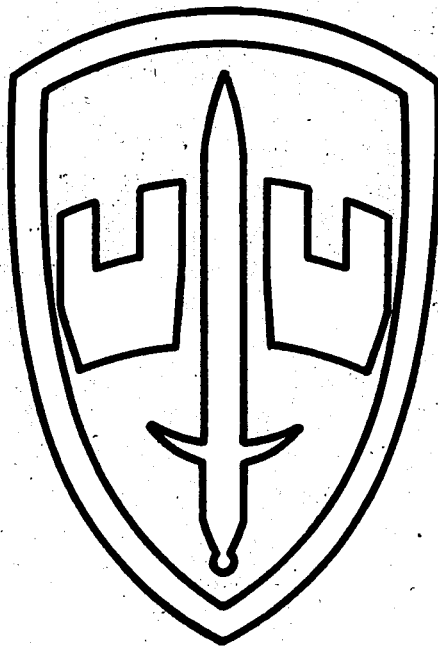




MACV
AUTOMATED DOCUMENT
STORAGE & RETRIEVAL
SYSTEM

MANUAL



SAIGON, VIET NAM

**INTELLIGENCE DATA HANDLING SYSTEMS
(IDHS)**

**AUTOMATED DOCUMENT STORAGE
AND
RETRIEVAL SYSTEM MANUAL
FOR**

U.S. MILITARY ASSISTANCE COMMAND, VIETNAM

A. E. PALMERLEE

**Second Edition
February 1968**

Prepared Under

U.S. Navy Contract N00600-67-C-0884

FMA Inc. 5730 Arbor Vitae Street, Los Angeles, California 90045

HEADQUARTERS
UNITED STATES MILITARY ASSISTANCE COMMAND, VIETNAM
Office of the Assistant Chief of Staff, Intelligence
APO San Francisco 96222

MACJ261

SUBJECT: Automated Document Storage and Retrieval System

TO: SEE DISTRIBUTION

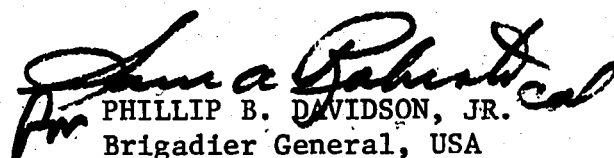
1. Forwarded for your information is the Automated Document Storage and Retrieval System Manual.

2. This system is part of the Department of Defense Intelligence Data Handling System's (IDHS) Automated Document Storage and Retrieval System, set up at major headquarters using the AN/FYQ-43 Document Data Processing System, manufactured by FMA, Inc., and known commercially as the FileSearch system.

3. This Manual has been prepared by Mr. A. E. Palmerlee, FMA's on-site Systems Analyst, as a comprehensive operating manual for personnel assigned to the system at MACV and for other commands retrieving documents from MACV-generated film. This Second Edition includes changes and additions to the procedures set forth in the First Edition. Individual pages are dated to show their latest revision.

4. Comments or recommendations for changes or additions to the Manual should be submitted to this Headquarters, ATTN: MACJ261 (CDEC).

1 Incl
as


PHILLIP B. DAVIDSON, JR.
Brigadier General, USA
Assistant Chief of Staff, J2

DISTRIBUTION:

6-DIA (DIAMS-1)
2-COMNAVINTCOM (NIC-2)
2-ACSI
5-CINCPAC (J216)
5-CINCPACAF (DIRI-P)
5-CINCUSARPAC (G2)
1-CINCUSAREUR (IDHS)

1-CG, USARV (G2)
1-CG, 7th AF (DI)
1-COMNAVFORV (N2)
2-CG, FMFPAC (G2)
1-COMUSMACTHAI (G2)
2-ACofS, J2
50-MACJ253
2-MACJ26
1-MACJ2 Library
50-MACJ261 (CDEC)

TABLE OF CONTENTS

	<u>Page</u>
Chapter I. THE FILESEARCH SYSTEM	1-1
A. Storage and Retrieval Processes	1-1
B. The MACV System	1-15
C. Related Publications	1-19
Chapter II. INDEXING	2-1
A. General	2-1
B. Indexing in Depth	2-2
C. Tags	2-6
D. Continued Descriptors	2-8
E. Factored and Phrased Descriptors	2-9
F. Subject Modifiers	2-11
G. Encoding Procedures	2-12
H. The Index Form	2-13
Chapter III. INDEXING PROCEDURES	3-1
A. Describing the Document (Factored Field)	3-1
B. Specific Document Types	3-9
C. Biographic and Personality Information	3-26
D. Geographic Areas	3-33
E. Installations and Organizations	3-46
F. Military Information	3-49
G. Medical Information	3-60
H. Việt Cộng Financial and Economic System	3-62
I. Việt Cộng Intelligence Reports	3-66
J. Letter Box Numbers	3-67
K. Keyword Indexing - The G-Tag	3-69
L. Modifiers for Special Types of Documents	3-71
Chapter IV. INPUT	4-1
A. Keypunching Index Cards	4-1
B. Recording	4-5
C. Film Developing	4-8
D. Checking Film Quality	4-21

TABLE OF CONTENTS (Continued)

	<u>Page</u>
Chapter V. FILE MAINTENANCE	5-1
A. Reel Identification System	5-1
B. Duplicating the Master Film	5-3
C. File Expansion	5-4
D. Film Dissemination	5-10
E. Film Cleaning	5-11
Chapter VI. RETRIEVAL TECHNIQUES	6-1
A. Introduction	6-1
B. Request Logic	6-3
C. The ISC	6-26
D. Geographic Requests	6-31
E. Requests Using Names of People	6-37
F. Date Descriptors	6-41
G. Programming Requests	6-44
H. The Hit Counter	6-46
I. The Role of the Retrieval Analyst	6-47
J. The Retrieval Check List	6-49
Chapter VII. MISCELLANEOUS PROCEDURES	7-1
A. Cleanliness of File Search Areas	7-1
B. Records and Logs	7-2
Appendix A. GLOSSARY	A-1
Appendix B. ADMINISTRATIVE DIVISIONS OF SOUTHEAST ASIA	B-1
Republic of Việt Nam – Geographical Listing	B-5
Republic of Việt Nam – Alphabetical Listing	B-25
Democratic Republic of Việt Nam – Geographical Listing	B-51
Democratic Republic of Việt Nam – Alphabetical Listing	B-69
Laos	B-90
Cambodia	B-91
Appendix C. LIST OF ACRONYMS	C-1
Appendix D. INDEXING GUIDE	D-1

LIST OF ILLUSTRATIONS

<u>Figure No.</u>		<u>Page</u>
1-1	Storage of Documents	1-3
1-2	FileSearch Flexowriter	1-4
1-3	FileSearch Index Card	1-4
1-4	FileSearch Record Unit	1-5
1-5	Typical Frames of FileSearch Film	1-5
1-6	File Maintenance	1-7
1-7	FileSearch Retrieval Unit	1-8
1-8	FileSearch Binary Code	1-9
1-9	Chronological Arrangement of Documents	1-11
1-10	Retrieval of Documents	1-13
1-11	MACV Input Cycle	1-16
1-12	MACV Retrieval Cycle	1-18
1-13	Flow Chart Symbols	1-21
2-1	Sample Document for Indexing	2-3
2-2	IDHS Tag Assignments	2-7
2-3	Sample Index Form (Reduced Size)	2-14
3-1	Summary Translation of Captured Documents	3-12
3-2	Indexing of Interrogation Reports	3-14
3-3	Intelligence Information Report	3-17
3-4	Agent Report	3-22
3-5	Translation of South Vietnamese Government Document	3-24
3-6	Vietnamese Telegraphic Code	3-30
3-7	Military Regions	3-38
3-8	Provinces of the Republic of Việt Nam	3-40
3-9	Việt Cộng Provinces & Military Regions	3-41
3-10	Coding Chart	3-48
4-1	Machine Room Input Flow Chart	4-4
4-2	Security Cover Sheets	4-6
4-3	Batch Cover Sheet	4-7
5-1	File Expansion Categories	5-5
5-2	Retrieval Time vs. Categories	5-6
5-3	Indexing of a File Expansion Leader	5-8
5-4	Identified Film Strip	5-9
6-1	Request Form	6-4
6-2	Request Card	6-4
6-3	Association Logic - Two Terms	6-8
6-4	OR Association Logic	6-10
6-5	Excerpt from the ISC	6-27
7-1	Statistical Log	7-4
7-2	Monthly Statistical Worksheet	7-7

CHAPTER I

THE FILESEARCH SYSTEM

A. STORAGE AND RETRIEVAL PROCESSES

The operation of FileSearch system at an installation involves three primary functions: storage of documents, file maintenance, and retrieval. The FileSearch System General Information Manual, available from FMA and its representatives, contains a general description of these processes and is recommended as supplementary reading.

1. Storage of Documents

The flow of processing documents into the system is shown in Fig. 1-1.

