

AD 727016

Information and Retrieval System
for the
AFCRL Research Library

by

Donald J. Armstrong
Frances H. Healey

Carolyn M. Mandell
David Stoll

Analysis & Computer Systems, Inc.
113 Terrace Hall Avenue
Burlington, Massachusetts 01803

Contract No. F19628-67-C-0397

Scientific Report # 1

March 1971

Contract Monitor: Austin A. Almon Jr., Analysis and Simulation Branch

This document has been approved for public
release and sale; its distribution is unlimited.

Prepared
for

AIR FORCE CAMBRIDGE RESEARCH LABORATORIES
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
BEDFORD, MASSACHUSETTS 01730

Revised by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va. 22151

D D C
RECEIVED
JUL 28 1971
RECEIVED

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Analysis & Computer Systems, Inc. 113 Terrace Hall Avenue Burlington, Massachusetts 01803		2a. REPORT SECURITY CLASSIFICATION Unclassified	
		2b. GROUP	
3. REPORT TITLE INFORMATION AND RETRIEVAL SYSTEM FOR THE AFCL RESEARCH LIBRARY			
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Scientific. Interim.			
5. AUTHOR(S) (First name, middle initial, last name) Donald J. Armstrong Carolyne M. Mandell Frances H. Healey David Stoll			
6. REPORT DATE March 1971		7a. TOTAL NO. OF PAGES 57	7b. NO. OF REFS 2
8a. CONTRACT OR GRANT NO. F19628-67-C-0397		9a. ORIGINATOR'S REPORT NUMBER(S) Scientific Report # 1	
b. Project, Task, Work Unit Nos. n/a n/a n/a			
c. DoD Element: 61102F		9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)	
d. DoD Subelement: 681300		AFCL-71-0210	
10. DISTRIBUTION STATEMENT 1 - This document has been approved for public release and sale; its distribution is unlimited.			
11. SUPPLEMENTARY NOTES TECH. OTHER		12. SPONSORING MILITARY ACTIVITY Air Force Cambridge Research Laboratories L. G. Hanscom Field (SUY) Bedford, Massachusetts 01730	
13. ABSTRACT This report is a scientific report prepared to describe the status and progress of the information and retrieval system implemented for use by AFCL Research Library personnel. This report describes, in detail, the design of a computer oriented system required to automate the AFCL Research Library and the computer programs completed to date to assist the library personnel in the Selection, Acquisition, Cataloging and Circulation phases.			

DD FORM 1473
1 NOV 65

Unclassified

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
MARC II Selection Acquisition Cataloging Circulation Statistics Supervisor Master Inventory File Authorization Permanent File Order Forms Storage Retrieval						

Information and Retrieval System
for the
AFCRL Research Library

by

Donald J. Armstrong
Frances H. Healey

Carolyn M. Mandell
David Stoll

Analysis & Computer Systems, Inc.
113 Terrace Hall Avenue
Burlington, Massachusetts 01803

Contract No. F19628-67-C-0397

Scientific Report # 1

March 1971

Contract Monitor: Austin A. Almon Jr., Analysis and Simulation Branch

This document has been approved for public
release and sale; its distribution is unlimited.

Prepared
for

AIR FORCE CAMBRIDGE RESEARCH LABORATORIES
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
BEDFORD, MASSACHUSETTS 01730

ABSTRACT

This report is a scientific report prepared to describe the status and progress of the information and retrieval system implemented for use by AFCRL Research Library personnel. This report describes, in detail, the design of a computer oriented system to automate the AFCRL Research Library and the computer programs completed to date to assist the library personnel in the Selection, Acquisition, Cataloging and Circulation phases.

TABLE OF CONTENTS

<u>SECTION</u>		<u>PAGE</u>
	ABSTRACT	i
	LIST OF ILLUSTRATIONS	v
	FOREWORD	vii
1	INTRODUCTION	1-1
2	BACKGROUND INFORMATION	2-1
3	SYSTEM DEVELOPMENT	3-1
	3.1 Initial File Building	3-1
	3.2 Selection Phase	3-3
	3.3 Acquisition Phase	3-5
	3.4 Cataloging Phase	3-5
	3.5 Update Phase	3-5
	3.6 Circulation Phase	3-7
4	CONVERSION	4-1
5	SUPERVISORY CONTROLLED AUTOMATED LIBRARY SYSTEM (SCALS)	5-1
6	CONTRIBUTORS	6-1
7	REFERENCES	7-1
8	APPENDIX	8-1

LIST OF ILLUSTRATIONS

<u>ILLUSTRATION</u>	<u>PAGE</u>
Exhibit 1 Table of Mnemonics	2-3 - 2-12
Exhibit 2 Octal Configuration Relationship	2-13
Figure 1. Initial File-Building Phase	3-2
Figure 2. Selection Phase	3-4
Figure 3. Update Phase	3-6
Figure 4. Circulation Phase	3-8
Exhibit 3 Library Form for Generating Paper Tape	3-10

FOREWORD

The design of the information and retrieval system for the AFCRL Research Library, together with the computer systems that implement these designs, are the result of analytical research performed for,

The Analysis and Simulation Branch (SUVA)
AFCRL Computation Center
Air Force Cambridge Research Laboratories
Bedford, Massachusetts 01730

The computer programs with complete documentation contained in this report may be obtained from the above organization upon request by referencing Problem Number 1680.

SECTION 1

INTRODUCTION

This report describes the work which has been done for purposes of automating the AFCRL Library. The requirements of the Library are to build a Master Inventory File using, if possible, information from Inforonics, Inc., as a data base, and using the Library of Congress MARC II system to add to the file, after using the MARC II tapes in an acquisition system; and primarily to develop a Circulation System, using the Master Inventory File as a data base, with an automatic capability for recording books charged out and returned, to recognize and prepare notices for overdue material, to provide statistical data showing which books are most frequently used and which groups use them, and also to provide greater accountability for the books in the Inventory. The initial system is for monographs only; serials and journals are to be incorporated later. The AFCRL IBM 7094 DCS and IBM 1460 computer systems were utilized for this work effort.

SECTION 2

BACKGROUND INFORMATION

Prior to any development for the Library, the MARC II system was reviewed. It was then decided that the general MARC II concept of tape records would be suitable for maintaining and utilizing a Master Inventory File. This concept consists of a tape record for each monograph with a directory in the front portion of the record and variable length data fields in the remainder of the record. The directory portion of the record has three entries for each data field. The first entry is a mnemonic for the data field, as TIL means Title, MEP means Author.

Exhibit 1, pages 2-3 to 2-12, is a table of these mnemonics. The second entry is the length in characters of the data field. The third entry is the location of the first character in the data field. This arrangement allows the programmer to read a record, search the directory for the mnemonic or mnemonics he needs, and then pick the data out of the record. It also facilitates the addition of data pertaining to a monograph of the file very readily without rewriting a lot of programs, since no program would be dependent on the location of data within a record.

Since the program for the Circulation System had to be run at the AFCRL Computer Center, it was decided that Fortran IV would be the language used for coding the programs. No major problems were foreseen in using Fortran for this type of work and indeed, no major problems were encountered. Ease of conversion from a 7094 DCS computer system to any other system that might be installed was the main factor in the decision to use Fortran.

The programming considerations that were troublesome were related to or inherent in the 7094 DCS. Namely, input tape records must have a multiple of thirty-six bits (6 characters) to be acceptable to the system. No means of recovery have been found if this condition is not met. The 7094 character set is limited and is not directly translated into either the 6-bit code used on MARC II tape or the PDP code used on the Inforonics tapes. The DCS Sort system is not amenable to processing Fortran output tapes if they exceed Fortran buffer allocations, normally 256 characters. Further, MARC II tapes are written in EVEN parity (BCD mode), which in Fortran I/O schemes, are limited normally to 132 characters per record. Although this restriction can be changed by the programmer, the Inforonics tapes are delivered in ODD parity (binary mode), so that it was easier to change the parity of the MARC tapes for the sake of uniformity.

The problem of 6-character multiples was solved by using the IBM 1460 as a preprocessor to check the records, padding records with blank characters where needed. At the time the MARC II records were checked and padded, they were also rewritten in ODD parity.

The sort problem was initially overcome by writing a sort routine when there were not many records to be sorted, but it became apparent that soon too many records would have to be sorted. Therefore, we would adapt the records to the DCS Sort system by generating tapes to be sorted by using the "SETBUF" function on the system to expand the buffer size, and by writing out fixed-length tape records, with the sort key data in the front of the records. These records were then easily sorted by the system, but the output tapes had to be returned to their original format before further processing. The only sorting so far has been for class numbers for a cross-index listing.

The character-set problem is handled by converting MARC II and PDP characters to 7094 characters. Each set has its own conversion, and the results of the conversion are not identical. Therefore, records in the Circulation System must carry an indicator to identify the source of the data so that the meaning of the 7094 characters can be truly determined. The basic reason for this problem is the diversity of the two original character sets. See Exhibit 2, page 2-13, for the conversion chart.

The output graphic characters on the 7094 DCS System are also limited. It would be desirable to print out upper and lower case characters, plus a wide variety of special characters. This, however, was not possible on the IBM 7094 DCS unless a special print chain were obtained. Alternately, use could be made of the CRT plotter for output. This CRT method was tried, and it works successfully. The problem with the CRT is that it is an extremely slow process and consumes too much tape to draw all the vectors for all the various characters.

TYPING TAG

FILE ITEM NO.

Inforonics (if different)	MARC	Tag	MARC Indicator	Inforonics Octal	
	lan	003 0		000300	Languages
	lanx	003 1		000301	Languages work is a translation
	crd	NUMBERS 010 -		002020	L.C. Card No.
	nbn	011 -		002120	NBN No.
	sbn	012 -		002220	Standard book no.
	pln	013 -		002320	Pl 480 no.
	sco	014 -		002420	Search Code
afn	bnb	019 -		003120	Local System Number
	ddc	020 -		004020	BNB Class No.
	cal	030 -		006020	Dewey Decimal Class No.
	xcal	050 0		012000	L.C. Call Number
	cop	050 1		012001	L.C. Indicator Book not in L.C.
	nln	051 -		012120	Copy Statement
	nal	060 -		014020	NLM Call No.
	asc	070 -		016020	NAL Call No.
		071 -		016120	NAL Subject Category No.

