Starting size, oversize 1 (in cm)	
ØVS2	Starting size, oversize 2	
ØVS3	Starting size, oversize 3	
ØVL1	Oversize 1 symbol	
ØVL2	Oversize 2 symbol	
ØVL3	Oversize 3 symbol	
ØVP	Position of oversize label[5]	
MHL	Maximum Headings Line d=3	
CPH1	Ident position, heading line 1	
CPII2	Ident position, heading line 2 - or	
CSH	Char. set headings 1=full, 2=caps, d=1	
TRC	Tracings present[6]	
SKLHM	Skip line between Heading+M.E. 0=no,1=yes,d=0	
SKLMT	Skip line between M.E.+Title	
SKLTC	Skip line between Title+coll.	
SKLCN	Skip line between coll.+notes	
PHOLD	Position of holdings data on card[7]	
SETI1	Set character position of Identure 1	
SETI2	Set character position of Identure 2	
SETI3	Set character position of Identure 3	
SETI4	Set character position of Identure 4	

Catalog Codes

1. Disposition code would indicate physical location of catalog data, i.e.

   1= Cards
   2-50 tape set n=1
   51-99 disk set n-50,etc.

2. 1= ATS   4=5   7=AS
   2= A   5=AT
   3= T   6=ST   0= no cards

3. Codes same as (2). Lists card types to be given continuations, i.e. CØNTS=1 will cause all cards to be continued.

4. Meaning depends on CTYPE. Refers to column of p16

5. 1= on right of call
   2= below call
   3= on left of call

6. Same code as (2). Only specified sets will have tracings present.

7. Position of holdings.

Exhibit E

Command List

    This is a list of all the commands available in LEMARC.
The minimum characters to represent the command is capitalized,
the rest is optional. For each command the chapter describing
it is listed.

Lcnumber	4
Isbook	4
PMentry	4
Cmentry	4
Title	4
Nset	4
And list	4
Display	3
Go	3
PAge	3
SDoc	3
-	5
Workspace	5
SAve	5
SEQuence	
D∅Cument	3
PUrger	6
STorec	6
Use	6
GEtrec	6
Message	7
SET p	7
DOne	7
ECHO	7

APPENDIX B

FIELD EDITING CRITERIA

This appendix includes the editing criteria for data
stored in retrieval keys and the '008' field in the MARC record.
The purpose of this editing is two fold.

     (1) Normalization of data to simplify retrieval.

        An example of this would be the LC Number

        supplement, which we desired to place separately

        from the LC number text.

     (2) Minimize Storage Requirements. An example would

        be the removal of very frequent (and non-sorting)

        words from the retrieval files.

MARCLC

1. LC Number 8 digits in R8 format

2. Supplement number (if present)

    Prefix

    Suffix   format

    Revision date

    These fields are not separated by slashes as in a MARC record

3. Document number

CARD	MARC	LEMARC
78-657093	ƀƀƀ78657093ƀ	78657093
78-657093 rev2 (revision date 470)	ƀƀƀ78657093ƀ//r702	78657093//r702
78-657093 AC	ƀƀƀ78657093ƀ/AC	78657093/Ac
78-657093 AC (revision date r70)	ƀƀƀ78657093ƀ/AC/r70	78657093/AC/r70

MARCLC first word	MARCLC second word
--78657093	_____
--78657093	r702--------
--78657093	AC----------
--78657093	ACr70-------

<u>008</u>   "Fixed" length data elements

"Type of Publication Date Code" (char. pos. 6) determines

if the DATE1 (7-10) and DATE2 (11-14) are stored.

If one of these fields is suppressed, all of the fields

following it are shifted left 4 characters.

<u>Code</u>	<u>Date 1</u>	<u>Date 2</u>
S	-	suppressed
C	-	-
n	suppressed	suppressed
r	-	-
m	-	-
q	-	-

<u>100</u>   Personal Main Entry

1. Up to 10 characters of last name   (L format)

2. Remaining character of last name (if any)

     00 byte   (L format)

     Initials

3. Document number

Last Name - All characters up to the first comma in a MARC

record except diacritical marks (shift 2) and any shift 1

characters

    <u>MARC</u>

                               Word 1   Word 2

LAⱢ SOR,Ɫ WILLIAMⱢALBERTLAⱢSOR---- -WA-------

RIMSKY- KØRSAKØV,Ɫ ILLYCH

                       RIMSKY-KØRSAKØV-I---

MARCCM

1. Stem   (L format)

2. Position, Length (of phrase), Tag, Doc#.

   1 origin

The "STEM" is a maximum of 10 characters in length. It does not

include shift 1 characters or any diacritical marks. Apostrophies

and the characters that follow it are included in a stem.

   A stem is delimited in a MARC record by:

      a) beginning or end of a subfield

      b) period, comma, hyphen, or blank

      c) any numeric character not in a date.

   A date is any string of four and only four numeric characters

in which the first character is not zero.   A date is considered to be

a stem.

   Some stems receive special treatment.   "United States" and "U.S."

are translated into the single stem "U-S"; "Great Britain" into

"GRTBR".   The position of the composite stem is that of the first

stem that produced it, but all stems preserve their original position

in the phrase.   For example, for the phrase:

      UNITED STATES C∅NGRESS

We generate the stems:

      U-S   position = 1   length = 3

      CONGRESS   position = 3   length = 3

   A stoplist is used to keep "insignificant" stems out of the

MARCCM file. This stoplist is generated by truncating all the

words in CISCØNFILE24STØPLIST to 10 characters and combining any duplicates that are produced. In addition to this stoplist, any stems of less than three characters are not considered significant (note that "U.""S." are exceptions). An insignificant stem does not change the position of stems that follow it in the phrase (a situation similar to "STATES" in the example above).

"Phrase" has the following meanings
a) 110 field - $A subfield

$B subfields (as many as there are)

$K subfields only if it immediately follows the $A subfield

b) 111 field - $A subfield only

c) 130 field - $A subfield

d) 245 field - $A subfield

VITA

Preston F. Marshall, Jr., son of Preston F. and Gwendoly S. Marshall was born in London, England on August 19, 1950. He graduated from Walpole High School in June 1968. He graduated from Lehigh University in 1972 with a B.S. in Electrical Engineering.

From 1972 to 1974 he was a student in Information Science and a research assistant in the Center for Information Science. Since 1974 he has been Assistant to the Director of the Center for Information Science.

Mr. Marshall is married to the former Elizabeth H. Schlerf of Kennett Square, Pennsylvania.