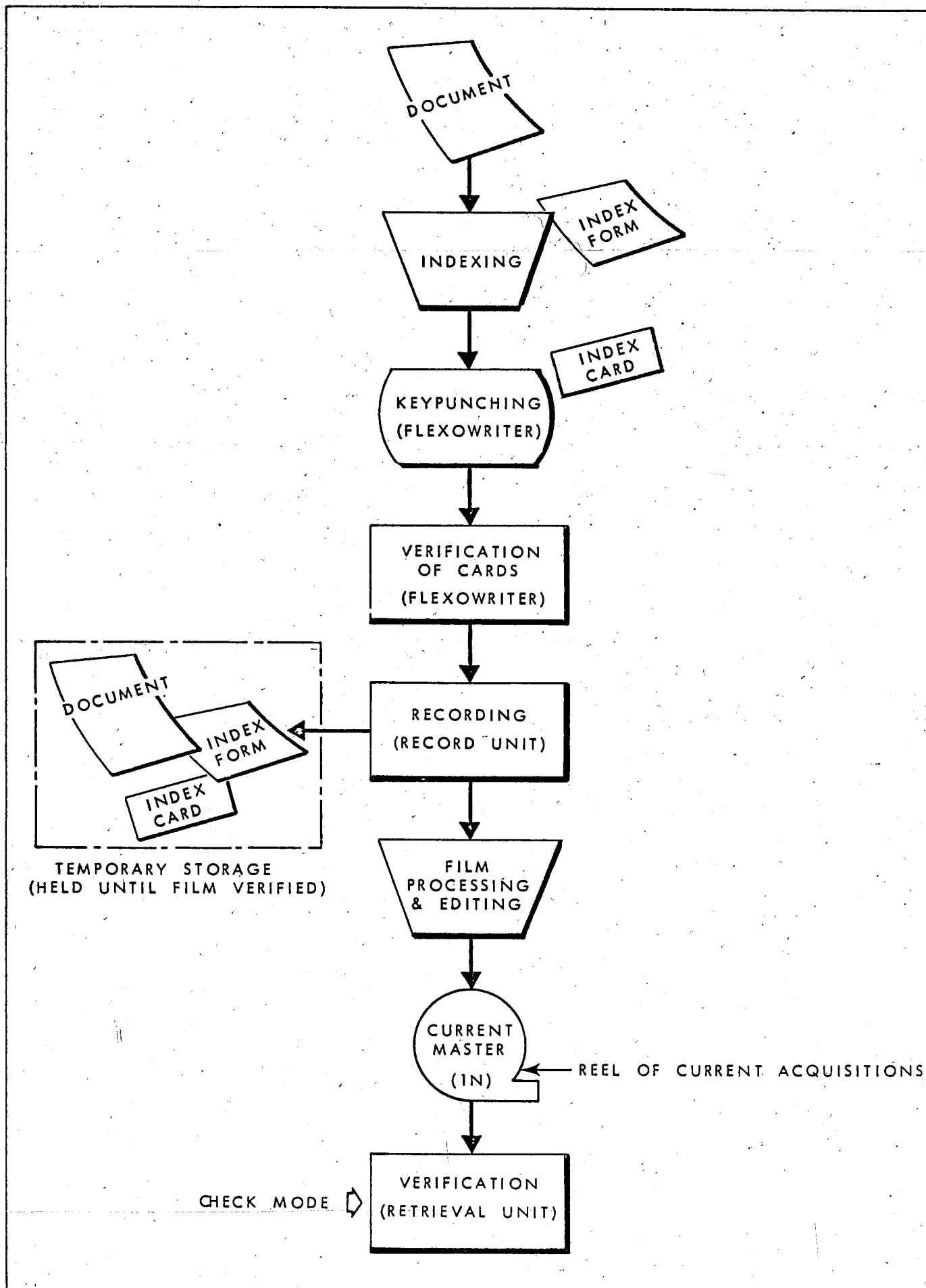
Documents are indexed by writing their descriptions or index terms on an index form. The indexing process is described in more detail in Chapter Two. The index terms, key-punched into one or more cards on the FileSearch Flexowriter (Fig. 1-2) are transcribed into machine-readable codes (Fig. 1-3). The keypunching can be verified by reading the cards on the Flexowriter and comparing the typed copy with the terms on the index form.

Next the document, index form and cards are recorded using the FileSearch Record Unit (Fig. 1-4). The first page of the document is positioned on the table of the Record Unit, and the first index card is inserted into the card reader, causing the document and code to be imaged onto 35mm microfilm (Fig. 1-5). The film is automatically advanced so that succeeding pages and index cards can be recorded. The last page to be photographed is the index form.* If there are more pages than cards, a blank card is used to trigger the camera. Recording is discussed in more detail in Chapter Four.

When the film in the Record Unit magazine is processed and edited (viewed and spliced), the master film, containing the current document acquisitions, is ready for processing into the file. It is a negative film, so called because the dark and light areas of the original appear reversed.

The documents, index forms and cards are held until the film has been verified.

* The index form can be used to monitor retrieval.