Exhibit 1. Table of Mnemonics

TYPING TAG		FILE ITEM NO.			Inforonics (if different)	MARC	Tag	MARC Indicator	Inforonics Octal	
		udc	080 -	020020				020020	Universal Decimal Classification No.	
afc			090 -	022020				022020	Local Call No.	
			MAIN ENTRY							
		mps	100 0	040000				040000	Personal Name Surname First	
		mpf	100 1	040001				040001	Personal Name Forename First	
		mpm	100 2	040002				040002	Personal Name Multiple Surname	
		mpn	100 3	040003				040003	Personal Name Name of Family	
		mcn	110 0	042000				042000	Corporate Name	
		mcp	110 1	042001				042001	Corporate Place	
		mcs	110 2	042002				042002	Corporate Surname	
		mmn	111 0	042100				042100	Meeting or Conference Name	
		mmp	111 1	042101				042101	Meeting or Conference Place	
		mms	111 2	042102				042102	Meeting or Conference Surname	
		mcf	120 -	044020				044020	Corporate Name with Form Sub-Heading	
		mut	130 -	046020				046020	Uniform Title Heading	
			Subject Is Main Entry							
mpss			100 4	040004				040004	Personal Name Surname First	
mp			100 5	040005				040005	Personal Name Forename First	
mpms			100 6	040006				040006	Personal Name Multiply Surname	

TYPING TAG		FILE ITEM NO.	
Inforonics (if different)	MARC	MARC Tag Indicator	Inforonics Octal
mpns		100 7	040007
mcns		110 4	042004
mcps		110 5	042005
mcSS		110 6	042006
mmns		111 4	042104
mmps		111 5	042105
mmSS		111 6	042106
		SUPPLIED TITLE	
	utio	200 0	100000
	utix	200 1	100001
	romo	210 0	102000
	romx	210 1	102001
	tra	220 -	104020
		TITLE PARAGRAPH	
	tilo	240 0	110020
	tilx	240 1	110001
	edn	250 -	112020
		IMPRINT	
	imp	260 -	114020

Personal Name Name of Family

Corporate Name

Place

Surname

Meeting or Conference Name

Place

Surname

Uniform Title

Uniform Title: Title Is On L.C. Printed Card

Romanized Title

Indicator: Receives Title Added Entry

Translated Title

Title

Title: Title Receives Added Entry

Edition Statement

Imprint Place

TYPING TAG		FILE ITEM NO.		
Inforonics (if different)	MARC	Tag	Indicator	Inforonics Octal
	#	261	0	114100
	#4	261	1	114101
	##	262	-	114220
	col	300	-	140020
	pri	350	-	152020
cpr1		360	-	154020
<u>SERIES NOTES</u>				
Series Traced the Same				
	rps	400	0	200000
	rpf	400	1	200001
	rpm	400	2	200002
	rpn	400	3	200003
	rcu	410	0	202000
	rcp	410	1	202001
	rce	410	2	202002
	rmn	411	0	202100
	rmp	411	1	202101
	rms	411	2	202102
	MAIN ENTRY			
rpsm		400	4	200004
rpfm		400	5	200005

Imprint Publisher
 Imprint Publisher Is Main Entry
 Dates
 Collation
 Bibliographic Price
 Converted Price

Personal Name Surname List
 Personal Name Forename First
 Personal Name Multiply Surname
 Personal Name Name of Family
 Corporate Name
 Corporate Name Corporate Place
 Corporate Name Corporate Surname
 Meeting or Conference Name
 Meeting or Conference Place
 Meeting or Conference Surname
 Personal Name Surname First
 Personal Name Forename First

TYPING TAG	MARC	MARC Tag Indicator	Inforonics Octal	FILE ITEM NO.
Inforonics (if different)				
rpmm		400 6	200006	Personal Name Multiply Surname
rpnm		400 7	200007	Name of Family
ronm		410 4	202004	Corporate Name
rcpm		410 5	202005	Corporate Place
rasm		410 6	202006	Corporate Surname
rmnm		411 4	202104	Meeting or Conference Name
rmpm		411 5	202105	Meeting or Conference Place
rmSm		411 6	202106	Meeting or Conference Surname
(following the above notes)	#	448 -	211020	Title
	rtd	440 -	210020	Title
	run	490 0	222000	Series Untraced
	rtd	490 1	222001	Series Traced Differently
BIBLIOGRAPHIC NOTES				
	bib	500 -	240020	Bibliography Note
	dis	510 -	242020	Dissertation Note
	con	520 -	244020	Contents Note
	wit	530 -	246020	"Bound With" Note
	lim	540 -	250020	"Limited Use" Note

TYPING TAG

FILE ITEM NO.

Inforonics (if different)	MARC	MARC Tag Indicator	Inforonics Octal	
	gcn	550 -	252020	General Notes (all others)
	ann	560 -	244020	Annotation or Abstract
		SUBJECT ADDED	ENTRIES	
	sps	600 0	300000	Subject Added Entry Surname First
	spf	600 1	300001	Forename First
	spm	600 2	300002	Multiple Surname
	spn	600 3	300003	Name of Family
	scn	610 0	302000	Sub. Added Entry Corporate Name
	scp	610 1	302001	Place
	scs	610 2	302002	Surname
	smn	611 0	302100	Sub. Added Entry Meeting or Conference Name
	smp	611 1	302101	Sub. Added Entry Meeting or Conference Place
	sms	611 2	302102	Sub. Added Entry Meeting or Conference Surname
	scf	620 -	304020	Corporate Name with Form Sub-Heading
	#	648 -	311020	Title
	sut	630 -	306020	Uniform Title Heading
	sto	650 -	312020	Topical
	sgc	651 -	312120	Geography Name

(Following items 600-620)

TYPING TAG

FILE ITEM NO.

Inforonics (if different)	MARC	MARC Tag	MARC Indicator	Inforonics Octal	
	spj	652	-	312220	Political Jurisdiction
	sno	653	-	312320	Proper Names Not Capable of Authorship
	sgn	655	-	312520	General Subdivision
	spr	656	-	312620	Period Subdivision
	sp1	657	-	312720	Place Subdivision
	msh	660	-	314020	NLM Subject Headings (MESH)
	abv	670	-	316020	NAL Agricultural/Biological Vocabulary
lsh		690	-	322020	Local Subject Headings
					OTHER ADDED ENTRIES
apsa		700	g	340040	Alternative (Personal Surname First) Entry
apfa		700	a	340041	Alternative (Personal Forename First) Entry
apma		700	b	340042	Alternative (Personal Multiple Surname) Entry
apna		700	c	340043	Alternative (Personal Name of Family) Entry
apss		700	d	340044	Secondary (Personal Surname First) Entry
apfs		700	e	340045	Secondary (Personal Forename First) Entry
apms		700	f	340046	Secondary (Personal Multiple Surname) Entry
apns		700	g	340047	Secondary (Personal Name of Family) Entry
apsn		700	h	340050	Analytical (Personal Surname First) Entry
apfn		700	i	340051	Analytical (Personal Forename First) Entry

TYPING TAG FILE ITEM NO.

Inforonics (if different)	MARC	MARC Tag Indicator	Inforonics Octal	
apmn		700 j	340052	Analytical (Personal Multiple Surname) Entry
apnn		700 k	340053	Analytical (Personal Name of Family) Entry
acna		710 @	342040	Alternative (Corporate Name) Entry
acpa		710 a	342041	Alternative (Corporate Place) Entry
acsa		710 b	342042	Alternative (Corporate Surname) Entry
acns		710 d	342044	Secondary (Corporate Name) Entry
acps		710 e	342045	Secondary (Corporate Place) Entry
acss		710 f	342046	Secondary (Corporate Surname) Entry
acnn		710 h	342050	Analytic Name Entry
acpn		710 i	342051	Analytic Place Entry
acsn		710 j	342052	Analytic Surname Entry
amna		711 @	342140	Alternative (Meeting or Conference Name) Entry
ampa		711 a	342141	Alternative (Meeting or Conference Place) Entry
amsa		711 b	342142	Alternative (Meeting or Conference Surname) Entry
ar.ns		711 d	342144	Secondary (Meeting or Conference Name) Entry
amps		711 e	342145	Secondary (Meeting or Conference Place) Entry
amss		711 f	342146	Secondary (Meeting or Conf. Surname) Entry

TYPING TAG

FILE ITEM NO.

Inforonics (if different)	MARC	MARC Tag Indicator	Inforonics Octal	
amnn		711 h	342150	Analytic (Meeting or Conf. Name) Entry
ampn		711 i	342151	Analytic (Meeting or Conf. Place) Entry
amsn		711 j	342152	Analytic (Meeting or Conf. Surname) Entry
(following the above added entry items)	acf	720 0	344020	Corporate Name with Form Subheading
	#	248 -	350020	Title
	aut	730 -	346020	Uniform Title Heading
	atd	740 -	350020	Title Traced Differently
	ano	753 -	352320	Proper Names Not Capable of Authorship
				SERIES ADDED ENTRIES
	dps	800 0	400000	Personal Name Surname First
	dpf	800 1	400001	Personal Name Forename First
	dpm	800 2	400002	Personal Name Multiple Surname
	dpn	800 3	400003	Personal Name Name of Family
dcn	810 0	402000	Corporate Name	
dcp	810 1	402001	Corporate Place	
dcs	810 2	402002	Corporate Surname	
dmn	811 0	402100	Meeting or Conference Name	
dmp	811 1	402101	Meeting or Conference Place	

TYPING TAG		FILE ITEM NO.		
Inforonics (if different)	MARC	MARC Tag Indicator	Inforonics Octal	
(following the above added entries)	dms	811 2	402102	Meeting or Conference Surname
	#	848 -	403020	Title
	dti	840 -	410020	Title

OCTAL CONFIGURATION RELATIONSHIP OF 7094 - PDP - MARC II

7094	PDP	MARC I	7094	PDP	MARC II
00	12	20	54	54	12
01	01	21	55	55	37
02	02	22	56	56	05,33
03	03	23	57	57	76
04	04	24	60	20,32,75	00,40
05	05	25		00,37	
06	06	26	61	21	17
07	07	27	62	22	63
10	10	30	63	23	64
11	11	31	64	24	65
12		74	65	25	66
13	13	35	66	26	67
14	14	07	67	27	70
15	15	32	70	30	71
16		36	71	31	72
17	74	73	72	76	02
20	35,60	13	73	33	14
21	61	41	74		10
22	62	42	75		75
23	63	43	76	17	
24	64	44	77	77	77
25	65	45			
26	66	46			
27	67	47			
30	70	50			
31	71	51			
32	72	06			
33	73	16			
34	34	11			
35		03			
36	16	34			
37					
40	40	15			
41	41	52			
42	42	53			
43	43	54			
44	44	55			
45	45	56			
46	46	57			
47	47	60			
50	50	61			
51	51	62			
52	52	01			
53	36, 53	04			

Exhibit 2. Octal Configuration Relationship of 7094 - PDP- MARC II

There are six (6) phases of system development.