JUL 1967

Figure 1-1. STORAGE OF DOCUMENTS

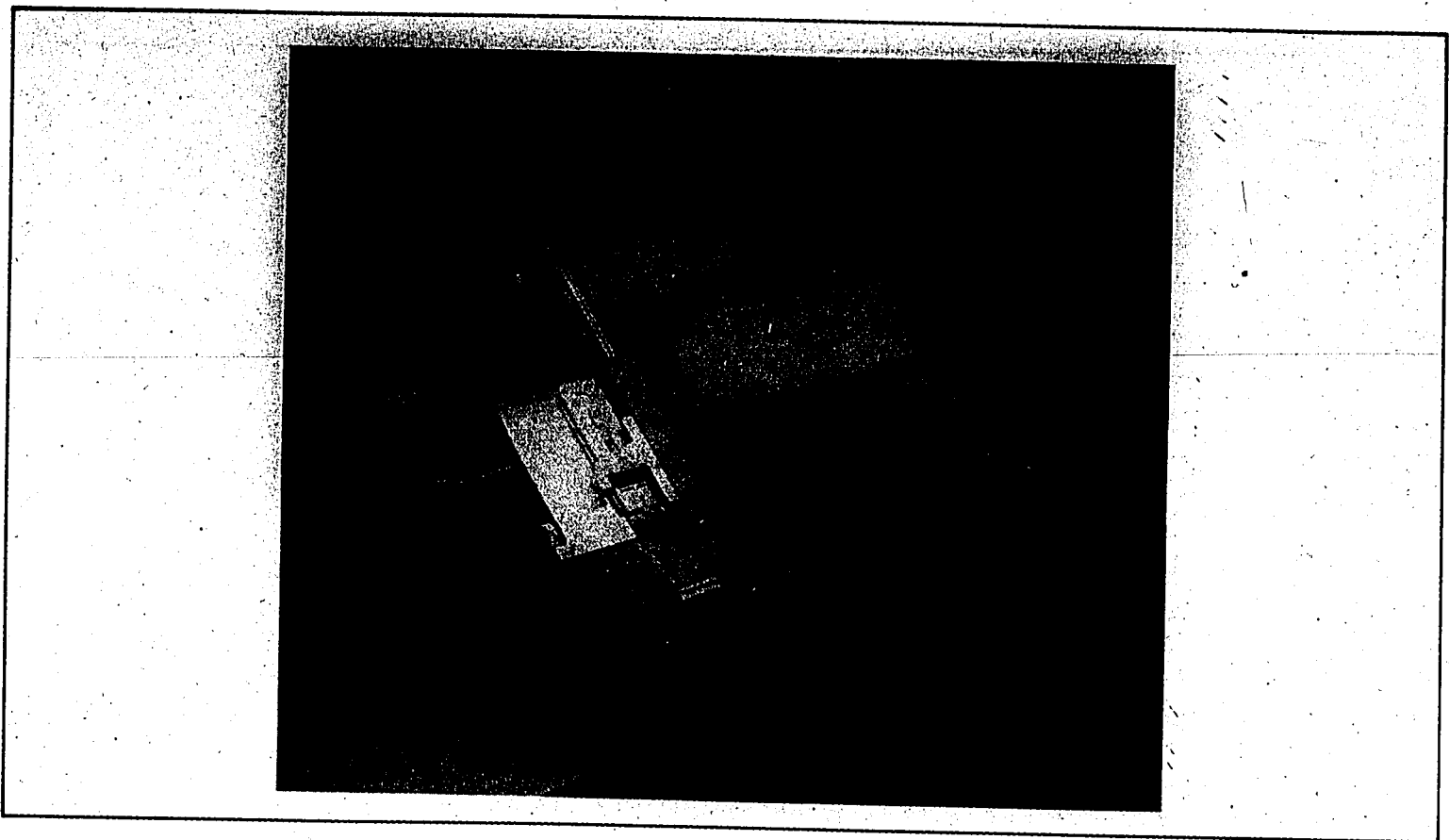


Figure 1-2. FILESEARCH FLEXOWRITER

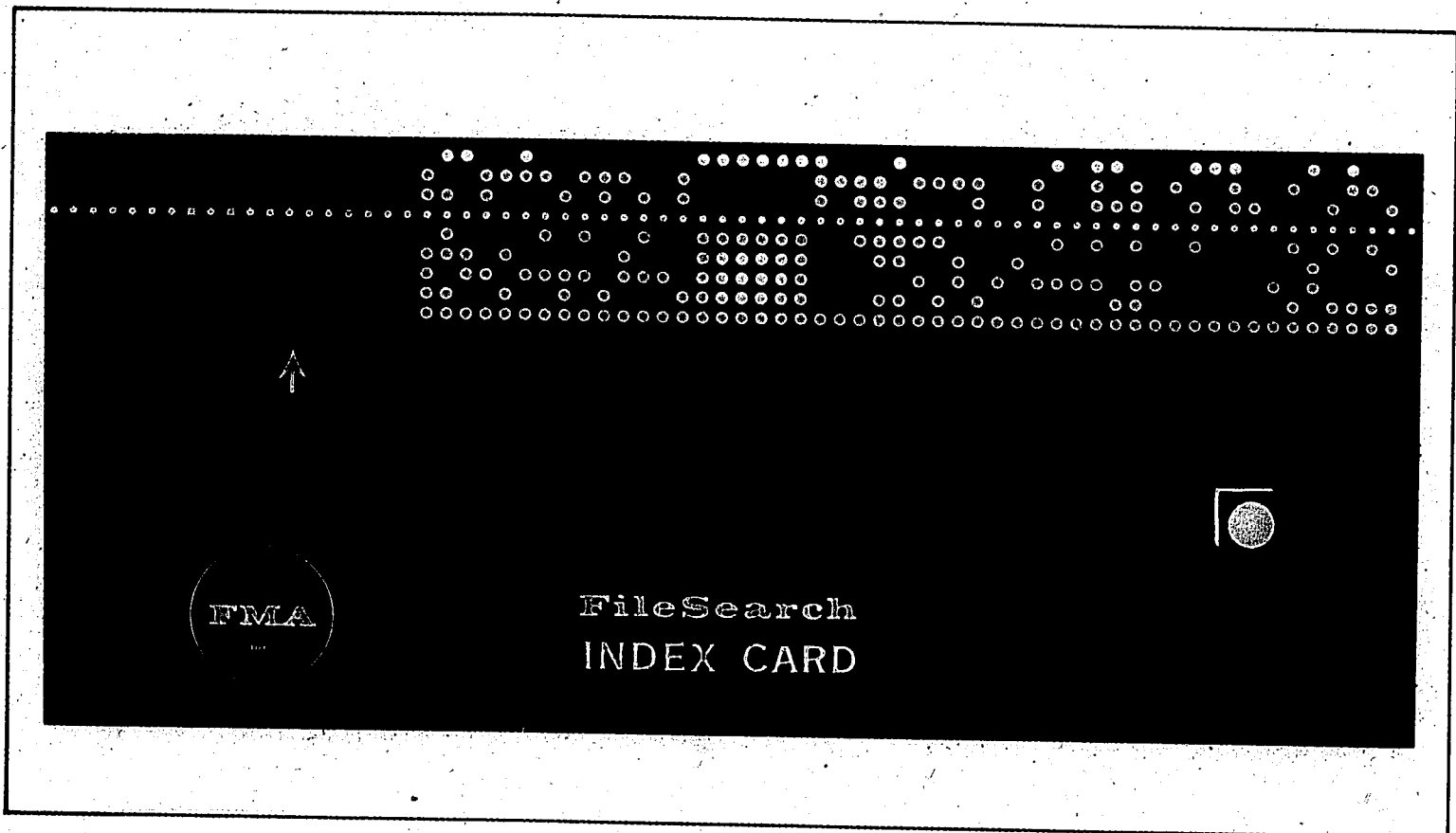


Figure 1-3. FILESEARCH INDEX CARD

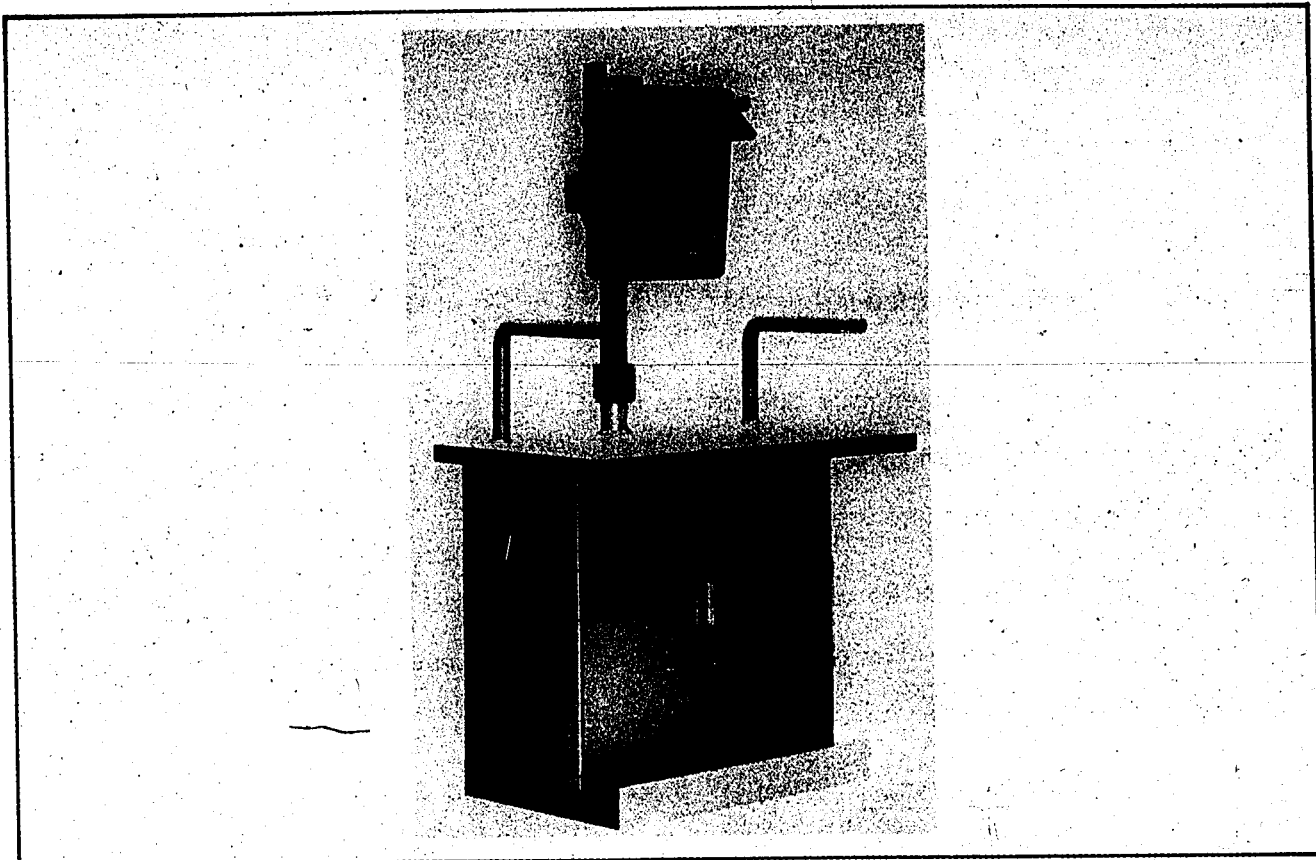


Figure 1-4. FILESEARCH RECORD UNIT

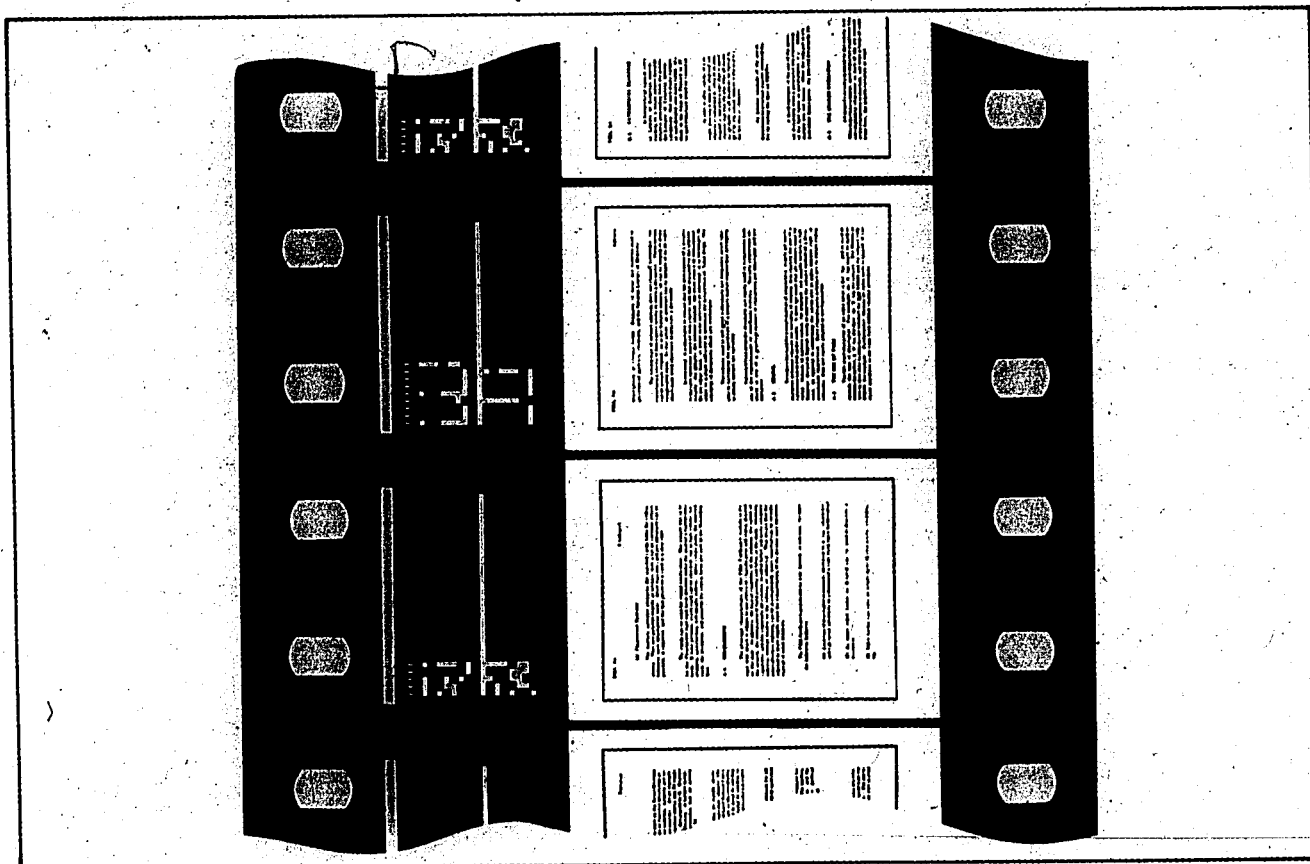


Figure 1-5. TYPICAL FRAMES OF FILESEARCH FILM

2. File Maintenance

The file maintenance operations (Fig. 1-6) include the distribution of File-Search film to other users, the formation of the working film files, and verification of the film. These procedures are discussed in more detail in Chapter Five.

The FileSearch Retrieval Unit (Fig. 1-7) is used to check the film (Check Mode) and duplicate the film either by copying every frame (Copy Mode) or by automatically selecting documents matching a request (Expand Mode).

In the Check Mode, the first page of a document is displayed on the viewer of the Retrieval Unit if the document contains an error. The error can be caused by:

- a. An improperly coded character; that is, each character must have an odd number of ones, or odd parity. (See Fig. 1-8).
- b. An index term having more than seven characters, that is, a tab error.
- c. A document whose indexing does not match the request. For example, the request used to check the master film could ask for all documents classified Secret or less which are indexed with a source, document number, date of document, date of indexing and indexing unit. Since it is assumed that all the current acquisitions should be so indexed, only error documents will be displayed on the screen.

In the Copy Mode, an exact duplicate (1:1 copy) of every frame on the master film is made. In the Expand Mode, all pages of each document matching the request are duplicated onto film. Thus the mode used to prepare film for distribution to other users will depend upon requirements.

The film duplicated from the master film is a second generation positive (2P). The code is positive and the image is the same as the original with respect to dark and light areas.

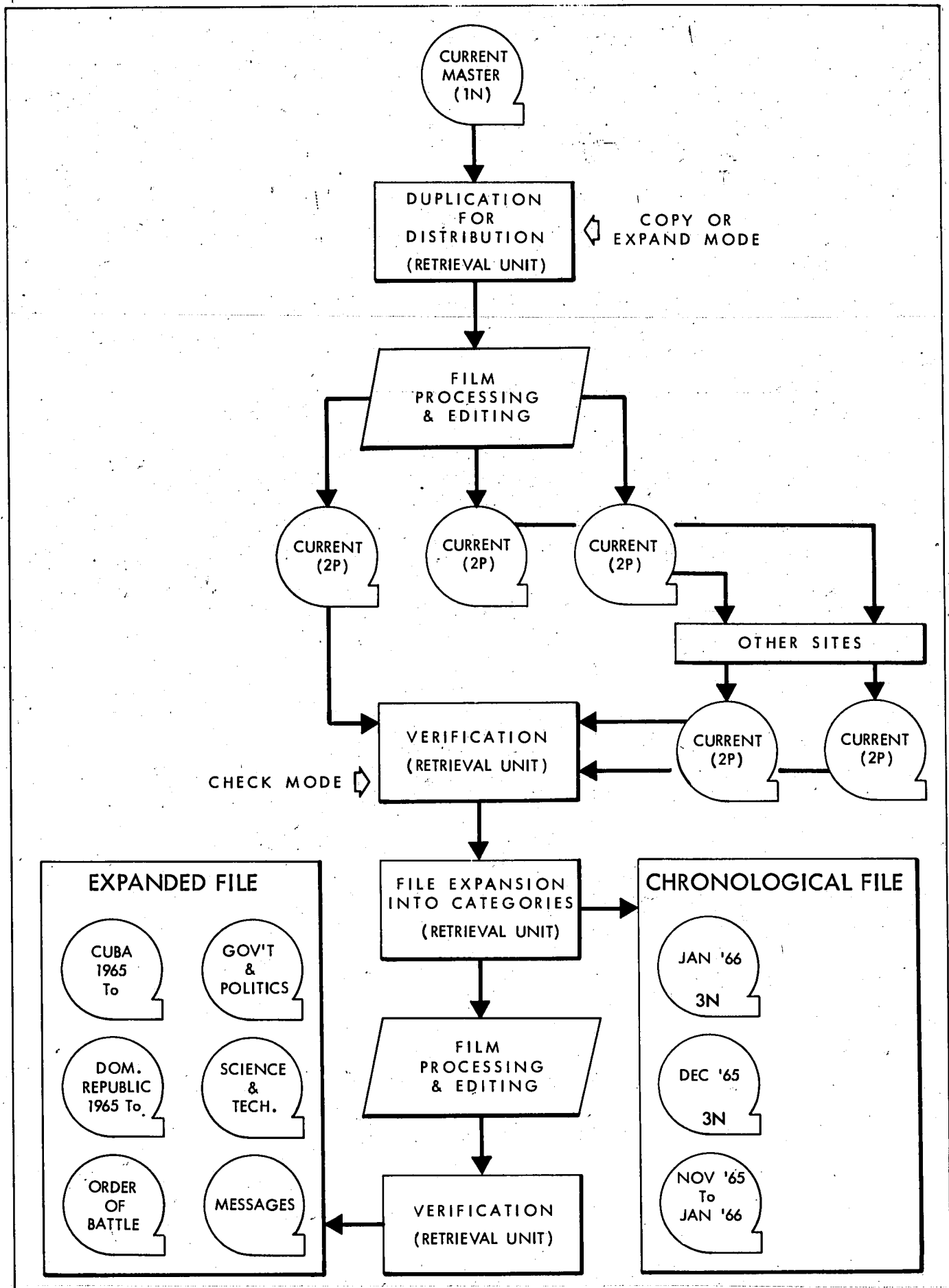


Figure 1-6. FILE MAINTENANCE

1-8

JUL-1967

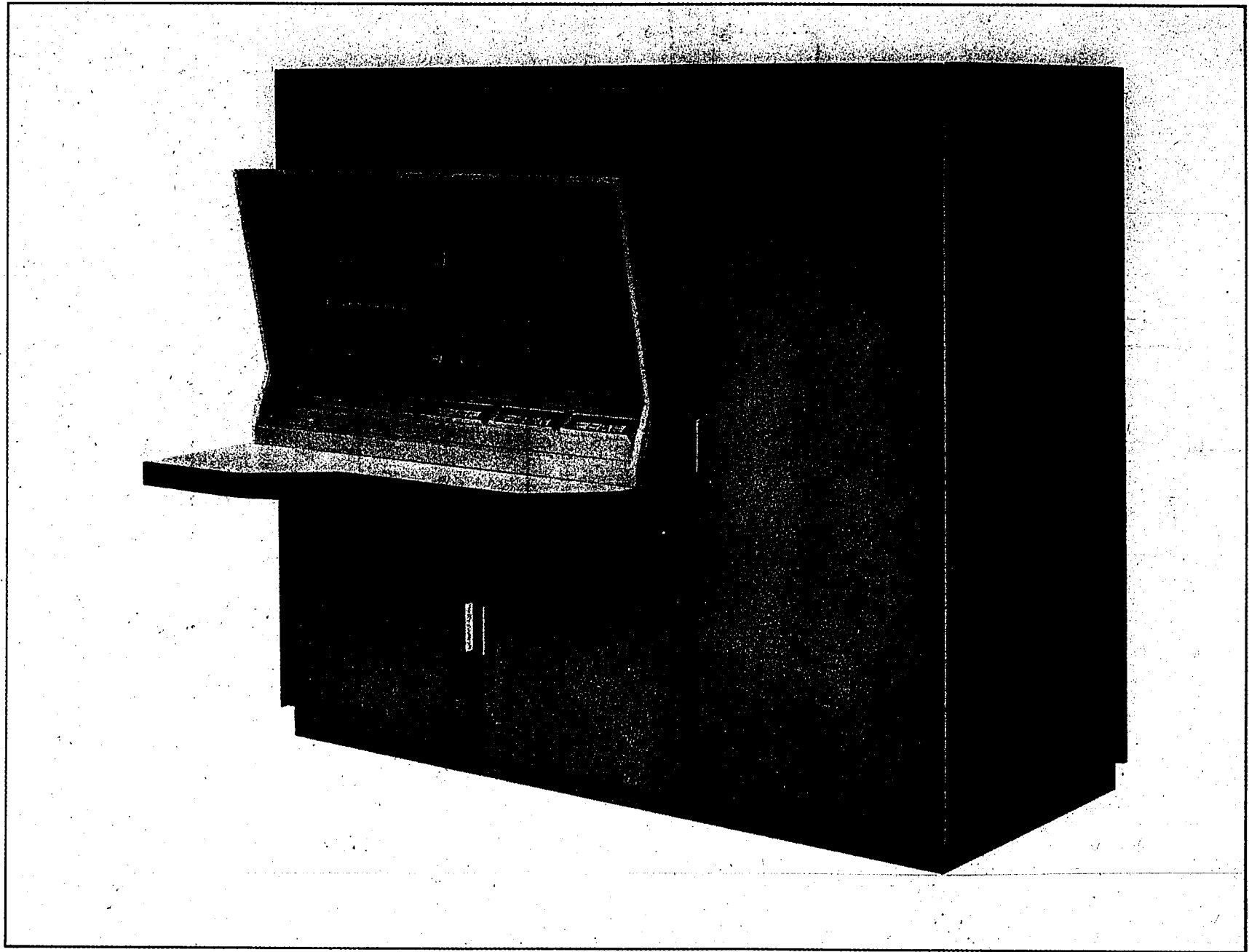


Figure 1-7. FILESEARCH RETRIEVAL UNIT

Character	Binary	Film Negative	Film Positive	Punched Data
0	000001			
1	000010			
2	0000100			
3	0000111			
4	0001000			
5	0001011			
6	0001101			
7	0001110			
8	0010000			
9	0010011			
-	0010110			
/	0011010			
.	0011111			
space	0100000			
A	0100011			
B	0100101			
C	0100110			
D	0101001			
E	0101010			
F	0101100			
G	0101111			
H	0110001			
I	0110010			
or	0110111			
J	1000011			
K	1000101			
L	1000110			
M	1001001			
N	1001010			
O	1001100			
P	1001111			
Q	1010001			
R	1010010			
-	1010100			
car. ref.	1011101			
and	1011110			
S	1100100			
T	1100111			
U	1101000			
V	1101011			
W	1101101			
X	1101110			
Y	1110000			
Z	1110011			
stop code	1110110			
*	1111001			
tab	1111100			
delete	1111111			

Figure 1-8. FILESEARCH BINARY CODE

Just as film is distributed to other FileSearch sites, film is received from other sites. This film, when combined with the positive copy of the film you have recorded, contains documents arranged in approximate chronological order. It is used to form the working files which are structured into categories so as to minimize access time to requested documents. These categories will initially reflect the same organization as your present manual files.

By requesting all documents pertaining to Cuba, for example, and searching the positive film in the Expand Mode, all documents concerning Cuba will be copied, via the 1:1 camera, onto film. This process is then repeated for other categories of the working files. The processed film is cut into strips, and each strip spliced onto the appropriate reel of the working library. The file expansion operation can be checked by searching each reel in the Check Mode with the corresponding request used for file expansion.

The term "file expansion" derives from the fact that the volume of film in the working files is increased. One document could appear on several reels such as Cuba, The Dominican Republic and Government and Politics. File expansion is analogous to cross filing duplicate documents in a manual system.

The working files contain third generation negative film. Within one reel, the documents are arranged in approximate chronological order, and latest acquisitions are the first searched during retrieval. (See Fig. 1-9.) The positive film and master negative are held in the retention files. These reels are not normally used in processing requests, but are used if it becomes necessary to reform a working reel.