- 3.1 Initial File Building
- 3.2 Selection
- 3.3 Acquisition
- 3.4 Cataloging
- 3.5 Update
- 3.6 Circulation

The following sections describe what has been completed and what remains to be done.

3.1 Initial File Building

The building of a Master Inventory File upon which the Circulation System operates is accomplished by using tapes from Inforonics as a data base.

These tapes from Inforonics contain data for all the monographs in the Library. Each monograph is identified by an F number in the format FYY-XXXXX, as F69-00035, where YY is the year of acquisition and XXXXX is the acquisition number. The tape format and description is in "Interim Report, 15 February 1968", by Inforonics. Briefly, each tape record contains all the data for a simple monograph, including F number, TITLE, AUTHOR, LANGUAGE, COLLATION, IMPRINT, CIASS NUMBER, etc.

The procedure for generating tapes is as follows:

The Library fills out forms (see Exhibit 3, page 3-10) for each volume. Inforonics takes the forms and types them onto paper tape. Magnetic tapes are generated from the paper tapes, and these are the tapes used for the initial file creation. Further processing is done by Inforonics to create labels and other outputs for the Library.

Please see Figure 1, page 3-2, for the flowchart of the Initial File Building Process.

Program "COMBO" is used to generate an input tape for the Master Inventory File from an "Inforonics" tape. This program consists of a series of subroutines, which originally were independent programs, developed as more was understood about what had to be done. These routines perform the following functions:

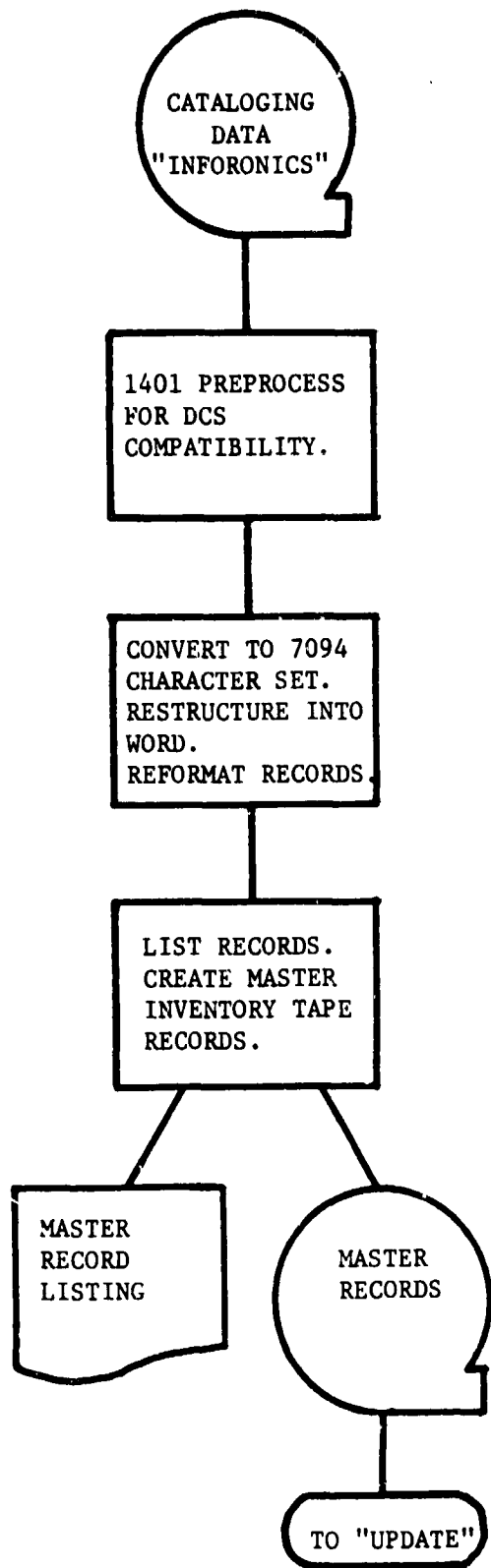


Figure 1. Initial File-Building Phase

1. Convert PDP characters to 7094 characters
2. Restructure character-oriented data into word-oriented data
3. Reformat each record into directory and variable field format
4. List the records

The output tape from "COMBO" is read by the "UPDATE" program, described later, to add to the Master Inventory File.

Library personnel are expected to review these listings. The records are generally in F-number sequence, but there is no guarantee. They should be sorted prior to update. This has not been provided for, yet. So far, several tapes from Inforonics have been processed. The program "COMBO" is working well on the DCS. For purposes of building the Master File, it would be wise to restart with tape 1 from Inforonics. Occasionally, a data tape with bad records raised havoc with the DCS System. This tape would be returned and another copy requested.

3.2 Selection Phase

MARC II tape records are used in the selection phase. The MARC II tapes are delivered weekly, by subscription, from the Library of Congress. These weekly tapes contain cataloging data for those monographs cataloged by the Library of Congress during a preceding period.

The monographs pertain to all subjects. There are only certain subjects of interest to the AFCRL Research Library. These monographs can be identified by the first two characters of the class number. It is desired to scan the incoming MARC tapes and list those records having a class number which is of interest. These records of interest are to be stored on a 90-day revolving file. If a book is ordered from a vendor, the record is removed and put into a holding file until it is received, and the record goes into the Master Inventory File.

Please see Figure 2, page 3-4, for a chart of the selection phase.

Program "MARC-COMBO" is used in this phase. Like program "COMBO", it is a series of subroutines which used to be independent programs. These routines perform the following functions:

1. Convert the ASCII 6-bit characters to 7094 characters.
2. Restructure character-oriented data into word-oriented data.
3. Extract desired records by checking first characters of class numbers.
4. List the selected records and write them on tape.

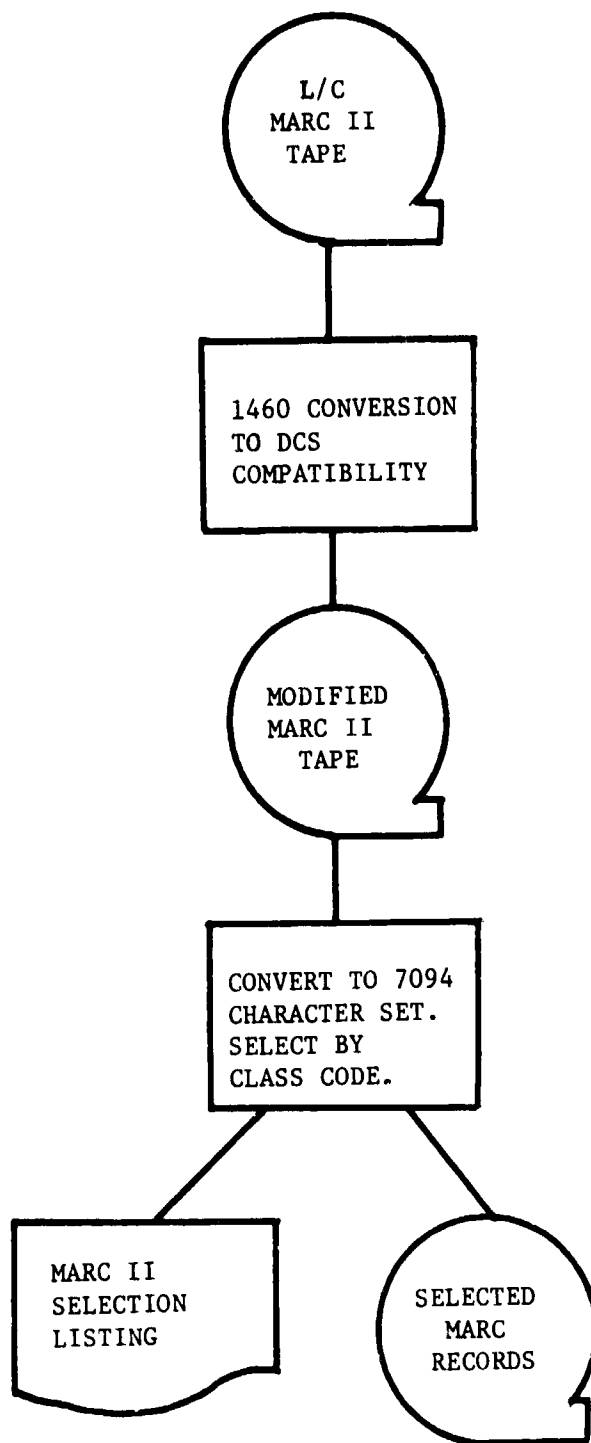


Figure 2. Selection Phase

This is the extent of the work done for this phase. Many MARC II tapes have been processed, and the program works fine. Errors encountered on MARC tapes are dumped without further processing.

It remains to develop routines to date the MARC II records and keep them in a revolving 90-day file. These records do not have F numbers, and will not have F numbers until, if and when, they are acquired. They do have "card numbers", and these are used as control or identity numbers. No provisions have been made for sorting any of these records. All of these programs must be converted to the CDC 6600 computer system.

3.3 Acquisition Phase

Nothing has been done in this area except for the preliminary MARC II processing in the selection phase.

The following things should be done:

1. Build routines for an ordering system.
2. Pick out MARC II records of books on order; place them in a holding file; add F numbers upon receipt of books; and add the records to the Master Inventory List.

3.4 Cataloging Phase

There has been no work done for this phase, and no discussions about requirements were had, except to state that the cataloging effort includes making labels and catalog cards.

3.5 Update Phase

This separate plan has been developed to allow for manipulating the Master Inventory File. Program "UPDATE" is used to add whole records to the file from cards or tape, and to change, add, or delete data within a record. "UPDATE" also prints out Master Listings.

Please refer to Figure 3, page 3-6, for a chart of "UPDATE".

The program performs the following functions:

1. Reads the current Master File.
2. Reads input tape records
3. Reads input records on cards
4. Checks for a match of F numbers on input cards or input tape against the Master File; files new numbers into the proper place; performs additions, changes, and deletions.
5. Writes new Master File.
6. Prints new Master File.

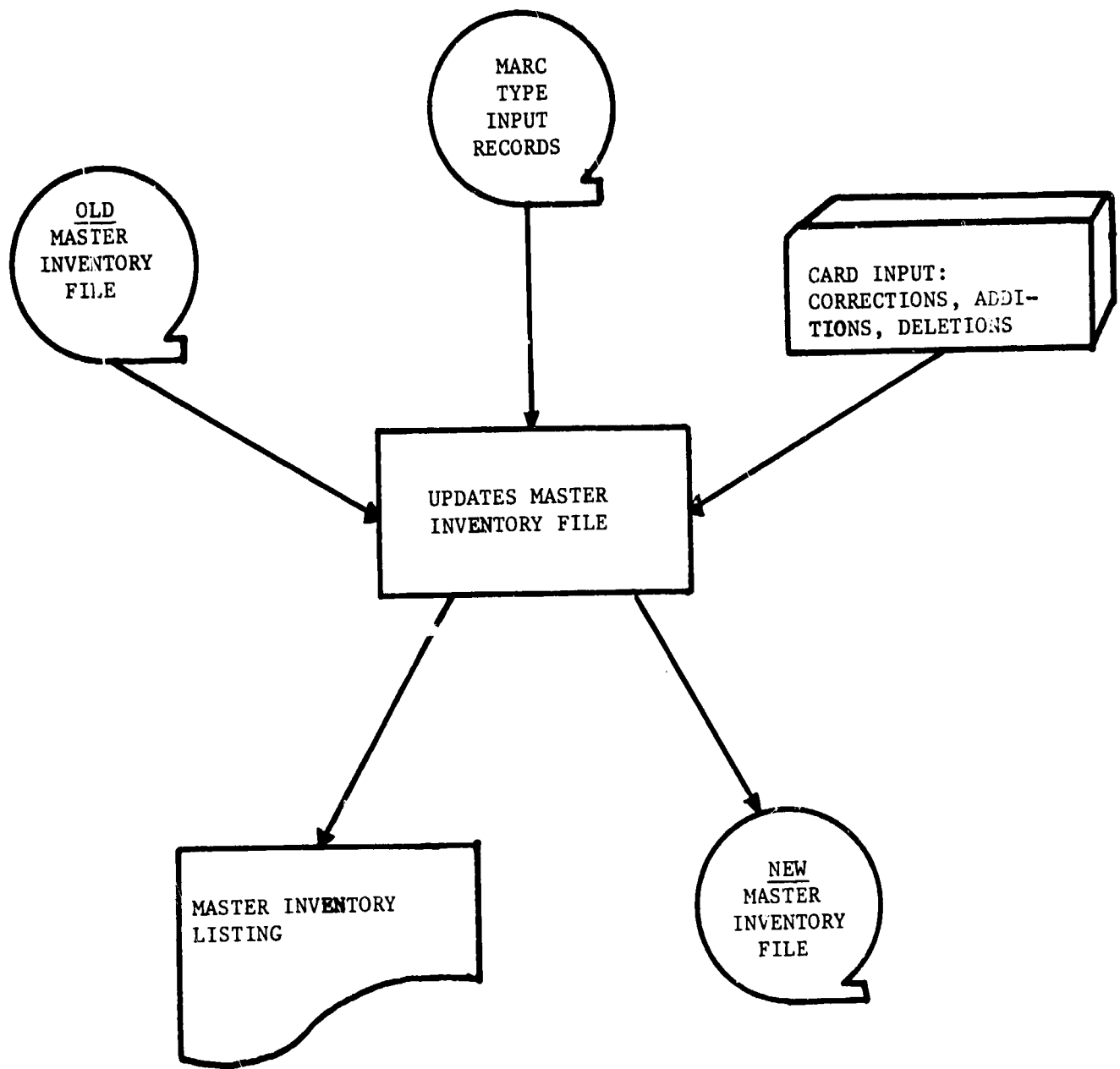


Figure 3. Update Phase

The Inforonics records have data for only one copy of each book in the Library. Since each copy of each book has a different F number, these must also be entered into the file. This is done by cards punched by the Library. These cards have an F number, class number (which shows the copy number, as QA 457 B1 C.2) and an REF entry, which contains a note to see another F number for the rest of the data.

There are two changes which should be made to this program.

1. Something must be done about printing out all the records. Although this program provides a complete record of the Master Inventory File, it will soon result in a massive pile of paper.

2. As of now, the program handles only "Inforonics" records. When MARC II tapes became available, it was discovered that the two character sets were not 100% compatible. In an effort to preserve as much of the original character sets as possible, it was decided to add an indicator to the MARC records to identify them as being a MARC character set. The Inforonics records and the Master Inventory records have no indicator. All that is required is to add a word to the front of these records. Then, the MARC, Inforonics and Master Inventory records will all be consistent.

3.6 Circulation Phase

In this phase, books borrowed and returned are monitored; the master borrower's list is printed; overdue books are indicated; and a history file of books withdrawn is prepared.

A card file of eligible borrowers is used. Each borrower has a borrower number. Each book in the library will eventually have an F number. When a book is withdrawn, a card is prepared with the borrower's number, the book number, and the date. A similar card will be used for books returned.

Please refer to Figure 4, page 3-8, for a flow diagram of the Circulation Phase.

Program "BORROWLST" prepares a listing of the Borrower's File. This is used to provide the library with current listings of the borrowers and their numbers. Library personnel keypunch new cards and assign the numbers.

Program "LIBCIRC" performs the following functions:

1. Reads in and stores all current transactions.
2. Reads borrower file and matches borrower number and adds borrower data to transaction records.

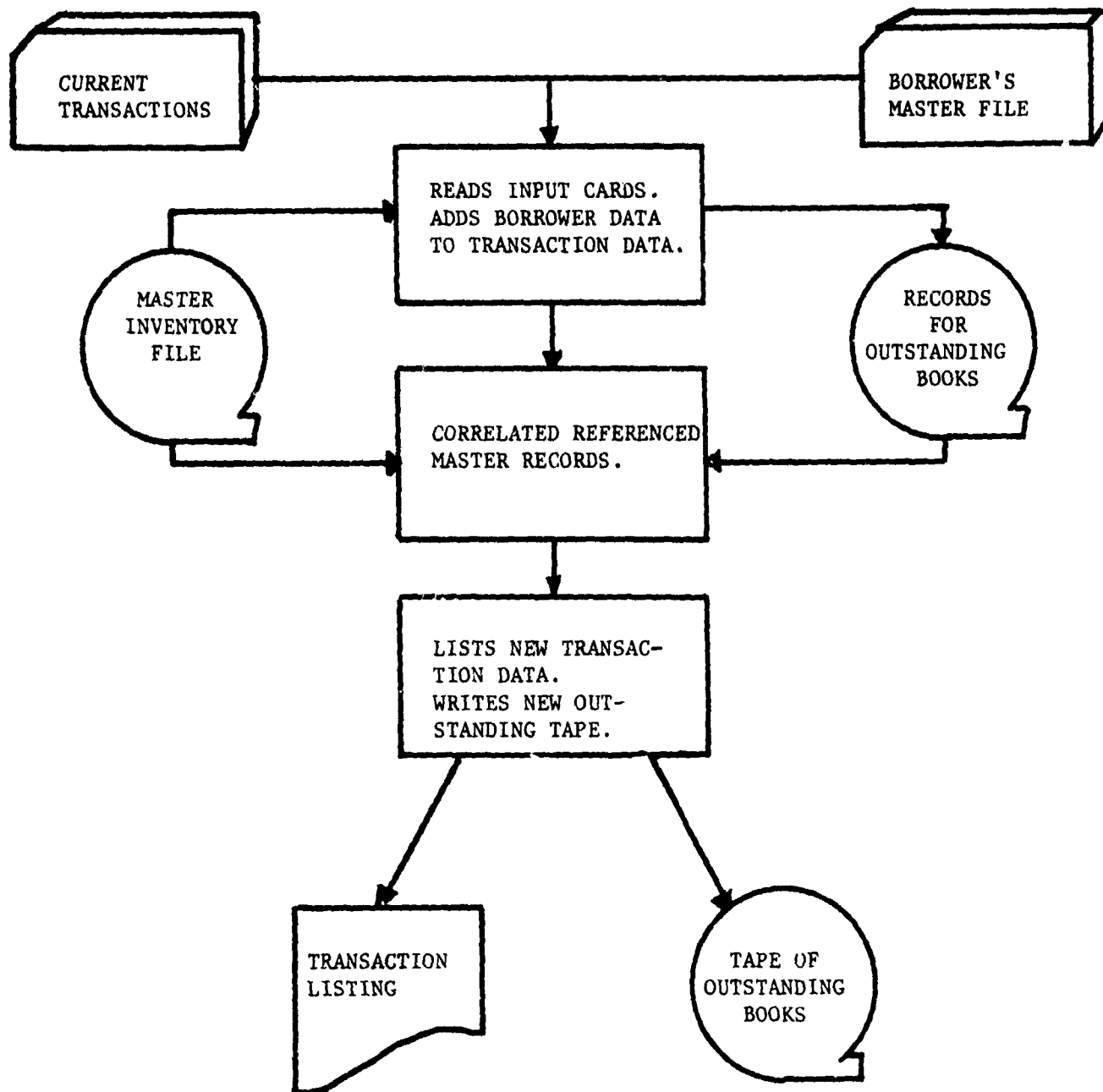


Figure 4. Circulation Phase

3. Reads Master Inventory for a match F number, and the data is written on unit 2. If another F number is referenced, that F number is saved for later processing.
4. Continues step 3 until all transactions have been read, and then rewinds the Master tape and reads the F number which was saved, and puts this data on unit 3.
5. Merges the data from units 2 and 3, and writes out these records for current books onto unit 4.