JUL 1967

1-11

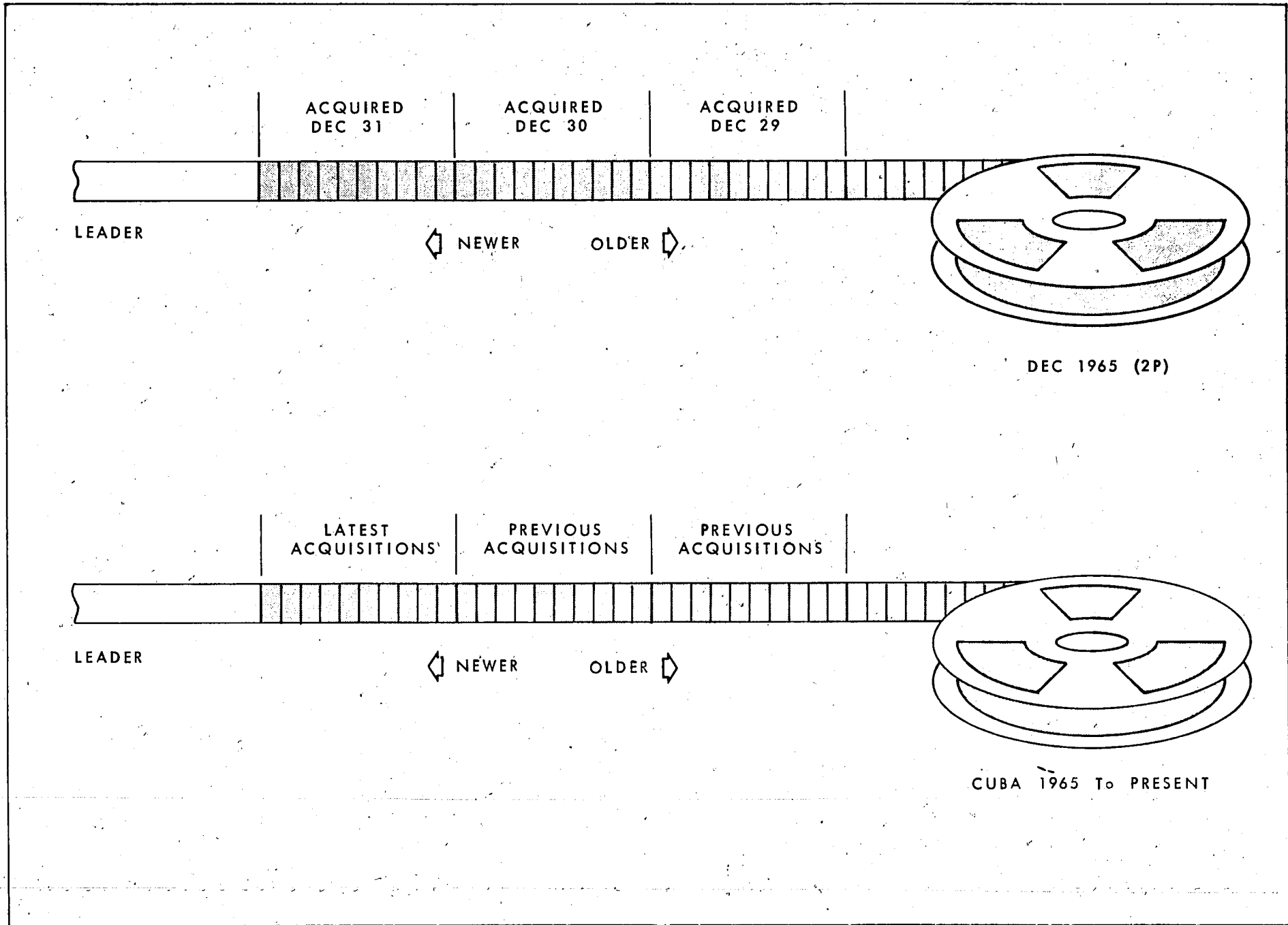


Figure 1-9. CHRONOLOGICAL ARRANGEMENT OF DOCUMENTS

3. Retrieval

The flow of processing a request for documents is shown in Fig. 1-10.

Retrieval techniques are presented in more detail in Chapter Four.

A request can be received in person, via the telephone, or in writing. It can take several forms: for example, it can ask for a specific document or a series of documents which require searching several reels; it can ask for a verbal answer (such as "No message has been received"), a hard copy of the documents, or a statistical count. Whatever its form, it must be transcribed into FileSearch language analogous to indexing, by writing the request terms on a form. The request is then entered into the Manual Entry Retrieval Console or a request card is punched, and the proper reel selected from the working files.

The operator of the Retrieval Unit mounts the reel and selects the output mode:

- a. In the Browse Mode, the first page of a document matching the request is projected on the viewing screen. By depressing the proper switch, the operator can view succeeding or previous pages, make a hard copy of any page he wishes, and continue searching or rewind. In this mode, the Unit operates as it does in the Check mode, except that matching (not erroneous) documents are recovered.
- b. In the Retrieve Mode, hard copies of each page of matching documents are automatically produced, and search is automatically continued. In this mode, the Unit operates as it does in the Expand Mode, except that hard copies (not film copies) are produced.
- c. In the Count Hits Mode, the Hit Counter displays the number of documents which match the request.

In all of these modes, the Retrieval Unit automatically rewinds the film when the reflective marker at the end of the film is sensed.

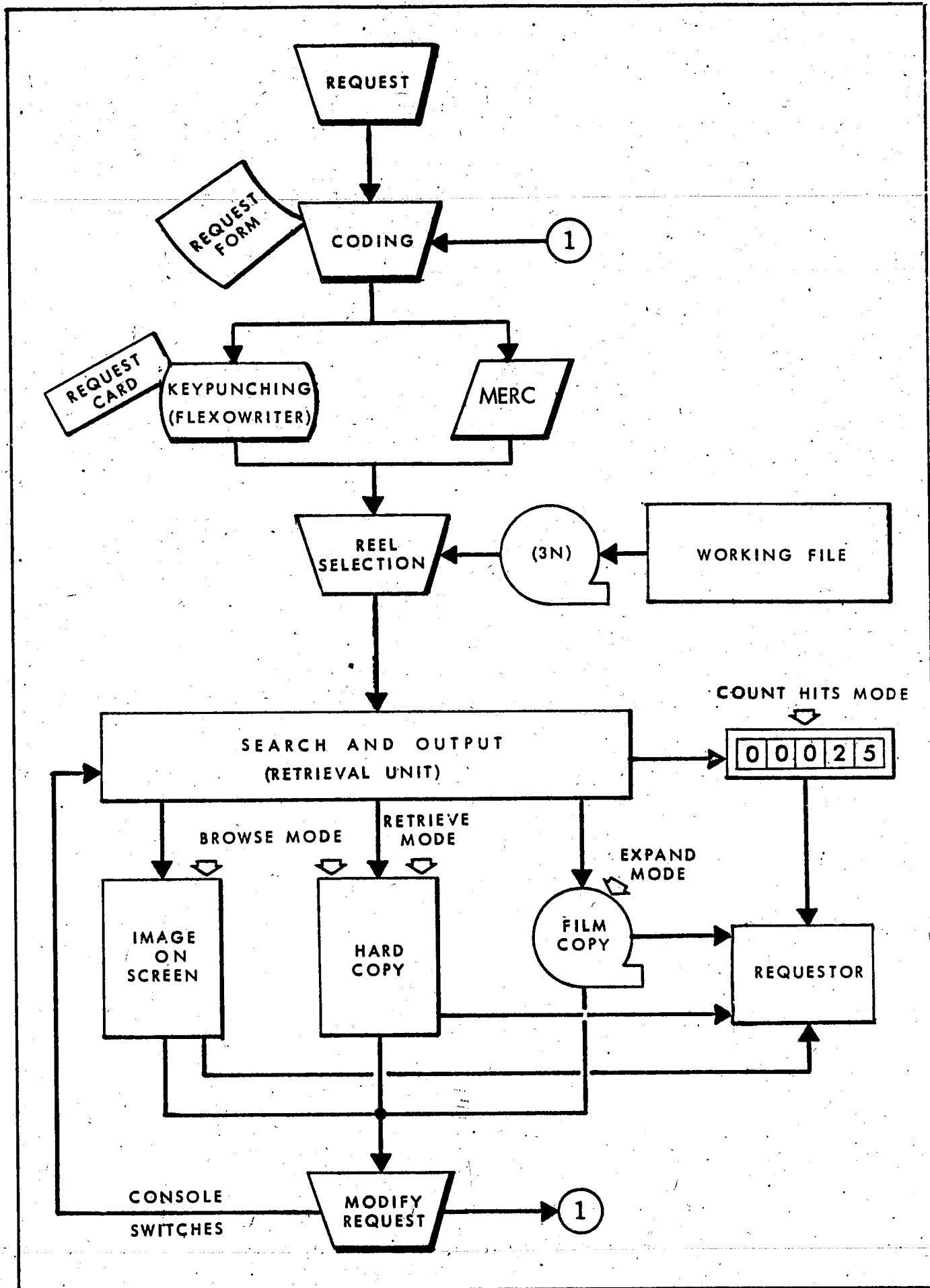


Figure 1-10. RETRIEVAL OF DOCUMENTS

Some requests may require modification in order to recover the requested documents. The request can be altered by changing switches on the console of the Unit, or by restating the request and repeating the retrieval procedures.

It is usually most desirable to search negative film; negative images of textual materials are less fatiguing to view and result in positive hard copies. This is a major reason for using negative film in the working files.

4. Control Procedures

During all of the operations involving FileSearch, system controls (e.g., verification, identification, security) are used. Some methods for applying these controls are discussed in the following chapters.

B. THE MACV SYSTEM

1. Mission

The Automated Document Storage and Retrieval System has been established as an automated archival repository for all raw intelligence documents. As such it receives these documents from the producing agencies, indexes them in depth, and records both the documents and the indexing on microfilm. This data base is then available to all intelligence organizations in Việt Nam for selective retrieval to meet their specific needs.

2. The Input Cycle

Documents are indexed as they are received by the Coding Section. They are then transferred to the Machine Room where the indexing information is keypunched on flexowriters into paper tape. As the paper tape is verified, it is also converted to edge-punched cards. These cards are recorded along with the original documents onto 35mm microfilm using the Record Unit. The paper tape is stored for future interface with digital computer systems.

3. The Retrieval Cycle

Two basic types of retrieval will be run, analyst queries and film dissemination. Figure 1-12 is a flow chart representing the complete retrieval cycle.

a. Queries. The standard output from the system will be in the form of specific requests run for individual analysts. These may be submitted by any intelligence organization, either in person, or by telephone, letter or message. Each question is coded as it is received. This will normally be done directly on the request console, although a request may be entered using an edge-punched card.