What remains to be done is to agree on a format for cards to denote books being returned so that these records will be dropped from the outstanding file; to develop a program to list the outstanding tape with flags for overdue books; and, to add all books withdrawn to a history tape of all books ever taken out.

As statistical data becomes required in the future, the history file will have all the data needed to develop these statistics, and appropriate programs will be needed. All computer programs described in this report were written under Problem Number 1680, and are available from the Analysis and Simulation Branch Problem Library.

SECTION 4

CONVERSION

During the contractual period, a number of 7094 DCS functional programs for handling serials were converted to the CDC 6600 system. The following SERIAL programs were converted and made operational:

LIBCOR
PRINTMORE
SUBSCRIBE
SCRIBE & CKLIST
LISORT 3
LISORT 4
CLERICAL
PUBLISHER

Any program from the above list which had been written in machine language was rewritten in Fortran IV prior to the total program conversion to the CDC 6600.

Some of the major problems encountered in converting programs from the Direct Couple System to the CDC SCOPE Operating System were as follows:

1. Word Size

Both the IBM 7000 series and the CDC 6600 computers are word-oriented machines; however, CDC 6600 word size is 60 bits long as opposed to the IBM 7000 series word size of 36 bits. This difference in word size causes some specific problems in shifting, masking and floating point formats.

A. Shifting

All shift functions written in Fortran for the DCS must be converted to the CDC shift functions or a logical procedure must be substituted to perform a comparable manipulation; and this modification must account for the larger word size.

B. Masking

All masking instructions written in FORTRAN on the DCS must be modified to conform with the CDC syntax. The actual masks themselves must also be modified to conform with the larger word size of the CDC 6600.

C. Floating-Point Format

The floating-point format on the CDC 6600 is composed of a 48-bit integer coefficient with an 11-bit biased exponent. The IBM 7000 series computer uses a 27-bit integer coefficient with an 8-bit biased exponent. Thus the CDC will show more significance than the 7094. IBM floating point words on tape must be converted to conform to the CDC floating point format; consequently, routines were written in FORTRAN and COMPASS to do this conversion.

2. BCD Octal Code

The octal character code on the IBM 7000 series computers and the CDC 6600 are different; therefore, it is necessary to convert all IBM BCD characters to the equivalent CDC octal character code. When there is no corresponding character for the IBM code in CDC, a substitute must be made. All FORTRAN DCS Data statements which express BCD characters in their octal equivalents must be converted to the CDC octal character equivalent.

3. Control Card Language

All DCS IBSYS control cards must be replaced by CDC SCOPE control cards for the program to operate under the SCOPE Operating System. The SCOPE Operating System has an extensive control card language and can perform functions on-line which with the DCS had to be done off-line using the IBM 1460. One advantageous option which the SCOPE control card language gives the programmer is the ability to perform predefined utility routines if his program should terminate in an error mode. CDC uses a FORTRAN Program card at the beginning of the main FORTRAN routine which specifies all I/O devices used during execution, and sets up the logical and physical linkages for I/O units.

4. Non-Standard Tapes

The IBM 1460 was used to preprocess non-standard DCS tapes so they could be processed by the IBM 7094 II/7044 DCS. The IBM 1460 was used to pack the data into 36-bit words and to ensure correct parity for the DCS system. The CDC 6600 eliminates the need for the IBM 1460, and can process non-standard CDC tapes directly. Routines have been written for the CDC 6600 to read the following non-standard CDC tapes on the CDC 6600:

1. Library data base tapes created on a PDP-9
2. Library of Congress MARC II tapes created on an IBM 360/30
3. Unblocked IBM 7094 II/7044 DCS tapes

The data on these tapes also had to be converted to CDC-character and format codes. Routines were written to process blocked IBM 7094 II/7044 DCS tapes utilizing the routine TCONVT.

5. Machine Language

The DCS Machine Language (MAP) and the SCOPE Operating System Machine Language (COMPASS) are not compatible; i.e. MAP cannot be easily converted into COMPASS. It was found to be much easier to convert MAP routines into CDC FORTRAN during the conversion of AFCRL Research Library DCS Programs to the CDC 6600. The use of CDC FORTRAN is possible in these conversions because of the availability of functions in CDC FORTRAN to perform the functions that were necessary in MAP on the IBM 7094.

6. Overlays

A program can be divided into parts for an overlay when it exceeds available memory if each part is independent and can be called and executed as necessary. Each part (overlay) consists of a single main program and any necessary subprograms. Differences in control cards, END-OF-FILE recognition in Input/Output routines, labelled COMMON statements and item 1- through 5 mentioned previously all need to be considered when converting from the IBM 7000 series to the CDC 6600.

Although the SERIAL programs presently are independent of the programs necessary to create and maintain a master inventory list of monographs using the MARC II format, they could be easily incorporated into the overall system for automating the research library.

SECTION 5

SUPERVISORY CONTROLLED AUTOMATED LIBRARY SYSTEM (SCALS)

During the contractual period a total system design to automate the AFCRL Research Library was presented which would reflect both the present requirements of the Research Library and future requirements when remote terminal real-time processing becomes available.

The overall view of the Supervisory Controlled Automated Library System (SCALS) was written and presented to the library personnel in August 1970, see Section 8, Appendix. The report described how the various library phases defined as Selection, Acquisition, Cataloging and Circulation could be automated through the use of a Supervisory Controlling Program. The report described in detail, from a data processing point of view, the various functions of each phase, along with the necessary input and projected output to implement each phase.

It included a list of data processing files needed and a brief description of each. Flow charts illustrated how each phase could be controlled through the use of an English-oriented language used as input to the controlling program. Various examples showing the numerous options available for use, as well as projected by-products of the system were given in the form of statistical and management-type reports. Finally, a tentative long-range schedule was provided to project the implementation of each phase of the total library system. This presentation is included as an Appendix to this report.

SECTION 6
CONTRIBUTORS

The writers express their thanks to Miss Eunice C. Cronin, Analysis and Simulation Branch (SUYA), AFCRL Computation Center, Air Force Cambridge Research Laboratories, and to Mr. Austin A. Almon, Jr., Contract Monitor, whose technical guidance and experience were invaluable in the preparation of this report.

SECTION 7

REFERENCES

1. "Interim Report, 15 February 1968," by Inforonics, Inc.
2. "Automated Library System" (unpublished presentation, Analysis and Computer Systems, Inc., August, 1970).

APPENDIX

AUTOMATED LIBRARY SYSTEM

TABLE OF CONTENTS

	PAGE NUMBER
LIST OF ILLUSTRATIONS	8-3
I. INTRODUCTION	8-4
II. AUTOMATED LIBRARY PHASES	8-5
III. GENERAL FLOW DIAGRAMS	8-7
IV. DESCRIPTION OF PHASES	8-6
. Supervisor	8-6
. Selection	8-15
. Acquisition	8-15
. Cataloging	8-15
. Circulation	8-15
. Statistics	8-15
V. AUTOMATED LIBRARY FILES	8-17
VI. YEARLY TASK DESCRIPTION	8-18

LIST OF ILLUSTRATIONS

<u>FLOW DIAGRAMS</u>		<u>PAGE NUMBER</u>
Figure 1	General Flow	8-7
Figure 2	Supervisor	8-8
Figure 3	Selection	8-9
Figure 4	Acquisition	8-10
Figure 5	Cataloging	8-11
Figure 6	Circulation	8-12
Figure 7	Statistics	8-13

AUTOMATED LIBRARY SYSTEM

I. INTRODUCTION

This report is an initial overall view of the automated library system, encompassing the various library phases defined as selection, acquisition, cataloging, and circulation.

This report will outline two data processing phases in addition to the library phases, i.e., the supervisor and statistical phases, which will enhance the operation of the automated library system from a user's standpoint.

This report will describe, from a data processing point of view, the various functions of each phase, along with the necessary input and projected output to implement each phase. A list of data processing files and a breakdown of each phase is included, along with a brief description of each.

Finally, this report will show tentative yearly task descriptions necessary to implement the automated library system over a five-year period. Future goals of the automated library system are included for consideration.

a. AUTOMATED LIBRARY PHASES

PHASE	PURPOSE	TYPE
SUPERVISOR	CONTROL PROGRAM	DATA PROCESSING
SELECTION ACQUISITION CATALOGING	CREATE AND EDIT MASTER INVENTORY LIST (MIL)	LIBRARY
CATALOG INQUIRY	UTILIZATION OF MIL	LIBRARY
STATISTICS	MANAGEMENT INFORMATION	DATA PROCESSING

IV. DESCRIPTION OF PHASES

Supervisor

The Supervisor is a control-type program with the following functions:

1. Accepts and interprets users' inquiries and/or requests concerning the data processing phases of the automated library system.
2. Retrieves and transfers control to routines required to execute these inquiries and/or requests.
3. Ensures legitimate access to the system safeguarding against unauthorized use.
4. Maintains proper recording of events and data necessary for providing management information service.