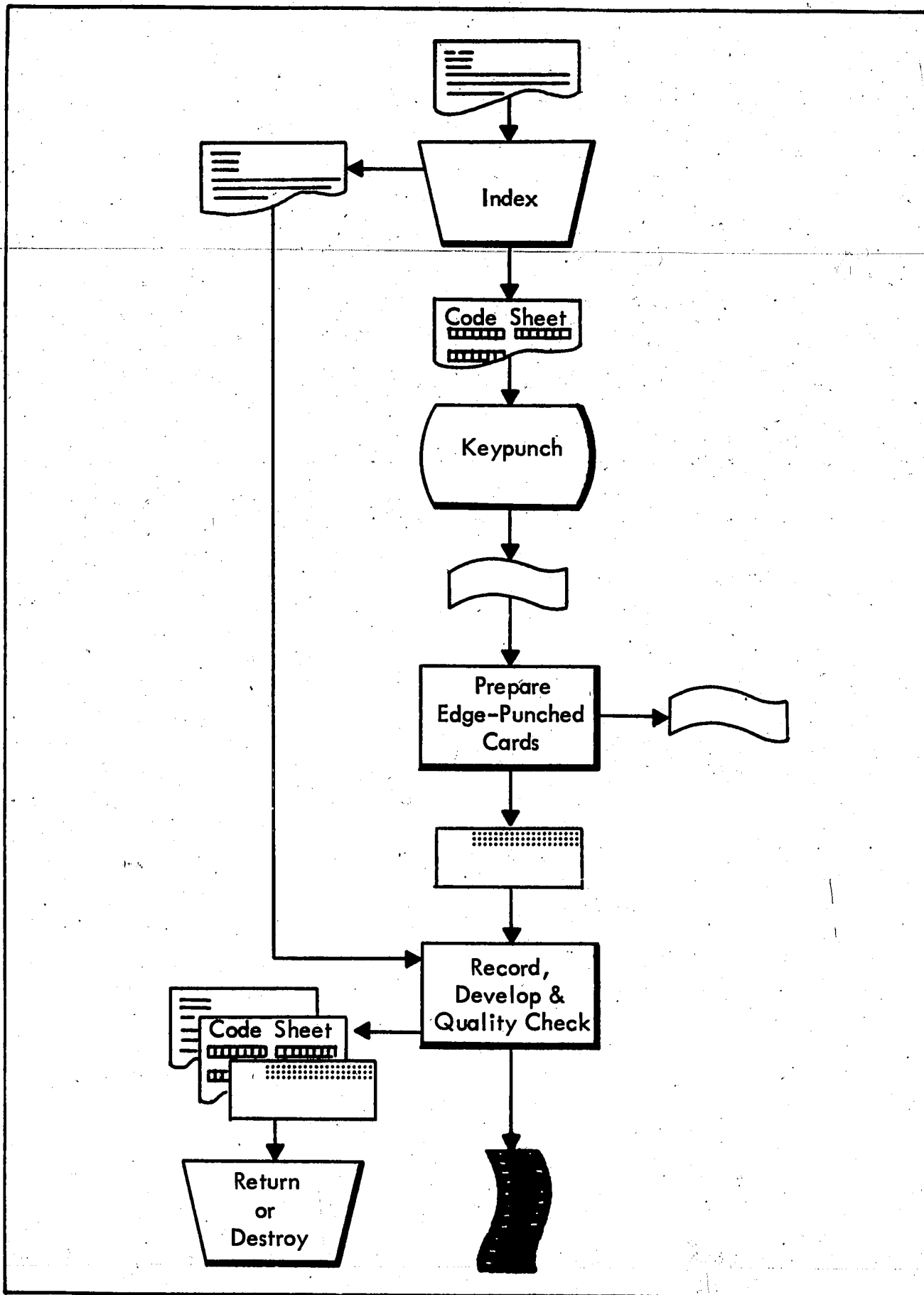


Figure 1-11. MACV Input Cycle

Requests with limited output will be run using the browse mode in which documents retrieved are displayed on the reading screen of the Retrieval Unit, and hardcopy may be made as desired. On the other hand, when extensive output is expected, the Retrieval Unit will be operated in the expand mode with the output in the form of 16mm microfilm for use on an off-line reader or reader-printer.

b. Dissemination. For installations that have microfilm readers available, routine dissemination of all documents on a particular area or subject of interest can be made at regular intervals. For this type of output a standard query will be prepared on an edge-punched card, and then run against all new film since the last dissemination was made. Output will be in the form of 16mm microfilm.

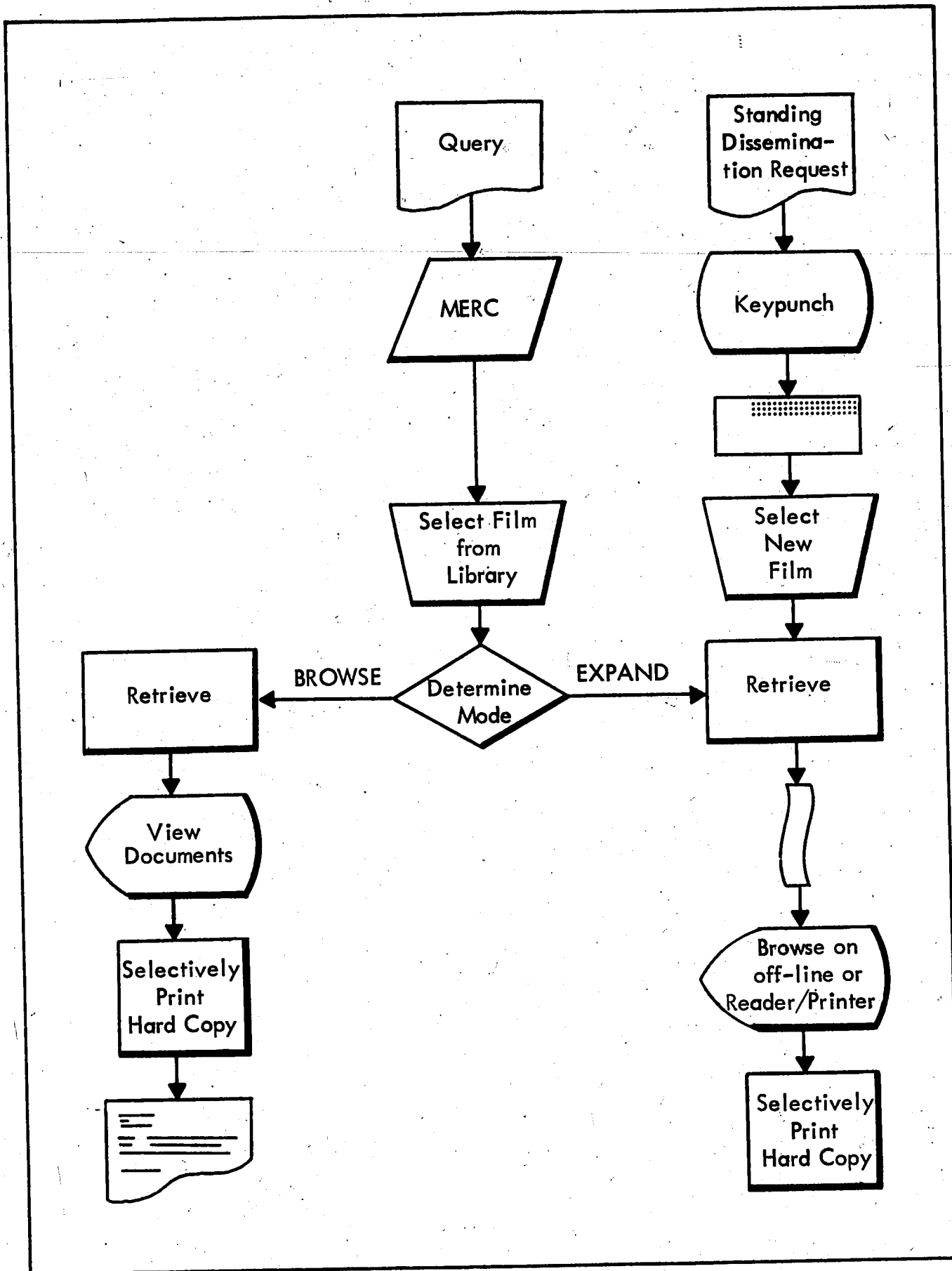


Figure 1-12. MACV Retrieval Cycle

C. RELATED PUBLICATIONS

1. The reader is referred to these FMA publications:

a. Operating Manuals

1) The FileSearch System General Information Manual - describes the storage, file maintenance and retrieval processes; includes pictures and specifications of the equipment; gives examples which do not use IDHS indexing; lists definitions of terms.

2) FileSearch Flexowriter Instructions - describes the operation of the Flexowriter; punching, reading, and duplicating cards.

3) FileSearch Record Unit Instructions for Models 122SA and 123 - describes the operation of the Record Unit; making leaders and trailers; choosing exposures; setting the frame counter; loading and unloading the magazine; correcting errors.

4) FileSearch Film Maintenance Instructions - describes the storage and processing of film; splicing; forming reels; special leaders and trailers.

5) FileSearch Retrieval Unit Instructions for Models 112SA and 113 - describes the operation of the Retrieval Unit; functions and terms related to retrieval; use of the console switches; mounting film; loading the printer; loading and unloading the file expansion transport module; procedures for interpreting and correcting indicated error or alarm conditions.

b. Application Notes

FMA has established a series of bulletins concerning various aspects of the application of the FileSearch system (Application Notes). At the present time, the following list of titles is available. Information concerning the latest additions can be obtained from FMA or its representatives.

1) Tags (ANPB65102) - describes the significance of the use of tags to identify index or request terms.

2) Variable Length Descriptors (ANPB65103) - describes the use of variable length index terms and the corresponding retrieval techniques.

3) The Uses of the Continued Function (ANPB65104) - describes the use of continued index terms and the corresponding function of the Retrieval Unit.

2. The reader is also referred to these government publications, which are referenced throughout this guide. They are available through appropriate channels from:

Defense Intelligence Agency
Washington D.C. 20301
ATTN: DIAMS-1

a. Automated Document Storage and Retrieval Guide, Intelligence Data Handling Systems (IDHS), June 1965, Defense Intelligence Agency: describes the Minicard and FileSearch systems; prescribes the standard codes and indexing rules; appends detailed indexing instructions for certain types of information; includes the Area Classification Code and Department of State Post Locator Numbers; classified Confidential. (Ref. 1)

b. The Intelligence Subject Code, latest edition, United States Intelligence Board, Committee on Documentation: describes the hierarchic classification of subjects; gives the six-digit codes; lists applicable modifiers; notes application of primary and secondary area rules. (Ref. 2)