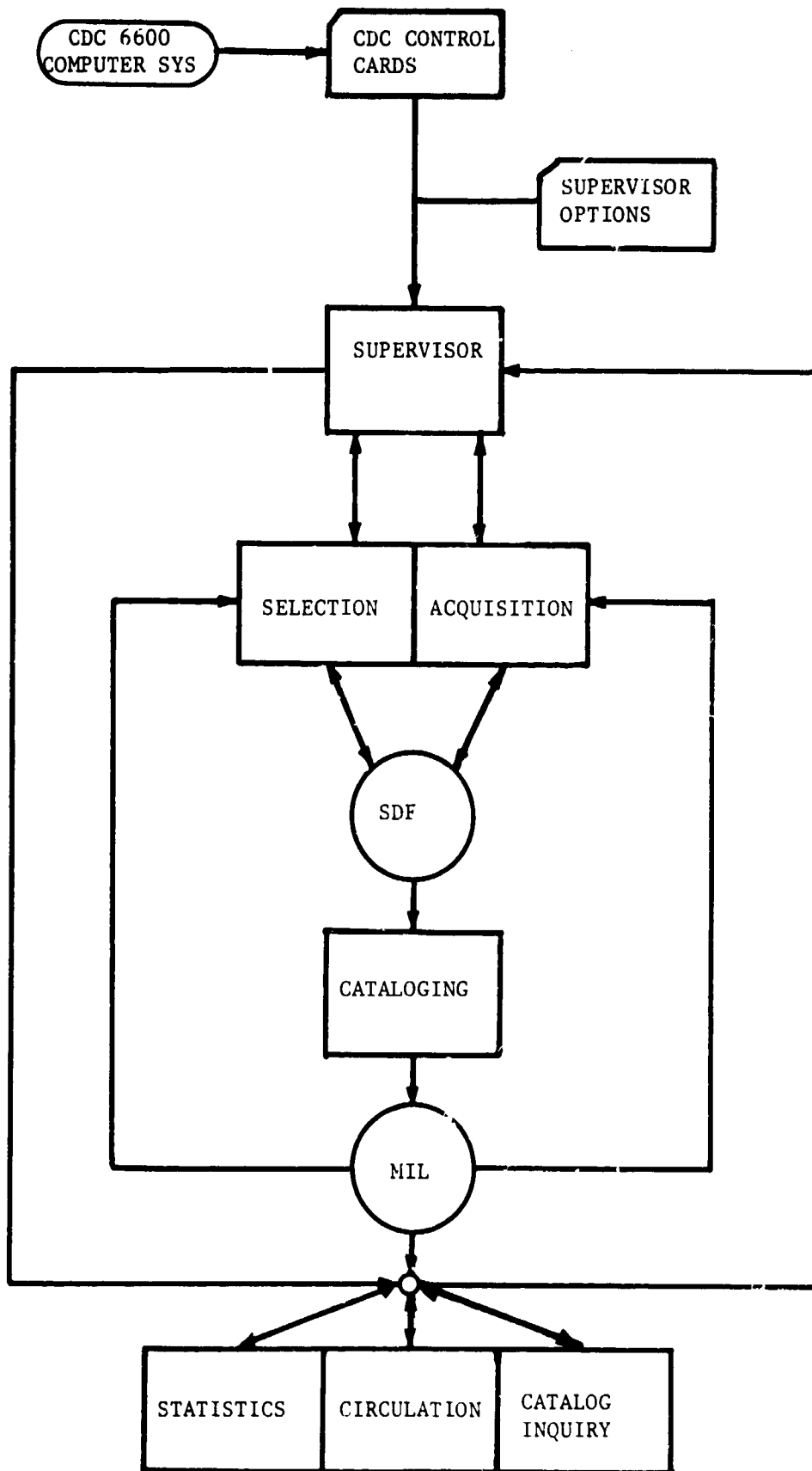


FIGURE 1

SUPERVISOR

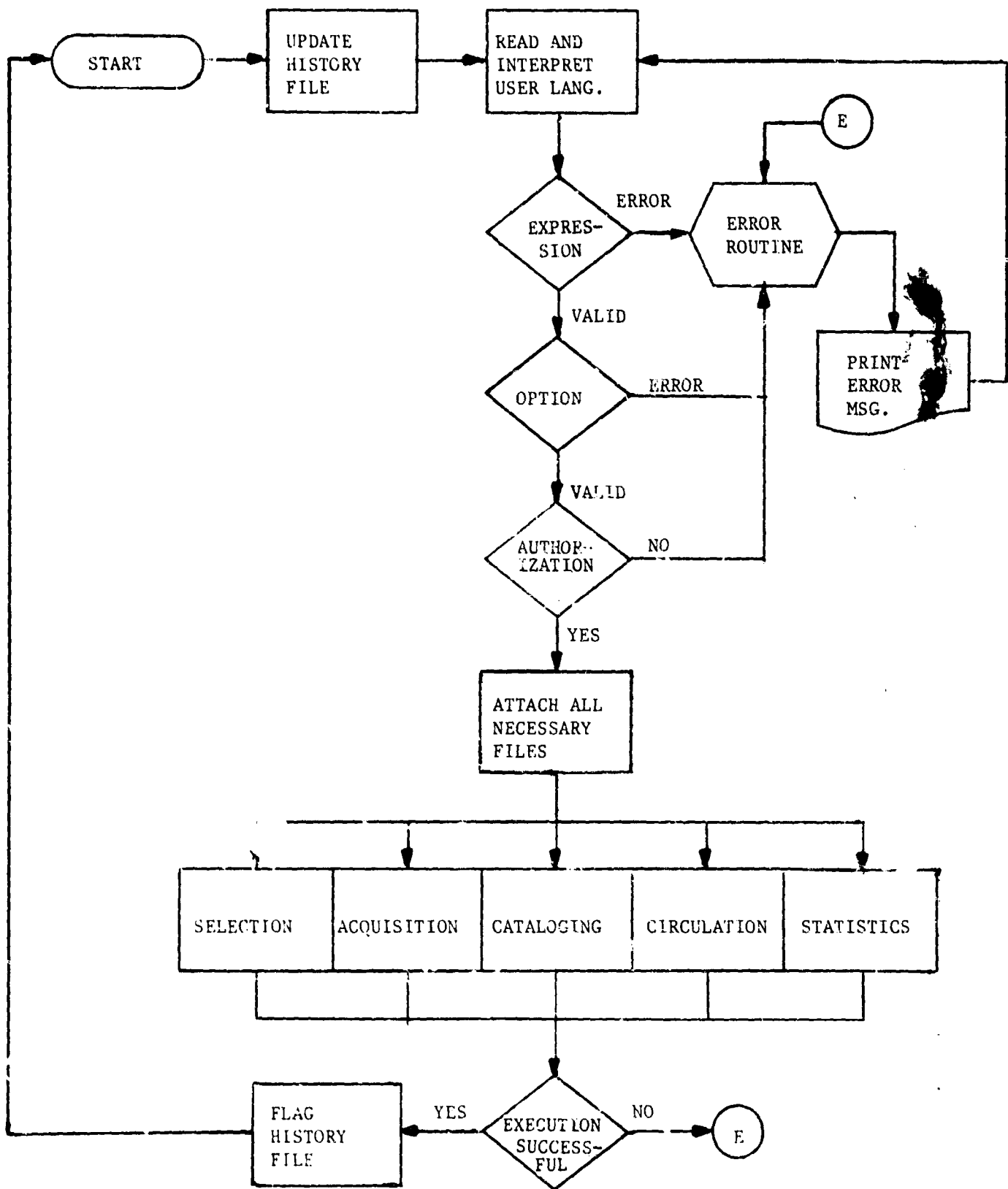


FIGURE 2

SELECTION

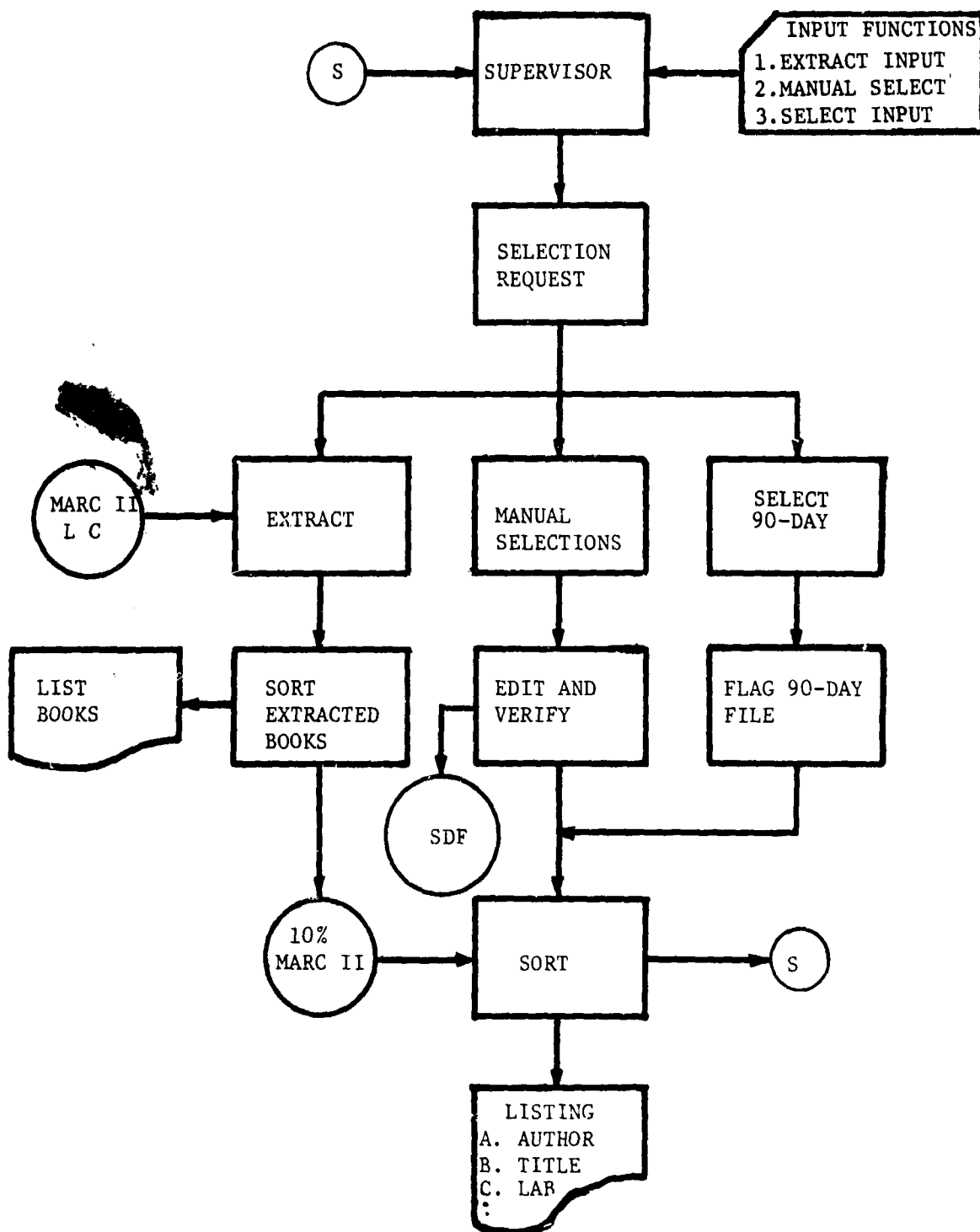


FIGURE 3

ACQUISITION

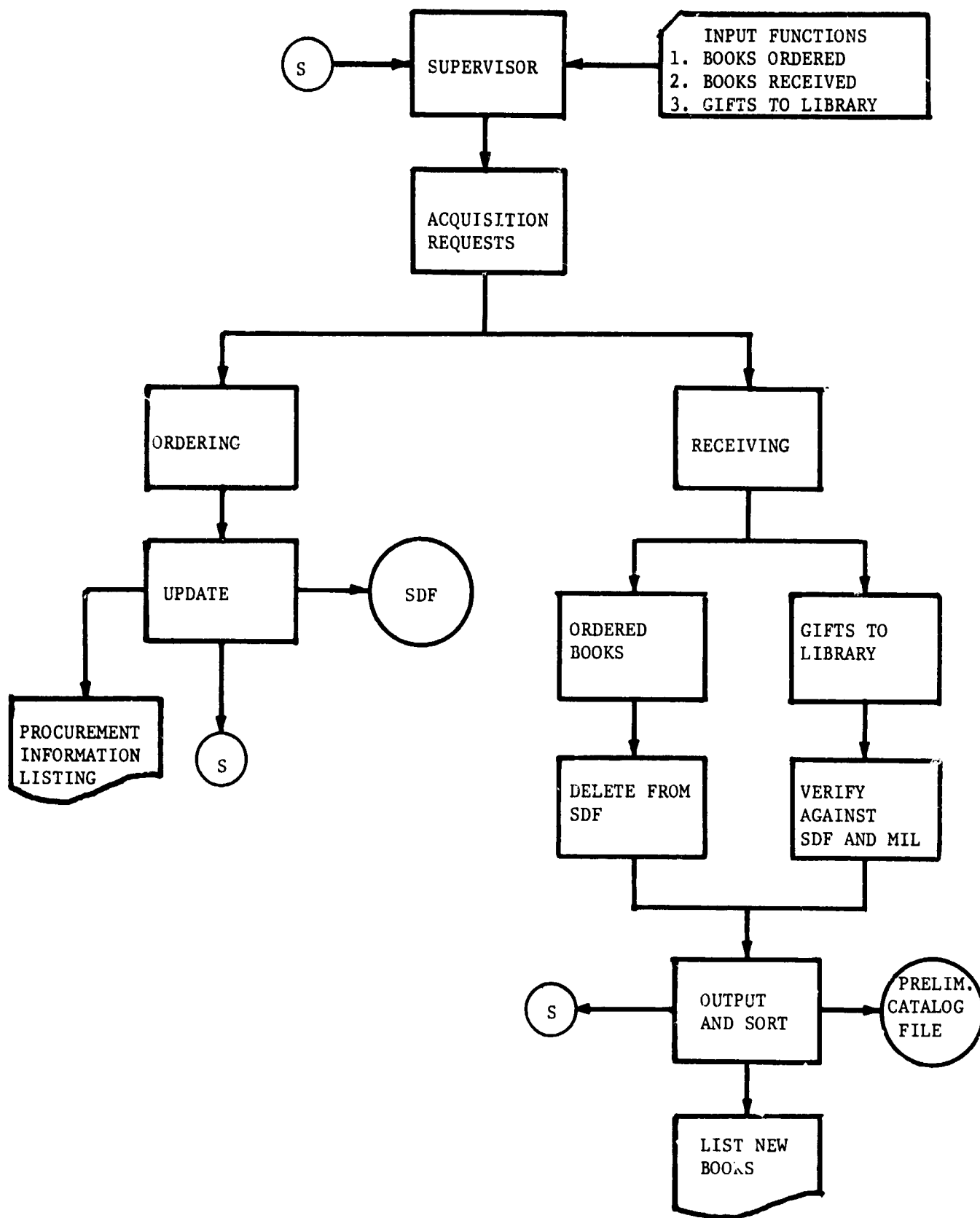


FIGURE 4

CATALOGING

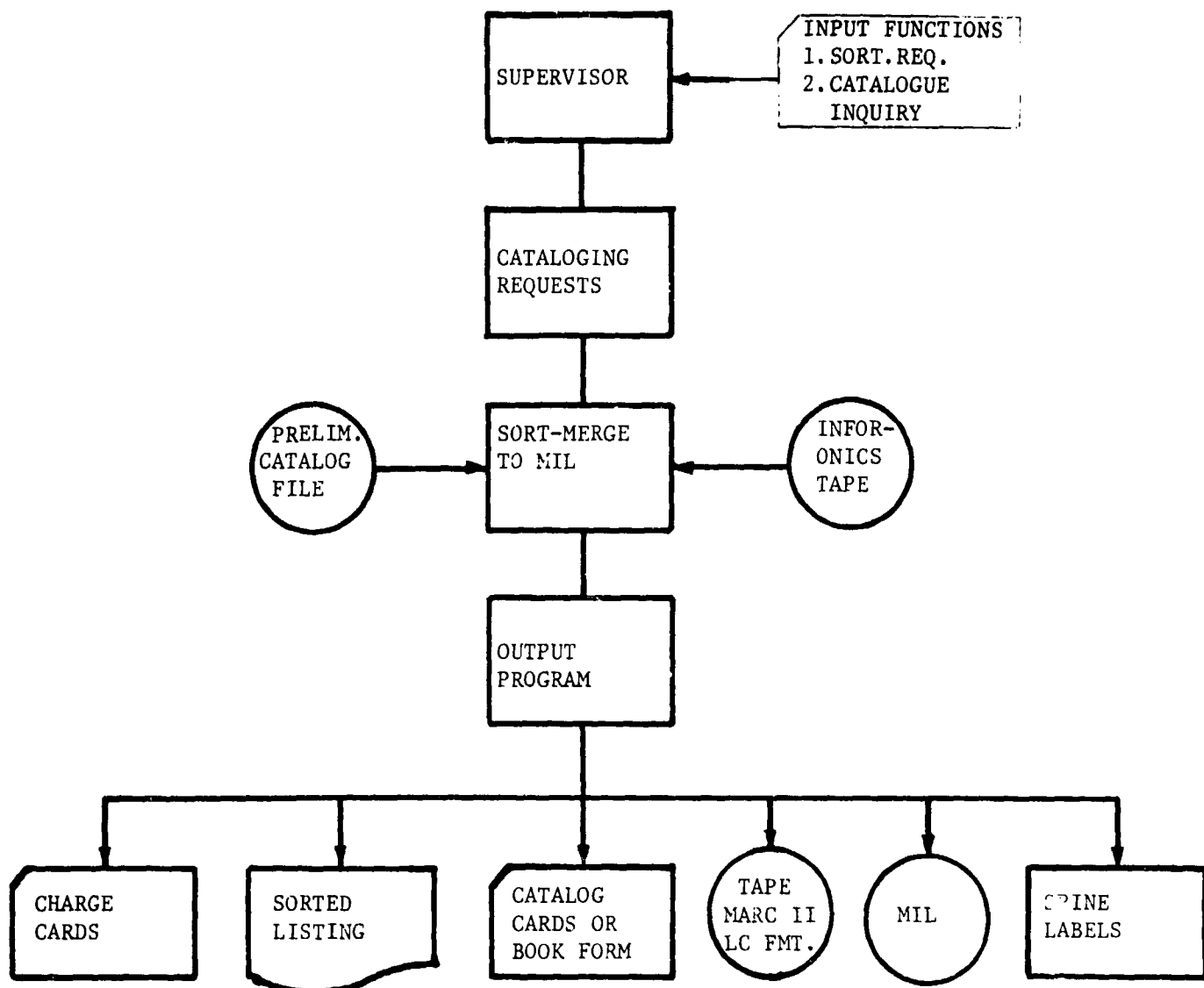


FIGURE 5

CIRCULATION

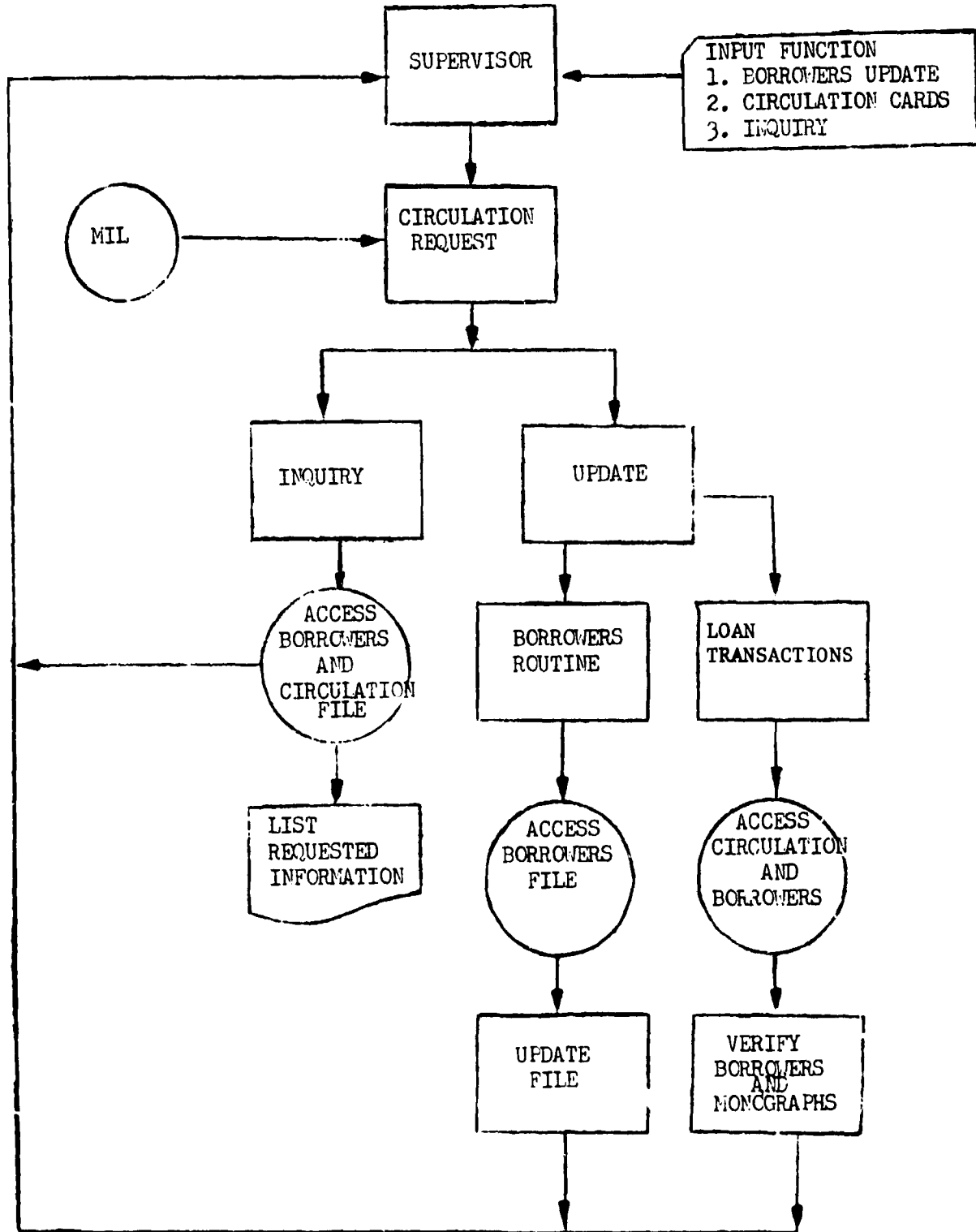


FIGURE 6

STATISTICS

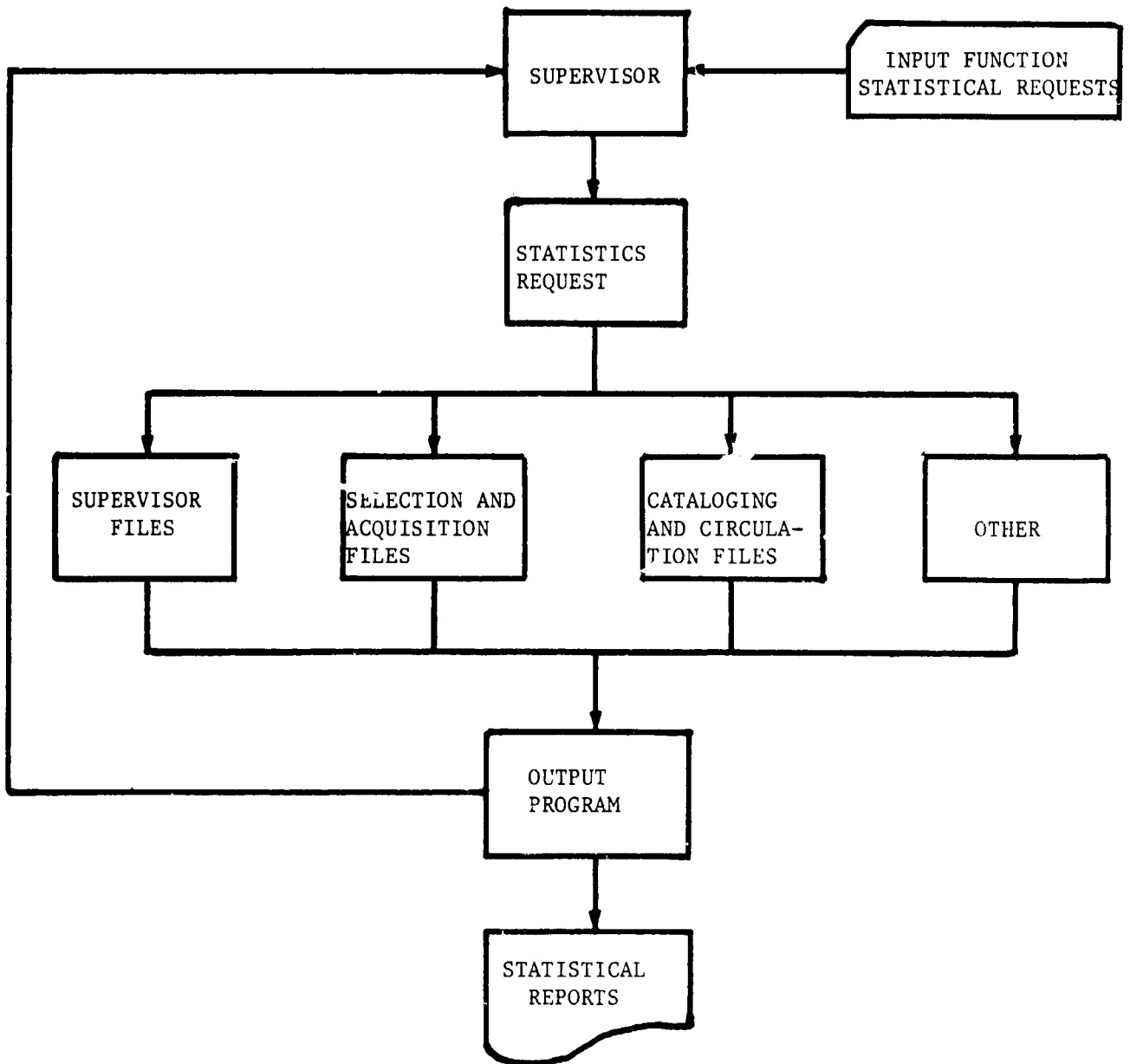


FIGURE 7

Example of Supervisor Use:

CDC 6600 Control Cards

Job control card
RPACK (operator to mount private disk pack)
REQUEST (attach all files used by library system)

.
.
.

7/8/9 (end of record)

Supervisor Control Cards

NAME	=	John Doe
DATE	=	062570
OPTION	=	INQUIR
REQUEST	=	FDCIRC
OUTPUT	=	LIST (author, A, M)

Selection

Selection of desired monographs from weekly Library of Congress Tape (MARC II), and the building of 90-day file for ordering and distribution to borrowers.

Acquisition

Once monographs have been selected and approved for purchase, we enter the acquisition phase, where we basically create and edit the master inventory file of received monographs.

Cataloging

Format, Sort, Inquire and Integrate cataloging data from master inventory list to be displayed by CRT, PRINTER and cards.

Circulation

Receive and interpret transactions with monographs and borrowers keeping the circulation and borrowers' files updated at needed periods with reference to current master inventory listing.

Statistics

The statistical phase has access to all automated library files and can produce reports as requested by user. Some examples of reports are as follows:

1. Frequency
 - a. Lab
 - b. Monograph
 - c. Borrower
2. Order by Subject Content
3. Monograph Acquisition
4. Circulation Per-Month, Year, Week . . .
5. Etc

V. AUTOMATED LIBRARY FILES

NAME	TYPE	DESCRIPTIONS
1. TRANSACTION HISTORY	PRINTED OUTPUT	History of all transactions in using automated library
2. MARC II	TAPE	Input from Library of Congress
3. 10% MARC II	DISK	Temporary file of extracted monographs from MARC II LC
4. EXTRACT LISTING	PRINTED OUTPUT	Sorted listing of all extracted monographs from MARC II
5. PROCUREMENT DOCUMENT	PRINTED OUTPUT	Order form for purchasing selected monographs
6. 90 DAY	DISK AND PRINTED OUTPUT	Weekly updated file of monographs extracted from Library of Congress
7. 83 DAY	DISK AND PRINTED OUTPUT	Flag on 90-day file, notifying user one week before removal from 90-day file
8. SDF	DISK AND PRINTED OUTPUT	Source data file - file of monographs selected from 90-day file
9. RECEIVED	PRINTED OUTPUT	File of monographs received after order
10. SORT-90	PRINTED OUTPUT	Sorted 90-day file
11. MIL	DISK	Master inventory list - file of all monographs on hand in library
12. NEW MIL	PRINTED OUTPUT	Listing of monographs as added to MIL
13. BORROWERS	DISK AND PRINTED OUTPUT	File of all eligible borrowers
14. CIRCULATION	DISK AND PRINTED OUTPUT	File of all monographs in circulation
15. SORT-MIL	DISK AND PRINTED OUTPUT	Sorted MIL by author, title, subject, Library of Congress No., etc....
16. STAT	DISK, TAPE, PRINTED OUTPUT	Statistics of all above files
17. ERRORS	PRINTED OUTPUT	Listing of error conditions as needed. Can occur throughout system.

VI. YEARLY TASK DESCRIPTIONS

YEAR	DESCRIPTION
1	<ol style="list-style-type: none">1. Basic MIL File - constructed from MARC II, Inforonics and local data.2. Supervisor - general structure to implement building of MIL.3. Cataloging - limited fixed listings of MIL.4. Documentation - start of detailed documents of programs and operating procedures.5. Detailed Documented Plans for next 3+ years.