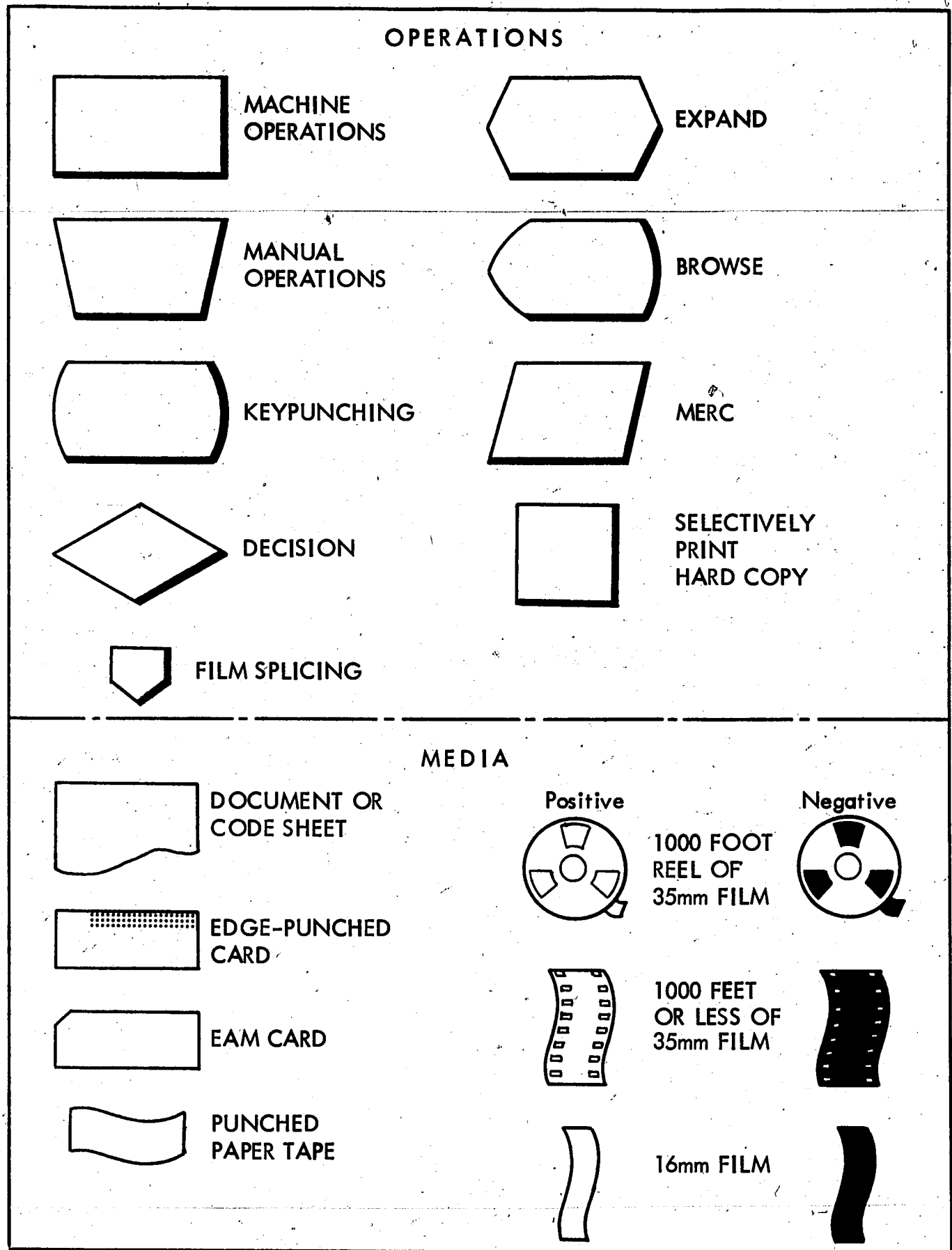


Figure 1-13. Flow Chart Symbols

CHAPTER II

INDEXING

A. GENERAL

Each document must be properly described or indexed in order for it to be properly retrieved. Indexing consists of analyzing a document to select the proper descriptors and encoding these descriptors as index terms. The guide issued by DIA (Ref. 1) contains a lucid description of the indexing process. Indexing is critical to accurate operation of the system. An incorrectly indexed document might be overlooked during search when it contains vital information, or it might be retrieved when it contains irrelevant information (a false drop).

B. INDEXING IN DEPTH

The FileSearch system permits indexing in depth; there is no restriction on the number of index terms (descriptors) which can be assigned to a document. To illustrate, a State Department publication on Cambodia (the first two pages are shown in Fig. 2-1) might be described in depth as follows.

DESCRIBING A DOCUMENT

<u>KIND OF DESCRIPTOR</u>	<u>DESCRIPTOR IN CLEAR TEXT</u>
About the document:	
Reporting Agency	Dep't. of State, Office of Media Services
Publication Number	7471
Date of Document	June, 1963
Enclosure (not shown)	Map of Southeast Asia
Security Classification	Unclassified
The document is about:	
Country	Cambodia, in Southeast Asia
City	Phnom Penh
City	Sihanoukville
Personality	Prince Norodom Sihanouk, chief of state
Installation	Khmer-American Friendship Highway
Subjects	Cambodian language and culture Constitutional form of government Agricultural economy rice crop rubber, exported Foreign aid
Countries	United States Communist China Russia

CAMBODIA

The Kingdom of Cambodia is the remaining portion of the once vast and powerful Khmer empire which, at its height some 6 or 7 centuries ago, extended across the breadth of Southeast Asia from the South China Sea, possibly as far west as the Bay of Bengal. Today Cambodia has an area of about 67,500 square miles (approximately the size of Missouri), is bounded on the south and east by the Republic of Viet-Nam, on the north by Laos and Thailand, and on the southwest by the Gulf of Siam.

Cambodia is unique in Southeast Asia in that it is relatively underpopulated and less than one-fourth of all the arable land is under cultivation. Over half of the country is covered by tropical hardwood forest. The central area is level and forms a basin for the Mekong River and the great lake of Tonle-Sap, creating ideal conditions for rice production. Near the borders of the country there are several low mountain chains. In general, the inhabited parts of the country are of a savanna-type terrain not unlike the flat inland parts of Florida. The climate is generally warm and humid.

Phnom Penh, the nation's capital and its political, economic, and transportation center, serves also as an inland port and distribution center for goods shipped overland to and from the new deepwater port of Sihanoukville on the Gulf of Siam.

2

More than 4 million of the Kingdom's 5 million people are Cambodians, or ethnic Khmers, showing, however, traces of Chinese, Malayan, and Indian physical characteristics. Minority groups consist of some 400,000 Vietnamese, 300,000 Chinese, 85,000 Chams (Moslems of early Indonesian stock), and 4,000 Europeans.

The Cambodians are a notably integrated and homogeneous people, unified by their long history and ancient culture, a common religion (Hinayana Buddhism), a single language, deep reverence for the throne, and a strong nationalism. The people speak Cambodian or Khmer, and French is the principal secondary language. ("Khmer" means "Cambodian" in the Cambodian language.)

Buddhism exerts a pervading influence and conditions the daily life of the people. There are over 2,400 Buddhist temples and monasteries in the kingdom. As in most parts of the world, many of the most common activities are carried out according to custom or ritual. The teachings of Buddhism are partly responsible for the Cambodian's tranquil and tolerant nature.

Most Cambodians are farmers and live in small villages, while the Chinese and Vietnamese are important in commercial activities. Educated Cambodians have traditionally entered government service, but an increasing number of Cambodians trained under the recently expanded educational system are now going into other fields as well. In the Cambodian social tradition manual labor of any kind, even skilled labor, was looked down upon. However, Cambodian workers hired by the American construction firm building the Phnom Penh-Sihanoukville highway proved efficient and quick to learn the use of modern tools and to operate heavy

3

Figure 2-1. SAMPLE DOCUMENT FOR INDEXING

JUL 1967

Several different kinds of descriptors are listed, some telling about the document, others telling what the document is about. One of the latter kinds is called a "subject" and denotes a topic or concept, as contrasted with names of countries, cities, people, etc.

1. Classification

Some kinds of descriptors can be classified to show hierarchic relationships by using codes. For example, Cambodia can be classified as a country in Southeast Asia by using positional notation:

Continent	Country		Continental Subdivision	
A	C	B	∅	S
(Asia)	(Cambodia)		(Southeast Asia)	

Another example encodes the relationships between the general category of commodities and rice crops:

Commodity	Agricultural Crops	Grain Crops	Rice	
6	∅ 1	1	4	5

Dates can also be arranged to show the hierarchy between year, month and day. For example, 16 September 1965 is coded:

Year		Month		Day	
6	5	∅	9	1	6

Retrieval is facilitated if certain kinds of index terms are coded to indicate hierarchical relationships. For example, the document on Cambodia could be retrieved in response to requests for Southeast Asian countries; for agricultural crops in Asian countries; for documents on Cambodia published in 1963; etc.

In the IDHS system, several kinds of descriptors are encoded. The classification codes for countries and areas are given in Appendix 12 of the DIA guide (Ref.1). Coding schemes used in indexing items such as dates and UTM coordinates are explained in this Manual, as well as the DIA guide (Ref.1).

2. Clear Text

Other kinds of index terms are names of people, places, and things. Nothing is gained by classifying these descriptors, so they are indexed in "clear text," e.g., Sihanoukville. The format of clear-text terms entered on the index form is **standardized**, however. Electronic brains are not smart enough to decide that a request for San Francisco, for example, is satisfied by a document indexed with Frisco. These formatting procedures are also given in this Manual and the DIA guide (Ref.1).

C. TAGS

1. Kind of Information

Each index term entered on the index form has a tag which indicates the kind of index. The tag is the first character of the encoded term and is used to differentiate categories of information. To illustrate, the ISC subject codes are numeric and are tagged with an "A." A document dealing with steam turbines would be indexed with the code A631025. The date of an occurrence is tagged with a "J"; an event occurring 25 October 1963 would be indexed J631025. During retrieval, tags are always compared for an exact match, so that the subject A631025 could not be confused with the date J631025.

2. Role Indicator

The document on Cambodia illustrates another use for tags. Certain kinds of subjects (such as foreign aid) involve two or more countries (such as Cambodia and the United States). Ambiguity can occur unless the indexing specifies the role played by the countries; that is, which country granted aid and which country received it. The IDHS document indexing system defines the two roles by assigning two different tags. The primary area (i.e., the granting country) is tagged with a "B" and the related area (i.e., the recipient country) is tagged with an "R." The classification code remains the same (i.e., Cambodia is indexed BACB00S or RACB00S). The use of the B- and R-tags is defined in the ISC (Ref. 2) for each of the subjects which might imply ambiguity.

3. Tag Assignments

The assignment of tags to a particular kind of descriptor is standardized throughout the DIA Automated Document Storage and Retrieval System. Most tags are used universally, while a few have been assigned for special or local use. Documents indexed with these local-use tags are not usually distributed to other commands. Fig. 2-2 shows the current tag assignments.

Character	Kind of Descriptor	Field	Use
∅			
1	(Originator's Office Symbol)	Factored	Special
2	(Addressee)	Factored	Special
3	(Administrative Subject Code)	Phrased	Special
4			
5			
6	(DGZ Designator)	Phrased	Special
7			
8			
9	Time (Date/Time)	Factored	Special
-	Continuation	Either	
/	Date of Key punching	Factored	
✓ A	Intelligence Subject Code	Phrased	
✓ B	Country/Area Code	Phrased	
✓ C	Personality	Phrased	
✓ D	Geographic Place Name	Phrased	
✓ E	Ship Name	Phrased	
✓ F	Organization	Phrased	
✓ G	Keyword	Phrased	
H			
I			
J	Date of Intelligence	Phrased	
K	Mission Number	Phrased	
✓ L	Geographic Coordinates	Phrased	
✓ M	Subject Modifier	Phrased	
N	BE Number/Chart Number	Phrased	
O			
P	Collection Project Number, SICR	Factored	Special
Q	Phrase Boundary Control		
✓ R	Related Country/Area Code	Phrased	
S	Source/Originator	Factored	
T	(Route Slip Number)	Factored	Special
U	Document Number	Factored	
V	Security classification, restriction	Factored	
W	Indexing unit, document type	Factored	
X	Report Date	Factored	
Y	Date of Capture	Factored	
Z	Batch number	Factored	Special

NOTE: Tags enclosed in parentheses will not be used at MACV.

Figure 2-2. IDHS TAG ASSIGNMENTS

D. CONTINUED DESCRIPTORS*

The maximum length of an encoded index term is seven characters, corresponding to the seven-character length of a request register. Sometimes it is necessary to continue a term into a second group of seven characters (word). For example, the tag for names of geographic places is a "D," and the indexing of the city of Copenhagen requires continuing into a second word, tagged with a dash (-):

DCOPENH -AGEN

Descriptors that do not use all available positions will terminate with a period.

DCOPENH -AGEN.

* The Minicard system and some DIA publications refer to these as extended descriptors.

E. FACTORED AND PHRASED DESCRIPTORS

1. The Factored and Phrased Field*

The document on Cambodia illustrates two classes of descriptors: those which tell about the document and those which tell what the document is about. The first class comprises the factored descriptors which must always appear first or in the factored field of a document's index. The second class comprises the phrased descriptors which must appear next, or in the phrased field of a document's index.

Each tag, identifying a type of descriptor, has been assigned to either the factored or phrased field. Although the factored descriptors must always appear before the phrased descriptors, within each field the descriptors can appear in any order—with the exception of continued descriptors. The order of the words constituting a continued index term must be preserved.

2. Phrase Boundaries

One tag, the "Q", has a special use. It does not identify a kind of descriptor; rather, it is a control code. A Q-tag, followed by a tab or carriage return, is a phrase boundary. It is used in the phrased field to group related index terms and disassociate them from unrelated terms. For example, the document dealing with a hospital in East Germany and a tracking station in France would be a false drop if retrieved in response to the request for tracking stations in East Germany. During indexing, a phrase boundary is used to separate the index terms hospital and East Germany from tracking station and France, thereby avoiding the false drop.

* The Minicard system and some DIA publications refer to these as fixed (factored) and open (phrased) field.

If a document's index terms are all related, no phrase boundaries are required in the indexing. A Q-tag is sometimes used to separate the factored field from the phrased field; this is not a requirement of the FileSearch system*, but can be done for convenience.

Factored descriptors, as contrasted with phrased descriptors, can always be associated with any of the other index terms. For example, assume the above document was originated by the Air Force on 17 November 1965. The factored terms Air Force and 17 November 1965 can be related to hospital and East Germany or tracking stations and France without causing false drops. They are analogous to a mathematical factor. For example, the factor (ab) when multiplied with (cd + ef) forms the product (abcd + abef).

* The Minicard system uses a Q-tag followed by six zeroes (a 'phrase gate'), and requires a phrase boundary following the fixed field.

F. SUBJECT MODIFIERS

Certain general topics discussed in documents can be used to modify subjects in order to more precisely describe a document's content. The document on Cambodia, for example, discusses crude rubber as the principal Cambodian export. In this case, crude rubber (a subject) is more precisely described if export of (a modifier) is included in the index. General modifying terms such as export (of), organization (of), construction (of), can be applied to a wide variety of subjects. The IDHS document indexing system uses an M-tag to identify these modifiers. Their codes and the subjects they can modify are listed in the ISC (Ref. 2). Modifiers may also be used to indicate the type of document (yearbook, agenda) or enclosures (photographs, maps).

G. ENCODING PROCEDURES

The encoding procedures used through the various commands are compatible, although some will vary slightly with individual needs. Some general rules for encoding can be stated:

1. Alphabetic codes are positioned to the left, next to the tag. Numeric codes are positioned to the right, and left-hand positions are filled with zeroes.

a. The name JONES is indexed:

C J O N E S .

b. Latitude 1°2'N is indexed:

L N 0 0 1 0 2

c. By positioning numeric codes in this manner, the position of the decimal point is implied. Thus the latitude in the above example cannot be confused with 10°20'N:

L N 0 1 0 2 0

2. Whenever an index term is continued, a dash-tag is used for the second or following words.

a. The name EISENHOWER is indexed:

C E I S E N H - O W E R .

H. THE INDEX FORM

1. Informational Entries

A sample index form is shown in Fig. 2-3. Informational entries can be made by the indexer at the foot of the form. These entries include instructions for photographing the document.

2. Index Terms

The coded index terms are entered in the seven-character format (lower portion of form). It is assumed that the first six index terms are always used, so the tags and identifications are pre-printed. Additional index terms with appropriate tags are entered as required. The Flexowriter operator punches a tab or carriage return following each word; the carriage of the Flexowriter moves accordingly.

3. Phrase Boundaries

Phrase boundaries are indexed by writing a "Q" in the tag position of a descriptor and leaving the remaining positions blank. The Flexowriter operator punches a "Q" followed by a tab or carriage return.

4. Symbols

In order to avoid confusion, certain symbols can be used on the index form:

<u>Enter</u>	<u>To Represent</u>
∅	zero
O	letter O
Œ	letter S (sometimes mistaken for a 5)
1	one
I	letter I
Z	letter Z (sometimes mistaken for a 2)

- Top Secret
- Secret
- Confidential
- Official Use Only
- Unclassified
- Other _____

No. of Doc. Pages
12

Page 1 of 1
Indexer [Signature]
Entry Date 15 Aug '65

Originator	Number	Class.	Yr. Mo. Day
\$ \$ T A T E D C R	U 0 0 7 4 7 1 C R	V 5 0 0 2 0 0 C R	X 6 3 0 6 0 0 C R
←	<i>Cambodia</i>		
W 4 L R P 0 4 C R	B A C B 0 0 \$ C R	Stop Code	
D P H N O M P C R	- E N H . 0 0 C R	D \$ I H A N O C R	- U K V I L L C R
		<i>Chief of State</i>	
C \$ I H A N O C R	- U K . 0 0 0 C R	A 1 1 5 1 0 0 C R	Stop Code
		<i>Highway Importance</i>	<i>Cambodian Culture</i>
D K H M E R A C R	- M E R E C A C R	A 5 0 2 2 1 0 C R	A 2 1 0 3 6 0 C R
<i>Constitutional Govt</i>	<i>Rice Crop</i>	<i>Map</i>	
A 1 1 3 1 0 0 C R	A 6 0 1 1 4 5 C R	M 5 0 0 2 2 0 C R	Stop Code
	<i>Cambodia</i>	<i>Rubber, Crude</i>	<i>Export of</i>
Q C R	B A C B 0 0 \$ C R	A 6 0 8 1 0 0 C R	M 6 0 0 0 6 6 C R
	<i>Foreign Aid, Loans</i>	<i>U.S.A.</i>	
Q C R	A 4 9 3 2 0 0 C R	B H U \$ 0 0 X C R	Stop Code
<i>ChiCom</i>	<i>USSR</i>	<i>Cambodia</i>	<i>Date of Key punching</i>
B A C H 0 0 C C R	B U U R 0 0 0 C R	R A C B 0 0 \$ C R	/ 6 6 1 1 2 3 C R
			Stop Code
			Stop Code

Figure 2-3. SAMPLE INDEX FORM (Reduced Size)

The indexing of the document on Cambodia is shown in Fig. 2-3. Three phrases have been used. The first phrase includes the descriptors Cambodia; Phnom Penh; Sihanoukville; Sihanouk, N.; chief of state; Hhmer-American Friendship; relative importance of highway; Cambodian culture; constitutional government; and rice crop. In order to avoid associating the modifier export with rice crop, a second phrase is used; it includes Cambodia; rubber, crude; and export of. The third phrase is needed because the subject foreign aid, loans requires that the lender (U.S.A, Red China and Russia) be tagged with a "B", the recipient (Cambodia) tagged with an "R".

It should be pointed out that the document on Cambodia is indexed in depth for illustrative purposes. This document presents a general discussion of Cambodia, serving only as a reference work. In practice, it might not be entered into the File-Search system.

CHAPTER III

INDEXING PROCEDURES

A. DESCRIBING THE DOCUMENT (FACTORED FIELD)

The first step in indexing is to describe the features that identify a specific document. These descriptors comprise the factored field. Some descriptors treated here (such as the P-tag for SICR numbers) do not truly describe the document, but are treated as factored descriptors in retrieval, and are considered here for convenience.

1. S-Tag. Originator

The S-tag will be used to identify the originator of a document. For all input except information reports (DD Form 1396) the abbreviations contained in a separate list will be used. The first character of the abbreviation indicates the nationality of the originating organization, and is taken from the following table:

Code	National Affiliation
A	South Việt Nam government, and Binational organizations
B	South Việt Nam, non-government
C	North Việt Nam government
D	North Việt Nam, non-government
E	Communist China
F	Việt Cộng
G	Union of Soviet Socialist Republics
H	Cambodia
J	Thailand
K	Korea
L	Laos
M	Malaysia
N	New Zealand
P	Phillippines
R	Australia
U	United States

For all but Việt Cộng documents this letter will be followed by an abbreviation for the specific agency that generated the document. A complete list of established S-tags has been issued separately.

Information reports carry a special coding for the originator of the report as part of the IR number. Thus, in MACV IR 6-075-5463-66, the 6-075 stands for MACV. The first four digits of an IR number will be entered in the S-tag, right justified:

S 006075

2. U-Tag. Document Number

The report number will be entered in the U-tag. For the exact procedures on indexing report numbers for specific documents see Section III-B below.

For IR's the report number will be the last six digits of the IR number. The cite number for messages will be entered in the U-tag.

The log number assigned by the Combined Document Exploitation Center, when given, will be entered in the U-tag. The month will be entered in the first two positions, and the serial number assigned within the month will be entered in U³⁻⁶. For example, log number 4-1765 would be entered

U 0 4 1 7 6 5

3. V-Tag

The six positions of the V-tag will be used independently for the following information.

a. Classification. A single numeric entry will be entered in V¹ to identify the classification of the document, using the following table:

Code	English	Vietnamese
1	Top Secret	Tối Mật
2	Secret	Mật
3	Confidential or Confidential, Modified Handling Authorized	Kín
4	For Official Use Only	
5	Unclassified	

b. Control Markings. A single alphabetic code will be entered in V² to indicate control markings appearing on the document. Codes will be taken from the following table:

Code	Control Marking
∅	No control marking
A	Warning Notice— Sensitive Sources and Methods Involved
B	Controlled Dissemination
C	No Dissemination Abroad
D	Background Use Only
	For codes for combinations of these control markings see Reference 1.

c. Release Restrictions. Enter any release restrictions appearing on the document in V³, using codes from the following table:

Code	Release Restriction
∅	No release restrictions
D	Department of Defense Only
N	No Foreign Dissemination

Documents releaseable to other Free World Military Assistance Forces in Việt Nam, but otherwise considered No Foreign Dissemination, will be coded with a "∅."

Documents not releaseable to these countries will be coded with an "N."

d. Type of Information. Enter the type of information in V⁴. Raw intelligence information will be coded with a "1"; finished intelligence will be coded with a "2."