YEAR	DESCRIPTION
2	<ol style="list-style-type: none">1. Supervisor - limited user language with a fixed input.2. Serials - start of serials programs with plans to incorporate into automated library.3. Cataloging - various additional sorted listings of MIL.4. Circulation - initial program to create and edit the borrowers' file.5. Documentation - detailed documents of programs, operating procedures and user language manuals.

YEAR	DESCRIPTION:
3	<ol style="list-style-type: none"> 1. Supervisor - expansion of user language with variable input and new options. 2. Serials - incorporate serials and monographs into one automated system. 3. Circulation - increase file to include a basic circulation file automated with borrowers' file. 4. Statistics - start of basic statistics files. 5. Documentation - detailed documents of programs, operating procedures and user manuals.

YEAR	DESCRIPTION:
4	<ol style="list-style-type: none"> 1. Supervisor - expansion of user language with variable input and new options. 2. Statistics - further expansion of statistical possibilities and implementation of results. 3. Procurement document - write procurement programs and build into automated system. 4. Remotes - start study of remote terminals for library, including real-time system. 5. Documentation - detailed documents of programs, operating procedures and user manuals.

YEAR	DESCRIPTION:
5	<ol style="list-style-type: none"><li data-bbox="720 339 1450 397">1. Supervisor - expansion of user language with variable input and new options.<li data-bbox="720 436 1450 529">2. Statistics - further expansion of statistical possibilities and implementation of results.<li data-bbox="720 564 1450 653">3. Documentation - detailed documents of programs, user manuals and write-ups of the system to date.